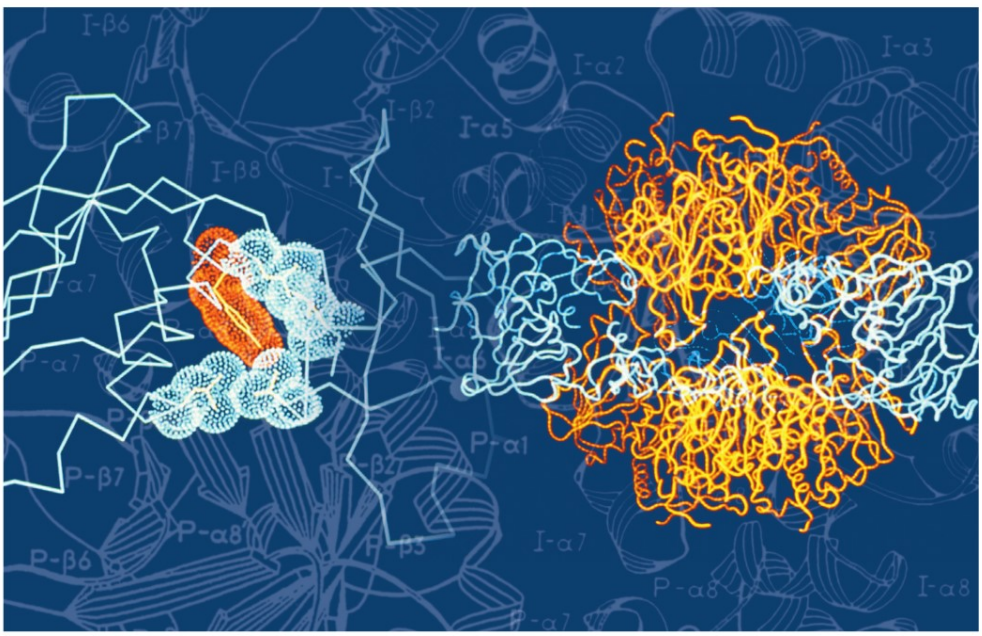
SELECTED BY GRENOBLE SCIENCES

Molecular and Cellular Enzymology

Volume I



Molecular and Cellular Enzymology

Volume I

Grenoble Sciences

Grenoble Sciences pursues a triple aim:
- to publish works responding to a clearly defined project, with no curriculum or vogue constraints,
- to guarantee the selected titles' scientific and pedagogical qualities,
- to propose books at an affordable price to the widest scope of readers.

Each project is selected with the help of anonymous referees, followed by an average one-year interaction between the authors and a Readership Committee whose members' names figure in the front pages of the book. Grenoble Sciences then signs a co-publishing agreement with the most adequate publisher.

Contact: Tél.: (33) (0)4 76 51 46 95 – Fax: (33) (0)4 76 51 45 79
E-mail: Grenoble.Sciences@ujf-grenoble.fr
Website: **http://grenoble-sciences.ujf-grenoble.fr/**

Scientific Director of Grenoble Sciences: Jean Bornarel, Professor at the Joseph Fourier University, France

Grenoble Sciences is supported by the French Ministry of Higher Education and Research and the "Région Rhône-Alpes".

"Molecular and Cellular Enzymology" is a translation of the original French book published by Grenoble Sciences in partnership with EDP Sciences (2005). The Reading Committee of the French version included the following members:

- Jean-Pierre Mazat, Professor at the Victor Ségalen University, Bordeaux 2
- Guy Lauquin, Professor at the Victor Ségalen University, Bordeaux 2
- Jacques Ricard, Professor at the Pierre et Marie Curie University, Paris 7
- Joël Lunardi, Professor at the Joseph Fourier University, Grenoble 1
- Jean Pelmont, Professor at the Joseph Fourier University, Grenoble 1
- Michel Van der Rest, Professor at the École Normale Supérieure de Lyon
- Bahram Houchmandzadeh, Senior Researcher at the CNRS, Laboratoire de Spectrométrie Physique, Grenoble

Translation performed by Philip Simister, Dr. in Biochemistry, University of Bristol, Kathleen Wood, Dr. in Physics for Life Sciences, Joseph Fourier University of Grenoble and Lisa Walter, Dr. in Pharmacology, University of Washington, Seattle.

Front cover illustration composed by Alice Giraud

Jeannine Yon-Kahn · Guy Hervé

Molecular and Cellular Enzymology

Volume I

Springer

Dr. Jeannine Yon-Kahn
Université Paris-Sud XI
UMR CNRS
Inst. Biochimie, Biophysique
91405 Orsay
Batiment 430
France
jeannine.yon-kahn@wanadoo.fr

Prof. Guy Hervé
Université Paris VI
CNRS UMR 7631
Labo. Biochimie des Signaux
96 bd. Raspail
75006 Paris
France
guy.herve@upmc.fr

Originally published in French: Jeannine Yon-Kahn et Guy Hervé, Enzymologie moléculaire et cellulaire, EDP Sciences, Collection Grenoble Sciences 2005, Tome 1, ISBN 978-2-86883-816-2

ISBN 978-3-642-01227-3 e-ISBN 978-3-642-01228-0
DOI 10.1007/978-3-642-01228-0
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2009934310

Cover design: Giraud, Alice

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

PREFACE

Enzymology lies at the boundary of several disciplines. It has benefitted from much progress: in physics, which has lead in particular to the determination of the three dimensional structures of biological macromolecules and to the study of their dynamics; in chemistry, with the development of analytical methods and processes of synthesis; and in biology, with the potential of genetic and cell biological techniques. The coincidence of these methods with those of molecular modelling today enables us to modify an enzyme in order to make it more stable or to alter its functional properties. These possibilities are of great importance not only for the precise knowledge of the mechanisms involved in catalytic events and in their regulation, but also to meet the need for applications in biotechnology. Furthermore, enzymology is of interest to diverse fields in biology and even in medicine where it contributes, amongst other things, to the development of diagnostics for, and soon the treatment of, genetic and metabolic diseases. Important developments are expected due to the entry of biology into the post-genomic era; it is now vital to process the information contained within genomes i.e. the structure and function of proteins coded for by genes. This is why we felt it necessary to bring together in a single work the ensemble of current knowledge relating to enzymology.

Molecular and Cellular Enzymology addresses not only experienced enzymologists but also all those, e.g. biologists, medical doctors, industrialists, who are confronted by enzymological problems during their fundamental or applied research. Equally, it addresses students who will find herein the basics, as well as the most recent developments in the subject. We hope that this work will be able to stimulate in these readers an interest in this field in which an urgent need for training exists in order to meet both the requirements of research and industrial endeavours. This book consists of several levels: practical aspects and elementary explanations are given for the benefit of non-specialists' understanding; specialists will also find topics more deeply expounded with the principal bibliographic references cited. The bibliography, however, is not exhaustive; the choice includes general books and review articles as well as some specialised articles. In order to facilitate the task of students, two typographies have been adopted. The main text corresponds to basic knowledge, whereas text in a smaller font, indented and indicated by ◤◢, provides more specialised information.

Part I describes the thermodynamics of enzymatic reactions. At the very beginning, in Chap. 1, there is a succinct reminder of the laws of thermodynamics at equilibrium, the experimental methods for determining the energetic parameters of reactions, as well as a study of coupled reactions. Chapter 2 treats protein-ligand association equilibria and their experimental study. It concerns as much enzymologists as biologists confronted with the problem of binding molecules to receptors. The third chapter is devoted to the study of living systems, as open systems far from equilibrium.

Part II concerns kinetic studies of enzymatic reactions in solution. Chapter 4 is a reminder of those elements of chemical kinetics indispensable for the understanding of enzyme kinetics. In Chap. 5, all aspects of the kinetics of enzyme reactions possessing Michaelian behaviour are worked through. The experimental methods for studying these reactions, along with procedures for data processing and interpretation, are presented in Chap. 6.

Part III covers the formation and structure of enzyme active sites. A discussion about the origin of enzymatic function is included (Chap. 7), as well as a presentation of the formation of the functional structure of enzymes (Chap. 8). The diverse approaches used to determine the topology of enzyme active sites are the subject of Chap. 9.

Part IV analyses catalytic function, detailing the mechanisms brought into play in the formation of enzyme-substrate complexes (Chap. 10), the principal catalytic mechanisms and their involvement in enzyme catalysis (Chap. 11). To illustrate these as an ensemble, some specific enzymatic systems are studied in Chap. 12, which represent diverse types of catalytic mechanism. They were chosen based on structural knowledge of the enzyme molecule and on the extent to which the relationship between the structure and functional properties has been determined.

Part V explores the regulation of enzymatic activity. It contains an analysis of the non-covalent regulation of allosteric enzymes, the kinetics of which are treated in Chap. 13 and their covalent regulation forms the focus in Chap. 14. Multi-functional enzymes and multi-enzymatic systems are discussed in Chap. 15.

Part VI is devoted to enzymology in a structured environment, whether it be artificially immobilised enzymes or enzymes linked to well-defined cellular structures. The theory of metabolic pathway control is also included.

This work thus brings together diverse aspects of enzymology, which only exist in separate books or articles. For practical reasons we have had no hesitation in recalling certain basic concepts in the different parts, so that the reader may –thanks to a detailed Index– consult the topic of interest without being obliged to study the preceding chapters. In a similar vain, the indexing will enable precise points of interest to be readily looked up.

ACKNOWLEDGEMENTS

The authors would like to thank deeply Drs Florence LEDERER, Charis GHELIS and Michel DESMADRIL for their careful reading of certain parts of this work and for their constructive suggestions, and in particular Dr Nicolas KELLERSHOHN who took great care in checking all the equations. These thanks also go to Profs Alain COZZONE, Bernard BADET, Christian CAMBILLAU and Jacques HAIECH for the documentation that they provided. The authors thank equally the members of the Reading Committee, Profs G. LAUQUIN, J.P. MAZAT, B. HOUCHMANDZADEH, M. VAN DER REST, J. LUNARDI, J. PELMONT and J. RICARD for having carefully read the manuscript and for communicating their constructive remarks and suggestions.

The authors wish also to thank Philip SIMISTER, Lisa WALTER and Katy WOOD for their contributions to the translation of the French version of this book and to Julie RIDARD and Sylvie BORDAGE from Grenoble Sciences for their skilful preparation and shaping of the two versions.

Jeannine Yon-Kahn

September 2009 Guy Hervé

TABLE OF CONTENTS

VOLUME I

VOLUME II

PART IV – THE CATALYTIC FUNCTION

PART V – REGULATION OF ENZYME ACTIVITY

PART VI – ENZYMOLOGY IN STRUCTURED ENVIRONMENT

GENERAL INTRODUCTION

THE IMPORTANCE OF PROTEINS IN THE STRUCTURE AND ENERGETICS OF THE CELL

The importance of proteins in living organisms embraces two aspects: structural and functional. From a structural viewpoint, the role of proteins in cellular morphogenesis is fundamental. Indeed, morphogenesis is a process of self-organisation involving self-assembly mechanisms whereby organelles, organs and even whole organisms develop in time and space as a function of genetic information. Such a process is only possible because **one-dimensional information encoded by DNA is translated into three-dimensional information in proteins**. In general, proteins fold spontaneously in their biological environment. In order for cytosolic proteins to attain an active structure, water plays a definitive thermodynamic role. As for membrane proteins, their structure depends on a multiphasic environment. The hydrophobic effect is the driving force in the formation of active protein structure. The stability of protein molecules is maintained principally by hydrogen bonding and hydrophobic interactions, which play a crucial role in the formation of the tertiary structure of soluble proteins and oligomeric structures. An additional stabilising effect is contributed by the association of subunits. This largely involves hydrophobic interactions, but also salt bridges, which can form between subunits and are stabilised by an apolar environment (see GHÉLIS & YON, *Protein Folding*). A higher level of complexity is encountered in multienzyme complexes, in which non-identical subunits bearing different enzyme activities may associate. One significant example of this type of organisation is given by fatty acid synthase, a complex of 7 associated enzymes, in which several levels of structural organisation occur. In yeast, this complex is composed of 2 polypeptide chains and has a molecular weight of 2 300 kiloDaltons (kDa). Chain A (185 kDa) possesses 3 of its enzyme activities: acyl carrier protein (ACP), β-ketoacetyl reductase and the condensing enzyme. Chain B (175 kDa) contains the four remaining enzyme activities: acetyl transacylase, malonyl transacylase, β-hydroxy-acetyl dehydratase and enoyl reductase. Macromolecular assemblies can be even more elaborate heterocomplexes comprising proteins, membranes, ribosomes and viruses.

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_1, © Springer-Verlag Berlin Heidelberg 2010

Proteins have very varied molecular masses and, therefore, sizes. Table 1 lists examples of the dimensions of some molecules and cellular components. We might ask whether, during the morphogenesis of higher order structures such as organelles and mitochondria, the same self-assembly processes take place. According to certain authors, in particular LEHNINGER, self-assembly alone cannot fully explain morphogenesis; certain pre-formed elements acting as scaffolds must be present. **Therefore, proteins are the basis of all cellular organisation since they translate one-dimensional information into a three-dimensional structure either in an aqueous environment or in the multiphasic environment of the biological membrane.**

Table 1 Approximate dimensions and weights of selected biomolecules and cellular components *(from LEHNINGER, 1972)*

	Dimension (Å)	*Weight (Daltons)*
Alanine	5	89
Glucose	7	180
Phospholipid	35	750
Myoglobin	36	16 900
Haemoglobin	68	65 000
Myosin	1 600	470 000
Glutamate dehydrogenase	130	1 000 000
E. coli ribosome	180	2 800 000
E. coli bacteriophage X174	250	6 200 000
Tobacco mosaic virus	3 000	40 000 000
Liver mitochondrion	15 000	1×10^{-12} grams
E. coli cell	20 000	2×10^{-12} grams
Spinach chloroplast	80 000	1.3×10^{-10} grams
Liver cell	200 000	2×10^{-9} grams

A great number of proteins exist, many of which are enzymes. A cell such as *E. coli* contains around 3 000 different proteins and 1 000 distinct nucleic acids. Plants contain even greater numbers of each and higher organisms yet more. In humans there are about 5×10^6 different proteins. The total number of proteins present in all living species can be estimated to be roughly 10^{11}. We thus find ourselves in the presence of a huge diversity of molecular species. This diversity can be simplified in two ways: by either functional or structural analogy. Discarding all analogous proteins that have the same function in different species and organisms, for example the diverse range of cytochromes c, the number then becomes approximately 10^5. If we ignore functional differences and consider only structurally analogous proteins, the number diminishes to a few hundred. Both of these methods of reduction, that is, by functional or structural analogy, correspond to two different aspects of evolution. The former corresponds to the evolution of species; the latter to the evolution of proteins, or in other words, differentiation.

In terms of function, proteins can support a range of activities depending on their nature and their degree of complexity. Thus, the basic function of all monomeric or oligomeric proteins is **the specific or selective binding of one or more ligands**. This mutual recognition involves a non-covalent interaction between the protein and its ligand, which may be a small molecule or another macromolecule. These interactions play a very important role in many biological processes, including sensory perception. Apart from this primary role, proteins fulfil many functions, for example:

- catalysis, in the case of enzymes,
- motility, exemplified by contractile proteins,
- transport, as is the case for haemoglobin and membrane transporters,
- photoreception, in the case of proteins involved in vision (rhodopsin).

Among these, we are interested principally in functional proteins and more specifically, enzymes, which form the subject of this work.

HISTORICAL VIEW OF THE DEVELOPMENT OF ENZYMOLOGY

The origin of Enzymology coincides with the origin of Biochemistry. In actual fact, biochemistry as a discipline really crystallised around enzymology. However, even though both their roots can be traced back to the beginning of the 19th century, biochemistry and enzymology only began to overlap after much development, and for a long while came up against the ideas of the vitalists. In fact, the existence of enzymatic activity had been known for nearly two centuries. But only at the beginning of the 20th century, after several attempts, was a quantitative theory of enzyme catalysis successfully developed; the theory was first established for a particular case: the mechanism of action of invertin (now called invertase). The law governing the rate of enzymatic reactions developed by Victor HENRI, then MICHAELIS and MENTEN, still remains valid today in its phenomenological form. Later developments represented increasingly wider generalisations of their initial hypothesis. In fact, these first kinetic and thermodynamic approaches remained purely phenomenological for a long time. Structural studies, which have enabled understanding of the relationship between structure and function, were only developed very much later. Nowadays, although enzyme catalysis may be considered as a particular case of chemical catalysis, we should remind ourselves that at one time, when it was not known how to obtain a purified enzyme, many scientists treated enzymatic activity as a manifestation of the vital force.

The first observations were carried out on vegetable and animal cell-free extracts. Thus, in the 18th century SPALLANZANI (1783) discovered the action of gastric juices on the liquefaction of meat. A little later, PLANCHE (1810–1820) showed that a root extract caused the dye gaiacol to turn blue in colour and named the agent responsible for this action cyanogen, although GAY-LUSSAC had already described

C_2N_2. In 1830, ROBINET and BOUTRON-CHALARD succeeded in the hydrolysis of amygdalin using an extract of bitter almonds, which, they suggested, contained an active principle. LIEBIG and WÖLHER (1837), and later ROBIQUET (1838), called this principle emulsin. A certain number of enzymes, termed "ferments" or "diastases", were then identified: salivary diastase, or ptyalin (LEUCHS, 1831), the diastase from malt, or maltase (PAYEN & PERSOZ, 1833), sinigrase (FAURÉ, 1835) and pepsin (SCHWANN, 1836). In parallel to the discovery of these diverse enzymes, the synthesis of urea by WÖLHER in 1828 showed that biochemistry was a branch of organic chemistry. In spite of this, following on from positivists certain authors, including LITTRÉ, maintained that there was no chemistry of the living but only of dead, organic substances. Thus, in the second half of the 19th century, the separation arose between chemistry and biology. Dead, organic substances belonged to the former; to the latter: living, organic substances.

With the work of Claude BERNARD, it became more and more clear that catalysts played a role in metabolic processes. In 1848, BERNARD discovered the role of pancreatic juices and described the action of an albuminous substance, pancreatin. Similarly, in the breakdown of glycogen into glucose, he invoked the activity of a diastase. Claude BERNARD accepted therefore the existence of "soluble ferments", most likely proteins catalysing breakdown reactions. Thus, he wrote in 1878 in *Leçons sur les phénomènes de la vie communs aux animaux et aux végétaux* (*Lessons on the phenomena of life common to animals and plants*): "The generality of diastasic action makes for the archetypal chemical process of living beings *par excellence*." However, he did not explain the syntheses. During the same period some new enzymes were identified and purification trials were attempted. In 1862, DANILEWSKI succeeded in separating trypsin from pancreatic amylase by adsorption of the latter to collodion. BÉCHAMP, in 1864, considered that the transformation of saccharose into glucose involved an enzyme, zymase. Not long after (1886), Raphaël DUBOIS discovered luciferase, the enzyme responsible for bioluminescence. Later, different authors, in particular Gabriel BERTRAND (1896), discovered several oxidative enzymes; G. BERTRAND isolated laccase and showed its specificity for polyphenols having their hydroxyl group in ortho or in para, but not in meta. Laccase is a metalloprotein that contains Mn^{++}, which led BERTRAND to introduce the concept of cofactor. Thus, from this period on, several enzymes started to be identified.

The diverse attempts to explain enzymatic activity could be grouped into two contrasting views. In the first, enzymatic activity was reduced to a simple chemical action, a hypothesis already put forward by BERZELIUS, who wanted to include it in ordinary chemical processes. In the second, enzymatic activity remained a property of living matter, a manifestation of the "vital force". The controversy between PASTEUR and LIEBIG, which sparked in 1870, illustrates these two attitudes. LIEBIG represented the purely chemical theory of enzymatic processes, whereas PASTEUR recognised and thought to have demonstrated that yeast must be living in order to ensure alcoholic fermentation (1871). PASTEUR established, therefore, a distinction

between the non-organised "ferment" like rennin or diastase, and an organised "ferment" like yeast and bacteria, which can lead to lactic acid production. He considered cellular structures indispensable for this action, which thus gave rise to the distinction between an essentially organised "ferment" and a non-organised enzyme. The term *enzyme*, from the Greek *zume* (zyme) meaning "leaven", was introduced by KÜHNE in 1878; enzymes are the principle found in leaven (*en zume*). The PASTEUR-LIEBIG controversy ended in 1897, when BÜCHNER showed that a yeast extract completely lacking cells could just as well stimulate alcoholic fermentation. All the facts provided by PASTEUR were correct, though incorrectly interpreted, and LIEBIG's hypothesis proved to be right. In the same era, in 1898, HILL discovered that maltase could catalyse a reversible reaction, i.e. **an enzyme catalyses equally well a reaction directed towards synthesis as towards degradation**.

In the second half of the 19th century tests to quantify enzymatic activity coincided with the emergence of physical chemistry and above all with chemical kinetics and and thermodynamics, with GUDBERG and WAAGE, VAN T'HOFF and ARRHENIUS. The study of reaction rates formed the rational basis of chemical kinetics and progressively of enzyme kinetics. The path of thought traced by BERZELIUS led O'SULLIVAN and TOMPSON (1880) to study the action of yeast saccharase by quantitative methods. This enzyme converts saccharose into glucose and fructose (laevulose), which is manifested by an inversion of the deviation of plane-polarised light. Digestion of saccharose, a dextrorotatory molecule, gives rise to one molecule of glucose and one molecule of fructose, both of which are laevorotatory. Thus, the authors were able to follow quantitatively the progress of the reaction with the help of a polarimeter. To explain this process, they proposed a first-order law in which the rate of appearance of the hydrolysis products is expressed by the relation:

$$dx/dt = K(a - x)$$

where a represents the total concentration of substrate, x the total concentration of the reaction products that appeared as a function of time, and K, a constant. Integrating this equation gives the following:

$$Kt = \ln a/(a - x)$$

which is characteristic of first-order reactions. **This study probably represents the first truly quantitative approach to the kinetics of enzyme reactions.**

◤ In 1898, E. DUCLAUX showed that even though this relation holds true for a predetermined substrate concentration, there is, however, no proportionality between the reaction rate and the substrate concentration. In other words, the constant K varies with the substrate concentration. DUCLAUX noticed that, conversely, for a very short period at the start of the reaction the quantity of inverted sugar was proportional to time. He suggested that the reaction products acted to slow down the reaction according to the equation:

$$dx/dt = K - K_1 a/x$$

so: $$t = \frac{a}{K_1} \ln \frac{K_a}{K_a - K_1 x}$$

The phenomenon of inhibition by the reaction products had previously been pointed out by TAMMANN.

In reality, neither first-order kinetics nor zero-order kinetics with the inhibitory activity of the hydrolysis products sufficed to explain entirely the experimental facts. In particular, as already indicated by TAMMANN, none of the previous expressions took into account the influence of the enzyme concentration on the reaction rate. Therefore, a more suitable schema remained to be found that would be capable of integrating all the known experimental results, i.e.:

- the logarithmic profile of the kinetics,
- the inhibitory activity of the reaction products,
- the influence of enzyme concentration,
- the influence of substrate concentration. ◢

The basis for the theory of enzymatic reactions rests on the temporary formation of an intermediate complex between the enzyme and substrate. This idea was found to be closely related to the concept of enzyme specificity, the origin of which may be attributed to an observation by PASTEUR, who showed that during the fermentation of DL-tartrate only the D isomer was destroyed. But it was in fact the convergence of early enzymology with organic chemistry that led Emil FISCHER to interpret specificity in terms of molecular structure, at a time when only the structure of the substrate could have been known. In 1894, E. FISCHER carried out a series of experiments proving the influence of the substrate's stereochemical configuration on the enzymatic activity. Furthermore, he showed that the stereospecificity even extends to the inhibition by compounds whose structures are analogous to that of the substrate, which was confirmed by ARMSTRONG (1904) and Victor HENRI (1905). FISCHER thus concluded that a temporary association formed between the enzyme and its substrate; the metaphor that he gave was that of a "*key in a lock*", which had a lasting impact on the notion of enzyme catalysis. The intermediate complex breaks down afterwards, regenerating the enzyme in its initial form –an idea previously put forward by WURTZ in 1881.

◤ The forerunners to V. HENRI, such as A. BROWN (1902), followed by H. BROWN and GLENDINNING (1902), diverged from FISCHER's hypothesis. BROWN accepted the formation of an intermediate complex between the enzyme and a part of the substrate – a complex that exists for a short time – but he was unable to come up with a quantitative expression for the reaction kinetics as he did not specify any relative rates for the formation or dissociation of the complex. BROWN and GLENDINNING included the hydrolysis of starch by amylase in the group of catalysed reactions that form an intermediate complex very rapidly and then break down slowly. The global reaction rate is proportional to the concentration of the complex, which gives a rate curve that starts off linear and then becomes logarithmic. But this explanation was still incomplete, since the authors did not take into account the inhibition by the hydrolysis products. ◢

In 1902, Victor HENRI applied to enzyme reactions the law of mass action as used in catalysis in general. He embarked upon a series of investigations into the action of invertin on saccharose and sought to integrate the ensemble of results into a single schema. For HENRI, the reaction took place in two steps. The first was the reversible formation of an enzyme-substrate complex. The concentration of the complex at any moment is given by the law of mass action. In other words, HENRI made the implicit assumption of a quasi-equilibrium. The second step was the irreversible breakdown of the complex to regenerate the enzyme and to give rise to the reaction products. By incorporating the inhibitory effect brought about by the reaction products, the law that governs the rate of saccharose inversion by the action of invertin could thus be written:

$$v = \frac{k_s m(E)(S)}{1 + m(S) + n(P)}$$

In this expression k_s represents the specific rate constant for the breakdown of the enzyme-substrate complex, m and n the association constants, respectively, for the enzyme and substrate, and for the enzyme and inhibitor, such that:

$$m = (ES)/(E)(S) \quad ; \quad n = (EP)/(E)(P)$$

Therefore, m and n represent the respective affinity constants of the enzyme for the substrate and for the reaction products. V. HENRI showed that this relation could be satisfactorily applied to the system studied when using the empirically determined values of m = 30 and n = 10.

The rate equation takes into account the general nature of the reaction and the diverse factors that influence it. It shows a linear relationship between the reaction rate and the enzyme concentration. Furthermore, it is reduced to a zero-order reaction with respect to the substrate when the substrate concentration is high enough. If there is no inhibition by the reaction products, the rate equation simplifies to:

$$v = \frac{k_s m(E)(S)}{1 + m(S)}$$

V. HENRI checked that this law could be applied to other enzymatic reactions, including the action of emulsin on salicilin and the hydrolysis of starch by amylase. In the case of the latter, he referred to the existence of intermediate compounds.

In 1913, MICHAELIS and MENTEN continued the experiments on the action of invertin with the aim to verify HENRI's hypothesis. They formulated two fundamental criticisms of HENRI's experimentation; namely, that he had neither taken into consideration the concentration of H^+ ions, nor the mutarotation of glucose, which appears firstly in its birotatory form before reaching its final form. MICHAELIS and MENTEN therefore worked under optimal pH conditions for the enzyme, as defined by the experiments of SÖRRENSEN, MICHAELIS and DAVIDSON, and in an adequately buffered solution to avoid any change that might be sensitive to pH during

the reaction. Additionally, they eliminated experimentally the inhibition by the reaction products and thus introduced the concept of **initial rate**. Bearing in mind the modifications arising during the experimental process, MICHAELIS and MENTEN were able to verify the theory of V. HENRI far more satisfactorily than he had ever managed to do by himself. The principal interest of the work of MICHAELIS and MENTEN lies, above all, in the development of a graphical method to determine the values of the kinetic parameters for a reaction, in particular, the value of 1/m. This has since been called the MICHAELIS constant. However, it was necessary to wait until the 1930s and even later for linear graphical representations (HANES, 1932; LINEWEAVER-BURK, 1934; EADIE, 1949; HOFSTEE, 1949).

The hypothesis of HENRI and MICHAELIS assumed that a reaction could only take place if it first formed an intermediate complex between one enzyme molecule and one substrate molecule. Additionally, this kinetic treatment assumed that the equilibrium is rapid relative to the chemical breakdown of the complex. A more general approach was later presented by BRIGGS and HALDANE (1925), and then HALDANE (1930), who made no hypotheses as to the relative values of the rate constants and applied the same schema under conditions of steady state. Furthermore, HALDANE applied the treatment of the steady state to enzyme reactions in which he considered all steps to be reversible. These diverse aspects are expanded further in Part II.

All these authors made the simplified assumption that the substrate concentration is much higher than that of the enzyme. This limitation, although legitimate for a great number of reactions in vitro, became subject to criticism from the year 1943. STRAUSS and GOLDSTEIN (1943) and GOLDSTEIN (1944), studying the acetylcholine-choline esterase system inhibited by prostigmin, rigorously re-worked the mathematical treatment by including none of the prior simplifications. They introduced a novel concept: the specific or reduced concentration of the enzyme and substrate, i.e. the concentration divided by the MICHAELIS constant, K_m:

$$s' = s/K_m \quad ; \quad e' = e/K_m$$

and by putting a = (ES)/e, they obtained the fundamental equations:

$$s' = a/(1 - a) + ae' \quad \text{and} \quad v = k_s a$$

This type of approach did not resurface for a long time. It was simply not applicable to the majority of enzyme reactions under in vitro study, in which the enzyme is present in catalytic concentrations. This is why it has been necessary to wait over 20 years for the problem to be addressed again, taking into account the respective concentrations of the enzyme and substrate within the cell. Thus, SRERE (1967), and later SOLS and MARCO (1970), underlined the fact that, during metabolism in vivo, the substrate concentration is often very low compared to that of the enzyme. This problem reveals its importance again today as we begin to tackle quantitively enzyme reactions in the cellular environment. The theoretical advances show the consequences of this situation in the regulation of the enzyme activities under

physiological conditions (LAURENT & KELLERSHOHN, 1984; KELLERSHOHN & LAURENT, 1985). These aspects of cellular enzymology, which are not classical, form the subject of the final part of this book.

Research into enzyme kinetics, which adopted HALDANE's approach for those enzymes possessing Michaelian behaviour, considered more complicated situations and branched in two directions. The first extended the kinetic treatment of those reactions that take place in the presence of several ligands, i.e. a substrate, an effector, activator or inhibitor, or a second or even several other substrates. Diverse formalisms and analyses were reported by ALBERTY (1956), DALZIEL (1957), WONG and HANES (1962), CLELAND (1963) and BLOMFIELD et al. (1963). The second direction concerned the treatment, at steady state, of enzyme reactions involving multiple intermediates (PELLER & ALBERTY, 1959). The general form of the equations derived in these different instances was practically always of the same type as the MICHAELIS-MENTEN equation; however, the experimental parameters have more complex meanings.

Alongside these studies, the application of fast kinetic methods to the study of enzyme reactions progressively developed. The use of flow methods under conditions of pre-steady state was introduced by CHANCE in 1943 during research on catalase and peroxidase, the aim of which was to find direct evidence for the MICHAELIS complex. In fact, the complexes observed did not correspond to the MICHAELIS complex because they appeared later. Today, we know that the life-span of the first enzyme-substrate complex is too short to be detected by flow methods. Thereafter, numerous enzymological works employing flow methods were described. Later in 1963, EIGEN and DE MAEYER developed chemical relaxation methods, which they applied to enzymatic reactions. These techniques, reaching a time-scale three orders of magnitude shorter than with flow methods, made it possible to detect intermediates having a life-span of the order of a microsecond. Ever more precise knowledge of the intermediate steps arising during enzymatic reactions has largely contributed to the conceptual development of enzyme catalysis – and all the more so with the progress in understanding enzyme structure.

While all these phenomenological aspects of studying enzymatic reactions were being advanced, progress in the identification, purification, and determination of the sequence – then later the three-dimensional structures – of proteins has gradually led to the interpretation of enzyme function in terms of structure. First of all, the crystallisation of urease by SUMNER in 1926 marked an important date in the history of enzymology. KUNITZ and NORTHROP followed immediately with the crystallisation of some pancreatic enzymes, which put an end to the controversy sustained by vitalists, as irrefutable proof was provided of the proteinaceous nature of enzymes.

Structural knowledge expanded from the determination of global properties, such as molecular mass, size, form and electrical charge on proteins thanks to the development of hydrodynamic methods and electrophoresis, to which we associate

principally the names SVEDBERG and TISELIUS. Later on, the development of optical methods permitted delving a little more into structural details. The resolution of the amino acid sequence of the insulin molecule by SANGER, from 1954 onwards – the first important date in primary structure determination – marked decisive progress in protein chemistry. Increasingly, the gradual identification of amino acid side chains participating in catalysis enabled a better understanding of how a few enzymes functioned. From this point on, enzymology developed in terms of the structure-function relationship, leading to a progressive evolution in the perception of enzyme reactions. The static image of a "lock and key" was substituted by the notion of molecular flexibility, which enables a protein to adapt to its substrate and thus contributes to the efficiency and selectivity of enzyme catalysis. This idea is illustrated by the "induced-fit" theory introduced by KOSHLAND around the 1950s. Also interpreted in terms of structure are the deviations from MICHAELIS' law observed for certain enzymes, following studies of the cooperative behaviour of haemoglobin in binding oxygen. As in the case of enzyme reactions having Michaelian behaviour, the initial development was purely phenomenological before the introduction of allosteric models by MONOD, WYMAN and CHANGEUX in 1965, and then by KOSHLAND, NÉMÉTHY and FILMER in 1966. These models and their variations are dealt with in Part V.

As a result of crystallographic studies, a new degree of precision was achieved in the interpretation of enzyme activity; descriptions of biological macromolecules were achievable at the atomic scale. It is necessary to cite the remarkable works of L. PAULING around the 1950s, lying at the origin of all these advances, which established rules for the formation of regular structure within biological macromolecules and, more particularly, in polypeptides and proteins. We will never be able to emphasise enough the pioneering role played by PAULING in the progress of all of modern biology. After haemoglobin by PERUTZ (1960) and myoglobin by KENDREW (1960), proteins that certain people considered to be *vestigial enzymes*, lysozyme was the first enzyme whose structure was solved; we owe this result to the group of PHILLIPS at Oxford (1965). Thereafter, numerous crystal structures were determined, and to date the three-dimensional structures of thousands of enzymes are known at atomic resolution. The works of crystallographers allow us to know the spatial positions of all atoms in a molecule as complex as an enzyme. The acquisition of such precise structural knowledge, permitting the visualisation of the topology of enzyme active sites, marked an important step in the understanding of their functional properties, although this has led, albeit over some time, to quite a static representation of protein architecture. The development of high-field Nuclear Magnetic Resonance (NMR) in the 1980s offered a new tool for structural studies that enabled protein structure to be probed in solution. This method has produced a wealth of information for the analysis of the catalytic mechanisms of diverse enzymes.

Protein structures determined by X-ray diffraction in fact represent time-averaged molecular views. In reality, proteins display varied internal motion covering a

time-scale spanning from the nanosecond to the second, or even longer depending on the amplitude of the movement (LINDERSTRØM-LANG, 1955; WEBER, 1975; CARRERI et al., 1975, 1979; COOPER, 1979; YON, 1982). It is clear that, were proteins rigid objects, an oxygen molecule would never penetrate as far as to the iron in haemoglobin and myoglobin, the haem group being buried deeply within the structure (KARPLUS et al., 1979). Nowadays, thanks to the methods of structural refinement, crystallographers are able to evaluate these movements by determination of temperature factors or B-factors. Thus, for proteins like myoglobin (FRAUENFELDER et al., 1979), hen egg-white lysozyme (ARTYMIUK et al., 1979) and the trypsin/trypsinogen system (FELHAMMER et al., 1977; BODE, 1979), it appears that the flexible parts of the molecule, where movements of the largest amplitude take place, are localised to the enzyme active site. This suggests that these movements are of primary importance for the expression of enzymatic activity. With this type of study and molecular dynamics simulations, the fourth dimension, i.e. time, was introduced to protein structure and, in particular, to enzyme catalysis and its regulation.

Aside from these ever more elaborate molecular aspects, the cellular aspects of enzymology are being progressively developed, and can now be addressed at a rigorous level of description. The technological progress made is beginning to permit the application of reasoning methods and molecular enzymological techniques to the study of enzymes either associated, or associating transiently, with cellular structures in the cell. The physico-chemical approach to enzyme behaviour *in situ* corresponds to the original and also very current trend in enzymology. The fact that enzymes are often found inside cells at high concentrations and that they are frequently associated to other enzymes in multifunctional complexes or to cellular structures having a polyanionic character, such as membranes or cell walls, modifies their kinetic behaviour with respect to that measured in solution *in vitro*. All of these problems in cellular enzymology are set to be considerably addressed in the coming years. It is nowadays possible from a rational basis to tackle the differences between the behaviour of enzymes in solution and when linked to cellular structures. Thus, the partisans of the cellular theory from the last century would find, in the near future, more rational justification of the role of the cellular environment in enzyme function, whereas their belief was the product of vitalist perceptions of the era. If some of their assertions contained a hint of truth, the principles on which they were based would be no less incorrect.

Modern enzymology is progressing, therefore, along two main branches. The first is the high-resolution molecular aspect, which includes the temporal dimension of structure. The second is the cellular aspect, which takes into account the cellular medium in which catalytic activity and its regulation occur, in all its complexity.

BIBLIOGRAPHY

BOOKS

CORNISH-BOWDEN A., JAMIN M. & SAKS V. –2005– *Cinétique enzymatique*, Grenoble Sciences Collection, EDP Sciences, Paris.

DEBRU C. –1984– *L'esprit des protéines*, Hermann, Paris.

DIXON M. & WEBB E.C. –1962– *Enzymes*, Longmans, London.

GHELIS C. & YON J. –1982– *Protein folding*, Acad. Press, New York.

HALDANE J.B.S. –1930– *Enzymes*, Longmans, London.

HENRI V. –1903– *Lois générales de l'action des diastases*, Ph.D. Thesis, Sorbonne - University of Paris.

LEHNINGER A.L. –1972– *Biochemistry*, Worth Pub, New York.

The references relating to specialised articles are not given in the Introduction. They feature in the corresponding chapters.

PART I

THERMODYNAMICS OF ENZYMES REACTIONS

1 – THE LAWS OF THERMODYNAMICS CONCEPT OF CHEMICAL EQUILIBRIUM

1.1. THE LAWS OF THERMODYNAMICS

Living organisms are centres of numerous transformations during which energy is converted from one form to another. Today, ***bioenergetic conversion*** constitutes an important chapter in Biochemistry. Energy can be transformed into chemical, mechanical or electrical work, or indeed radiated as heat. These bioenergetic conversions are accompanied by a loss of energy, which must be compensated for by an external supply or a transformation of reserve substances. Metabolism as a whole is under the control of thermodynamic laws. It is important to understand how living beings both extract energy from the world around them and how they use this energy. Without going too much into detail here, it is essential however to recall the principal thermodynamic laws. We will concentrate in particular on the practical aspects necessary for understanding those fundamental energetic principles that underlie the chemical reactions involved in metabolism.

Every phenomenon in the universe is governed by the energy content of a system (in thermodynamics, a "system" is made up of all matter within a well-defined space) and by the exchange of energy between this system and its surroundings.

- The **first law** of thermodynamics teaches us that the total energy of a system and its surroundings is constant. In other words, the total energy content of the universe is constant, which is the *law of energy conservation*.
- The **second law** of thermodynamics states that the entropy of the universe increases continually: the *law of energy loss* (degradation).
- The **third law** states that there is zero entropy at absolute zero (–273°C).

1.1.1. FIRST LAW

The first law describes a thermodynamic relationship, which is obeyed by a system being transformed from an initial to a final state:

$$\Delta E = E_A - E_B = Q - W$$

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_2, © Springer-Verlag Berlin Heidelberg 2010

E_A is the energy of the system in its initial state; E_B, the energy of the system in its final state. Q is the heat absorbed by the system and W, the work done by the system. By convention, Q is positive when heat is absorbed by the system; W is positive when work is carried out by the system on its surroundings.

From the first law, **the change in internal energy of the system depends solely on the initial and final states, but is independent of the path taken by the energy transformation.** The first law, the law of energy conservation, is a universal principle.

If W represents some mechanical work at constant pressure P, with the system changing from volume V_A to V_B, we have:

$$\Delta W = P\,\Delta V \qquad \text{where } \Delta V = V_A - V_B$$

and consequently: $$Q = (E_B + P\,V_B) - (E_A + P\,V_A)$$

giving: $$\Delta H = \Delta E + P\,\Delta V$$

The quantity E + PV defines a function describing the state of the system and is the heat content or **enthalpy**. The enthalpy change associated with a transformation is equivalent in magnitude to the heat of the reaction at constant pressure, but with the opposite sign.

If a transformation not only takes place at constant pressure but also in a constant volume without work done in the system, the previous relationship simplifies to:

$$\Delta H = \Delta E$$

The internal energetic change is equal to the change in heat content at constant pressure and volume. The enthalpy change of a chemical reaction is expressed in kilocalories . mol^{-1} (kcal . mol^{-1}) or now more commonly in kilojoules . mol^{-1} (kJ . mol^{-1}). Indeed, a mechanical equivalent of heat exists, where a calorific energy of 1 cal . g^{-1} corresponds to a mechanical energy of 4.185 J.

The enthalpy change of a chemical reaction at a given temperature T may result from two contributing factors: chemical and thermal.

$$\Delta H = \Delta H_0 \int_0^T \Delta C_p dT$$

ΔH_0 represents the chemical contribution, or in other words, the enthalpy change that would be observed were the reaction to take place at absolute zero. The thermal contribution depends on the difference between the heat capacities of the system in its initial and final states at constant pressure, ΔC_p. If $\Delta C_p = 0$, ΔH is constant.

1.1.2. SECOND LAW

The first law of thermodynamics does not enable us to predict if a reaction could arise spontaneously. Now, certain reactions do occur spontaneously even if ΔE is

positive. The thermodynamic parameter allowing such a prediction is known as **entropy**.

The second law states that **a process may occur spontaneously when the sum of the entropy of the system and the external entropy increases**. In a spontaneous process:

$$\Delta S_i + \Delta S_e > 0$$

In order to illustrate this notion of entropy, let us consider two metal blocks placed together in an isolated enclosure; initially each block has a different temperature (Fig. 1.1a). Over time the temperatures will adjust until they become equal. Similarly, we shall consider the chemical example in Fig. 1.1b, where a 1 M salt solution is placed in a chamber separated from a chamber of pure water by a semi-permeable membrane. The molecules of NaCl from the left compartment will diffuse into the right compartment until the salt concentrations are equal. Another example is provided by gas molecules diffusing from a region of high pressure to one of low pressure. These examples show that such processes have a direction, yet they cannot be predicted from the first law.

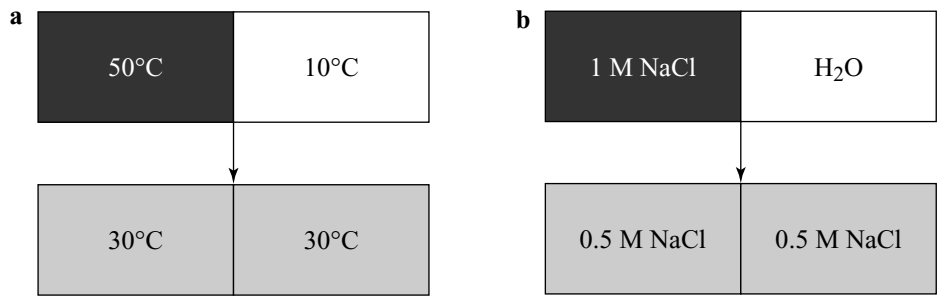


Fig. 1.1 (a) temperature equilibration of two identical metal blocks placed together (b) diffusion of NaCl across a semi-permeable membrane until the concentrations in the two compartments are equal

All systems tend towards a state of equilibrium, or rest, where pressure, temperature and all parameters of the state are uniform. This equilibrated state cannot be reversed by itself; systems at equilibrium do not spontaneously revert to a non-uniform state.

Entropy is a property of state. It represents the state of disorder of a system. When disorder increases:

$$\Delta S > 0$$

Thus, entropy is characteristic of a disordered state of energy incapable of carrying out work. The criterion for the spontaneity of a transformation is defined by the change in entropy associated with it.

If the transformation of a system is due to the reversible absorption of a quantity of heat, dQ, at temperature T the corresponding gain in entropy is:

$$dS = \frac{dQ}{T}$$

Entropy has dimensions of calories . mol^{-1} . K^{-1} or, in other words, dimensions of a **capacity factor**. Multiplying by T (an **intensity factor**) gives the energy. **Extensive variables**, like volume or the heat capacity of chemical bodies, are proportional to the quantity of matter in question; **intensive variables**, such as pressure, temperature or molal heat capacity, are independent of material quantity. At a given temperature, entropy values are relatively low for solids, intermediate for liquids and much higher for gases, as the gaseous state is the most disordered.

During a spontaneous process, the entropy of the system can decrease if the external entropy increases such that the sum of both is a positive value. Thus, it is thermodynamically possible for a highly ordered biological structure to form if the external entropy rises.

The second law establishes that the entropy of the universe continually grows with irreversible evolving processes. It indicates the impossibility to return to an ordered state without the external environment compensating. Theoretically, if the entropy of an evolving system remains constant, the transformation is reversible. In practice, real phenomena are nearly always irreversible due to friction (see Chap. 3).

1.1.3. THIRD LAW

The third law teaches us that the entropy of a substance having a perfect crystalline structure is nil at absolute zero. At absolute zero, thermal motion does not exist; all atoms possess perfect order. In absolute values, the entropy at a temperature T is essentially, therefore, the entropy of passing reversibly from absolute zero to T. The entropy increases with temperature because thermal motion becomes greater.

If, at absolute zero, a substance exists in only one conformation, the entropy at temperature T will be:

$$S_T = \int_0^T \frac{dQ}{T} = \int_0^T C_p \frac{dT}{T} = \int_0^T C_p d\ln T$$

If the material undergoes a sudden change of phase between absolute zero and T, the heat of this transition must be taken into account, and:

$$S_T - S_0 = \int_0^T C_p d\ln T + \sum_i \frac{\Delta H_i}{T_i}$$

ΔH_i is the enthalpy of the transition produced at temperature T_i.

Statistical thermodynamics gives a more concrete idea of entropy. Physical parameters and thermodynamic quantities are **macroscopic** quantities. Statistical thermodynamics relates entropy to microscopic states of the system.

$$S = k \log W$$

It expresses the evolution of the arrangement of these microscopic states towards the most probable arrangement. W is the number of possible arrangements of those states in which the system energy is distributed: the larger the number, the greater the entropy.

1.1.4. FREE ENERGY

All phenomena tend to reach a state of maximum entropy. This is the origin of all spontaneous transformations. Clearly, the entropy of both the system and its surroundings increases. The total entropy of the system with its surroundings is, in fact, difficult if not impossible to measure. Without a precisely determined value, the total entropy can only really provide a measure of whether or not a process could occur spontaneously. This limitation was overcome by the introduction of another thermodynamic quantity, namely **free energy**, or **G**, which was defined by Josiah Willard GIBBS in 1878 and combines the first and second laws of thermodynamics:

$$\Delta G = \Delta H - T\Delta S$$

In this expression, **ΔG is the change in free energy** of the system that is transformed at constant pressure and at temperature T. ΔG and ΔS are, respectively, the change in enthalpy and entropy of the system. Indeed, if we consider the expression:

$$(\Delta S)\ \text{system} + (\Delta S)\ \text{surroundings} > 0$$

or:

$$\left(\frac{\Delta Q}{T}\right)\text{system} + \left(\frac{\Delta Q}{T}\right)\text{surroundings} > 0$$

and the first law:

$$\left[\frac{\Delta Q}{T}\right]\text{surroundings} = \left[\frac{\Delta(E+PV)}{T}\right]\text{system} = \left[\frac{-\Delta H}{T}\right]\text{constant pressure system}$$

and bearing in mind the previous relationship, we have:

$$\frac{\Delta Q}{T} - \frac{\Delta H}{T} \geq 0$$

giving:

$$\Delta H_{system} - T\Delta S_{system} \leq 0$$

This expression comprises only properties of the system and not the surroundings. Gibbs free energy can be considered to be the component of the total energy that can perform work under isothermal conditions.

The transformation of the system is spontaneous when $\Delta G < 0$, termed **exergonic**. If $\Delta G > 0$, the reaction is **endergonic**. This is only possible when coupled to another reaction that is sufficiently exergonic that the sum of their ΔG values is negative. In the case of a reaction at equilibrium, $\Delta G = 0$, and the change in entropy, ΔS, of the system is given by:

$$\Delta S = \frac{\Delta H}{T}$$

During irreversible transformations the free energy decreases, whereas the entropy increases. **Free energy, G, and enthalpy, H, as well as entropy, S, are all functions of the system's state.**

A reaction may be accompanied by a large, positive enthalpy change (a strongly endergonic reaction) and, however, be spontaneous if the gain in entropy is sufficient to compensate for ΔH in such a way that ΔG is negative. A great many reactions exist for which large changes in enthalpy are compensated by a significant increase in entropy. This will be later reemphasised with the help of a few examples (see Part VI).

Thus, the free-energy change, ΔG, is a criterion that enables us to predict if a reaction will occur spontaneously. To summarise:

- a reaction can only occur spontaneously if $\Delta G < 0$;
- a system is at equilibrium and will not undergo a transformation if $\Delta G = 0$;
- a reaction cannot occur spontaneously if $\Delta G > 0$; it must be coupled to another reaction sufficiently exergonic to give a combined $\Delta G < 0$.

The free-energy change is independent of the reaction pathway and depends only on the initial and final states.

1.2. CONCEPT OF EQUILIBRIUM – STANDARD FREE ENERGY

Let us consider the following reaction:

$$A + B \rightleftharpoons C + D$$

The change in free energy of the reaction is given by the following expression:

$$\Delta G = \Delta G_0 + RT \ln \frac{[C][D]}{[A][B]}$$

where ΔG_0 is the change in standard free energy. [A], [B], [C] and [D] are the activities of the reaction components. ΔG is a function that changes continuously until equilibrium is attained and represents the affinity, having a positive value (see the definition of this affinity pp. 69). In a very dilute solution, the activities of the reaction components are interchangeable with their molar concentrations.

The relationship between concentration, c, and activity, a, is given by the equation:

$$a = \gamma c$$

where γ, the activity coefficient, is a function of the ionic strength of the solution and is described by the DEBYE equation:

$$\log\gamma = \frac{-Az^2\sqrt{\mu}}{1+Br\sqrt{\mu}}$$

The coefficients A and B depend on the temperature and the dielectric constant of the medium; r is the radius of the molecule and μ, the ionic strength of the medium. In water at 25°C, this equation becomes:

$$\log\gamma = \frac{-0.505z^2\sqrt{\mu}}{1+0.347\times10^8\sqrt{\mu}}$$

In dilute solutions, the above formula simplifies to:

$$\log\gamma = -0.505z^2\sqrt{\mu}$$

The change in standard free energy is the free energy of the reaction under standard conditions i.e. when the concentrations of both reactants and products are 1 M.

At equilibrium, $\Delta G = 0$ and the free energy equation becomes:

$$\Delta G_0 = -RT\ln\frac{(C)_e(D)_e}{(A)_e(B)_e}$$

Since the equilibrium constant, K_{eq}, is defined as follows:

$$K_{eq} = \frac{(C)_e(D)_e}{(A)_e(B)_e}$$

the free energy equation may be written:

$$\Delta G_0 = -RT\ln K_{eq}$$

or alternatively:

$$K_{eq} = e^{-\Delta G_0/RT}$$

Converting to base-10 logarithms, we obtain the common expression for calculating the standard free-energy change:

$$\Delta G_0 = -2.3\log K_{eq}$$

Thus, the standard free-energy change is related in a simple way to the equilibrium constant of a reaction. This is the most frequently used parameter in bioenergetics.

◤ It is more rigorous to demonstrate this relationship in terms of chemical potentials. The **chemical potential is defined by the partial molar free energy:**

$$\mu = \frac{\partial G}{\partial n}$$

n being the number of reacting molecules and μ is related to the activity of a component by the equation:

$$\mu = \mu_0 + RT \ln a$$

The reference state is that in which the activity equals 1. If a component switches from an activity a to an activity a′, the chemical potential is:

$$\mu' = \mu'_0 + RT \ln a'$$

and the change in chemical potential:

$$\Delta\mu = -RT \ln \frac{a}{a'}$$

or in molality:

$$\Delta G_0 = -RT \ln \frac{n'_A a'}{n_A a}$$

Let us consider the following relationship:

$$n_A A + n_B B = n_C C + n_D D$$

At equilibrium, we can write in terms of chemical potentials:

$$n_A \mu_A + n_B \mu_B = n_C \mu_C + n_D \mu_D$$

giving:

$$\Delta G_0 = n_C \mu_C + n_D \mu_D - n_A \mu_A - n_B \mu_B$$

and thus:

$$\Delta G_0 = -RT \ln \frac{(C)_e^n (D)_e^n}{(A)_e^n (B)_e^n} = -RT \ln K_{eq}$$ ◢

1.3. EXPERIMENTAL DETERMINATION OF THERMODYNAMIC PARAMETERS

Knowledge of the thermodynamic parameters of the major biochemical reactions forms the basis of all bioenergetics and for understanding the mechanisms by which metabolic reactions take place.

1.3.1. ENTHALPY CHANGE

The change in enthalpy, or the heat of a reaction at constant pressure, can be measured directly by microcalorimetry. From this technique we obtain ΔH from which ΔH_0 can be deduced. In the case of reversible reactions, it is typical to study the change in equilibrium constant with temperature.

The VAN T'HOFF equation: $$\frac{-d \ln K}{dT} = \frac{\Delta H_0}{RT^2}$$

also written: $$\frac{d \ln K}{d(1/T)} = \frac{-\Delta H_0}{R}$$

enables determination of ΔH_0 from knowledge of the equilibrium constant, K, at several temperatures. The graphical representation of VAN T'HOFF gives a linear relationship between ln K and 1/T; the slope of the line is equal to $-\Delta H/R$ (Fig. 1.2). In this expression R, the gas constant, is equal to 1.98×10^{-3} kcal . mol^{-1} . K^{-1}.

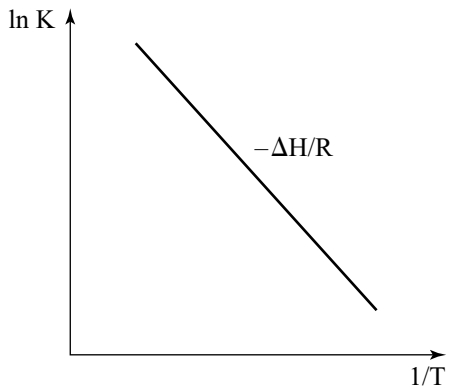


Fig. 1.2 VAN T'HOFF representation

This relationship shows that, when the temperature rises, K increases in endergonic reactions but decreases in exergonic reactions. Table 1.1 gives values of ΔH_0 for some important biochemical reactions.

Table 1.1 Values for the enthalpy of combustion of some biological molecules

Molecule	*ΔH_0 (kcal . mol^{-1})*
Glycine	–234
Lactic acid	–326
D-glucose	–673
Palmitic acid	–2 380

1.3.2. FREE-ENERGY CHANGE

ΔG_0 values for chemical reactions can be obtained in three different ways: from thermochemical data; by analysing the concentrations of reaction components at equilibrium; or by measuring the work done by a system.

1.3.2.1. THERMOCHEMICAL ANALYSIS

ΔG_0 can be obtained from the basic equation:

$$\Delta G_0 = \Delta H_0 - T\Delta S_0$$

when thermochemical data are available that allow the determination of ΔH_0 and ΔS_0. These values can be determined by calorimetric measurement when certain conditions are met. Thus, PRIVALOV and collaborators applied these methods to the study of the thermal denaturation of proteins. They showed that, for most proteins studied, the enthalpy change varies linearly with temperature, i.e. ΔC_p is constant. Under these conditions:

$$\Delta H_D = \Delta H_m + \Delta C_p(T - T_m)$$

ΔH_D being the enthalpy of denaturation, and ΔH_m its value at the melting temperature, T_m. The change in entropy is given by:

$$\Delta S_D = \Delta S_m + \Delta C_p \ln T/T_m$$

1.3.2.2. EQUILIBRIUM STUDY

Table 1.2 Standard free-energy change of some chemical reactions at pH 7.0 and 25°C

Reaction	*$\Delta G_0'$ (kcal . mol^{-1})*
Hydrolysis	
Acid anhydrides	
Acetic anhydride + H_2O ⟶ 2 acetates	–21.8
Pyrophosphate + H_2O ⟶ 2 phosphates	–8.0
Esters	
Ethyl acetate + H_2O ⟶ ethanol + acetate	–4.7
Glucose-6-phosphate ⟶ glucose + phosphate	–3.3
Amides	
Glutamine + H_2O ⟶ glutamate + NH_4^+	–3.4
Glycylglycine + H_2O ⟶ 2 glycines	–2.2
Glycosides	
Saccharose + H_2O ⟶ glucose + fructose	–7.0
Maltose + H_2O ⟶ 2 glucoses	–4.0
Rearrangement	
Glucose-1-phosphate ⟶ glucose-6-phosphate	–1.7
Fructose-6-phosphate ⟶ glucose-6-phosphate	–0.4
Elimination	
Malate ⟶ fumarate + H_2O	+0.75
Oxidation	
Glucose + $6O_2$ ⟶ $6CO_2 + 6H_2O$	–686.00
Palmitic acid + $23O_2$ ⟶ $16CO_2 + 16H_2O$	–2338.00

This method involves calculating ΔG_0 from the equation:

$$\Delta G_0 = -RT \ln K_{eq}$$

and is without a doubt the most often used in biochemistry. The technique requires a precise estimation of the activities or concentrations at equilibrium of each reaction component. Depending on the properties of each, various methods of measurement are employed.

The majority of biochemical reactions occur near to neutral pH and so standard free-energy change, designated by $\Delta G_0'$, generally refers to reactions at pH 7 and at 25°C. Table 1.2 opposite gives values of $\Delta G_0'$ for a few biochemical reactions and Table 1.3, $\Delta G_0'$ values for the formation of several biological molecules.

Table 1.3 Standard free-energy change for the formation of some components
(the values shown refer to a 1 M aqueous solution at pH 7.0 and 25°C)

Component	*$\Delta G_0'$ (kcal . mol^{-1})*
H^+ (standard reference)	0.00
NH_4^+	–19.00
OH^-	–37.60
Ethanol	–43.39
Water	–56.69
L-alanine	–88.75
$Acetate^-$	–88.99
Carbon dioxide (gaseous)	–94.45
$Pyruvate^-$	–113.44
Glycerol	–116.76
Lactate	–123.76
HCO_3^-	–140.33
$Fumarate^{2-}$	–144.41
$Succinate^{2-}$	–164.97
L-$aspartate^-$	–166.99
$Oxaloacetate^{2-}$	–190.53
α-$ketoglutarate^{2-}$	–190.62
L-$malate^{2-}$	–201.98
α-D-glucose	–219.22
cis-$aconitate^{3-}$	–220.51

1.3.2.3. DIRECT MEASUREMENT OF WORK SUPPLIED BY THE SYSTEM

For biochemical reactions involving an oxidation-reduction, the free-energy change for the reaction can be evaluated by measuring the potential difference between two electrodes. Indeed, it is possible to set up cells where the electrical work of the reaction is practically equal to the maximum work.

In the reaction: $\mathbf{A_{red} \rightleftharpoons A_{ox} + n\ e}$

n is the number of electrons, e.

The equilibrium constant is: $$K = \frac{(A)_{ox}(e)^n}{(A)_{red}}$$

If in the reaction an electron acceptor exists such that:

$$\mathbf{B_{ox} + ne \rightleftharpoons B_{red}}$$

then the oxidation-reduction equation is as follows:

$$\mathbf{A_{red} + B_{ox} \rightleftharpoons A_{ox} + B_{red}}$$

These two reactions are coupled (see below); one cannot take place without the other. The total free-energy change is equal to the sum of the free-energy changes of each individual reaction. The free-energy change can be measured based on the oxidation-reduction (also called redox) potential, Eh. This value can be determined from an oxidation-reduction cell as illustrated in Fig. 1.3. The cell comprises two separate compartments linked by conducting material that allows the flow of electrons between the two. A_{red} and B_{ox} are isolated in each of the compartments, which contain platinum electrodes. The potential difference is measured across the terminals under conditions in which the current is nearly zero i.e. in conditions close to reversibility.

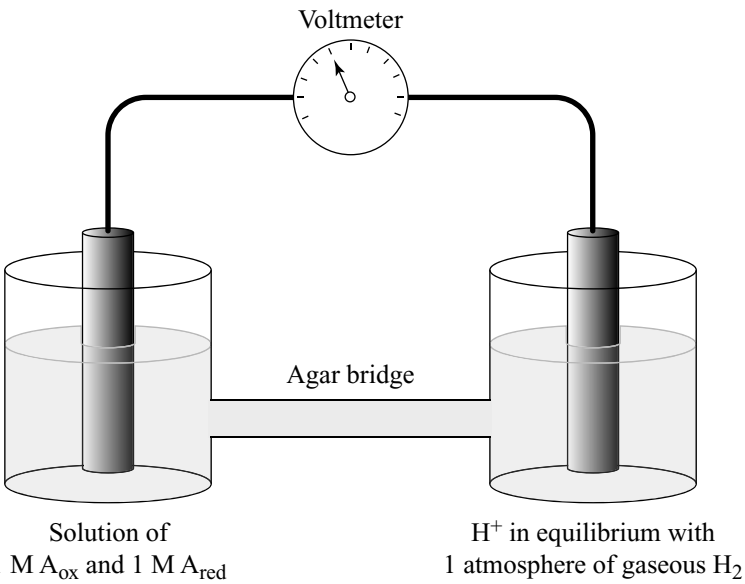


Fig. 1.3 Schematic representation of an oxidation-reduction cell

In one of the compartments the following reaction takes place:

$$H^+ + e \rightleftharpoons \tfrac{1}{2} H_2 \text{ (1 atmosphere)}$$

Therefore the potential difference between the electrodes, or:

$$Pt : A_ox, A_{red} : : H^+(a = 1), H_2(1 \text{ atm.}) : Pt$$

is the redox potential of the system. The change in free energy is related to this potential difference by the formula:

$$\Delta G + nRT \ln \frac{1}{e_a} = nFE_h$$

F is the Faraday equal to 96 500 coulombs, or 23.062 kcal . V^{-1}. The redox potential of the system is therefore:

$$E_h = \frac{RT}{nF} \ln K + \frac{RT}{nF} \ln \frac{A_{ox}}{A_{red}}$$

By definition:

$$E_0 = \frac{RT}{nF} \ln K$$

where E_0 is the normal potential obtained when $A_{ox} = A_{red}$. Thus, at 25°C we have:

$$E_h = E_0 + \frac{0.0586}{n} \log \frac{A_{ox}}{A_{red}}$$

Table 1.4 below lists the redox potentials for half-reactions at pH 7 of several biological systems. NAD^+ ($NAD^+ \rightleftharpoons$ NADH) and FAD (FAD $\rightleftharpoons$ $FADH_2$) are essential electron transporters during electron transfer within living organisms. NADPH is the principal electron donor in reductive biosynthesis.

1.4. COUPLED REACTIONS

1.4.1. DEFINITION OF ENERGETIC COUPLING

A thermodynamically unfavourable (endergonic) reaction can take place when coupled to a thermodynamically favourable reaction if in total the result is exergonic. For example, let us consider these reactions:

$$A + B \rightleftharpoons C + D \qquad \Delta G_1 = +5 \text{ kcal}$$

$$D + E \rightleftharpoons F + G \qquad \Delta G_2 = -7 \text{ kcal}$$

The first is endergonic and so would be impossible if not coupled to the second to give an exergonic reaction, overall:

$$\Delta G_{total} = \Delta G_1 + \Delta G_2 = -2 \text{ kcal}$$

The additivity of free-energy relationships stems directly from the law of mass action. Indeed, the equilibrium constants of these reactions are, respectively:

$$K_1 = (C)(D)/(A)(B) \quad \text{and} \quad K_2 = (F)(G)/(D)(E),$$

The global equilibrium constant is thus:

$$K = (C)(D)(F)(G)/(A)(B)(D)(E) = K_1 K_2$$

The two reactions are coupled via a **common intermediate**, D. The principle of the common intermediate forms the basis of all biological energy transfers. ATP acts as a common intermediate in consecutive reactions of this type.

Table 1.4 Half-reaction redox potentials, E_{m7}, at pH 7 of various systems at around 30°C

System	*E_{m7} (volts)*
Flavins and flavoproteins	
Riboflavin	–0.208
Riboflavin 5′ phosphate (FMN)	–0.219
Flavin adenine dinucleotide (FAD)	–0.219
Xanthine oxidase	–0.350
D-lacticodehydrogenase (anaerobic yeast)	–0.178
Vitamin B2	–0.123
D-amino acid oxidase	–0.110
Glucose dehydrogenase	+0.080
Haems and haemoproteins	
Horseradish peroxidase	–0.271
Haem	–0.120
Cytochrome b_6 (chloroplast)	–0.060
Metmyoglobin	+0.046
Cytochrome b (heart)	+0.068
Methaemoglobin	+0.144
Cytochrome c (heart)	+0.250
Cytochrome oxidase	+0.285
Cytochrome a (heart)	+0.290
Metabolites and others	
Cysteine-cystine	–0.400
Thioglycolic acid	–0.340
NADH-NAD^+	–0.320
NADPH-$NADP^+$	–0.316
Propanol-acetone	–0.286
L-β-hydroxybutyrate-acetoacetate	–0.284
Glutathione	–0.230
Ethanol-acetaldehyde	–0.200
D-lactate-pyruvate	–0.190
L-malate-oxaloacetate	–0.157
Glycerate-hydroxypyruvate	–0.150
Succinate-fumarate	+0.024
Ascorbic acid-dehydroascorbic acid	+0.058

1.4.2. ROLE OF ATP

The main supplier of energy in living organisms is adenosine triphosphate or ATP. From 1905, HARDEN and YOUNG noticed that alcoholic fermentation could only occur in the presence of phosphate. A certain number of phosphate esters were identified. In 1930, MEYERHOFF and LIPMANN showed that, through phosphate ester intermediates, the cell is able to harness part of the energy of chemical bonds contained in nutrients. Isolated for the first time around 1930 from muscle tissue, ATP was considered for a long while to be a specific component of muscle. Later it was discovered that ATP exists in cells from all species of animal, plant and microbe. Figure 1.4 displays the chemical structure of an ATP molecule. It is strongly charged: at pH 7.0 there are four negative charges due to the complete ionisation of each of the three phosphate groups.

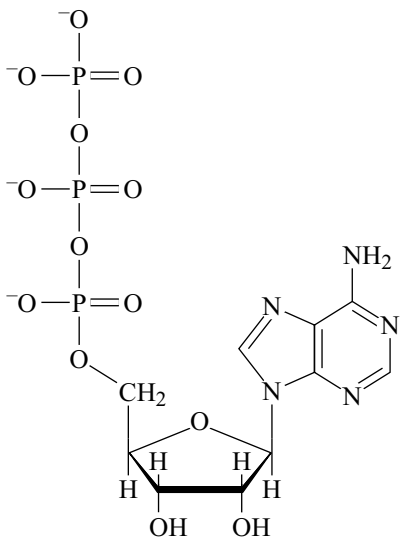


Fig. 1.4 The ATP molecule

The energetic role of ATP was discovered in 1941 by LIPMANN and KALCKAR. ATP hydrolysis can take place in two different ways:

$$ATP^{4-} + H_2O \rightleftharpoons ADP^{3-} + P_i + H^+$$

$$ATP^{4-} + H_2O \rightleftharpoons AMP^{2-} + PP_i + H^+$$

Each of these reactions is accompanied by a free-energy change of $\Delta G_0' = -7.3\ \text{kcal} \cdot \text{mol}^{-1}$ under standard conditions. Compared to simple esters, glycosides and amides, this is significantly more exergonic. For this reason, ATP is classed as an "energy-rich substance". Its specific bonds, whose hydrolysis leads to large, negative ΔG values, are also described as being "rich in energy". In fact, this notion is incorrect and ambiguous. **The so-called energy of the phosphate bond does not apply to the energy of the covalent bond, but to the difference in the energy content of the molecule and its hydrolysis products**.

ATP hydrolysis is utilised to facilitate processes that require energy. Thus, chemical work such as biosynthesis and active transport across membranes and mechanical work, for instance, muscle contraction, all make use of the energy liberated by ATP hydrolysis. Some examples of energetic coupling will be examined later.

ATP is reconstituted in the course of the oxidation of combustible molecules during metabolism. However, the hydrolysis energy of the phosphodiester bond is much less than the quantity of energy released by the catabolism of a single glucose molecule (686 kcal . mol^{-1}). Were this to occur in a single metabolic step, there would be insufficient energy to form several "energy-rich bonds" and hence glucose catabolism requires several steps. In this way, each metabolic step releases an amount of energy equivalent to the free energy of hydrolysis of a single "energy-rich bond".

Coupling a chemical reaction to ATP hydrolysis shifts the equilibrium of the coupled reaction by a factor of the order of 10^8. So let us consider the following reaction:

$$A \rightleftharpoons B$$

with an equilibrium constant $K_{eq} = 10^{-3}$, and $\Delta G_0' = 4$ kcal . mol^{-1}. In the presence of ATP:

$$A + ATP \rightleftharpoons B + P_i + ADP + H^+$$

the value of $\Delta G_0' = -3.3$ kcal . mol^{-1}.

So
$$K'_{eq} = \frac{(B)_{eq}}{(A)_{eq}} \frac{(ADP)_{eq}(P_i)_{eq}}{(ATP)_{eq}} = 10^{3.3/1.36}$$

$$K'_{eq} = 2.67 \times 10^2$$

In cells, the ATP-regenerating system maintains the $(ATP)/(P_i)$ ratio near to 500. Under these conditions:

$$K_{eq} = \frac{(B)_{eq}}{(A)_{eq}} = 2.67 \times 10^2 \times 500 = 1.35 \times 10^5$$

which is about 10^8 times greater than the equilibrium constant of the uncoupled reaction.

1.4.3. FREE ENERGY OF HYDROLYSIS OF SOME PHOSPHORYLATED COMPOUNDS

The most important "energy-rich" substance is ATP, however other molecules also store energy. Some are directly formed during catabolic reactions, others are synthesised as a result of the energy produced from ATP hydrolysis. Table 1.5 opposite shows the free-energy changes of the principal phosphorylated molecules

found in living organisms; phosphates are transferred from molecules at the top of the table towards those at the bottom.

Table 1.5 Change in free energy of hydrolysis of some major phosphorylated compounds

Compound	$\Delta G_0'$ *(kcal . mol^{-1})*
Phosphoenol pyruvate	–14.8
1,3-diphosphoglycerate	–11.8
Phosphocreatine	–10.3
Acetyl phosphate	–10.1
Phosphoarginine	–7.7
ATP	**–7.3**
Glucose-1-phosphate	–5.0
Fructose-6-phosphate	–3.8
Glucose-6-phosphate	–3.3
Glycerol-1-phosphate	–2.2

The substances listed at the top of the scale have a propensity to lose their phosphate groups; conversely, those lower down tend to retain their phosphate. It is important to note that ATP possesses by no means the greatest free energy of hydrolysis amongst phosphate esters. It has an intermediate value and may be considered to be in the middle of a thermodynamic scale. **ATP occupies an intermediate position on a thermodynamic scale of phosphorylated biomolecules.** The ATP-ADP system acts as a bridge between phosphorylated compounds that have a high phosphoryl transfer potential and those whose transfer potential is low. Furthermore, its particular role is explained by the fact that the many ATP-dependent processes take place with the assistance of enzymes that harbour a binding site for ATP and ADP. For instance, the reaction:

$$\text{ATP} + \text{AMP} \rightleftharpoons 2\ \text{ADP}$$

is catalysed by adenylate kinase. ATP and ADP behave like a shuttle for phosphate groups, always in the general direction away from "energy-rich" substances towards "energy-poor" substances. In cells, the terminal phosphate group of ATP is reformed very quickly, estimated to be within a fraction of a second.

Many metabolic reactions are controlled by the energetic state of the cell. The concept of **energy charge** is commonly used (ATKINSON). The energy charge is proportional to the molar fraction of ATP plus a half-molar fraction of ADP, since ATP contains two anhydride bonds whereas ADP has only one; it is expressed by the following relationship:

$$\text{energy charge} = \frac{(\text{ATP}) + \tfrac{1}{2}(\text{ADP})}{(\text{ATP}) + (\text{ADP}) + (\text{AMP})}$$

The phosphorylation potential is defined thus:

$$\frac{(ATP)}{(ADP)(P_i)}$$

The metabolic pathways that produce ATP are inhibited by a high energy load. Conversely, those pathways that consume ATP are activated by a high energy load. As such, this mechanism provides an "energetic buffering effect".

1.4.4. SOME EXAMPLES OF ENERGETIC COUPLING

1.4.4.1. FORMATION OF ATP FROM THE OXIDATION ENERGY OF NUTRIENTS

The first step in the energy conversion of nutrients is the transformation of an energy-rich substance able to transfer its energy to ADP in order to produce ATP. The formation of a carboxylic acid from free aldehyde releases energy:

$$R—CHO + H_2O \rightleftharpoons 2\,H + R—COO^- + H^+$$

where $\Delta G_0' = -7$ kcal . mol^{-1}.

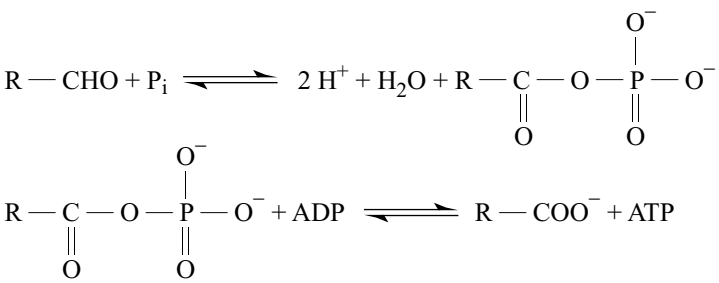
In the cell, this reaction is catalysed by an enzyme in such a way that the energy is not dispersed but stored in the form of phosphate:

$$\text{3-phosphoglyceraldehyde} \rightleftharpoons \text{3-phosphoglycerate}$$

This reaction is linked to the formation of ATP from ADP:

$$R—CHO + P_i + ADP \rightleftharpoons 2\,H^+ + RCOO^- + ATP$$

and the overall energy change, $\Delta G = 0$ kcal . mol^{-1}. The drop in free energy resulting from the aldehyde oxidation (–7 kcal . mol^{-1}) is exactly compensated by the formation of ATP from ADP (+7 kcal . mol^{-1}) and made possible only by coupling the reactions. In fact, the reaction comprises two distinct steps catalysed by two different enzymes, yet involving a common intermediate: 1,3-diphosphoglycerate. To restate, the conservation of oxidative energy from ATP is only possible because the oxidative reaction and the phosphorylating reaction share this common intermediate.



1.4.4.2. USE OF THE ENERGY FROM ATP FOR CHEMICAL WORK

The biosynthesis of complex macromolecules within cells absolutely requires energy. Indeed, the synthesis of a single protein molecule involves assembling correctly a sequence of hundreds of amino acids linked by peptide bonds. Peptide-bond formation is an endergonic process. Similarly, for the construction of polysaccharides such as cellulose, starch and glycogen, hundreds of glucose molecules must be linked by glycosidic bonds, which is also endergonic. The same holds true for the formation of nucleic acids. The general equation for biosynthesis may thus by written:

$$\text{building blocks} \rightleftharpoons \text{macromolecules} + H_2O$$

which represents a highly endergonic reaction. Table 1.6 displays the energies of formation of the major biological macromolecules.

Table 1.6 Chemical work during biosynthesis

Macromolecule	*Element*	*Type of bond*	*$\Delta G_0'$ (kcal . mol^{-1})*	*Number of bonds*
Protein	Amino acid	Peptide	+4	1.2×10^2 to 10^4
Nucleic acid	Mononucleotide	Phosphodiester	+5	3×10^3 to 10^6
Polysaccharide	Monosaccharide	Glycosidic	+4	2×10^3 to 10^4

Each time that a large, complex structure is formed from simple but disordered elements, the system entropy decreases. Given that all complex, organised systems have a natural tendency to return to a state of disorder, it is vital to supply energy to the system in order to counteract this effect.

The reactions are mostly coupled to the hydrolysis of ATP. For instance, the formation of saccharose, a disaccharide, from its building blocks, glucose and fructose, is coupled to ATP hydrolysis by the following sequential reactions:

$$\text{ATP} + \text{glucose} \longrightarrow \text{ADP} + \text{glucose-1-phosphate}$$

$$\text{glucose-1-phosphate} + \text{fructose} \longrightarrow \text{saccharose} + P_i$$

The first reaction is endergonic ($\Delta G_0' = +5.5$ kcal . mol^{-1}) and alone, therefore, is highly unlikely. When coupled to ATP hydrolysis, however, the overall $\Delta G_0'$ is -1.5 kcal . mol^{-1} and hence is exergonic, the common intermediate being glucose-1-phosphate.

Similarly, during protein synthesis, we find the following reaction sequence:

$$\text{amino acid} + \text{ATP} \longrightarrow \text{AA—AMP} + PP_i$$

The formation of an activated amino acid or an adenylated amino acid is then followed by a transfer reaction to the corresponding tRNA:

$$\text{AA—AMP} + \text{ARNt} \longrightarrow \text{AA—ARNt} + \text{AMP}$$

Then: AA—ARNt + polypeptide chain of n amino acids

$\longrightarrow$ ARNt + polypeptide chain of (n + 1) amino acids

These diverse reactions, which always involve a common intermediate, are catalysed by specific enzymes: amino-acyl tRNA synthetases.

1.4.4.3. OSMOTIC WORK

All cells are able to transport and store certain essential substances from the external environment. This results in their much higher concentration within the cell. On the contrary, cells expel or secrete unwanted or degraded substances, even if their internal concentrations are much lower than outside of the cell.

Now, the movement of molecules against a concentration gradient is not spontaneous. This process is assisted by the energy supplied, in general, from ATP hydrolysis and is referred to as active transport. The work of transport pumps is critical in enabling the cell to maintain constantly an appropriate, internal composition, within an environment that may be extremely different or even life-threatening. Besides, thanks to these pump mechanisms, cells are able to extract from the environment the required molecules even if they are present in only very dilute quantities.

The electrical work carried out by certain cells is a form of osmotic work. When certain charged ions, Na^+ or K^+, are transported across the membrane, an unequal distribution of charge, or potential difference, develops. The potential difference is essential in nerve and muscle cells, for example, to create an active state by excitation and for transmitting the signal.

1.4.4.4. MECHANICAL WORK

Mechanical work is a type of biological work, the most striking example of which in higher animals results from muscle contraction. However, muscle contraction is simply the visible manifestation of a property that is possessed by all cells, but which has been perfected. In other cells, this property has the more general effect of enabling the exertion of intracellular traction forces by means of contractile filaments. During cell division, contractile filaments are responsible for the separation of cellular components.

Structures possessing movement such as cilia and flagella also carry out mechanical work, i.e. propulsion. It is worth noting that the mechanical work done in cells is directly sustained by chemical energy. Systems designed by humans to supply mechanical work are generally powered by thermal or electrical energy, whereas in living organisms chemical energy is directly used. These diverse processes are, of course, linked to ATP hydrolysis.

BIBLIOGRAPHY

BOOKS

BANERJEE R.P. –1974– in *Biochimie*, F. CHAPEVILLE & H. CLAUSER eds, Hermann, Paris.

KLOTZ I.M. –1967– *Energy changes in biochemical reactions*, Acad. Press, New York.

LEHNINGER A.L. –1970– *Biochemistry*, Worth Pub., New York.

PRIVALOV P.L. & KHECHINASHVILI N.N. –1974– *J. Mol. Biol.* **86**, 665.

ROCARD Y. –1952– *Thermodynamique*, Masson, Paris.

STRYER L., BERG J.M. & TYMOCZKO J.L. –2002– *Biochemistry*, 5th ed., Freeman Pub., San Francisco.

2 – PROTEIN-LIGAND ASSOCIATION EQUILIBRIA

A common function of all proteins bestowed with biological activity is the non-covalent binding of specific or non-specific ligands. The association of substrates or diverse effector molecules – inhibitors or activators – to an enzyme, and the association of biologically active molecules to a membrane receptor or to a soluble transporter, are just some of the more familiar examples. The binding of protons, H^+, to a protein is another particular case of protein-ligand interaction.

The binding of diverse compounds to proteins has been the subject of numerous studies. Such associations lie within the scope of the theory of multiple equilibria developed by VON MURALT in 1930. The studies presented by SCATCHARD (1949), KLOTZ (1953), EDSALL and WYMAN (1958) are among the most classic. Different cases of protein-ligand association equilibria will be discussed, including those for proteins that possess one or more binding sites, either independent or dependent, and equivalent or non-equivalent. These equilibria obey the law of mass action (see Chap. 1).

2.1. PROTEINS POSSESSING A SINGLE LIGAND-BINDING SITE

When a protein P possesses only a single binding site for a ligand L, for example a monomeric enzyme having a single substrate-binding site, the association takes place according to a simple equilibrium:

$$P + L \rightleftharpoons PL$$

with the association constant:

$$K_{eq} = \frac{(PL)}{(P)(L)}$$

and the standard free-energy change, ΔG_0, for the formation of the complex:

$$\Delta G_0 = -RT \ln K_{eq}$$

In order to determine the equilibrium constant, it is sufficient to measure the concentration of free ligand when equilibrium is reached:

$$K_{eq} = \frac{\text{(bound ligand)}}{\text{(protein-bound ligand)(free ligand)}}$$

and:

$$\text{total ligand} = \text{bound ligand} + \text{free ligand}$$

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_3, © Springer-Verlag Berlin Heidelberg 2010

By simply determining the free ligand concentration at equilibrium we can obtain the association constant. The experimental methods used to achieve this are described further in the chapter.

2.2. *PROTEINS POSSESSING SEVERAL EQUIVALENT AND INDEPENDENT SITES*

Now let us consider the most general case where a protein possesses n sites that are equivalent and independent. This is frequently encountered with oligomeric enzymes having a substrate-binding site on each protomer. There is a system of multiple equilibria, such that:

$$\begin{aligned} P + L &\rightleftharpoons PL \\ PL + L &\rightleftharpoons PL_2 \\ PL_2 + L &\rightleftharpoons PL_3 \\ &\cdots \\ PL_{i-1} + L &\rightleftharpoons PL_i \\ &\cdots \\ PL_{n-1} &\rightleftharpoons PL_n \end{aligned}$$

with the following association constants:

$$K_1 = \frac{(PL)}{(P)(L)}; K_2 = \frac{(PL_2)}{(PL)(L)} \ldots K_i = \frac{(PL_i)}{(PL_{i-1})(L)} \ldots K_n = \frac{(PL_n)}{(PL_{n-1})(L)}$$

Since all the sites are equivalent and independent, a ligand binding to one site can be defined by a single **microscopic or intrinsic association constant, K**, which is the same for all receptor sites. The successive equilibrium constants, K_1, K_2, K_3… K_i… K_n, only differ from the microscopic constant, K, by a probability factor; the probability of binding to the first molecule being different from binding to the second and so on. Thus, we have the relationships:

$$\mathbf{K_1 = nK}$$

$$K_2 = \frac{(n-1)}{2}K$$

$$\cdots$$

$$K_i = \frac{[n-(i-1)]}{i}K$$

$$\cdots$$

$$\mathbf{K_n = \frac{K}{n}}$$

Indeed, there are n different possibilities for the binding of the first ligand molecule and only one for the dissociation of the complex PL (from where the probability factor n for the first equilibrium is derived). There are (n – 1) ways to bind the second ligand molecule and two ways to dissociate the complex PL_2, which gives the probability factor (n – 1)/2 for the second equilibrium and so on.

The average number of sites on the protein molecule occupied by the ligand may be defined as:

$$\bar{\nu} = \frac{\sum_{i=1}^{n} iPL_i}{\sum_{i=0}^{n} PL_i}$$

We can expand this expression in terms of the individual species:

$$\bar{\nu} = \frac{(PL) + 2(PL_2) + 3(PL_3) + \ldots + n(PL)_n}{(P) + (PL) + (PL_2) + \ldots + (PL_n)}$$

Introducing the different equilibrium constants:

$$\bar{\nu} = \frac{K_1(L) + 2K_1K_2(L)^2 + \ldots + nK_1K_2 \ldots K_n(L)^n}{1 + K_1(L) + K_1K_2(L)^2 + \ldots K_1K_2 \ldots K_n(L)^n}$$

We then substitute the equilibrium constants for their values, which are functions of the microscopic constant multiplied by the probability factor:

$$\bar{\nu} = \frac{nK(L)\left[1 + (n-1)K(L) + \ldots + \frac{(n-1)(n-2)\ldots 1 \times K^{n-1}(L)^{n-1}}{2 \times 3 \ldots (n-1)}\right]}{1 + nK(L) + \frac{n(n-1)K^2(L)^2}{2} + \ldots + \frac{n(n-1)(n-2)\ldots 1 \times K^n(L)^n}{2 \times 3 \ldots n}}$$

Both numerator and denominator contain binomial series expansions, so:

$$\bar{\nu} = \frac{nK(L)[1 + K(L)]^{n-1}}{[1 + K(L)]^n}$$

giving the relationship:

$$\bar{\nu} = \frac{nK(L)}{1 + K(L)}$$

The value $\bar{\nu}$ is experimentally measurable. This relationship corresponds to a hyperbola and can therefore be linearised. Two methods of linearisation have been suggested, one by KLOTZ and the other by SCATCHARD. The KLOTZ expression is written thus:

$$\frac{1}{\bar{\nu}} = \frac{1}{n} + \frac{1}{nK(L)}$$

The KLOTZ or inverse plot involves plotting $1/\bar{\nu}$ as a function of 1/(L). This gives a straight line with slope 1/nK that intersects the y-axis at 1/n (Fig. 2.1). The x-axis intercept gives the value –K. This representation requires a sufficient number of points corresponding to high ligand concentrations in the zone close to saturation, which is not always experimentally feasible. Furthermore, experimental precision in this zone is not reliable. An imprecise value for the extrapolated point of intersection on the vertical-axis will lead to quite a large error for n since it is its inverse: the larger the value of n, the smaller the value of 1/n and the greater will be the risk of error. This representation is practically no longer used these days.

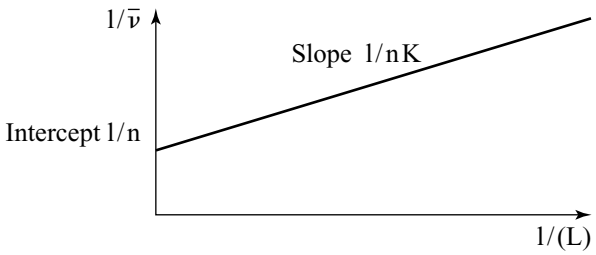


Fig. 2.1 KLOTZ plot

The SCATCHARD equation is expressed as follows:

$$\frac{\bar{\nu}}{(L)} = K(n - \bar{\nu})$$

The SCATCHARD plot, where $\bar{\nu}/(L)$ is plotted as a function of $\bar{\nu}$ is frequently used for ligand-binding studies involving soluble or membrane proteins. Figure 2.2 shows the linear relationship between $\bar{\nu}/(L)$ and $\bar{\nu}$. The horizontal-axis intercept gives n, the number of sites, and from the slope we obtain K, the intrinsic association constant. The vertical-axis intercept is equal to nK. This diagram enables greater precision in the estimation of the parameters n and K, which are obtained directly, rather than from their inverses. A sufficient number of points are required, however, to cover a range of ligand concentrations wide enough to include the value of K.

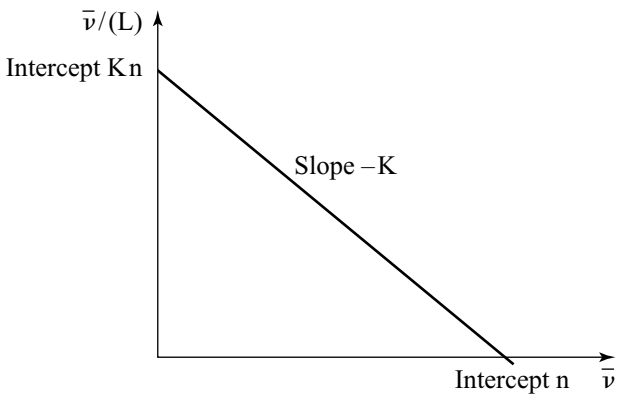


Fig. 2.2 SCATCHARD plot

Equally, we can define a function for the saturation of a protein by the ligand Y_L:

$$Y_L = \frac{\bar{\nu}}{n} = \frac{K(L)}{1+K(L)}$$

During saturation, $\bar{\nu}$ varies from 0 to n while Y_L varies from 0 to 1; Y_L represents the degree of saturation with respect to one site.

2.3. *PROTEINS POSSESSING N INDEPENDENT AND NON-EQUIVALENT SITES*

A protein may contain several categories of binding site having different affinities for the ligand, but independent of each other. Let us suppose that m categories of binding site exist; in each category i the number of sites n_i, defined by their intrinsic association constants K_i, is equivalent and independent. The total number of sites, n, capable of binding the ligand is:

$$n = \sum_{i=1}^{m} n_i$$

The average number of sites occupied by the ligand is given by:

$$\bar{\nu} = \sum_{i=1}^{m} \frac{n_i K_i(L)}{1+K_i(L)}$$

When several categories of site exist, the SCATCHARD plot deviates from linearity (Fig. 2.3). The value at the horizontal-axis intercept is equal to:

$$\sum_{i=1}^{m} n_i$$

in other words $(n_1 + n_2)$, if m = 2. The value of the vertical-axis intercept is:

$$\sum_{i=1}^{m} n_i K_i$$

i.e. $(n_1K_1 + n_2K_2)$ for m = 2.

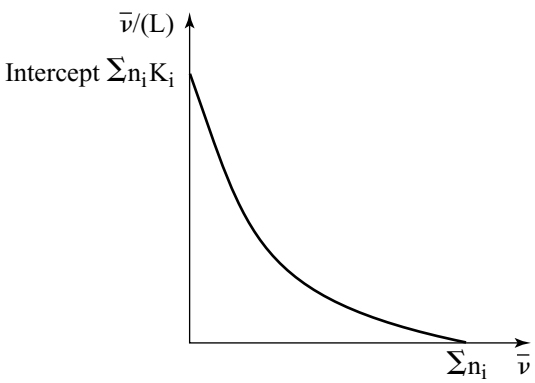


Fig. 2.3 SCATCHARD plot for a protein possessing n independent but non-equivalent sites

In the case of a protein that contains two categories of independent site, such that n_1 sites are defined by their microscopic constant K_1 and n_2 sites by their microscopic constant K_2, the relationship becomes:

$$\bar{\nu} = \frac{n_1 K_1(L)}{1 + K_1(L)} + \frac{n_2 K_2(L)}{1 + K_2(L)}$$

It is possible, on the condition that the intrinsic constants are sufficiently different, to deconstruct the SCATCHARD plot and estimate the binding parameters. If the constant $K_1 >> K_2$ by at least a factor of 50, we may estimate n_1 by the intercept of the line of steepest slope on the horizontal-axis (values of $\bar{\nu}$) and n_1K_1 by its vertical-axis intercept. After taking the difference, we can obtain n_2. By subtracting point by point the first saturation polynomial, we can estimate the parameters K_2 and n_2 using a new plot:

$$\bar{\nu} - \frac{n_1 K_1(L)}{1 + K_1(L)} = \frac{n_2 K_2(L)}{1 + K_2(L)}$$

The values of n_2 and K_2 can be reintroduced into the equation and after progressive refinement we may obtain the four system parameters with satisfactory precision.

When there are more than two categories of site, or even with only two but where the constants K_1 and K_2 are not sufficiently different, it becomes difficult, indeed impossible, to determine the binding parameters. For a complex system, an approximate solution may be found by curve-smoothing based on an initial hypothesis, starting with the simplest possible. Initial estimates for each parameter are used, which are subsequently refined in an iterative manner eventually revealing the solution.

2.4. PROTEINS POSSESSING N EQUIVALENT BUT DEPENDENT SITES

There are proteins that possess n equivalent sites for which interactions exist between the sites. This dependence may occur for various reasons, for instance as a result of electrostatic interactions or steric effects, where the binding of one molecule interferes with the binding of a second, and so on. Additionally, it may be related to the existence of several conformational states of the protein, either induced or pre-existing, where ligand binding induces a conformational change or shifts the equilibrium between several forms of the protein.

2.4.1. EQUIVALENT SITES PRESENTING AN ELECTROSTATIC DEPENDENCE

When a ligand molecule is electrically charged and its binding to the protein involves electrostatic interactions, the net charge of the protein varies according to

the extent of saturation. Consequently, there is an increase in work necessary to bind a second charged molecule and the law of saturation includes an additional term to represent the electrostatic contribution. The average number of occupied sites per protein molecule is described by the relationship:

$$\bar{\nu} = \frac{nK(L)e^{-2wZ}}{1+K(L)e^{-2wZ}}$$

where w is the electrostatic interaction coefficient, defined by the formula:

$$w = \frac{\varepsilon^2}{2DkT}\left(\frac{1}{b}-\frac{\chi}{1+\chi a}\right)$$

ε is the charge carried by one electron (4.8×10^{-10} electrostatic units); k, the BOLTZMANN constant; D, the dielectric constant of the medium; T, the absolute temperature; b is the radius of the protein approximated to a sphere whose charge is distributed uniformly over the surface; a is the exclusion radius (the combined radius of the protein plus ligand); χ is a factor that depends on the dielectric constant of the medium, in water it is equal to $0.38\times10^{-8}\sqrt{\mu}$, μ being the ionic strength of the aqueous solution. The factor 2wZ represents the electrical work that would be needed to discharge the sphere. In fact, everything proceeds as though the equilibrium constant were related to a macroscopic equilibrium constant extrapolated to zero charge:

$$K = K_0e^{-2wZ}$$

Due to the existence of electrostatic interactions the SCATCHARD plot deviates from linearity, but using the formula:

$$\frac{\bar{\nu}}{(L)e^{-2wZ}} = K(n-\bar{\nu})$$

it is possible to obtain a linear graph by plotting $\bar{\nu}/(L)e^{-2wZ}$ as a function of $\bar{\nu}$. In order to do this, both the electrostatic interaction factor and the protein charge must be known.

A classic example of multiple binding equilibria with electrostatic interactions between the sites is the coupling of protons to a category of basic groups on a protein, for example carboxylates or tyrosinates, during acid titration. We may divide all titratable groups on a protein into m categories of n_i titratable groups that contribute to electrostatic interactions. If the protein approximates an evenly charged sphere, the expression corresponding to proton binding becomes:

$$\bar{\nu} = \sum_{i=1}^{m}\frac{n_iK_i(L)e^{-2wZ}}{1+K_i(L)e^{-2wZ}}$$

Here, the ligand is the proton; K_i is the inverse of the ionisation constant of those groups in category i, where:

$$K_{i,diss} = 1/K_i$$

and:

$$K_{i,diss} = \frac{\alpha_i(H^+)}{(1-\alpha_i)}$$

α_i being the average degree of dissociation:

$$K_{i,diss} = K_{i0,diss}\, e^{2wz}$$

giving the expression for titration curves, known as the LINDERSTRØM-LANG equation:

$$\log\frac{\alpha_i}{1-\alpha_i} = pH - pK_{i0,diss} + 0.865wZ$$

2.4.2. EQUIVALENT SITES PRESENTING STERIC OR CONFORMATIONAL INTERACTIONS

2.4.2.1. PHENOMENOLOGICAL ASPECT

Interactions between sites can lead to an association that is either cooperative or anti-cooperative. For enzymes comprising several subunits and therefore several substrate-binding sites, cooperative effects are generally observed, which result essentially from variations in the conformational state of the protein. Purely steric effects, which would be mainly anti-cooperative, are rare in enzymatic systems and for proteins in general. Anti-cooperativity, however, can also arise from conformational effects.

The interactions between ligand-binding sites on a protein have been discussed by many authors: HILL (1910), ADAIR (1925 and 1949), WYMAN (1948 and 1964), SCATCHARD (1949), (NOZAKI et al., 1957), EDSALL and WYMAN (1958), in terms of their phenomenological aspect; the allosteric models suggested later will be discussed in Part V.

In the case where interactions exist between sites, the events take place as if the microscopic constant varies as a function of the degree of saturation of the protein. We can define an apparent microscopic constant that is actually variable:

$$K' = \frac{\bar{\nu}}{(n-\bar{\nu})(L)}$$

K' varies as a function of $\bar{\nu}$. As for electrostatic interactions, we can write:

$$K' = K_0 e^{-f(\bar{\nu})}$$

If $f(\bar{\nu})$ is an increasing function of $\bar{\nu}$, K' increases gradually with saturation and there is cooperativity. Conversely, if $f(\bar{\nu})$ is a decreasing function, there is anti-cooperativity. K_0 is the intrinsic constant extrapolated at zero saturation.

By way of example, let us recall the system studied by NOZAKI et al. in 1957, which was not a biological system but nonetheless a good model for the interactions between ligand-binding sites. Their system concerned the interactions of copper and zinc metal ions with 4-methyl imidazole.

In the case of copper ions, K′ increased as a function of $\bar{\nu}$ (cooperative interaction); with zinc ions, this "constant" was reduced (anti-cooperative interaction). Figure 2.4 illustrates these results.

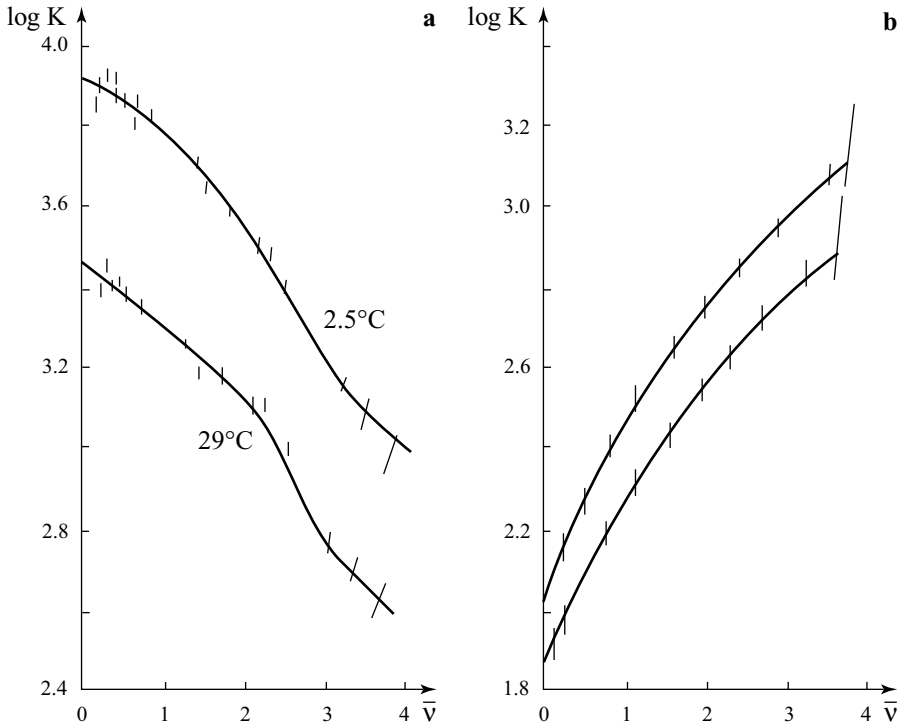


Fig. 2.4 Interaction of 4-methyl imidazole
(a) *with Cu^{++}: an example of a negative interaction*
(b) *with Zn^{++}: an example of a positive interaction*

When the sites are equivalent but show cooperative effects, the SCATCHARD plot deviates from linearity (Fig. 2.5 below). With cooperative binding, the plot is concave; with anti-cooperative binding, it is convex. However, the latter cannot be distinguished from the situation described previously when several categories of site exist. Where cooperativity is observed it is possible, by extrapolating the linear part of the graph, to obtain the number of binding sites and the microscopic constant corresponding to the binding of the first molecule.

If we consider, for example, the case of a tetrameric enzyme that has four binding sites for its substrate (or another specific ligand), the successive equilibrium constants are K_1, K_2, K_3 and K_4, and the corresponding microscopic association

constants: $K_1/4$, $2K_2/3$, $3K_3/2$ and $4K_4$. If $K_1/4 < 2K_2/3 < 3K_3/2 < 4K_4$, there is cooperativity between the sites. On the contrary, when $K_1/4 > 2K_2/3 > 3K_3/2 > 4K_4$, there is anti-cooperativity. In effect, it is as though the microscopic constant gradually increases (cooperativity) or, conversely, decreases (anti-cooperativity) during the course of saturation.

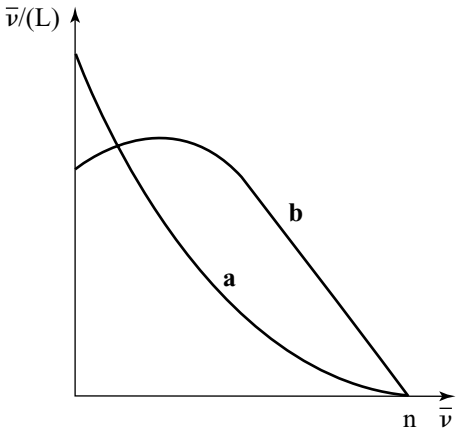


Fig. 2.5 SCATCHARD plot of a protein possessing equivalent but non-independent sites
(a) *anti-cooperative effects*
(b) *cooperative effects*

2.4.2.2. INTERACTION ENERGY BETWEEN SITES

The interaction energy between two binding sites can be determined. The apparent free-energy change for the ith ligand molecule is:

$$\Delta G_0' = -RT \ln K_i$$

The constant K_i contains a statistical factor, indeed:

$$K_i = \frac{[n-(i-1)]K}{i}$$

In order to determine the intrinsic free-energy change corresponding to the binding of the ith ligand the contribution of this statistical factor must be taken into account.

$$\Delta G'_{i,0} = -RT \ln K_i + RT \ln \frac{[n-(i-1)]}{i}$$

The interaction energy $\Delta G_{I,i,j}$ between two sites is the difference in association energy of the ith and the jth ligand. Thus:

$$\Delta G_{I,i,j} = \Delta G_{j,0} - \Delta G_{i,0}$$

or:

$$\Delta G_{I,i,j} = -RT \ln \frac{K_j}{K_i} + RT \ln \frac{[n-(j-1)]i}{[n-(i-1)]j}$$

If the jth ligand binds more strongly than the ith ligand, and $\Delta G_{I,i,j} < 0$, there is cooperativity. If $\Delta G_{I,i,j} > 0$, there is anti-cooperativity. If $\Delta G_{I,i,j} = 0$, the sites are equivalent and independent.

2.4.2.3. EMPIRICAL EQUATIONS

When interactions exist between sites the saturation function can no longer be represented linearly. The classic representations (see Sect. 2.6.9) have curvilinear profiles. Diverse empirical equations have been proposed to describe this phenomenon; the HILL equation (1910), introduced to account for the cooperative binding of oxygen to haemoglobin, is shown below:

$$\frac{Y_L}{1-Y_L} = K'(L)^{\alpha_H}$$

where α_H is the HILL number and K′, the apparent association constant. The coefficient α_H reflects the interactions between the sites. If $\alpha_H > 1$, there is cooperativity. If $\alpha_H < 1$, there is anti-cooperativity. When $\alpha_H = 1$, all sites are equivalent and independent. α_H cannot be greater than the number of sites and is, at most, equal when the cooperativity is optimal. Figure 2.6 shows the form of the saturation curve for different values of α_H. The HILL number is defined by the equation:

$$\alpha_H = \frac{d\ln[Y_L/(1-Y_L)]}{d\ln(L)}$$

The HILL equation can be written in logarithmic form:

$$\log\frac{Y}{1-Y} = \alpha_H \log(L) + \log K'$$

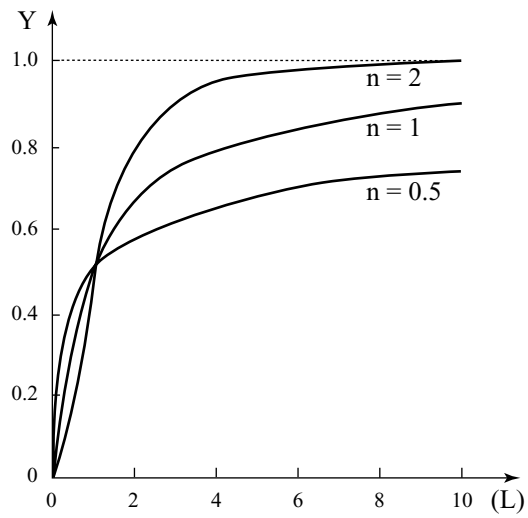


Fig. 2.6 Saturation curves for different values (0.5, 1 and 2) of the HILL coefficient

By plotting logY/(1 – Y) as a function of log(L), we obtain a curve better known as the **HILL plot**, which is often used to describe phenomena displaying cooperativity or anti-cooperativity. If there is an interaction between the sites, we obtain a curve that has an inflexion point where α_H is a maximum (Fig. 2.7 below). The slope at the origin is 1; the slope at saturation is also 1. Thus, the HILL plot is a tangent to two straight lines each with a gradient of 1. In Part V (Chap. 13), we will analyse

the HILL plot, in particular the way in which the plot is used to determine the interaction energy between sites.

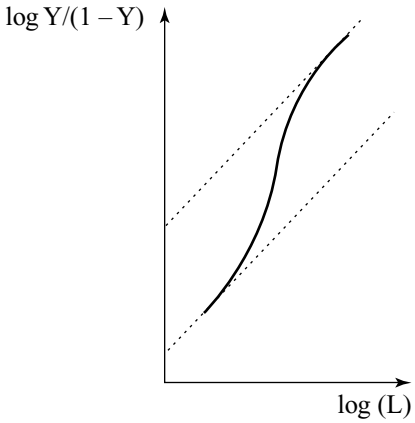


Fig. 2.7 HILL plot

In 1925, ADAIR proposed a general equation to describe cooperative binding; the purpose of this study was also to explain the "abnormal" binding of oxygen to haemoglobin. The ADAIR equation assumes that all sites are equivalent at the beginning, but that the binding of one molecule modifies the intrinsic affinity of a protein for the following molecule, and so on. This amounts to attributing different microscopic constants to the saturation of successive sites on the protein. The ADAIR equation is expressed as follows:

$$\bar{\nu} = \frac{\sum_{i=1}^{n} i(L)^i \prod_{j=1}^{i} K_j}{1 + \sum_{i=1}^{n} (L)^i \prod_{j=1}^{i} K_j}$$

Thus, the binding of oxygen to haemoglobin is described by four microscopic constants, namely $K_1/4 < 3K_2/2 < 2K_3/3 < 4K_4$.

2.5. LINKED FUNCTIONS

Before embarking on the formalism of a general case, let us consider the simple binding to a protein of two ligand molecules L_1 and L_2. We have the following equilibria:

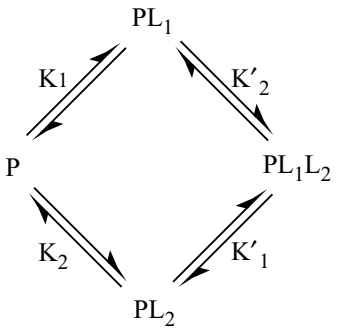
$P + L_1 \rightleftharpoons PL_1$ with the constants K_1 and $\Delta G_{1,0}$

$P + L_2 \rightleftharpoons PL_2$ with the constants K_2 and $\Delta G_{2,0}$

$PL_2 + L_1 \rightleftharpoons PL_1L_2$ with the constants K'_1 and $\Delta G'_{1,0}$

$PL_1 + L_2 \rightleftharpoons PL_1L_2$ with the constants K'_2 and $\Delta G'_{2,0}$

which may be written in the following form:



These four equilibria are not independent. **They are linked** by the free energy relationship:

$$\Delta G_{0,1,2} = \Delta G_{0,1} + \Delta G'_{0,2} = \Delta G'_{0,1} + \Delta G_{0,2}$$

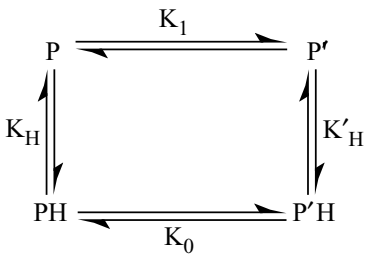
with the constants related as follows:

$$K_1 K'_2 = K_2 K'_1$$

We will have the opportunity to study equilibria of this type in the case of an enzyme having two substrates that bind randomly to their specific sites (see Chap. 5).

An example of linked functions is provided by proteins that change conformation when certain groups are ionised, in particular, groups that possess unusual ionisation pKs.

Let P and P′ be two conformations of a protein in an unprotonated form, and PH and P'H the corresponding states when the protein is protonated. We have the following equilibria:



A theoretical treatment of such a schema has been given by TANFORD and coworkers (1960). The apparent constant of the protein's conformational change is:

$$K_{app} = \frac{(P') + (P'H)}{(P) + (PH)} = \frac{K_0(1+K_1)}{K_1(1+K_0)} K'_H = \frac{1+K_1}{1+K_0} K_H$$

This shows that the conformational change is linked to the protonation of the ionisable group. Reciprocally, the pK of the ionisable group is an apparent pK that

comprises a contribution from the true ionisation constant and from the conformational aspect:

$$K_{H,app} = K_H \frac{1+K_1}{1+K_0}$$

These two examples have been developed to give a simple and practical illustration of the idea of linked equilibria.

◤ Let us now consider the more general case as developed by WYMAN in 1964 when several different ligands bind to a protein. It is useful, first of all, to introduce a few definitions given by WYMAN with respect to the binding of a single ligand type. Let P be a protein capable of binding a ligand L. For the binding of i ligand molecules, the global equilibrium is:

$$P + iL \rightleftharpoons PL_i$$

with the microscopic association constant:

$$K_{Li} = \frac{(PL_i)}{(P_0)(L)^i}$$

and P_0 being the concentration of free protein, (L), the concentration of free ligand. The following expression relates the total protein concentration to P_0:

$$P_{total} = \sum_{i=0}^{n}(PL_i) = P_0\sum_{i=0}^{n}K_{Li}(L)^i = P_0X_L$$

where X_L is what WYMAN defined as the **binding polynomial**. A similar type of expression can be applied to the total ligand:

$$L_{total} = \sum_{i=0}^{n}i(PL_i) = P_0\sum_{i=0}^{n}iK_{Li}(L)^i = P_0\frac{LdX_L}{dL}$$

The saturation function may be written:

$$Y_L = \frac{1}{n}\frac{L_{total}}{P_{total}} = \frac{1}{n}\frac{LdX_L}{nX_L dL}$$

In the presence of two ligands, L and L′, we have the equilibrium:

$$P + iL + jL' \rightleftharpoons PL_iL'_j$$

with the apparent microscopic association constant:

$$K_{LiL'j} = \frac{(PL_iL'_j)}{(P_0)(L)^i(L')^j}$$

The saturation functions for L and L′ are, respectively:

$$Y_L = \frac{1}{n}\left(\frac{d\ln X_L X_{L'}}{d\ln L}\right)_L$$

$$Y_{L'} = \frac{1}{m}\left(\frac{d\ln X_L X_{L'}}{d\ln L'}\right)_{L'}$$

and the fundamental relationship for linked functions is written:

$$n\left(\frac{Y_L}{\ln L'}\right)_L = m\left(\frac{Y_{L'}}{\ln L}\right)_{L'}$$

2.6. METHODS TO STUDY LIGAND BINDING

A diverse range of methods is available to measure ligand-binding to proteins. The choice depends essentially on the nature of the ligand, its properties and the change in its properties upon protein binding. In certain cases, the spectral properties of the ligand or the protein (absorption or fluorescence) can be used; alternatively, binding studies may be made possible by employing radioactive ligands. In order to facilitate such experiments, many rapid methods have been developed. However, it is important to bear in mind that some techniques are prone to giving erroneous results, due to the fact that they disturb the thermodynamic equilibrium of ligand association.

2.6.1. EQUILIBRIUM DIALYSIS

Dialysis takes place across a semi-permeable membrane whose pores are relatively small and hence permeable to the small ligand molecules, but impermeable to the protein. In general, membranes are produced from cellulose, cellophane or collodion. These membranes are sold commercially in the form of cylinders, which are closed at each end to create a **dialysis bag** (Fig. 2.8).

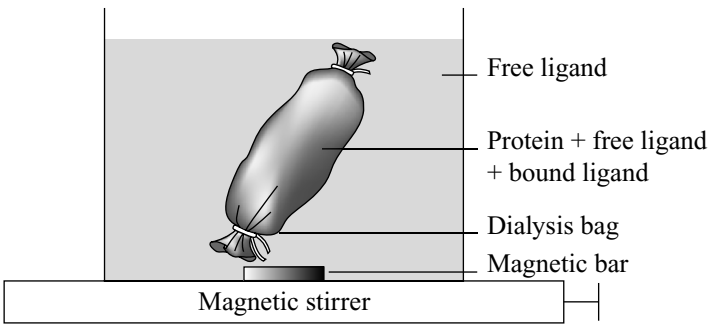


Fig. 2.8 Equilibrium dialysis experiment

The protein solution is placed inside the bag, which also contains free ligand and any ligand bound to the protein. Free ligand alone is found in the external solution. When equilibrium is reached, the free-ligand concentration inside and outside the bag is practically the same. In reality, this is not always possible due to the existence of the **DONNAN effect**. Indeed, if we consider a dilute salt solution of NaCl, the

concentrations of Cl^- and Na^+ ions inside and outside of the dialysis bag are not identical because of the existence of charges on the protein. In the interior of the bag (i), the following equality holds true:

$$(Cl^-)_i + (OH^-)_i = (P)Z + (Na^+)_i + (H^+)_i$$

where (P)Z is the charge on the protein. In the external solution (e), we have:

$$(Cl^-)_e + (OH^-)_e = (Na^+)_e + (H^+)_e$$

The DONNAN equilibrium law leads to the following formula:

$$\frac{(Na^+)_i}{(Na^+)_e} = \frac{(Cl^-)_i}{(Cl^-)_e}$$

This effect can be overcome by working at higher ionic strength (for example, 0.1 M NaCl) so that the protein charge becomes negligible relative to the charges present in the solvent.

In order to study a ligand-binding equilibrium it is important to choose carefully the protein concentration and to vary the ligand concentration in such a way that there is significant variation in the concentration of free ligand. Clearly, this depends on the equilibrium constant. The use of dialysis bags requires quite large quantities of protein. When there is only a limited quantity of protein available (50–100 μL), it is beneficial to use small dialysis cells comprising two compartments separated by a semi-permeable membrane.

2.6.2. DYNAMIC DIALYSIS

◤ Equilibrium dialysis is quite a slow process; in order to reach equilibrium completely it may be necessary to wait several hours or, indeed, several days. This experimental requirement is a major limitation when using a particularly unstable protein. COLOWICK and WOMACK (1969) developed a continuous-flow dialysis method or dynamic dialysis, which enables a complete dialysis experiment to be carried out in 20 min. The principle of the method is outlined as follows. A dialysis cell is used whose dimensions have been precisely calculated. The cell contains two compartments separated by a semi-permeable membrane. The lower compartment receives a constant flow of buffer solution, which is directed towards a fraction collector (Fig. 2.9 opposite).

The upper compartment contains the protein in equilibrium with a radioactively labelled ligand. The protein-ligand solution rapidly equilibrates; free radioactive ligand passes into the lower compartment and is carried along by the flow towards the fraction collector. The fractions are then recovered and their radioactivity content measured. Very quickly a steady state is reached wherein the rates of entry and exit of the ligand to and from the lower compartment become equal. Under these conditions, the radioactive ligand concentration in the flow-through is exactly proportional to the concentration of free ligand in the lower compartment.

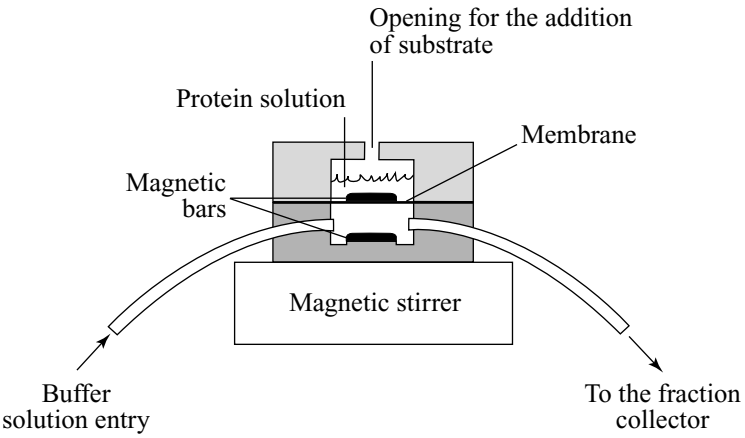


Fig. 2.9 Diagram of the apparatus for measuring ligand binding by dynamic dialysis

The volume of buffer needed to reach this steady state is approximately four times the volume of the lower compartment. The diffusion rate is therefore constant and it remains so for a long time, as the quantity of substrate diffusing is low compared to its concentration in the upper compartment. In the device described by COLOWICK and WOMACK the lower compartment has a volume of 2 mL and the flow rate is 8 mL/min; the steady state is therefore reached after one and a half minutes. The radioactive ligand is forced out incrementally with cold ligand, which is introduced in the upper compartment. First of all, ligand alone is placed in the upper compartment in order to determine the proportionality constant between the quantity of radioactive ligand diffusing out and the concentration of free ligand left in this compartment. More precisely, the experiment is performed by placing the enzyme in the presence of the ligand in the same conditions. A first plateau is reached from which the concentration of free ligand in these conditions may be deduced. A small volume of cold substrate solution is then added; a new steady state is reached as indicated by a new plateau and so on until the final plateau is reached.

This should be identical to the plateau measured in a control experiment lacking enzyme in the upper compartment. The corresponding diagram (Fig. 2.10 below) may be used to estimate the concentration of free ligand for each concentration of total ligand, represented by each plateau in the radioactivity level. This simplification results from the fact that the ligand concentration is exactly compensated by the isotopic dilution of the ligand. One consequence of this phenomenon is that, in the absence of enzyme, the addition of a non-radioactive ligand does not affect the plateau. The diagram allows us to obtain directly the free-ligand concentration by dividing each value corresponding to a plateau by the maximum value. From the difference, we can calculate the concentration of bound ligand. In this manner, from the experimental curve we can construct a SCATCHARD plot, from which the intrinsic affinity constant and the number of binding sites may be determined.

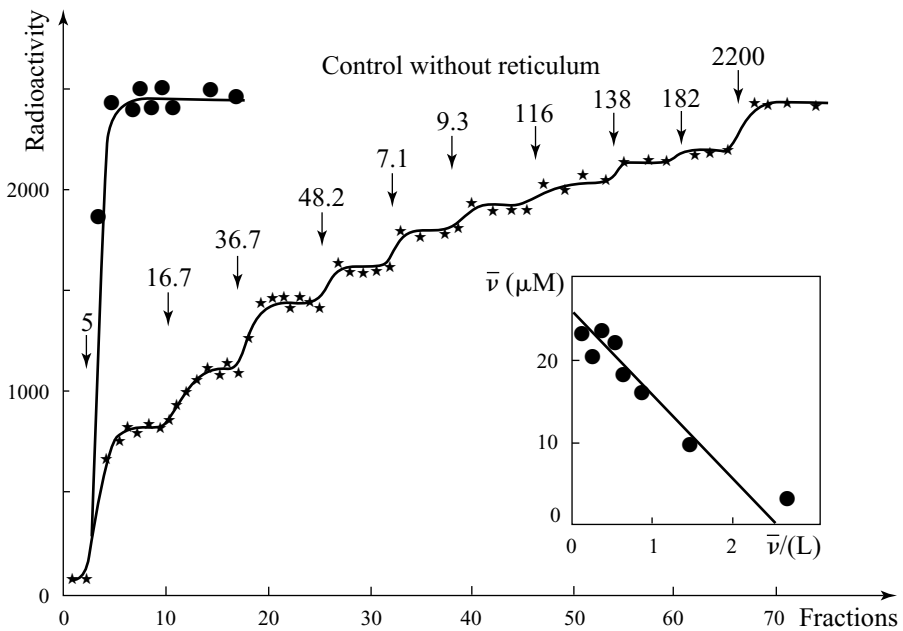


Fig. 2.10 Binding of radioactive ADP [^{14}C] to sarcoplasmic reticulum vesicles (shown as stars)

The control curve without sarcoplasmic reticulum is also shown (black circles). Inset*: the binding parameters are determined from the* SCATCHARD *plot,* $\overline{\nu}$ */ (L) the concentration of ADP bound (in μM) as a function of the concentration of free ADP.*
(Reprinted from *Biochimie*, **58**, TENU J.P. *et al.*, Binding of nucleotides ATP and ADP to sarcoplasmic reticulum: study by rate of dialysis, 513. © (1976) with permission from Elsevier)

The quantitative aspect of the method when the steady state is reached can be easily demonstrated. If N is the number of radioactive ligand molecules in the lower compartment at time t, after addition of substrate to the upper compartment, we have:

$$\frac{dN}{dt} = S_1 D - N\frac{f}{V} \qquad [1]$$

S_1 is the concentration of radioactive free ligand in the upper compartment, f the buffer flow rate (the buffer volume that crosses the lower compartment per unit time), V, the volume in the lower compartment, D is a constant that depends on the diffusing molecules and the characteristics of the apparatus. At steady state, dN/dt = 0 and:

$$\frac{N}{V} = S_1\frac{D}{f} \qquad [2]$$

N/V is the concentration of radioactive ligand in the lower compartment.

The time needed to approach the steady state is given by integrating the expression [1]:

$$\frac{N(t)}{V} = \frac{S_1 D}{f}(1 - e^{-(ft/V)}) + \frac{N_{(t=0)}}{V}e^{-(ft/V)}$$

Thus, when the volume crossing the lower compartment is less than four (or more) times the compartment volume, the exponential term $e^{-(ft/V)}$ is less than 0.018 and the concentration is greater than 98% of the value at steady state.

In order for this experiment to be well controlled quantitatively it is important that the flow rate, the mixing rate and the dimensions of the chamber are accurately calculated and that these rates are maintained perfectly constant.

2.6.3. MEASURING PROTEIN-LIGAND INTERACTIONS IN A BIPHASIC WATER-POLYMER SYSTEM

GRAY and CHAMBERLIN (1971) presented an alternative to the method of COLOWICK, with two advantages: speed of execution and the need for only small quantities of protein. The method involves partitioning in a two-phase system in which dextran (MW 500 000) and polyethylene glycol are mixed with the protein and ligand. The two phases are then separated; samples of each phase are removed and the concentrations of protein and ligand are measured in each. Proteins tend to partition preferentially in the dextran phase at weak ionic strength. Small ligands are found in both phases. The equilibrium of ligand and protein partitioning in the two phases is reached after a few minutes or even less. The technique requires only low quantities (microlitre volumes) of reactant; the validity of the method rests on the fact that the ligand molecules bound to protein partition into the same phase as the protein. In order to analyse the results, the fractions of protein and ligand molecules in both phases must be known. To achieve this, precise determinations are carried out prior to ligand-binding experiments. For this study, it is essential that the protein partitions preferentially into a phase in such a way that the concentration of bound ligand is always negligible compared to the free-ligand concentration in the other phase.

The dextran/polyethylene glycol system was chosen because the two phases can be quickly separated and protein molecules partition preferentially in the dextran phase if certain conditions are respected. However, numerous controls must be carried out as this type of technique has two important limitations. The first is that certain proteins, particularly small proteins, do not have a huge tendency to partition into a single phase; theoretical considerations show that this phenomenon depends on the size and on the surface properties of the molecule. The second limitation results from the fact that the environment of a protein in a biphasic system is not the same as in dilute aqueous solutions, in which ligand-binding studies are generally performed. Thus, this environment may affect the association properties of the ligand and protein. ◢

2.6.4. SIZE-EXCLUSION CHROMATOGRAPHY

Molecular sieves, for example Sephadex, which retard the elution of smaller molecules and let larger molecules pass more quickly down a chromatography column, can be used to separate free ligand from bound ligand that is carried along during

the elution of the protein. The column used must be equilibrated with a buffer solution containing the ligand at a given concentration. The protein, in the presence of the same ligand concentration, is then loaded onto the column and eluted by the buffer solution including ligand. The ligand must have a property (radioactivity, absorption band) that allows its total concentration to be measured in the eluate (Fig. 2.11). This process is repeated for each ligand concentration (HUMMEL & DREYER, 1962). The introduction of techniques of high performance liquid chromatography (HPLC) has improved the precision and speed of obtaining results.

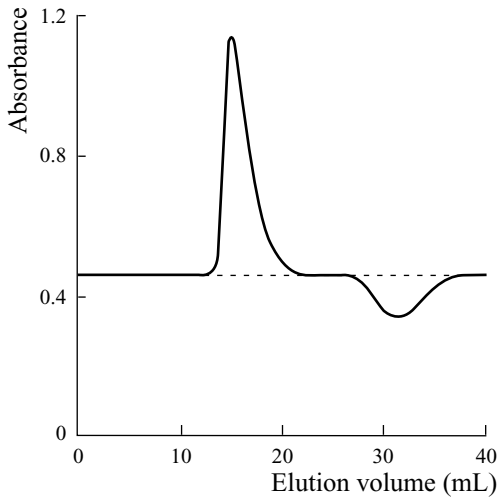


Fig. 2.11 Elution profile monitored by absorbance at 285 nm following the passage of ribonuclease down a Sephadex G25 column, previously equilibrated with 2'cytidilic acid

(Reprinted from *Biochimica Biophysica Acta*, 63, HUMMEL J.P. & DREYER W.J., Measurement of protein-binding phenomena by gel filtration, 530. © (1962) with permission from Elsevier)

2.6.5. ULTRAFILTRATION

Ultrafiltration uses the property of certain filters to absorb proteins while free ligand is not retained on them. In principle, it suffices to filter an homogenous, mixed solution of protein and ligand through these filters (e.g. nitrocellulose, semi-permeable membranes etc.). Subsequently, the proportions of ligand retained on the filter and free ligand in the filtrate are measured. Importantly, it must be stressed that this method may lead to errors because it disturbs the association equilibrium. Equally, it should be noted that the filters used might become saturated by relatively weak protein concentrations (YARUS & BERG, 1970).

2.6.6. ULTRACENTRIFUGATION

In this method the homogenous protein-ligand solution is subjected to a gravitational field. After a suitable period of time, at the top of the centrifugation tubes the solution contains only free ligand whose concentration can be determined (STEINBERG & SCHACHMAN, 1966). Originally, the method required analytical or preparative ultracentrifuges. It has since been adapted for use with ultra-rapid tabletop microcentrifuges (HOWLETT et al., 1978).

2.6.7. DIRECT SPECTROPHOTOMETRIC METHODS

If a substrate, or more generally, a ligand possesses a chromophore that absorbs in a region not interfering with the absorption spectrum of the protein, then spectroscopic methods offer a rapid and easy means to study its association to the protein. The absorption spectrum of the protein-ligand complex needs to shift with respect to the spectrum of the free ligand and is often the case. Furthermore, a wavelength must be found where the difference in absorbance is large enough for any variation to be measurable in a relatively large zone of ligand concentrations (Fig. 2.12).

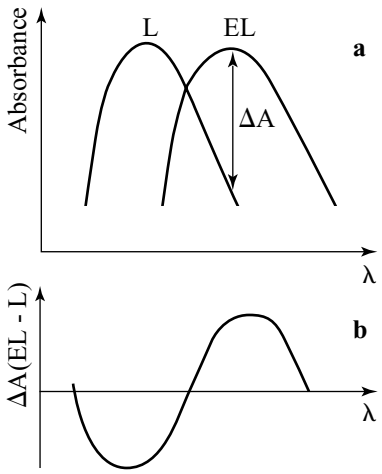


Fig. 2.12 Spectrophotometric method *(**a**) absorbance spectrum of a ligand L and of the enzyme-ligand complex, EL (**b**) difference absorbance spectrum. The difference is significant, so the conditions are favourable for the application of this method*

It is often preferable to employ much more sensitive fluorescence methods where possible. Such methods are applicable when, upon binding to an enzyme, the fluorescence signal of the substrate changes sufficiently for detection. Since fluorescence methods are highly sensitive it is crucial to ensure that buffer solutions do not contain fluorescent impurities.

In certain cases, the fluorescence of the enzyme itself may be monitored if ligand binding significantly affects its signal.

2.6.8. DIRECT TITRATION OF A NUMBER OF ACTIVE SITES

When a ligand possesses a very strong affinity for a protein, the protein concentration can be much higher than the dissociation constant. Under these conditions, virtually all the added ligand binds and the number of binding sites may be directly titrated. In this way, for example, it was possible to titrate directly the catalytic sites of aspartate transcarbamylase with N-phosphonacetyl-L-aspartate (PALA), an analogue of a substrate transition state (KERBIRIOU et al., 1977). Figure 2.13 below shows below the results of the titration.

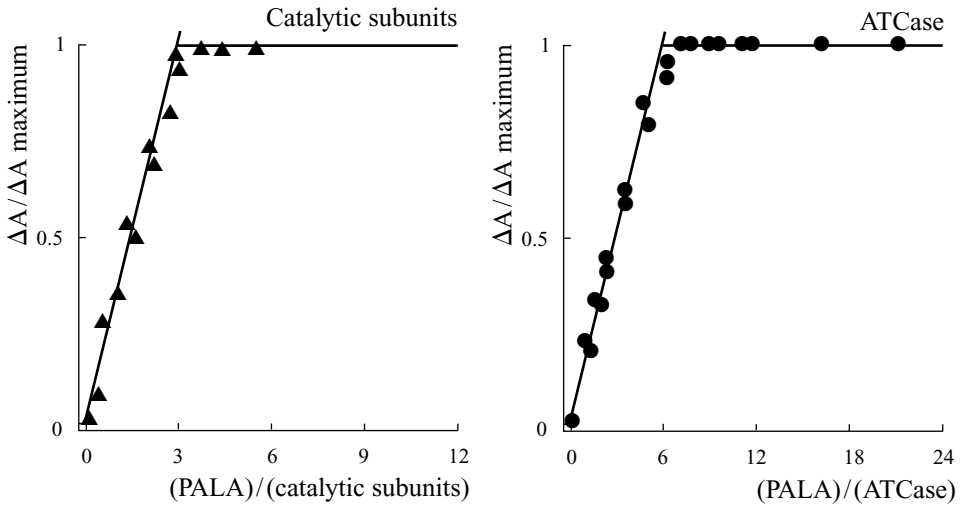


Fig. 2.13 Spectrophotometric titration of the catalytic sites of aspartate transcarbamylase (ATCase) by N-phosphonacetyl-L-aspartate (PALA)
*(**a**) using isolated catalytic subunits (catalytic trimer) – (**b**) using the full-length enzyme, which possesses six catalytic subunits.* (From *J. Biol. Chem.*, **252**, KERBIRIOU D. *et al.*, 2881. © (1977) with permission from The American Society for Biochemistry and Molecular Biology)

2.6.9. INTERPRETATION OF EXPERIMENTAL DATA

The analysis of experimental data from ligand-binding studies is often very sensitive and may lead to errors in interpretation, notably when a protein possesses several categories of binding site, or when it displays cooperative or anti-cooperative interactions. If only one category of equivalent and independent binding site exists the situation is simpler, although certain precautions must still be taken. Firstly, the experimental study must span a wide ligand concentration range. If possible this should vary by two orders of magnitude and around the intrinsic affinity constant (or constants). Indeed, if the concentration range covered in the experiment is too narrow the conclusions may sometimes lead to the perception of only a single category of binding site, when in fact several exist.

In order to illustrate this last point, we will consider one of the most discussed examples in articles regarding the analysis of ligand binding: that of carbamyl phosphate binding to aspartate transcarbamylase. SUTER and ROSENBUSH (1976) carried out several experiments in a concentration range of carbamyl phosphate varying between 0.5 and 200 μM. The corresponding SCATCHARD plot is clearly not linear (Fig. 2.14 opposite). The results were interpreted on the basis of two classes of site, that is, three (2.9 ± 0.2) high-affinity sites ($K_{1,diss}$ = 2.3 μM) and three (2.9 ± 0.2) sites with lower affinity ($K_{1,diss}$ = 62.5 μM), the enzyme concentration being 1.7 μM. If the authors had explored a narrower concentration range, they might have concluded that four equivalent and independent sites existed with an

affinity corresponding to $K'_{diss} = 3.2\ \mu M$. This type of error arising through experiment flaws is still frequently encountered in the literature. Furthermore, it is not always easy to reach sufficiently high ligand concentrations (due to low solubility) or, conversely, sufficiently low concentrations (due to inadequate sensitivity). It is therefore essential to bear in mind that in some experiments the presence of certain sites may not be detectable and this must be taken into account during their interpretation.

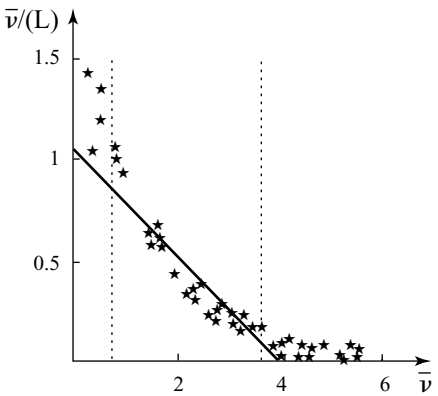


Fig. 2.14 Binding of carbamylphosphate to aspartate transcarbamylase
Dashed lines mark a narrower ligand concentration range.
(From *J. Biol. Chem.*, **251**, SUTER P. & ROSENBUSH J.P., 5986. © (1976) with permission from The American Society for Biochemistry and Molecular Biology)

Amongst the errors found in the literature, it is useful to point out the erroneous interpretations of the non-linear SCATCHARD plots mentioned by NØRBY et al. (1980). **Figure 2.15 illustrates this type of error**. Figure 2.15a represents the false interpretation, which involves comparing the slope of the linear parts of the diagram to the dissociation constant and simply extrapolating in order to have the number of sites of each category. Figure 2.15b represents the rigorous deconvolution of this diagram according to the procedure previously indicated (see Sect. 2.3.3).

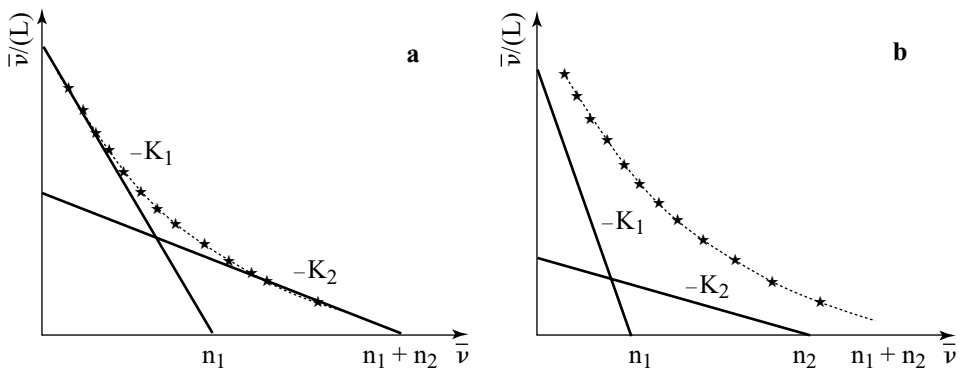


Fig. 2.15 Theoretical SCATCHARD diagrams for the binding of a ligand to two categories of different sites
(a) *incorrect resolution of the data –* ***(b)*** *correct resolution of the data*
(Reprinted from *Anal. Biochem.*, 102, NØRBY J.G. *et al.*, Scatchard plot: common misinterpretation of binding experiments, 318. © (1980) with permission from Elsevier)

When the experiment is conducted in a correct manner and there is no ambiguity in the graphical plot, it is still important to select the best procedure for data analysis. This has been the object of numerous discussions in the literature. It has been previously underlined that among the linear representations, the SCATCHARD plot enables a better determination of the experimental parameters than the KLOTZ plot. This will be demonstrated in Part II (Chap. 5, Sect. 5.2.3) for the EADIE and LINEWEAVER-BURK plots, which, in the case of enzyme reactions, are equivalent to the SCATCHARD and KLOTZ plots, respectively.

KLOTZ (1982) criticised the use of the SCATCHARD plot, in particular when extrapolating to deduce the number of sites in biological systems, where several categories of receptor site with different affinities often exist; so too for non-specific binding, since often these experiments do not allow values close to saturation to be reached. KLOTZ proposed the use of another type of representation in which the concentration of bound molecules is plotted as a function of the logarithm of the free ligand concentration. If all sites are equivalent and independent, this representation has the following properties: the inflexion point corresponds to half saturation, the sigmoidal curve is symmetrical about the inflexion point, the plateau corresponding to n, the number of sites, is reached asymptotically for very large values of free-ligand concentration. KLOTZ emphasised that in many cases extrapolation of the SCATCHARD plot is carried out when actually the inflexion point is not yet reached. Figure 2.16 illustrates KLOTZ's argument. In fact, the same arguments can be made for the semi-logarithmic plot, because it is often very difficult to locate the inflexion point on such a diagram and the symmetry that enables deduction of the saturation is only valid when the sites are both equivalent and independent.

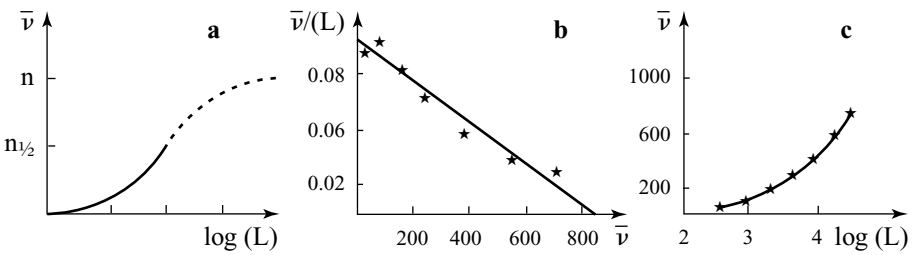


Fig. 2.16 (a) semi-logarithmic representation of bound ligand as a function of the logarithm of the free-ligand concentration, for a receptor with n identical sites (b) SCATCHARD plot for the binding of diazoprane to benzodiapine receptors from rat cerebral cortex membranes (c) changes in bound ligand as a function of the free-ligand concentration (same representation as in (a) with the data from (b))

(From *Science*, **217**, KLOTZ, Numbers of receptor sites from Scatchard graphs: facts and fantasies, 1248. © (1982) reprinted with permission from American Association for the Advancement of Science)

Incidentally, MUNSON and ROBARD (1983) raised a "constructive criticism" of both the SCATCHARD and KLOTZ representations and underlined that the semi-logarithmic plot is not superior to the SCATCHARD plot; the important point being that they should be used and interpreted correctly. They recommended using a statistical method to analyse the data. These authors (1980) suggested a mathematical procedure for the analysis of ligand-binding data. They employed a weighted least-squares

method for the determination of binding parameters by considering different multiple-binding models likely to be encountered in several areas of biology such as endocrinology, neurology, immunology, enzymology and the physico-chemistry of proteins. The SCATCHARD plot, despite its advantages, does not lend itself to statistical analysis, since the variables plotted on the horizontal and vertical axes are not independent. However, it is frequently used as an initial means to obtain provisional estimates for the constants. A detailed treatment of the statistical analysis methods is given in Part II (Chap. 5, § 5.2.3). Other calculation procedures have been suggested such as the simulation method based on the binding polynomial of WYMAN (see § 2.5) proposed by JOSÉ (1985). **Several programs for the statistical data analysis of protein-ligand binding have been developed and are available**. The program CMFITT, developed by M. DESMADRIL, is currently commercialised by the CNRS.

The fact remains that ligand-binding equilibria involving proteins that comprise different classes of binding site or interactions between the sites are always difficult to interpret and may lead to false conclusions even if the experiments have been conducted correctly and a statistical analysis of the data carried out. This notion has been emphasised by several authors, in particular LIGHT (1984) (see also the response of PAUL et al., 1984). Therefore, it is worth independently obtaining complementary information such as structural data, for example. Thus, in the case of the binding of carbamyl phosphate to aspartate transcarbamylase, the statistical treatment of FELDMAN (1983) and the comments by KLOTZ and HUNSTON (1984) clearly indicate that structural data were able to remove the ambiguity. ◢

BIBLIOGRAPHY

BOOKS

ADAIR G.S. –1949– in *Haemoglobin,* F.J.W. ROUGHTON & J.C. KENDREW eds, Butterworths, Cambridge.

EDSALL J.T. & WYMAN J. –1958– *Biophysical chemistry*, Acad. Press, New York.

YON J. –1959– *Structure et dynamique conformationnelle des protéines*, Hermann, Paris.

REVIEWS

KLOTZ I.M. –1953– in *The proteins*, H. NEURATH & K. BAILEY eds, 1st ed., Acad. Press, New York.

WYMAN J. –1964– Linked functions and reciprocal effects in hemoglobin, a second look. *Adv. Prot. Chem.* **19**, 223–286.

SPECIALISED ARTICLES

ADAIR G.S. –1925– *J. Biol. Chem*. **63**, 493.

COLOWICK S.P. & WOMACK F.C. –1969– *J. Biol. Chem.* **244**, 774.

FELDMAN H.A. –1983– *J. Biol. Chem.* **258**, 12865.
GRAY C.W.& CHAMBERLIN M.J. –1971– *Anal. Biochem.* **41**, 83.
HILL A.V. –1910– *J. Physiol. Lond.* **40**, 4.
HOWLETT G.J., YEH E. & SCHACHMAN H.K. –1978– *Arch. Biochim. Biophys.* **190**, 809.
HUMMEL J.P. & DREYER W.J. –1962– *Biochim. Biophys. Acta* **63**, 530.
JOSÉ M. –1985– *Anal. Biochem.* **144**, 494.
KERBIRIOU D., HERVÉ G. & GRIFFIN J.H. –1977– *J. Biol. Chem.* **252**, 2881.
KLOTZ I.M. –1982– *Science* **217**, 1247.
KLOTZ I.M. –1983– *Science* **220**, 981.
KLOTZ I.M. & HUNSTON D.L. –1984– *J. Biol. Chem.* **259**, 10060.
LIGHT K.E. –1984– *Science* **223**, 76.
MUNSON P.J. & ROBBARD D. –1980– *Anal. Biochem.* **170**, 220.
MUNSON P.J. & ROBBARD D. –1983– *Science* **220**, 979.
MURALT A. (VON) –1930– *J. Am. Chem. Soc.* **52**, 3518.
NØRBY J.G., OTTOLENGHI P. & JENSEN J. –1980– *Anal. Biochem.* **102**, 318.
NOZAKI Y., GURD F.R.N., CHEN R. & EDSALL J.T. –1957– *J. Am. Chem. Soc.* **79**, 2123.
PAUL S.M., HANGER R.L., HULIHAN-GIBLIN B.A. & SKOLNICK P. –1984– *Science* **223**, 77.
SCATCHARD G. –1949– *Ann. N.Y. Acad. Sci.* **51**, 660.
STEINBERG I.Z. & SCHACHMAN H.K. –1966– *Biochemistry* **5**, 3728.
SUTER P. & ROSENBUSH J.P. –1976– *J. Biol. Chem.* **251**, 5986.
TANFORD C., DE P.K. & TAGGART V.G. –1960– *J. Am. Chem. Soc.* **82**, 6028.
TENU J.P., GHELIS C., YON J. & CHEVALLIER J. –1976– *Biochimie* **58**, 513.
WYMAN J. –1948– *Adv. Prot. Chem.* **4**, 407–531.
YARUS M. & BERG P. –1970– *Anal. Biochem.* **35**, 450.

3 – *LIVING BEINGS, OPEN SYSTEMS*

3.1. LIVING BEINGS ARE OPEN SYSTEMS, FAR FROM EQUILIBRIUM

The thermodynamic laws, which were defined in Chap. 1, apply to closed systems that do not exchange matter with the surrounding medium and that reach a true thermodynamic equilibrium. Now, the fact that classical thermodynamics is based on the concepts of reversible processes and real equilibrium states leads to serious limitations in the macroscopic description of a number of physico-chemical processes. In classical thermodynamics at equilibrium it is assumed that changes take place in ideal systems in the absence of friction. In reality, a state of equilibrium is reached only under exceptional circumstances; real phenomena involve friction, as a function of time. In order to apply the laws of classical thermodynamics, the phenomenon of friction is normally considered to have an infinitesimally small rate and that any losses as a result are negligible. This approximation is not always valid, however. In chemistry, most systems are irreversible. Substances used by chemists are practically never in a state of equilibrium and many reactions continue until a complete conversion of the reactants is reached. Chemistry is a science of discontinuity and irreversibility.

Living beings are open and irreversible systems, for which both time and rate are important factors; **for living beings, a state of equilibrium means death**. They are in a steady state, or rather, they are the basis of multiple steady states with the rate of entry compensating the rate of exit of a given substance. In classical thermodynamics CARNOT's law represents a law of evolution veering towards continual disorganisation. Evolution in the vicinity of a thermodynamic equilibrium leads to the destruction of a system. On the contrary, in biology the idea of evolution is associated with an increase in organisation, leading to a gain in structures of increasing complexity. Now, dissipative structures can appear in systems that are ***far from equilibrium***, beyond the limit of thermodynamic stability.

While living beings in their entirety are governed by the thermodynamics of irreversible systems, it still remains possible to analyse by means of classical thermodynamics the energy exchange in isolated chemical reactions, even if these are irreversible, and all the previous arguments remain valid. Only when we want to analyse a long series of reactions with entry and exit rates in conditions of steady

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_4, © Springer-Verlag Berlin Heidelberg 2010

state does the system no longer correspond to a thermodynamic analysis of isolated systems and it is necessary to apply the thermodynamics of irreversible phenomena.

The thermodynamics of irreversible processes was introduced around the 1920s by Théodore DE DONDER, who founded the School of Brussels, and to whom we owe many important developments. Before DE DONDER, chemical thermodynamics concerned only reversible processes and systems at equilibrium, yet many chemical reactions are irreversible. DE DONDER underlined the irreversible nature of certain reactions in pure chemistry. Separating irreversibility, a characteristic of the reaction itself, from other possible phenomena that accompany it, he rigorously defined **a function of state that is associated with the chemical reaction itself: affinity**. Affinity may be broadly described as a measure of the distance of a reaction relative to the equilibrium position. More precisely, as we shall see later on, it represents the partial derivative with respect to the reaction progress of the non-compensated heat of CLAUSIUS correlated to the reaction that takes place in a medium in physical equilibrium with the surroundings. Firstly, we distinguish between an affinity at constant pressure and that at constant volume, while the existence of different affinities in the same reaction is contrary to thermodynamic laws. We shall return in detail to the notion of affinity in the sections that follow. By the **introduction of this real function of state, the affinity, and by using the reaction rate**, both linked to the irreversibility of a process and to the creation of entropy as a result, DE DONDER is truly at the origin of chemical thermodynamics.

Following on from DE DONDER, different aspects of the thermodynamics of irreversible systems were developed. To ONSAGER we owe the treatment of irreversible phenomena close to equilibrium; to PRIGOGINE and his school, irreversible phenomena far from equilibrium. It is useful, therefore, to distinguish between the three fields in thermodynamics:

- reversible phenomena where classical thermodynamics may be applied, as discussed in Chap. 1,
- irreversible phenomena close to equilibrium (ONSAGER),
- irreversible phenomena far from equilibrium (PRIGOGINE and co-workers).

In the following sections, the concepts presented in Chap. 1 for closed systems will be re-examined in the context of open systems.

3.1.1. CONSERVATION OF MASS IN OPEN SYSTEMS

- The change in mass of a compound j during the course of a reaction can be expressed by the following equation:

$$dm_j = \nu_j M_j d\xi$$

M_j is the mass of compound j, ν_j its coefficient of stoichiometry; $d\xi$ is the degree of progress of the reaction defined by DE DONDER. The reaction rate is:

$$v = \frac{d\xi}{dt}$$

For a closed system, the law of mass conservation is written:

$$dm_j = \left(\sum_j \nu_j M_j \right) d\xi = 0$$

$\sum_j \nu_j M_j$ is the stoichiometric equation of the chemical reaction.

In an open system, the change in mass of a component j is the sum of two terms: one due to an internal change of the system, $d_i m_j$, and the other due to the contribution from the surroundings, $d_e m_j$:

$$dm_j = d_e m_j + d_i m_j$$

which may be also written as:

$$dm_j = d_e m_j + \left(\sum_j \nu_j M_j \right) d\xi$$

Taking into account the law of mass conservation in a closed system, this expression simplifies to:

$$dm_j = d_e m_j$$

This can be extended to n simultaneous reactions, each designated by the suffix ρ (ρ = 1, 2, 3… n). The total change in mass, dm_j, is the sum of the changes of each individual reaction:

$$dm_j = d_e m_j + M_j \sum_{\rho=1}^{n} \nu_{j\rho} d\xi$$

Bearing in mind the equations of stoichiometry:

$$\sum_{\rho=1}^{n} \nu_{j\rho} M_j = 0$$

and by adding these equations for all constituents (j = 1, 2, 3… n), we obtain the following equation:

$$\mathbf{dm = d_e m}$$

This expression represents **the law of mass conservation in an open system; it indicates that the total change in mass is equal to the mass exchanged with the surroundings**.

3.1.2. ENERGY CONSERVATION IN OPEN SYSTEMS: EXPRESSION OF THE FIRST LAW

The general expression for the law of energy conservation in a closed system has been given in Chap. 1, namely:

$$dE = dQ - dW$$

For work done at constant pressure:

$$dE = dQ - pdV$$

As for mass conservation, in open systems we can write:

$$dE = d_eE + d_iE \qquad \text{with } d_iE = 0$$

The general formula to express energy conservation in open systems may be written:

$$dE = d\Phi - pdV$$

Instead of simple heat transfer there is a flow of energy, $d\Phi$, lasting for a time interval, dt, that conveys both heat transfer and exchange of matter. The enthalpy, which was defined in Chap. 1, may thus be written:

$$dH = d\Phi + pdV$$

3.1.3. ENTROPY PRODUCTION IN OPEN SYSTEMS: SECOND LAW

As we saw in Chap. 1, the second law of thermodynamics calls upon an extensive variable, entropy, that allows us to predict whether a reaction may occur spontaneously. This is described by the relationship:

$$d_iS + d_eS \geq 0$$

d_eS encompasses the terms due to the exchange of matter. The production of entropy due to internal changes in the system d_iS is never negative. It is equal to zero in a reversible process ($d_iS = 0$) and positive if the system comprises an irreversible process ($d_iS > 0$). **This is the only general criterion for irreversibility.**

Let us suppose that we have two systems, with system 1 inside system 2, in such a way that together they form an isolated system. From the second law of thermodynamics we may write:

$$dS = dS_1 + dS_2 > 0$$

Now if we consider each system separately, we have:

$$d_iS_1 > 0 \quad \text{and} \quad d_iS_2 > 0$$

Consequently, entropy absorption by a part of the system being compensated by entropy production is a **thermodynamically impossible situation**. Thus, in each macroscopic region, i.e. in each region containing a sufficiently high number of molecules that the microscopic fluctuations are negligible, the entropy production due to irreversible processes is always positive.

If we reconsider the two systems, system 1 inside system 2, we have:

$$dS = dS_1 + dS_2 > 0$$

The expression for entropy defined in Chap. 1 can be applied:

$$dS = \frac{dQ}{T}$$

It follows that:

$$dS = \frac{dQ_1}{T_1} + \frac{dQ_2}{T_2}$$

The heat gained by each system can be deconstructed into two parts:

$$dQ_1 = dQ_{i1} + dQ_{e1}$$
$$dQ_2 = dQ_{i2} + dQ_{e2}$$

dQ_{i1} is the heat gained by phase 1 from phase 2 and dQ_{e1}, the heat passed to phase 1 from the surroundings. For the whole system, which is closed, the law of energy conservation by the system is written:

$$dQ_{i1} + dQ_{i2} = 0$$

giving the equation: $$dS = \frac{dQ_{e1}}{T_1} + \frac{dQ_{e2}}{T_2} + dQ_{i1}\left(\frac{1}{T_1} - \frac{1}{T_2}\right)$$

The formula: $$d_eS = \frac{dQ_{e1}}{T_1} + \frac{dQ_{e2}}{T_2}$$

represents the heat exchanged with the surroundings, and the term:

$$dQ_{i1}\left(\frac{1}{T_1} - \frac{1}{T_2}\right)$$

is the irreversible heat flow inside the system. Entropy production

is positive: $dQ_1 > 0$ when $\frac{1}{T_1} - \frac{1}{T_2} > 0$

and negative: $dQ_2 < 0$ when $\frac{1}{T_1} - \frac{1}{T_2} < 0$

Entropy production may be expressed per unit time:

$$\frac{d_iS}{dt} = \frac{dQ_{i1}}{dt}\left(\frac{1}{T_1} - \frac{1}{T_2}\right)$$

This represents the product of the speed of an irreversible process dQ_{i1}/dt by a function of state $(1/T_1 - 1/T_2)$. Its sign denotes the direction of heat flux.

Let us consider another example, where matter and energy are exchanged between two compartments separated by a membrane. Compartment 1 at temperature T_1 contains n_{j1} molecules and the chemical potential is μ_{j1}. Compartment 2, at temperature T_2, contains n_{j2} molecules and has a chemical potential, μ_{j2}. From the thermodynamic laws we can write:

$$dE = TdS - pdV + \sum_j \mu_j dn_j$$

The total energy of the system is equal to the sum of the energies of each compartment:

$$dE = dE_1 + dE_2$$

with $$dE_1 = T_1dS_1 + \sum_j \mu_{j1} dn_{j1}$$

$$dE_2 = T_2dS_2 + \sum_j \mu_{j2} dn_{j2}$$

giving the entropy production of each compartment:

$$dS_1 = \frac{1}{T_1} dE_1 - \frac{1}{T_1} \sum_j \mu_{j1} dn_{j1}$$

$$dS_2 = \frac{1}{T_2} dE_2 - \frac{1}{T_2} \sum_j \mu_{j2} dn_{j2}$$

From the laws of energy and mass conservation, we have:

$$dE_1 = -dE_2$$

$$dn_{j1} = -dn_{j2}$$

The formulas for entropy production for each compartment then become:

$$dS_1 = -\frac{1}{T_1} dE_2 + \frac{1}{T_1} \sum_j \mu_{j1} dn_{j2}$$

$$dS_2 = \frac{1}{T_2} dE_2 - \frac{1}{T_2} \sum_j \mu_{j2} dn_{j2}$$

from which we may express the total entropy production:

$$dS = dS_1 + dS_2 = \left(\frac{1}{T_2} - \frac{1}{T_1} \right) dE_2 + \sum_j \left(\frac{\mu_{j1}}{T_1} - \frac{\mu_{j2}}{T_2} \right) dn_{j2}$$

and the entropy produced per unit time:

$$\frac{dS}{dt} = \left(\frac{1}{T_2} - \frac{1}{T_1} \right) \frac{dE_2}{dt} + \sum_j \left(\frac{\mu_{j1}}{T_1} - \frac{\mu_{j2}}{T_2} \right) \frac{dn_{j2}}{dt}$$

This expression shows that **the transport of matter against a concentration gradient may arise if the quantity of energy supplied by the system is sufficient, and enables us to understand active transport by biological molecules**.

3.1.4. ENTROPY PRODUCTION DUE TO CHEMICAL REACTIONS: CHEMICAL AFFINITY

Chemical affinity has been defined by DE DONDER. For a chemical reaction, it is represented by the formula:

$$A = -\sum_j \nu_j M_j \mu_j$$

μ_j is the chemical potential per unit mass, M_j, the molecular mass of the component j, and n_j its stoichiometric coefficient. Relative to one mole:

$$A = -\sum_j \nu_j \mu_j$$

In a chemical reaction there is a relationship between the entropy flow and the entropy production due to the chemical system:

$$dS = \frac{dQ}{T} + \frac{A d\xi}{T}$$

Entropy is therefore the sum of two terms: the change in entropy due to interactions with the surroundings:

$$d_eS = \frac{dQ}{T}$$

and the entropy produced by the chemical reaction:

$$d_iS = \frac{Ad\xi}{T} > 0$$

Td_iS represents the non-compensated heat of CLAUSIUS and:

$$A = T\frac{d_iS}{d\xi}$$

Therefore, the affinity represents the partial derivative of the non-compensated heat of CLAUSIUS relative to the reaction progress.

The affinity can also be expressed as the partial derivative of the GIBBS free energy relative to the reaction progress at constant pressure and temperature:

$$A = -\sum v_j \frac{\partial G}{\partial n_j} = -\frac{\partial G}{\partial \xi} \quad \text{at constant pressure and temperature}$$

At equilibrium, the affinity is nil: $A = 0$.

The affinity may be considered as measuring the distance of the system from equilibrium. Its sign determines the direction of chemical reactions, which may lead the system towards a state of equilibrium, this being an attractor state. And to quote the imaginative definition presented by I. PRIGOGINE and I. STENGERS in *La Nouvelle Alliance*: "The affinity translates into the modern language of attractors –in other words preferences of nature, invincible tendencies as demonstrated by irreversibility– the ancient affinity where chemists would decipher the specific relationships between chemical bodies, molecular 'loves' and 'hates'."

If a transformation involves a component, j, passing from phase 1 to phase 2, the equilibrium condition implies equality of chemical potentials in the two phases:

$$\mu_{j1} = \mu_{j2}$$

Entropy production per unit time is related to the chemical reaction rate:

$$\frac{d_iS}{dt} = \frac{1}{T}Av > 0$$

This inequality, introduced by DE DONDER, expresses the most characteristic property of chemical affinity. It shows **the direct relationship that exists between the affinity and entropy production**.

The previous reasoning can be extended to the case of several simultaneous reactions. In this case we have:

$$d_iS = \frac{1}{T}\sum A_\rho d\xi_\rho > 0$$

$A\rho$ is the affinity of the ρth reaction such that:

$$A_\rho = \sum_j v_{j\rho}\mu_j$$

Entropy production per unit time is written:

$$\frac{d_i S}{dt} = \frac{1}{T}\sum_{\rho} A_\rho v_\rho > 0$$

Under equilibrium conditions, all the affinities equal zero:

$$A_1 = A_2 = \ldots = A_n = 0$$

If $A \neq 0$ and $v = 0$, the system is under conditions of false equilibrium. If $A = 0$, then $v = 0$ since $Td_iS = A$ and $d_iS = 0$, entropy production is nil and there is equilibrium. Consequently, the reaction rate is zero.

If we consider several simultaneous reactions, the second law requires that entropy production is positive. If we consider two reactions where $A_1v_1 < 0$ and $A_2v_2 > 0$, and if the sum:

$$A_1v_1 + A_2v_2 > 0$$

then these two reactions take place; they are coupled. Thermodynamic coupling allows a reaction to progress in a direction opposite to that corresponding to its affinity. As we have already mentioned in Chap. 1, coupled reactions are hugely important in living organisms.

3.1.5. ENTROPY PRODUCTION AND RATE OF IRREVERSIBLE PHENOMENA

The previous expressions show that entropy production is the sum of two terms: general forces X_k (such as affinities) and general fluxes J_k (such as their corresponding rates):

$$\frac{d_i S}{dt} = \sum_k J_k X_k > 0$$

For a chemical reaction, $J_k = v$ and $X_k = A/T$.

This relationship is valid at the macroscopic level where local fluctuations are weak enough that each infinitesimal region of a system may be described by a limited number of macroscopic parameters e.g. pressure and temperature.

From entropy production, we are able to distinguish three thermodynamic fields:

- **at equilibrium**, where entropy production is nil; the flux ($J_k = 0$) and the forces are zero ($X_k = 0$);
- **close to equilibrium**, where the thermodynamic forces are weak; the flux can be considered as a linear function of the force;
- **far from equilibrium**, where the thermodynamic forces are considerable; the flux is no longer a linear function of the force.

3.1.5.1. IRREVERSIBLE PHENOMENA IN THE VICINITY OF THE EQUILIBRIUM

In the vicinity of the equilibrium, the relationship between the general forces and the general fluxes can be considered to be linear. This field of the thermodynamics of irreversible processes is often called **linear thermodynamics of irreversible phenomena**. Therefore, a system of linear and homogenous relationships is evident near to equilibrium between the general fluxes and forces. In order to illustrate

this, we note that this includes the empirical laws such as FICK's law of diffusion and FOURIER's law of heat flow. The general formula for these relationships is as follows:

$$J_j = \sum_k L_{jk} X_k \qquad (j, k = 1, 2, 3 \ldots n)$$

in the case of n fluxes and n forces. Expressed as a matrix, this becomes:

$$\begin{bmatrix} J_1 \\ J_2 \\ \cdots \\ J_n \end{bmatrix} = \begin{bmatrix} L_{11} L_{12} \ldots L_{1n} \\ L_{21} L_{22} \ldots L_{2n} \\ \cdots \\ L_{n1} L_{n2} \ldots L_{nn} \end{bmatrix} \times \begin{bmatrix} X_1 \\ X_2 \\ \cdots \\ X_n \end{bmatrix}$$

The coefficients L_{jk} are called **phenomenologic coefficients**. The diagonal coefficients of the matrix, (i.e. L_{11}, L_{22}, $L_{jj} \ldots L_{nn}$) are called the **proper coefficients**. They represent, for example, thermal or electrical conductivity. The other coefficients L_{jk} ($j \neq k$) are known as **mutual coefficients** or **coefficients of interference**, for example, the thermodiffusion coefficient.

Let us consider two simultaneous, irreversible processes; the corresponding phenomenological relations are as follows:

$$J_1 = L_{11} X_1 + L_{12} X_2$$
$$J_2 = L_{21} X_1 + L_{22} X_2$$

Entropy production may be written:

$$\frac{d_i S}{dt} = L_{11} X_1^2 + (L_{12} + L_{21}) X_1 X_2 + L_{22} X_2^2 > 0$$

and must be positive for all positive or negative values of X_1, X_2, except when $X_1 = X_2 = 0$. The phenomenological coefficients must satisfy the following relationships:

$$L_{11} > 0 \quad \text{and} \quad L_{22} > 0$$
$$(L_{12} + L_{21})^2 < 4 L_{11} L_{22}$$

The proper phenomenological coefficients are positive. The mutual coefficients can be positive or negative assuming they satisfy the above conditions.

ONSAGER (1931) stated the reciprocity relations whose general form is shown below:

$$L_{jk} = L_{kj} \qquad (k \neq j)$$

They signify that, when the flux J_j corresponding to an irreversible phenomenon j is influenced by the force X_k of an irreversible phenomenon k, the flux J_k is influenced by the force X_j. For example, the existence of a temperature gradient might determine the process of diffusing matter and subsequently the appearance of a concentration gradient; reciprocally, a concentration gradient brings about a heat flux into the system with the same temperature coefficient. These relationships translate mathematically by symmetry in the matrix of the L_{kj} values with respect to the diagonal.

This important property may be considered as an addition to the second law of thermodynamics. As a result of the second law, within the scope of the validity of ONSAGER's relations, the system evolves towards a steady state characterised by a minimal entropy production compatible with the constraints imposed on the system. In the case of chemical reactions, these constraints might be the concentrations imposed by the surrounding medium on the initial and final products. The steady state, towards which the system evolves, is characterised by non-zero rates for its irreversible processes. These rates are such that the system-derived rate values are maintained independently of time. In the steady state, the transport of matter is zero ($J_m = 0$), whereas energy and entropy transport are not equal to zero.

The total entropy of the system is constant:

$$d_e S = -d_i S < 0$$

The flow of heat or matter coming from the surroundings leads to a change in entropy that exactly compensates the entropy change due to irreversible processes.

Let us consider, for example, two compartments separated by a membrane and that differ in temperature. The entropy production per unit time is given by the expression:

$$\frac{d_i S}{dt} = J_t X_t + J_m X_m > 0$$

The phenomenological equations are:

$$J_t = L_{11} X_t + L_{12} X_m$$

$$J_m = L_{12} X_t + L_{22} X_m$$

At steady state: $$J_m = L_{12} X_t + L_{22} X_m = 0$$

Taking into account ONSAGER's reciprocity relations:

$$L_{12} = L_{21}$$

we can write: $$\frac{d_i S}{dt} = L_{11} X_t^2 + 2L_{21} X_t X_m + L_{22} X_m^2 > 0$$

The derivative with respect to X_m for constant X_t is:

$$\frac{\partial}{\partial X_m}\left(\frac{d_i S}{dt}\right) = 2(L_{21} X_t + L_{22} X_m) = 2J_m = 0$$

Consequently, the two conditions:

$$J_m = 0 \quad \text{and} \quad \frac{\partial}{\partial X_m}\left(\frac{d_i S}{dt}\right) = 0$$

are equivalent as long as the linear relationships are valid. Under these conditions **entropy production is minimal at steady state**. ONSAGER's equations are valid for small fluctuations around the equilibrium position or even for larger deviations as long as there is a linear relationship between the fluxes and the forces.

In order to illustrate the linear approximation near to equilibrium let us consider a simple system, that of a bimolecular chemical reaction:

$$A + B \underset{k'}{\overset{k}{\rightleftharpoons}} C$$

The reaction rate is written as: $v = k(A)(B) - k'(C)$

$$v = k(A)(B)\left[1 - \frac{k'(C)}{k(A)(B)}\right]$$

If we state: $$Q = \frac{(C)}{(A)(B)}$$

away from equilibrium, the equilibrium constant being $K = k'/k$ (see Chap. 1), the rate equation then becomes:

$$v = k(A)(B)\left(1 - \frac{Q}{K}\right)$$

Q/K describes the approach to thermodynamic equilibrium; at equilibrium Q = K. If we now introduce the term for the affinity, we have:

$$-A_\rho = \Delta G_{0st} + RT \ln Q$$

$$\Delta G_{0st} = -RT \ln K$$

ΔG_{0st} is the variation in free energy at steady state. The expression for the affinity becomes:

$$-A_\rho = RT \ln Q/K$$

$$Q/K = e^{-A\rho/RT}$$

giving the rate equation as a function of the affinity:

$$v = k(A)(B)(1 - e^{-A\rho/RT})$$

This expression describes the relationship between the flux (the rate) and the forces $A\rho/T$. The linear approximation applies when the deviation from equilibrium is small. Under these conditions, in effect, the exponential term may be simplified to the first terms of its expansion as a series:

$$e^{-A\rho/RT} = 1 - \frac{A_\rho}{RT}$$

and the rate equation simplifies to:

$$v = \frac{k}{R}(A)(B) \times \frac{A_\rho}{T}$$

$$|J| = |L| \times |X|$$

The constant L is equal to k(A)(B)/R.

Another example will be given in the last part of this book, which refers to enzymes fixed to a solid matrix. In this instance, it is not merely the reaction catalysed by the enzyme that must be taken into account but also the diffusion of substrates and products of the reaction and, as a consequence, the coupling between the reaction and diffusion.

The law of microscopic reversibility follows from ONSAGER'S reciprocity equations. It indicates symmetry in the equations of individual particle movement with respect to time; these are invariant when there is a transformation from t to –t.

Linear thermodynamics describes the stable behaviour of systems that evolve towards stable steady states; it "forgets" the particular initial conditions.

3.1.5.2. IRREVERSIBLE PHENOMENA FAR FROM EQUILIBRIUM

Once the thermodynamic force becomes sufficiently great, i.e. in moving further from equilibrium, it is no longer possible to describe such processes by phenomenological linear equations; the approximations are no longer valid. For transport processes, in the majority of cases the phenomenological linear laws do remain valid. However, in certain cases it may be necessary to take into account the changes in the phenomenological coefficients, for example, when the thermal conductivity coefficient varies with temperature. In chemical reactions, the thermodynamics of irreversible linear processes is applied as long as the reaction rate is slow enough not to affect to an appreciable extent the equilibrium of BOLTZMANN's rate distribution of each component. This excludes those reactions that have an unusually weak activation energy.

The thermodynamics of irreversible systems involves evaluating the entropy production from GIBBS's law, which, in fact, has been entirely demonstrated solely under equilibrium conditions:

$$dS = \frac{dE}{T} + \frac{P}{T}dV - \sum_j \mu_j dn_j$$

Thermodynamics far from equilibrium rests, therefore, on a novel postulate that leads to the use of GIBBS's law under conditions of non-equilibrium. This brings us back to the notion that entropy depends on the same independent variables as for processes at equilibrium, which is certainly not true very far from equilibrium.

To PRIGOGINE and GLANSDORFF (1971) we owe an analysis relating to entropy production, P, in irreversible processes:

$$P = \frac{d_i S}{dt} = \sum_k J_k X_k > 0$$

The change in P as a function of time can be divided into two contributions, one relating to changes in forces, the other relating to changes in fluxes:

$$\frac{dP}{dt} = \frac{d_x P}{dt} + \frac{d_j P}{dt} = \sum_k J_k dX_k + \sum_k X_k dJ_k$$

Two behaviours may be described:

- under conditions where the phenomenological equations are linear:

$$\frac{d_x P}{dt} = \frac{d_j P}{dt} = \frac{1}{2}P$$

Indeed, from the following relationship:

$$\frac{d_x P}{dt} = \sum_k J_k dX_k = \sum_{kl} L_{kl} X_l dX_l$$

the conditions of reciprocity allow us to write:

$$\frac{d_x P}{dt} = \sum_{k1} X_1 (L_{k1} dX_k) = \sum_{1} X_1 J_1 = \frac{d_j P}{dt}$$

- PRIGOGINE and GLANSDORFF showed that in all valid areas of the thermodynamics of irreversible systems, the contribution to entropy production by changes in force over time is negative:

$$\frac{d_x P}{dt} < 0$$

In the case of a chemical reaction this expression becomes:

$$T\frac{d_x P}{dT} = vdA = dD \leq 0$$

dD is a symbolic expression of this differential, which is not a full differential.

$\frac{\partial D}{\partial A} < 0$ for a time-dependent evolution.

$\frac{\partial D}{\partial A} = 0$ at steady state.

The condition for stability of this steady state is that D must be a minimum, or:

$$\frac{\partial^2 D}{\partial A^2} > 0$$

which is equivalent to: $$\frac{\partial}{\partial t}\left(\frac{\partial^2 S_{ix}}{\partial A^2}\right) > 0$$

This expression represents the derivative, with respect to time, of the second derivative of entropy production with respect to the force (in this case the chemical affinity). If the condition is not satisfied, the lightest fluctuation would enable the system to depart from equilibrium. **Fluctuations can become determining factors in regions of instability**.

Under equilibrium conditions, the probability of a fluctuation in an isolated system is given by EINSTEIN's relation:

$$Pf \sim e^{\Delta S/k}$$

ΔS represents the increase in entropy of the fluctuation, from equilibrium; k is BOLTZMANN's constant. If we develop the expression for entropy around the equilibrium value as far as second-order terms, we obtain:

$$S = S_e + (\partial S)_e + \tfrac{1}{2} (\partial^2 S)_e$$

For an isolated system $(\partial S)_e = 0$, giving:

$$Pf = e^{(\partial^2 S_e)} / 2k$$

The conditions for the validity of this expression when departing from equilibrium have been studied by NICOLIS and BABLOYANTZ (1969). They imply that the time-scales associated with fluctuating systems and their surroundings are completely

separated. The time-scales of fluctuating systems must be much shorter than those of the surroundings such that the latter are practically independent of the instantaneous state of the fluctuating system.

The fact that the phenomenological equations are no longer linear gives rise to expressions for the steady state possessing several solutions. Among these multiple steady states, some are stable, others unstable. The system is unstable if the fluctuations, instead of regressing, are amplified and extend throughout the system, which evolves to a new behaviour. The appearance of a whirlpool in a fluid (a phenomenon of cavitation), when the flow becomes too great, is one example of this; there are others, in both chemical and biochemical reactions. In order to illustrate this situation, let us now consider the simple schema proposed by EDELSTEIN (1970):

$$A + X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} 2X$$

$$X + E \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} C$$

$$C \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} E + B$$

which corresponds to the autocatalytic production of X followed by its enzymatic transformation according to MICHAELIS's law (see later). The rate equations are written:

$$\frac{dX}{dt} = k_1(A)(X) - k_{-1}(X)^2 - k_2(X)(E) + k_{-2}(C)$$

$$\frac{dE}{dt} = -k_2(X)(E) - k_{-3}(B)(E) + (k_{-2} + k_3)(C)$$

$$(E) + (C) = (E)_t$$

enabling the equation for the steady state to be written:

$$f(X)\ aX^3 + [b + c(B) - d(A)]X^2 + \{(E)t - (A)[e + f(B)]X - g\ (B)(E)t\} = 0$$

The constants a, b, c, d, e, f and g are combinations of the specific rate constants.

Figure 3.1 opposite shows the multiple steady states of X as a function of (A). If we move from the lower branch towards increasing concentrations of A, we end up in the zone of stable steady states until reaching M. Then at this critical point the system changes behaviour abruptly by a sort of "catastrophe" effect and finds itself in the upper branch, which corresponds to a new zone of stability. Similarly, starting from the upper branch going towards decreasing concentrations of A, there is a zone of instability until a new critical point N where the system abruptly changes behaviour once again, passing to the lower branch. Thus, this describes a hysteresis cycle. **The system keeps a memory of the previous states**. This diagram also shows that, for a given value (e.g. A_1) of concentration (A), there are two stable steady states and one unstable steady state.

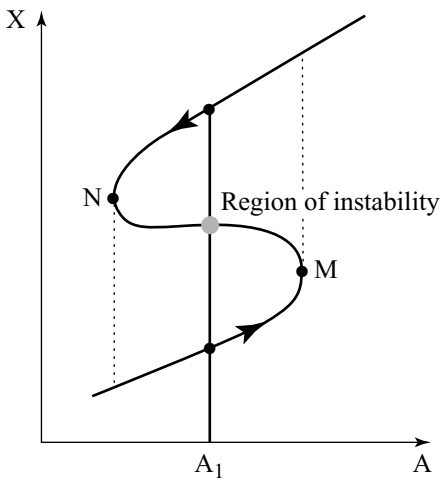


Fig. 3.1 Multiple steady states of X as a function of (A)
When the concentration of A increases, the system lies in a zone of stable steady states up to the critical point M (lower branch)*, then the system changes behaviour abruptly. Similarly, when (A) decreases, the system is in a zone of stable steady states up to the point N* (upper branch)

A system can evolve in the vicinity of the steady state in several different ways. In the phase plane, i.e. in the concentration plane, following the evolution of the system, different types of trajectory around the steady state may be described. There may be a **stable or unstable knot** whereby a weak perturbation would tend to force the system closer to, or further from, the steady state; similarly, a **stable or unstable centre**, or a **saddle point**, and lastly a **limit cycle** (Fig. 3.2 below). In the case of a stable centre, the system will approach the steady state by means of gradually dampened oscillations; if the centre is unstable, the system will move away from the steady state in an amplified oscillatory manner. Where a limit cycle exists, the system maintains a revolving path around the steady state. In this chapter, we will keep the description purely qualitative; we will provide a mathematical description, i.e. the differential equations leading to these different situations, in Part VI regarding the regulation of metabolic pathways. Indeed, different oscillatory processes have been described for biochemical systems. Living beings are centres of rhythmic phenomena such as the cardiac rhythm or circadian rhythms. These rhythms have different time-scales. There are periodic phenomena at the molecular level; generally, these correspond to those systems comprising autocatalytic or auto-inhibitory processes. Phenomena of temporal oscillations have been described for multi-enzymatic systems such as in glycolysis where periodic variations in the concentration ratio of NAD/NADH have been observed; similarly, the concentration of cyclic AMP in *Dyctiostelium discoïdeum* varies periodically during the course of differentiation.

A classic example of oscillation about the steady state is the biological problem of the fight for life and the coexistence of prey and predator species studied by LOTKA (1920) and VOLTERRA (1931). In a time-independent environment two species A and B exist where B behaves as a predator and eats individuals from species A. Kinetic equations are formulated as follows:

$$\frac{dA}{dt} = \varepsilon_1 A - AB$$

$$\frac{dB}{dt} = AB - \varepsilon_2 B$$

A and B represent the number of animals of each species; ε_1 and ε_2 are constants representing the rate of appearance and disappearance of each species. The following conditions determine the steady state: $\varepsilon_1 = B$ and $\varepsilon_2 = A$. However, if the system is not at steady state, it will follow a closed trajectory around the steady state. Let us note also, as we have previously underlined, that this example concerns two very different time-scales: the biological time of the species, A and B, and the "geological" time of the environment.

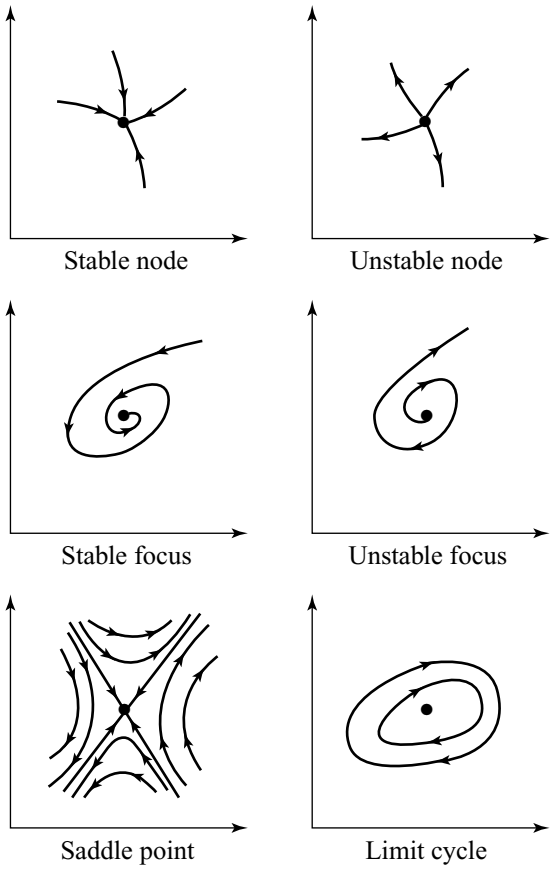


Fig. 3.2 Evolution of a system in the vicinity of the steady state in the phase plane

There are also certain chemical reactions that can be described by the LOTKA-VOLTERRA model. Among the chemical reactions giving rise to oscillatory phenomena, the ZHABOTINSKI reaction represents one of the most spectacular examples. This reaction involves the oxidation of a solution of malonic acid in the presence of cerium, bromate and potassium ions. In a homogenous solution, i.e. stirred, the reaction leads to temporal oscillations of the ions, Ce^{3+} and Ce^{4+}, which are made visible by their different colours; the solution alternates between a red and blue colour.

The diverse chemical or biochemical systems that have just been described are systems in homogenous phase. Under these conditions, the perturbations that give rise to

limit cycles generate temporal oscillations. **In a heterogenous medium, new types of spatial organisation can appear beyond the instability point giving rise to dissipative structures in space**. Thus, the ZHABOTINSKI reaction could be carried out in a heterogeneous medium in a test-tube by imposing a concentration gradient. The temporal oscillations are immediately generated starting at one point in the solution, they then self-propagate causing alternate rings of blue and red colour to appear. ◢

Thus, far from thermodynamic equilibrium, the spatial localisation of perturbations within chemical processes and the free diffusion of the components, which tends to homogenise the medium, can compete to create instabilities and lead to a stable but non-uniform distribution of matter. A new organisation is created by a "rupture of symmetry" (GLANSDORFF & PRIGOGINE, 1971). New types of organisation of matter, such as the spatio-temporal dissipative structures may appear beyond the instability point for systems sufficiently far from equilibrium. The field of thermodynamics beyond equilibrium therefore opens up important paths to the understanding of a great number of phenomena, including biological phenomena, by including the evolution of processes over time and by emphasising the importance of fluctuations.

It seems important to conclude this presentation by quoting GLANSDORFF and PRIGOGINE (1971): "One of the most remarkable aspects of the stability theory is due to its position at the interface between a deterministic description of the behaviour of matter, with the help of macroscopic equations, and the theory of random processes. The actual existence of spontaneous fluctuations is a simple manifestation of the atomic nature of the systems considered. However, in stable systems the occurrence of fluctuations is unimportant, as they regress. Their only effect on the average development is to add a sort of statistical background noise. The situation is radically different when instabilities occur. In this case, the fluctuations are amplified and reach a macroscopic level. When a new stable state is thus reached, a macroscopic description becomes valid once again, whether the state is steady or not. However, even if this be the case, the statistical character of the evolving process through time stays essential because the new stable state that appears may depend on the initial random type of the fluctuation. **We see that the evolution through time of similar systems may only be understood with the help of methods that are both deterministic and stochastic.**"

3.2. EXCHANGE OF MATTER AND ENERGY WITH THE ENVIRONMENT

Living beings exchange not only energy but also matter with the environment. In fact, there is a cycling of matter and energy between the mineral, vegetable and animal worlds. The general cycle can be represented by the schema in Fig. 3.3 below. The sun's light is the source of all biological energy. Thanks to plant photosynthesis, which requires the participation of light quanta, CO_2 and H_2O are

transformed into carbohydrates and oxygen, and then into other nutrient molecules. The biotopes that develop around hydrothermal sources from volcanic faults at the bottom of the ocean are an exception. The life of organisms that constitute these biotopes rests entirely on the capacity of autochemolithotropic bacteria to extract energy from chemical reactions, such as sulphurous oxidation, in order to synthesise organic molecules.

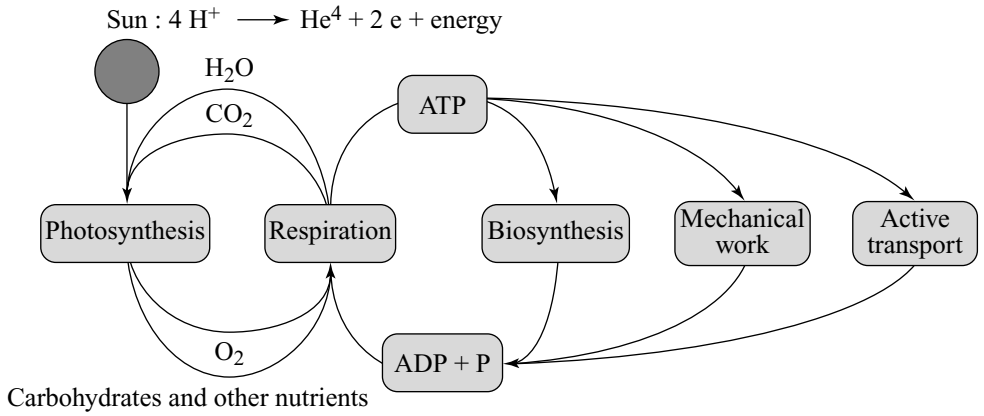


Fig. 3.3 General cycle of matter and energy between the mineral, vegetable and animal worlds

Visible light is a form of radiation or electromagnetic energy. At the extreme temperature found in the sun (6 000 K), a fraction of the enormous quantity of trapped energy within the nuclei of hydrogen atoms is liberated while being transformed by nuclear fusion into helium atoms and electrons:

$$4H \longrightarrow He^4 + 2e + h\nu$$

During this reaction, a quantum of energy is emitted in the form of radiation.

After a series of transformations during which the radiation is absorbed by electrons and atoms, much of the radiation is re-emitted in the form of photons, or quanta of light energy. Therefore, the nuclear fusion reactions in the sun are at the origin of all biological energy on our planet. Also, fuel used by man e.g. coal and petrol, originates from the sun.

During photosynthesis, diverse phenomena occur:

- absorption of radiated energy by chlorophyll,
- conversion of this energy to chemical energy,
- use of chemical energy to reduce CO_2 drawn into the atmosphere and to synthesise glucose. Photosynthesis corresponds broadly to the reaction:

$$6CO_2 + 6H_2O + h\nu \longrightarrow C_6H_{12}O_6 + 6O_2$$

The formation of glucose from CO_2 and H_2O requires considerable energy (686 kcal . mol^{-1}) and hence is a very strongly endergonic process. Thus, a large

input of energy is needed in order for the reaction to occur. The input energy comes from the light energy captured by chlorophyll. From glucose, plants are capable of synthesising a host of other molecules, including proteins, lipids and polysaccharides.

It is interesting to take stock of the quantity of carbon fixed each year on the Earth's surface by photosynthesis: approximately 16×10^9 tons (Table 3.1). This corresponds to the consumption of 45×10^9 people; the current world population is of the order of 6.3×10^9. Yet in the 1970s, coal and petrol consumption due to industrialisation corresponded to about 60×10^9 tonnes of coal per year and this figure has risen considerably since.

Table 3.1 Quantity of carbon fixed in the form of CO_2 by photosynthesis per year on the surface of the globe *(from LEHNINGER, 1969)*

Region	*Surface (10^6 km^2)*	*C fixed (tons/km^2)*	*Tons/an*
Forests	44	250	11×10^9
Cultivated land	27	149	4.1×10^9
Meadows	31	43	1.1×10^9
Deserts	47	7	0.2×10^9

When glucose is metabolised by the animal world, oxygen is used up. Respiration is the inverse operation of photosynthesis:

$$C_6H_{12}O_6 \longrightarrow 6CO_2 + 6H_2O + \text{energy}$$

Each of these processes actually represents a long series of oxidoreduction reactions.

Most biological oxidations take place without the direct participation of oxygen. Since oxygen has an indispensable role in animal function, it was natural to assume that it was directly involved in the oxidation of all carbon-containing substances. In fact, most biological oxidations occur in the absence of oxygen; essentially, only dehydrogenation reactions take place in the presence of an acceptor, which is reduced. Glucose is metabolised by a series of oxidoreduction reactions. The enzymes involved in dehydrogenation reactions are specific dehydrogenases that require a coenzyme to function, such as NAD, FMN or FAD as has already been mentioned (see Chap. 1). The oxidation reactions involved in metabolism form a long chain of dehydrogenation reactions assisted by coenzyme acceptors. For aerobic organisms, the final acceptor is oxygen. For anaerobic organisms, in particular certain bacteria and certain simple organisms, oxygen is poisonous. However, they still use oxidation reactions as an energy source but they replace oxygen by other oxidising agents. The process of oxidoreduction in the absence of oxygen is better known as *fermentation*.

The principal steps in intermediate metabolism leading to the oxidation of a glucose molecule in aerobic organisms are essentially the glucose degradation pathway, or

glycolysis, the tricarboxylic cycle or the KREBS cycle and the respiratory chain or oxidative phosphorylation, which lead to the formation of ATP and the fundamental metabolites. All metabolic reactions use these elements to synthesise all molecules and macromolecules involved in cellular structure and function. Although catalysed by enzymes, the reactions involved in metabolism are not the focus of this book. This brief review aims simply to show that **living organisms are open systems, exchanging both energy and matter with their surroundings and that the study of metabolic fluxes results from the thermodynamics of open systems.**

Thus, the collection of processes that enable energy to be produced from a glucose molecule ensures the cycling of matter and energy between living organisms, open systems, and the environment.

BIBLIOGRAPHY

BOOKS

DE DONDER Th. –1936– *L'Affinité*, Gauthier-Villars, Paris.

DODE P. –1956– *Bases fondamentales et applications de la thermodynamique chimique*, Sedes, Paris.

EDELSTEIN B. –1970– *Doctoral thesis in chemistry-physics*, Université Libre de Bruxelles (cited by GLANSDORFF & PRIGOGINE, 1971).

GLANSDORFF P. & PRIGOGINE I. –1971– *Structure, Stabilité, Fluctuations*, Masson, Paris.

LEHNINGER A.L. –1969– *Bioénergétique*, French edition: M. DUQUESNE, Ediscience, Paris.

ONSAGER L. –1931– *Phys. Rev.* **37**, 405.

PRIGOGINE I. –1967– *Introduction to thermodynamics of irreversible processes*, John Wiley, New York.

PRIGOGINE I. & STENGERS I. –1979– *La Nouvelle Alliance*, Gallimard, Paris.

STRYER L., BERG J.M. & TYMOCZKO J.L. –2002– *Biochemistry*, 5th ed., Freeman and C°, San Francisco.

VOLTERRA V. –1931– *Traité de Mathématiques de la lutte pour la vie*, Gauthier-Villars, Paris.

SPECIALISED ARTICLES

NICOLIS G. & BABLOYANTZ A. –1969– *J. Chem. Phys.* **51**, 2632.

LOTKA A.J. –1920– *J. Am. Chem. Soc.* **42**, 1595.

PART II

KINETICS OF ENZYMES REACTIONS IN SOLUTION

4 – CHEMICAL KINETICS

While thermodynamics provides us with information about the initial and final states of a reaction, kinetics is concerned with the way in which we pass from one to the other, and what is produced between these two states. It describes the reaction pathways and consequently the molecular mechanisms by which a reaction happens. Before tackling enzyme kinetics, it is crucial to remind ourselves of the essential elements of chemical kinetics. We shall only examine the background necessary to understand enzyme kinetics.

4.1. ORDER OF CHEMICAL REACTIONS

4.1.1. FUNDAMENTAL LAW OF CHEMICAL KINETICS

Let us consider the following chemical reaction:

$$nA + mB \longrightarrow pC + qD$$

The rate of a chemical reaction is defined by the rate of appearance or disappearance of one of the molecular species in the reaction. Now, let c be the concentration of the species C that appears during the course of the reaction. If we plot a graph with c on the vertical axis against time on the horizontal axis, and if we consider a time interval Δt, the **average rate** of the reaction is $\Delta c/\Delta t$. At time t_1, the reaction rate is the limit dc/dt of the ratio $\Delta c/\Delta t$, which is therefore the slope of the tangent to the curve at time t (Fig. 4.1a below). Experimentally, more often than not the value measured is the concentration of the product formed. However, in certain cases it is more convenient to follow the disappearance of either the species A or B as a function of time (Fig. 4.1b). The slope is thus negative.

The fundamental law of chemical kinetics expresses the relationship that exists between the reaction rate and the activities of the reactants. Let us recall that activity is related to concentration by means of the equation: $a = \gamma c$; a being the activity, c the concentration and γ the activity coefficient, which is equal to unity in the case of ideal solutions or those at low concentrations, as these tend to ideality (see Chap. 1).

The reaction rate is proportional to the product of the reactants' activities, with each activity raised to a power equal to the number of molecules, or stoichiometric coefficient, of the corresponding reactant featuring in the reaction

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_5, © Springer-Verlag Berlin Heidelberg 2010

equation. This expresses the **law of mass action** (see Chap. 1). The activities can, for the most part, be regarded as equivalent to concentrations:

$$v = k\, a^n\, b^m$$

$a^n b^m$ represents the product of the concentrations modified by the stoichiometric coefficients; k is the specific rate constant. The **global order of the reaction** is (n + m), which is the number of molecules reacting in the chemical event. The **partial order**, i.e. relating to only a single molecular species, is n with respect to A, and m with respect to B.

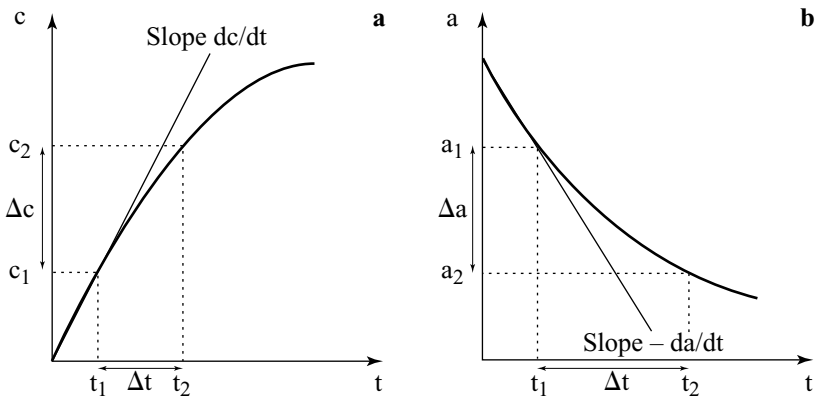


Fig. 4.1 Determination of the reaction rate from (a) the appearance of C over time (b) the disappearance of A over time

This law applies to a stoichiometric reaction or an elementary action. The experimenter who carries out a study of the rate of a given reaction is situated at the macroscopic scale and the reaction order that he or she determines may not correspond to the molecularity of this reaction. Some examples of this shall be given later on. The experimental data sometimes lead to fractional orders, which always indicates that the reaction is complex. For many simple reactions, however, order and molecularity coincide. The aim of kinetic studies is to establish, by experiment, the diverse elementary events that take place in a reaction. This involves the determination of both the reaction order and the specific rate constant (or constants), k.

4.1.2. DETERMINING THE ORDER OF A REACTION

Many chemical reactions fit into one of the two simplest categories: they are either first or second order. However, when a study is initiated, the order of the reaction is, of course, unknown. There are two ways to consider the order of a reaction:

- with respect to time,
- with respect to concentration.

The **determination of the reaction order with respect to time** involves following the change in reaction rate over time for given, initial concentrations of reactants (Fig. 4.2a). At the start, the rate depends on the initial conditions, but the concentrations alter over time and therefore the rate varies over time.

Let us consider the reaction: $nA \longrightarrow n'A'$

The rate equation is: $v = -da/dt = k_n a^n$

and in its logarithmic form:

$$\log v = \log(-da/dt) = \log k_n + n \log a$$

Thus, we first determine the rates at different times of the reaction, and then on a graph plot the logarithm of the reaction rates as a function of the logarithm of the reactant concentration over time. In this way, we obtain a straight line whose slope n gives the order of the reaction; the vertical-axis intercept gives the value of $\log k_n$ (Fig. 4.2b).

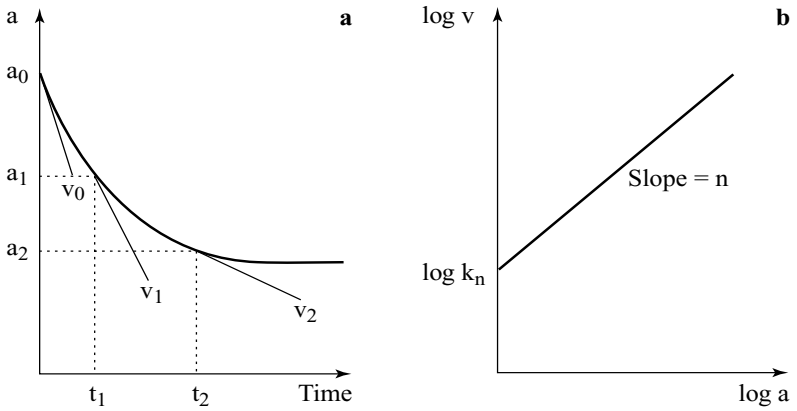


Fig. 4.2 Determination of the reaction order with respect to time

*(**a**) change in concentration of A over time and rate determination at time t_0 (v_0), t_1 (v_1) and t_2 (v_2) – (**b**) plot of the logarithm of the rate as a function of the logarithm of the concentration a*

In order to determine the **order with respect to the concentration or the initial reaction order,** it is necessary to determine the initial reaction rate for different concentrations of reactants:

$$v_0 = -(da/dt)_{initial} = k_n a_0^n$$

and in logarithmic form: $\log v_0 = \log k_n + n \log a_0$

a_0 being the initial concentration of the species A.

The expression so obtained is similar to the previous one, but in this case, by varying the initial concentration, the initial rate is determined experimentally rather than by obtaining the rates at different times of the reaction. Figure 4.3a below

shows the determination of the initial rates and Fig. 4.3b, the order and rate constant for the reaction.

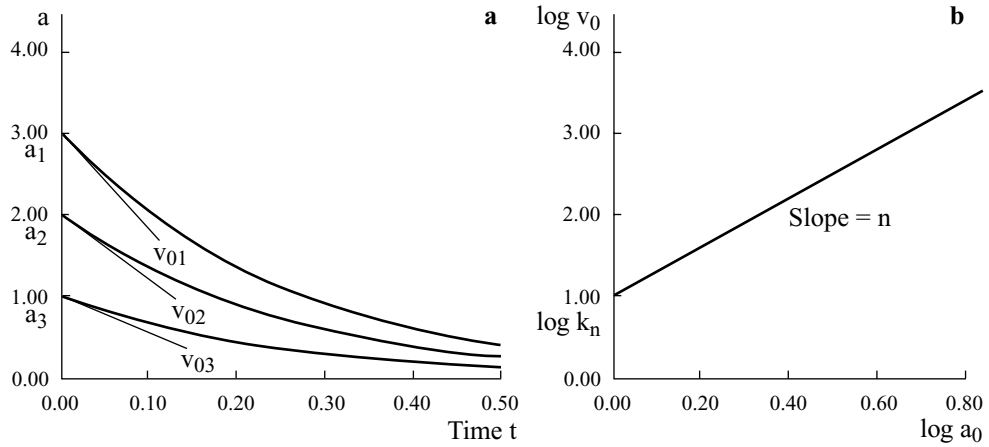


Fig. 4.3 Determination of the reaction order with respect to the concentration, or initial order

*(**a**) determination of the initial reaction rates for different concentrations of A – **(b)** logarithmic representation, as in Fig. 4.2b*

It may happen that the reaction order with respect to the concentration and the order with respect to time do not coincide. This is observed if the reaction products, or the intermediate products that may be formed, interfere with the reaction. This indicates that the reaction scheme is more complex than a reaction of order n. It is therefore equally important to determine the order with respect to time as to the concentration.

If, in a first approach, the experimenter wishes simply to determine the reaction order with respect to time, an essential precaution should be taken: the reaction should be left sufficiently long (at least 80% of the reaction progress) to be certain of its order. However, it is preferable to study at least two initial concentrations of reactant, which will enable concurrent determination of the order with respect to both time and concentration.

4.1.3. FIRST-ORDER REACTIONS

A first-order reaction is characterised by the disappearance of a molecular species A, the chemical event involving only a single molecule. The reaction is either an intramolecular transformation:

$$A \xrightarrow{k_1} A'$$

or a decomposition:

$$A \xrightarrow{k_1} A' + B$$

By applying the law of mass action, we can write the reaction rate, which at each instant is proportional to the concentration of the component undergoing transformation:

$$v = -da/dt = k_1 a$$

After integrating this equation, we obtain:

$$-\ln a = k_1 t + C$$

The constant of integration, C, can be determined from the initial conditions: for $t = 0$, $a = a_0$ (a_0 being the initial concentration of the component A). Thus, we can write the equation for first-order reactions:

$$k_1 t = \ln a/a_0 = 2.3 \log a/a_0$$

The law expressing the change in the concentration of A over time is:

$$a = a_0 e^{-k_1 t}$$

Sometimes it is preferable to express the kinetics of a reaction by the law of product appearance, in particular when the concentration of this product is obtained directly by experiment. Its concentration is $(a_0 - a)$, giving the relationship:

$$a_0 - a = a_0 (1 - e^{-k_1 t})$$

Figure 4.4 illustrates the different graphical representations corresponding to first-order reactions. The first-order constant, k_1, is independent of the units in which concentrations are expressed, and has the dimension of inverse time, expressed as s^{-1}. It represents the inverse of the time necessary for a fraction a_0/a of the initial substance to be converted, such that $\ln (a_0/a) = 1$.

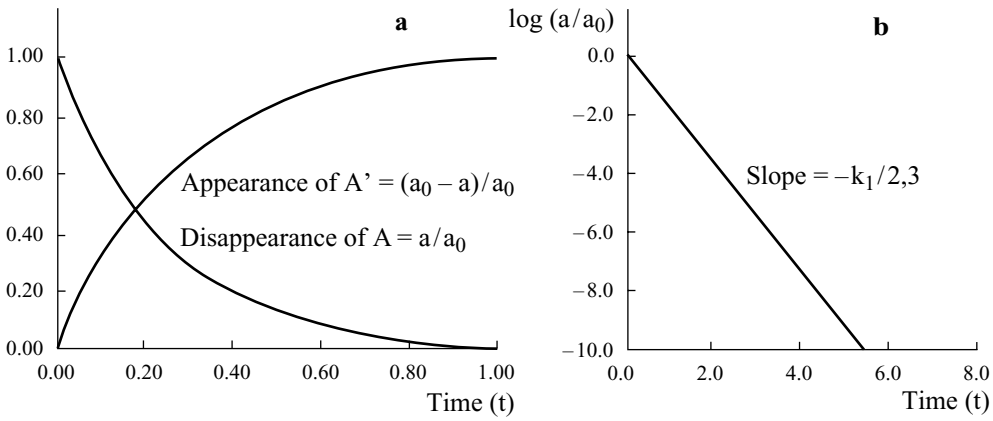


Fig. 4.4 First-order reactions

*(**a**) graphs showing the disappearance of A and the appearance of A' as a function of time – (**b**) semi-logarithmic plot enabling the determination of the rate constant k_1*

We call the **initial rate of a first-order reaction the product of the constant k_1 and the initial concentration of the substance being transformed.**

First-order reactions are very common in biochemistry. Protein denaturation (thermal denaturation, for example) generally occurs by a first-order process. If E_a is the concentration of active enzyme remaining (in native form) after either heating at a given temperature or incubation in the presence of a denaturing agent, for a time t, then the denaturation rate is given by:

$$-dE_a/dt = k_1 E_a$$

giving:

$$E_a = E_0 e^{-k_1 t}$$

and in logarithmic form:

$$2.3 \log E_a/E_0 = -k_1 t$$

There are other first-order reactions that are of great importance to the biochemist: those that use radioactively labelled molecules. The decay rate of radioactive molecules is expressed by first-order kinetics:

$$-dn/dt = k n$$

where n is the number of labelled molecules at time t and n_0 their initial number:

$$k t = 2.3 \log n_0/n$$

Generally, it is necessary to know the corresponding half-life periods. One half-life is the average time necessary for the number of radioactively labelled molecules to equal half the initial number, $n = n_0/2$ and:

$$t_{½} = (2.3 \log 2)/k = 0.693/k$$

It is important to bear this in mind when performing experiments involving labelled molecules whose half-lives are not infinitely longer than the duration of the experiment, in particular ^{32}P-labelled substrates, but not those using ^{14}C. Table 4.1 lists the half-life times for the isotopes routinely used in biochemistry.

Table 4.1 Half-lives of some radioelements

Radioisotope	*Half-life*	*Maximum energy of the particles (MeV)*
^{3}H	12.3 years	0.018
^{14}C	5 570 years	0.155
^{35}S	87 days	0.165
^{32}P	14.2 days	1.71

4.1.4. REVERSIBLE FIRST-ORDER REACTIONS

Reversible first-order reactions correspond to an intramolecular transformation, thus:

$$A \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} A'$$

From the law of mass action we can determine the reaction rate at any moment:

$$v = -da/dt = k_{12}a - k'_{21}a' = k_{12}a - k_{21}(a_0 - a)$$

When equilibrium is reached, this rate cancels itself out and the equilibrium constant can be defined as being equal to the ratio of the concentrations, at equilibrium, of A and A':

$$K_{eq} = a'_e/a_e = (a_0 - a_e)/a_e = k_{12}/k_{21}$$

By introducing the value of a_0 obtained from this expression into the rate equation, we have:

$$-(k_{12} + k_{21})\,t = 2.3 \log (a_0 - a_e)/(a - a_e)$$

and in its exponential form:

$$(a - a_e)/(a_0 - a_e) = e^{-(k_{12} + k_{21})t}$$

It appears, therefore, that from the logarithmic plot (Fig. 4.5) the sum of the two constants k_{12} and k_{21} can be determined. Furthermore, by knowing the equilibrium constant, we may then calculate the individual constants k_{12} and k_{21}.

Reversible first-order reactions occur very frequently in biochemistry. Previously, we gave the example of protein denaturation, which is mostly a reversible process. The appearance of reversible structural transitions within proteins may be influenced by various factors: raising the temperature, changes of pH, action of denaturing agents, diverse effectors etc. Different methods are available to follow the time-course of conformational changes in proteins. The change in a physical parameter may be studied; alternatively, in the case of an enzyme, the change in catalytic activity over time may be monitored. When the structural transition is reversible, equilibrium will be reached after a certain period of time. At equilibrium, part of the protein will remain in the native state. From this moment on, no further changes with time are observable. The primary concern of a kineticist, therefore, is to determine whether or not the reaction reaches equilibrium. Once the existence of an equilibrium has been established, we may deduce the apparent first-order constant, $k_{app} = k_{12} + k_{21}$. It is essential to know the equilibrium constant, K_{eq}, in order to calculate the individual constants.

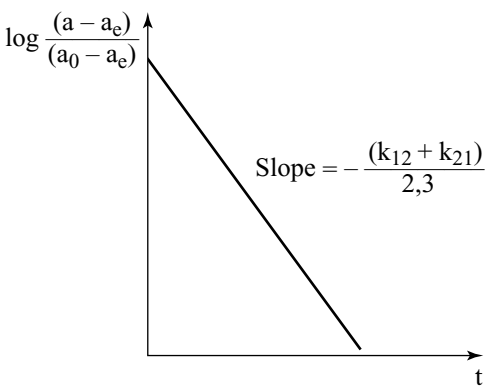


Fig. 4.5 Reversible first-order reaction Determination of the apparent rate constant, k_{app}

4.1.5. SIMULTANEOUS FIRST-ORDER REACTIONS

Reactions that cannot be described by simple first-order kinetics, but which correspond to the sum of several first-order reactions, are often encountered in biochemistry:

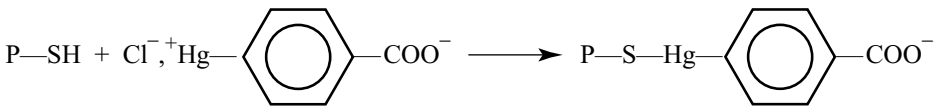
$$a = \sum_{i=1}^{n} A_n e^{-k_i t}$$

for the most general case.

When the above expression is limited to two terms, and if k_2 is small compared to k_1, we have:

$$A = Ae^{-k_1 t} + B\, e^{-k_2 t}$$

This kinetic behaviour is frequently observed during the action of a chemical reagent on protein groups as, for example, with parachloromercuribenzoate (PCMB) activity on enzyme sulphydryl (SH) groups. The reaction is as follows:



Several different SH groups may exist in a protein depending on their positions in the structure. The most accessible groups react the most readily.

Let us give an example: the titration of the SH groups on aspartate amino transferase (or glutamate oxaloacetate transaminase), an enzyme that catalyses the transfer of amine groups from glutamate to oxaloacetate according to the reaction:

L-glutamate + oxaloacetate ⟶ α-ketoglutarate + L-aspartate

The kinetics of the reaction of PCMB with the SH groups on the enzyme are presented in Fig. 4.6 opposite, which shows clearly the biphasic nature of the reaction. The curve can be decomposed into two first-order kinetic phases. In fact, **the reaction kinetics are pseudo-first order since PCMB is in excess and its concentration is not significantly modified during the reaction.**

Now, let n_0 be the number of SH groups in the first category reacting with a constant k_1. The number of free SH groups as a function of time would be given by the equation:

$$n = n_0\, e^{-k_1 t}$$

If m_0 is the number of SH groups in the second category reacting with a constant k_2, similarly, we have:

$$m = m_0\, e^{-k_2 t}$$

This is the total number of free SH groups measurable over time. The experimental curve corresponds to:

$$n + m = n_0\, e^{-k_1 t} + m_0\, e^{-k_2 t}$$

We can determine k_2 and k_1 if k_2 is small enough compared to k_1. The curve becomes linear after a certain period indicating that all of the "fast" SH groups have already reacted; only the "slow" SH groups contribute to the linear part of the curve. From the slope of this line, k_2 may be derived and we notice that only a single group belongs to this category. The changes in m as a function of time are therefore determined. By subtracting them from the experimental curve we obtain a new straight line corresponding to the changes in n, and thus can deduce k_2 and n_0. In this way, we find that three groups react with a rate constant $k_1 = 4\ h^{-1}$ and one group with a rate constant $k_2 = 0.2\ h^{-1}$.

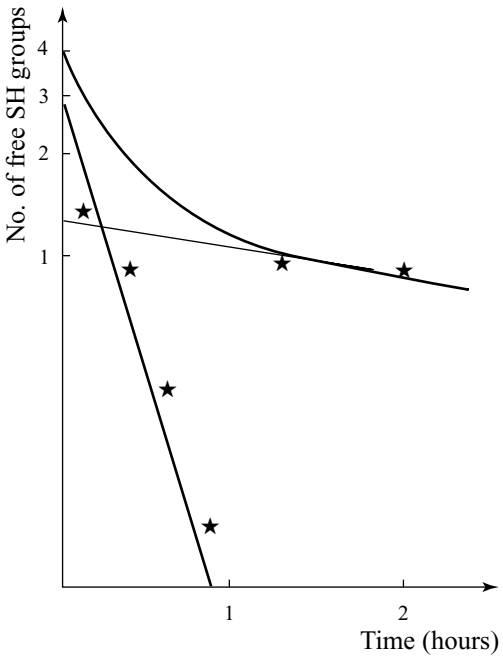


Fig. 4.6 Titration of the SH groups on aspartate amino transferase
An experimental curve *showing the number of free SH groups per protomer (*vertical axis*,* logarithmic scale*) plotted against time, in hours (*horizontal axis*). The* straight lines *correspond to the decomposition of this curve for two categories of group*

4.1.6. SECOND-ORDER REACTIONS

A second-order reaction has the form: $A + B \xrightarrow{k_2} C$.

By using the law of mass action, the reaction rate may be written:

$$v = -da/dt = k_2\, ab$$

which, at any given moment, is proportional to the concentrations of the two components A and B. The concentrations a and b of the two molecular species A and B diminish over the course of the reaction and give rise to a new species C whose concentration is c at all times. The quantity of A that disappears is equal to the quantity of B that disappears and to the quantity of C that appears:

$$a_0 - a = b_0 - b = c$$

In order to solve this system it is convenient to consider two cases in which the concentrations of A and B either differ initially ($a_0 \neq b_0$), or are equal ($a_0 = b_0$).

4.1.6.1. FIRST CASE: $a_0 \neq b_0$

Let d_0 be the difference in concentration of the initial conditions, where $d_0 = b_0 - a_0$. The rate equation may be written in the following form:

$$-da/dt = k_2\, a\, (a + d_0)$$

or:

$$-da/a\,(a + d_0) = k_2\, dt$$

To integrate, we first rearrange to give the sum of two fractions:

$$\frac{1}{d_0}\left[\frac{da}{a} - \frac{da}{d_0 + a}\right] = -k_2\, dt$$

By integrating and taking into account the initial conditions for the determination of the constant of integration, we finally obtain the expression corresponding to second-order reactions:

$$k_2 t = \frac{1}{(b_0 - a_0)} \ln \frac{a_0 b}{b_0 a}$$

In a second-order reaction, the specific rate constant has dimensions of the inverse of the product of time and concentration.

$$k_2 = \text{concentration}^{-1} \times \text{time}^{-1}$$

It is generally expressed as $\text{mol}^{-1} \,.\, \text{s}^{-1}$.

4.1.6.2. SECOND CASE: $a_0 = b_0$

If the concentrations of the two reactants are identical, the previous method of calculation cannot be used as it would lead to an indeterminate value, since $d_0 = 0$. In any case, the calculation is simpler as the rate equation becomes:

$$-da/dt = k_2\, ab = k_2\, a^2$$

Integrating this expression gives:

$$\frac{1}{a} - \frac{1}{a_0} = k_2 t$$

the constant of integration is defined as before by the initial conditions.

4.1.7. REVERSIBLE SECOND-ORDER REACTIONS

Let:

$$A + B \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} C$$

The reaction in the direction from A + B leading to C is second order. The reverse reaction, in the direction from C leading to A + B is first order. The reaction rate is given by the relationship:

$$v = -da/dt = dc/dt = k_{12}\, ab - k_{21}\, c$$

At equilibrium, $dc/dt = 0$ and the concentrations of the reacting species are a_e, b_e and c_e. At any point, the following relations hold:

$$a = a_0 - c \quad \text{and} \quad b = b_0 - c$$

As the equilibrium relationship is:

$$c_e/a_e b_e = k_{12}/k_{21} = K_{eq}$$

the rate expression becomes:

$$v = dc/dt = k_{12}[ab - (ca_e b_e/c_e)]$$

which is equivalent to:

$$\frac{dc}{(c_e - c)(a_0 b_0 - c_e c)} = k_{12}\frac{dt}{c_e}$$

Now, the integral of $dx/[(x - b)(x - a)]$ is:

$$\int \frac{dx}{(x-a)(x-b)} = \frac{1}{(a-b)} \ln \frac{x-a}{x-b}$$

giving:

$$k_{12}t + Ct = \frac{c_e}{(a_0 b_0 - c_e^2)} \ln \frac{(a_0 b_0 - c c_e)}{c_e (c_e - c)}$$

by taking into account the constant of integration, C, obtained when $t = 0$, we derive the final expression:

$$k_{12}t = \frac{c_e}{(a_0 b_0 - c_e^2)} \times 2.3 \log \frac{c_e (a_0 b_0 - c c_e)}{a_0 b_0 (c_e - c)}$$

4.1.8. DIMERISATION EQUILIBRIUM

We feel it is important to treat separately here the particular case of dimerisation equilibrium, which often arises with proteins.

$$2\,M \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} D$$

The reaction rate is given by the relationship:

$$v = -\tfrac{1}{2}\, dm/dt = dd/dt = k_{12}\, m^2 - k_{21}\, d$$

At equilibrium we have: $k_{12}\, m_e^2 = k_{21}\, d_e$

or: $k_{12}/k_{21} = K_{eq} = d_e/m_e^2$

By carrying over to the rate equation:

$$\frac{-dm}{dt} = 2k_{12}\left[m^2 - \frac{1}{2K_{eq}}(m_0 - m)\right]$$

with $(m_0 - m)/2 = d$. The equation to the second degree has two roots:

$$m = \frac{1}{4K_{eq}}\left(-1 \pm \sqrt{1 + 8m_0K_{eq}}\right)$$

the rate equation can now be put in the form:

$$\frac{-dm}{\left[m + \frac{1}{4K_{eq}}\left(1 + \sqrt{1 + 8m_0K_{eq}}\right)\right]\left[m + \frac{1}{4K_{eq}}\left(1 - \sqrt{1 + 8m_0K_{eq}}\right)\right]} = 2k_{12}dt$$

By integrating, we obtain:

$$k_{12}t = 2.3\frac{1}{4K_{eq}\sqrt{1 + 8m_0K_{eq}}}\log\frac{m + \frac{1}{4K_{eq}}\left(1 - \sqrt{1 + 8m_0K_{eq}}\right)\left[m_0 + \frac{1}{4K_{eq}}\left(1 - \sqrt{1 + 8m_0K_{eq}}\right)\right]}{m + \frac{1}{4K_{eq}}\left(1 + \sqrt{1 + 8m_0K_{eq}}\right)\left[m_0 - \frac{1}{4K_{eq}}\left(1 + \sqrt{1 + 8m_0K_{eq}}\right)\right]}$$

4.1.9. ZERO-ORDER REACTIONS

The kinetics are zero order when the reaction rate is independent of the concentration of the molecular species being converted, and consequently, independent of time. The rate is constant:

$$-dc/dt = k_0$$

These kinetics appear to be contrary to the fundamental law of the effect of concentrations. In reality, **when a reaction is zero order, the measured rate is actually the rate of conversion of an intermediate component whose concentration stays constant**.

Zero-order kinetics are important in enzymology. As we shall see, enzymatic reactions are zero order under certain conditions. It is the same for chemical reactions that take place on a surface to which the reactants are adsorbed. The molecular species, whose conversion controls the reaction rate, is in this case the adsorbed molecule and its superficial concentration is appreciably constant if, through the course of the reaction, practically all of the reactive surface points are occupied. When a protein undergoes surface denaturation, the reaction is zero order. In an enzymatic reaction, when all of the enzyme is saturated by the substrate, the reaction is zero order with respect to the substrate concentration, as we shall discover later on.

4.1.10. SIGNIFICANCE OF THE REACTION ORDER: ORDER AND MOLECULARITY

We have seen that, for zero-order reactions, the order does not correspond to the molecularity. Similarly, when a reaction is first order, this does not mean that it is monomolecular, i.e. that only a single molecular species exists, whose conversion is measured during the reaction. The classic example of the acid hydrolysis of saccharose is often given, the kinetics are first-order although the reaction is bimolecular:

$$C_{12}H_{22}O_{11} + H_2O \longrightarrow 2\ C_6H_{12}O_6$$

But water, which is the second molecule taking part in the reaction, is in such excess in the medium that its concentration does not vary significantly during the hydrolysis. Consequently, the kinetics are first order, since they are solely dependent on the saccharose concentration. This type of reaction is said to be **pseudo-first order**. Similarly, when a group G (SH or histidine, for example) of a protein is titrated with an excess of a chemical reagent R:

$$G + R \xrightarrow{k} G'$$

The reaction rate is given by the equation:

$$v = k\,r\,g$$

where g is the concentration of the titratable group. If r, the concentration of reagent, remains constant throughout the reaction, we define experimentally a constant of pseudo-first order k' such that:

$$k' = k\,r$$

and the rate becomes: $v = k'g$

4.2. ACTIVATION OF MOLECULES

4.2.1. ACTIVATION ENERGY

The rate of chemical reactions increases with rising temperature. At absolute zero, no reaction is possible. As VAN T'HOFF remarked previously, if all molecules were in the same state, they would be converted at the same instant. Now, among all the molecules in an evolving system, only a fraction is capable of reacting at any given moment; these are **activated molecules**.

In order for molecules to react, they must first pass through an unstable transition state, which is **the activated state**. The molecules must acquire sufficient energy to overcome a barrier of potential energy, termed the **activation energy**. The

difference in potential energy between the activated state and the initial state represents the experimental activation energy, E.

Let us consider the following intramolecular conversion:

$$A \longrightarrow A'$$

The reaction pathway would then be:

$$A \rightleftharpoons A^* \longrightarrow A'$$

As deduced from statistical considerations, the number of molecules having acquired the necessary energy to become activated is proportional to $e^{-E/RT}$. This number increases as the temperature rises. R is the ideal gas constant (equal to 1.98 kcal . mol^{-1} . K^{-1}) and T, the absolute temperature (273 + t°C).

According to kinetic theory, in order for molecules to react together, they must meet; thus, activation results from the collisions between molecules. However, not every collision is efficient. The collision theory relates the reaction rate to the activation energy:

$$k = ZP\, e^{-E/RT}$$

where P is the probability of an efficient collision and Z the number of collisions. The number of collisions per second can be calculated, in a first approximation, by adopting the results from the kinetic theory of gases:

$$Z = \frac{\sigma_1 + \sigma_2}{2}\left[2\pi kT\left(\frac{1}{M_1} + \frac{1}{M_2}\right)\right]^{1/2}$$

In this equation σ_1 and σ_2 are the diameters of the colliding molecules; M_1 and M_2 are their respective molecular masses. However, this expression – developed for ideal gases – does not take into account the effect of solvent and therefore introduces a substantial source of error.

Due to the inadequacy of the collision theory, it is preferable to use the **theory of absolute rates developed by EYRING. The theory of absolute rates assumes the existence of a statistical equilibrium between molecules in an activated state and those in the initial state.** The reaction rate is determined by the number of activated molecules that cross a certain critical region, which represents the peak of the energy barrier (Fig. 4.7a opposite). We can thus define an equilibrium constant between the two forms of the molecules:

$$K^* = A^*/A$$

The theory of absolute rates provides a relationship between the equilibrium constant, K*, and the specific rate constant, k, for the reaction:

$$k = \kappa \frac{k_B T}{h} K^*$$

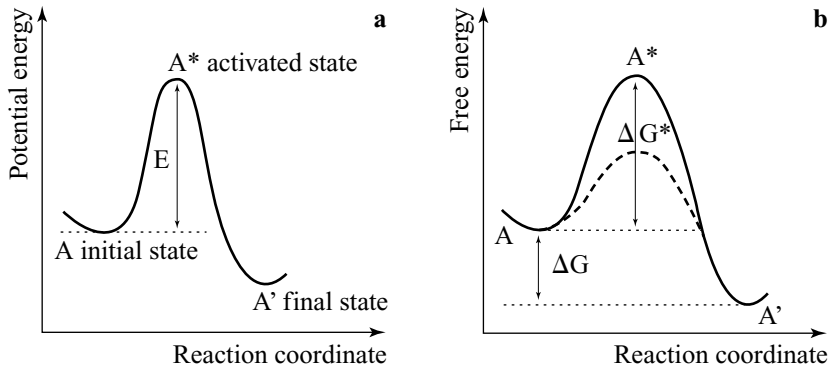


Fig. 4.7 (a) potential energy change as a function of the reaction coordinate (b) free energy profile;* dashed line*: the reaction takes place in the presence of a catalyst, there is a lowering of the energy barrier

In this expression k_B is BOLTZMANN's constant; h, PLANCK's constant, T the absolute temperature and κ, the transmission coefficient. Equal to unity in the majority of cases, this coefficient was introduced to account for the fact that, in certain reactions, the activated molecules may return to the initial state. k_BT/h is a universal constant, which, at 25°C (298 K), is equal to 6.25×10^{12} and represents the rate at which the activated molecules can surmount the peak of the energy barrier. The inverse of this value represents the lifetime of the activated molecules at the energy-barrier peak, roughly 10^{-13} s at about 25°C. The activated molecules, therefore, have an extremely short lifetime, which is immeasurable by experiment. The rate at which the energy barrier is exceeded corresponds to the vibrational frequency of the bond to be broken. This frequency is given by the quantum mechanical equation, $E = h\nu$; the corresponding energy in classical physics is $E = k_BT$, giving: $\nu = k_BT/h$.

Having defined activation equilibrium, we can now apply to it the laws of thermodynamics:

$$\Delta G^* = -RT \ln K^* = -RT \ln kh/k_BT$$

giving the value of k:

$$k = \frac{k_BT}{h} e^{-\Delta G^*/RT} = \frac{k_BT}{h} e^{-\Delta H^*/RT} e^{\Delta S^*/R}$$

ΔH^* is the change in activation enthalpy and ΔS^* the change in activation entropy of the reaction.

This expression clearly shows that the reaction rate depends on terms that encompass both the enthalpy and entropy of activation, and which carry different signs. By taking the logarithm of the rate constant with respect to the inverse of the absolute temperature, we obtain:

$$d \ln k/d(1/T) = -(\Delta H^* + RT)/R = -E/R$$

The experimental activation energy differs from the enthalpy by the quantity RT, which, at temperatures typical in biochemical reactions, is of the order of 0.6 kcal . mol^{-1}.

Experimentally, it is straightforward to determine the energetic parameters of a chemical reaction. ΔG^* can be calculated from the previous relation by using the value of the constant k obtained experimentally at a given temperature.

In order to derive ΔH^* it is necessary to determine the value of the rate constant, k, at different temperatures. We then make use of the ARRHENIUS plot (Fig. 4.8a), which involves plotting log k as a function of the inverse of the absolute temperature 1/T. The slope gives –E/2.3 R. From the value of the activation energy E, we can deduce ΔH^* ($\Delta H^* = E - RT$). The value E is independent of temperature. Knowing ΔG^* and ΔH^*, we can calculate ΔS^* by means of the relationship:

$$\Delta G^* = \Delta H^* - T\,\Delta S^*$$

The values of the energetic parameters of a particular reaction can provide us with important information regarding the reaction mechanism. We shall provide a few examples for enzymatic reactions.

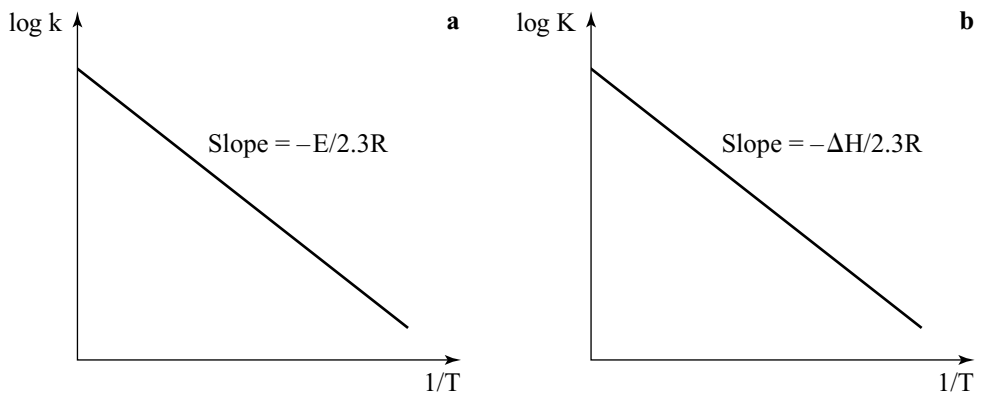


Fig. 4.8 (a) ARRHENIUS plot, from which the activation energy of a reaction can be determined (E = ΔH* + RT) - (b) VAN T'HOFF plot, for determining the enthalpy change of an equilibrium from the changes in the equilibrium constant K as a function of temperature

4.2.2. CATALYSED REACTIONS – ROLE OF THE CATALYST

Up to this point, we have only considered uncatalysed reactions. Before embarking upon enzyme catalysis, we shall briefly recapitulate the concept of catalysis. **A catalyst accelerates the reaction rate.** Furthermore, a catalyst is present in very low quantity relative to the reactants and at the end of the reaction it is found in the

same concentration and state as at the beginning. Even if a temporary bond forms between the catalyst and one of the molecular species, the catalyst only increases the rate of chemical conversion, but does not modify the initial or the final states. **A catalyst does not change the thermodynamic equilibrium.** If the reversible transformation:

$$A \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} A'$$

takes place in the presence of a catalyst, the equilibrium constant $K_{eq} = k_{12}/k_{21}$ will not be modified; k_{12} and k_{21} increase by the same proportion.

The catalyst neither changes the initial nor the final states, but merely lowers the energy barrier that must be overcome in order to pass from one state to the other. It provides the reaction with a more accessible pathway than would be available in its absence. **The catalyst lowers the activation energy of a reaction** (see Fig. 4.7b).

Different types of catalysis exist: homogenous catalysis, which takes place in solution, and heterogenous catalysis, which takes place on the surface of a disperse substance. In the latter case, the reacting molecules are adsorbed to the catalyst's surface. Whatever the exact mechanisms of catalysis, in general, catalysts either weaken strong substrate bonds or stabilise certain partially formed bonds within activated complexes.

5 – KINETICS OF ENZYMATIC REACTIONS WITH MICHAELIAN BEHAVIOUR

We shall now consider more specifically the kinetic behaviour of reactions catalysed by enzymes. As we have previously remarked for the general case of catalysts, enzymes are not consumed during the course of the reactions that they catalyse and they do not alter the equilibrium constant.

Let us study the conversion of A into B:

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} B$$

whose constants k_1 and k_{-1} are equal to 10^{-3} and 10^{-4} respectively. At equilibrium:

$$K_{eq} = \frac{(B)}{(A)} = \frac{k_1}{k_{-1}} = \frac{10^{-3}}{10^{-4}} = 10$$

B is ten-fold more concentrated than A whether or not the enzyme is present. The enzyme accelerates the reaction by the same factor in one direction as it does in the reverse direction. The efficiency of enzymes is very variable; some are capable of accelerating a chemical reaction by a factor of 10^8 or even up to 10^{11}.

Enzyme catalysis facilitates the conversion of a substrate S into a product P:

$$S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} P$$

This conversion proceeds through the association of a substrate molecule to the enzyme, i.e. by the formation of at least one enzyme-substrate complex, ES. Substrate binding takes place in a precise location within the protein, at the active site of the enzyme. The first authors on the subject (see the introduction) considered that an enzyme-catalysed reaction comprised two steps, which are the reversible formation of a stereospecific enzyme-substrate complex, followed by the decomposition of this complex, the appearance of the reaction products and the regeneration of the active enzyme, according to the simple scheme:

$$E + S \rightleftharpoons ES \longrightarrow E + P$$

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_6, © Springer-Verlag Berlin Heidelberg 2010

Before going into the details of the kinetic mechanisms, let us first consider the general phenomenological aspect of these reactions.

5.1. EVOLUTION OF ENZYMATIC REACTIONS: PHENOMENOLOGICAL ASPECTS

5.1.1. VARIATION IN THE QUANTITY OF PRODUCT FORMED AS A FUNCTION OF TIME

Figure 5.1 shows how the quantity of product formed evolves over time. The kinetics reveal several phases.

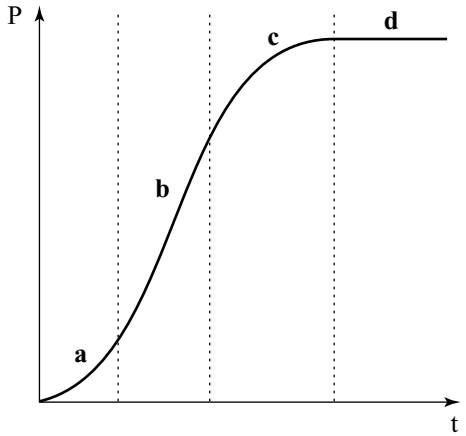


Fig. 5.1 Kinetics of the appearance of the reaction products as a function of time showing the different phases
(a) *pre-steady state phase*
(b) *steady state phase*
(c) *phase of inhibition by the products*
(d) *equilibrium phase.*

5.1.1.1. PRE-STEADY STATE PHASE

This is a very short phase during which the first molecules of the ES complex are formed until the concentration of this intermediate complex reaches a constant value (steady state phase). The pre-steady phase only lasts a fraction of a second and is not detectable during kinetic experiments carried out under classic conditions. Study of this phase requires the use of fast techniques, which we shall look at later.

5.1.1.2. STEADY STATE PHASE

During the steady state phase the rate of appearance of the product P is constant. The MICHAELIS-MENTEN theory predicts that under these conditions the concentration of the enzyme-substrate complex, ES, is constant.

5.1.1.3. PHASE OF INHIBITION BY THE REACTION PRODUCTS

In this phase the concentrations of the reaction products are no longer negligible and, consequently, the reverse reaction tends to lower their concentrations. Of

course, if the equilibrium constant is large, i.e. hugely in favour of the formation of P, the reverse reaction will be negligible.

5.1.1.4. EQUILIBRIUM PHASE

During this phase equilibrium is reached. The quantities of S and P are constant. In these conditions:

$$k_1(S) = k_{-1}(P)$$

The amount of P formed is equal to the amount of S converted.

5.1.2. MICHAELIS-MENTEN THEORY

The MICHAELIS-MENTEN theory accounts for these phenomena in conditions where the enzyme concentration is very low compared to the substrate concentration. Consider the following simple reaction:

$$\begin{array}{ccccc} & & k_1 & & k_2 & \\ E + S & \rightleftharpoons & ES & \rightleftharpoons & E + P \\ & & k_{-1} & & k_{-2} & \\ e \quad s & & (ES) & & p \end{array}$$

The respective concentrations of the species E, S, ES, and P are e, s, (ES) and p.

While the concentration of P is low relative to that at equilibrium, i.e. while in the initial conditions of the steady state phase, we may consider the rate of catalysis to be equal to the product of the concentration of the ES complex and the rate constant for its conversion:

$$v = \frac{dp}{dt} = -\frac{d[s+(ES)]}{dt} = k_2(ES) \qquad [1]$$

The rates of formation and disappearance of ES are as follows:

- rate of formation: $k_1 es$,
- rate of disappearance: $(k_{-1} + k_2)(ES)$.

Throughout the steady state phase, the concentration of the intermediate ES is constant, thus:

$$\frac{d(ES)}{dt} = 0$$

Consequently, the rate of product appearance is equal to the rate of substrate disappearance:

$$\frac{dp}{dt} = -\frac{ds}{dt}$$

and the rates of formation and disappearance of ES are equal.

Thus, we obtain: $k_1 es = (k_{-1} + k_2)(ES)$

which can be written: $(ES) = \dfrac{k_1 es}{k_{-1} + k_2}$

The MICHAELIS constant is defined by the relation:

$$K_m = \frac{k_{-1} + k_2}{k_1}$$

giving: $(ES) = \dfrac{es}{K_m}$

Under conditions where the enzyme concentration is low compared to the concentration of substrate to be transformed, the concentration of free substrate in the initial phase of the reaction is: $s = s_0$. The free-enzyme concentration is:

$$e = e_0 - (ES)$$

s_0 and e_0 are, respectively, the total concentrations of substrate and enzyme. By substituting into the previous equation, we have:

$$(ES) = \frac{[e_0 - (ES)]}{K_m} s$$

By solving this equation we obtain the value of (ES):

$$(ES) = e_0 \frac{\dfrac{s}{K_m}}{1 + \dfrac{s}{K_m}} = e_0 \frac{s}{s + K_m}$$

Substituting into equation [1], gives:

$$v = k_2 e_0 \frac{s}{s + K_m} \qquad [2]$$

The maximum rate of the reaction V_m is reached when the active sites of the enzyme are saturated by the substrate, which is ensured when s is much larger than K_m. In this case, the relationship $s/(s + K_m)$, which represents a function of saturation of the enzyme Y_s, tends towards 1 (Fig. 5.2 opposite). The maximum rate is $V_m = k_2 e_0$. By substituting this value into equation [2], we obtain:

$$v = V_m \frac{s}{s + K_m} \qquad [3]$$

This is the MICHAELIS-MENTEN equation.

The saturation function is thus:

$$Y_s = \frac{s}{s + K_m} = \frac{v}{V_m}$$

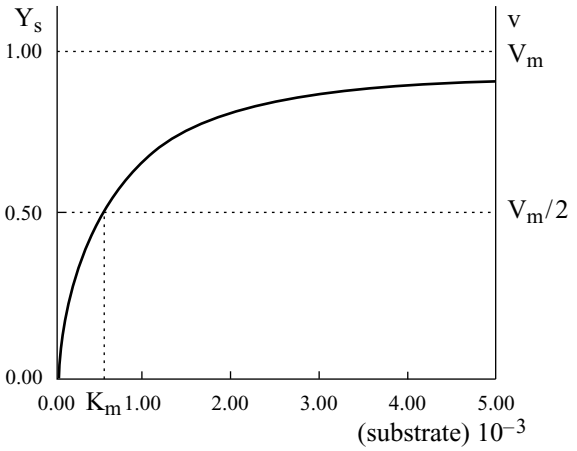


Fig. 5.2 Change in the saturation function Y_s of the enzyme by its substrate

Thus, the fraction Y_s of active sites occupied by the substrate is equal to v/V_m. The maximum rate V_m is therefore the maximum capacity of the enzyme to catalyse the reaction, i.e. the maximum quantity of substrate that can be converted per unit time. **The ratio V_m/e_0 represents the molar activity or turnover number.** Nevertheless, depending on the reaction and the substrate(s), it is possible that this value is never attained experimentally (due to K_m being too high, poor substrate solubility etc.). The MICHAELIS constant, K_m, represents in all cases the substrate concentration at which the rate is half maximal; indeed, for $s = K_m$, $v = V_m/2$ (Fig. 5.2). However, the significance of these kinetic parameters depends essentially on the reaction scheme.

Enzyme reactions are, in fact, more complex than the original authors on the subject were able to foresee. Generally, they comprise more than one substrate and involve more than one intermediate complex. We shall tackle the study of kinetic schemes of increasing complexity and for each, the significance of the kinetic parameters will be discussed.

5.2. ENZYMATIC REACTIONS WITH A SINGLE SUBSTRATE AND A SINGLE INTERMEDIATE COMPLEX

Most enzymatic reactions involve at least two substrates. However, there are borderline cases in which the concentration of one of the substrates is in excess and the system behaves mechanistically like a single-substrate enzymatic reaction. Let us take the example of enzymatic hydrolysis reactions in aqueous solution. Water, the second substrate, is in vast excess and plays no part in the reaction kinetics.

A great number of concepts that remain valid today were established after studying enzymatic reactions with a single substrate and a single intermediate complex. This is why this topic is particularly well developed.

5.2.1. REVERSIBILITY OF ENZYMATIC REACTIONS

Let us consider the simplest borderline case involving a single substrate and a single intermediate complex:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P$$

Enzymatic reactions, like all chemical reactions, are reversible. Nevertheless, the equilibrium can be strongly shifted in favour of the formation or, conversely, the degradation of a given metabolite, such that the reaction is practically irreversible. In certain cases, the reaction products undergo transformations (ionisations for example) that render the reaction quasi-irreversible. Thus, in peptide bond hydrolysis reactions at a pH conducive to protease action:

$$R–CO–NH–R' + H_2O \longrightarrow R–COOH + R'–NH_2$$

where R—COOH is dissociated and R′–NH_2 becomes protonated:

$$R–COOH \rightleftharpoons RCOO^- + H^+$$

and:

$$RNH_2 + H^+ \rightleftharpoons RNH_3^+$$

and so the reverse reaction becomes impossible.

If we consider the reaction scheme under equilibrium conditions, we can write the following equations:

$$k_1(E)(S) = k_{-1}(X)$$

and:

$$k_2(X) = k_{-2}(E)(P)$$

The equilibrium constant, K_{eq}, is defined as follows:

$$K_{eq} = (P)/(S) = k_1k_2/k_{-2}k_{-1}$$

We can show that a relationship exists between the equilibrium constant and the parameters of both the forward reaction (S ⟶ P) and the reverse reaction (P ⟶ S). Indeed, the kinetic parameters of the forward reaction are written:

$$V_{mf} = k_2e_0 \quad ; \quad K_{mf} = (k_{-1} + k_2)/k_1$$

The kinetic parameters of the reverse reaction are:

$$V_{mr} = k_{-1}e_0 \quad ; \quad K_{mr} = (k_{-1} + k_2)/k_{-2}$$

Therefore:

$$\mathbf{K_{eq} = K_{mr}V_{mf}/K_{mf}V_{mr}}$$

This expression is known as the **HALDANE relationship**.

5.2.2. *RATE OF ENZYMATIC REACTIONS: APPROXIMATION TO THE STEADY STATE, APPROXIMATION TO A QUASI-EQUILIBRIUM SIGNIFICANCE OF THE KINETIC PARAMETERS*

Let us now consider the previous scheme but under conditions of quasi-irreversibility:

$$E + S \rightleftharpoons ES \longrightarrow E + P$$

These conditions are practically attained when the equilibrium strongly favours the formation of P, or in the initial reaction conditions when the concentration of P is zero or practically zero, or even when P undergoes subsequent conversions precluding the reverse reaction, as described above.

This scheme, a highly simplified case, was historically the first scheme proposed (V. HENRI, L. MICHAELIS and M. MENTEN), and which J.B.S. HALDANE resolved in 1925 for the general case that assumes a steady state. Indeed, the general equation for enzymatic reactions, i.e. the MICHAELIS-MENTEN equation described above, rests on two assumptions:

- a low enzyme concentration,
- the presence of a steady state.

When the enzyme concentration becomes too high, (ES) is no longer negligible relative to s. There is no longer a first-order reaction with respect to the enzyme, but a more complex form that leads to a second-degree equation with respect to the enzyme, as analysed by STRAUSS and GOLDSTEIN in 1943. The rate equation can no longer be put in a linear form. The simplest procedure, whenever possible in *in vitro* studies, is to work with an enzyme concentration that ensures linearity between the reaction rate and this concentration. Again, it is worthwhile checking this (see later). In reactions that take place in a cellular environment (see Part VI), however, frequently the enzyme concentration becomes high relative to the substrate concentration.

The MICHAELIS equation assumes a steady state, but the steady state is not reached immediately; everything depends on the time-scale of the reaction. If it is possible to work in a zone of small time constants, we can follow the establishment of a steady state system. This topic will be explored in the following sections.

5.2.2.1. *KINETICS IN THE PRE-STEADY STATE*

The study of reaction kinetics in the pre-steady state involves estimating the reaction product during the very short time preceding the steady state, which is characterised by a constant concentration of the ES complex. There is an initial acceleration that lasts only a fraction of a second before the system reaches a steady state.

When the appearance of the reaction product is measured using a rapid kinetics device, we obtain a curve that has the form indicated in Fig. 5.3 below.

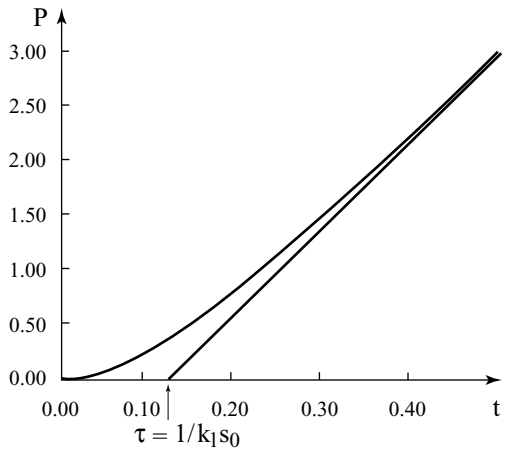


Fig. 5.3 Appearance of the reaction product over time
The initial curvature represents the pre-steady state; the linear part that follows corresponds to arrival at the steady state

From the scheme: $E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$

the rate equations for the formation of ES and P can be written:

$$d(ES)/dt = k_1[e_0 - (ES)]s - (k_{-1} + k_2)(ES)$$

$$v = dp/dt = k_2(ES)$$

and the acceleration: $d^2p/dt^2 = k_2 d(ES)/dt$

By combining these different expressions, we derive a second-order differential equation:

$$d^2p/dt^2 + dp/dt\,(k_1s + k_{-1} + k_2) - k_1k_2e_0s = 0$$

The solution to this equation is valid for both the pre-steady state and the steady state in the first period of the reaction as long as the concentration of s is little different from s_0, the initial substrate concentration:

$$p = \underbrace{\frac{k_2e_0s_0}{s_0 + \dfrac{k_{-1} + k_2}{k_1}}t}_{\text{I}} + \underbrace{\frac{k_1k_2e_0s_0}{(k_1s_0 + k_{-1} + k_2)^2}\left[e^{-(k_1s_0 + k_{-1} + k_2)t} - 1\right]}_{\text{II}}$$

If we expand (II) as an exponential series, which has the form:

$$e^{-ax} = 1 - ax + (a^2x^2/2) - (a^3x^3/3) + \ldots$$

and by ignoring all the terms for low values of t beyond the third degree, we obtain an expression that simplifies to:

$$p = k_1k_2e_0s_0\,t^2/2$$

This simplified equation is only valid for the initial part of the acceleration period. Thus, by measuring p in this initial period, it is possible to obtain $k_1k_2e_0$. As k_2e_0 is determined by measuring V_m at the steady state by the previously described

methods, we know the value of k_2 and consequently we can calculate k_1. Having the value of K_m by the same study at the steady state, we can also calculate k_{-1} and afterwards the true dissociation constant for the enzyme-substrate complex.

The complex equation obtained by integration can also be used in a different way. If we plot p as a function of time, we obtain a curve whose initial curvature is due to the exponential term. As t increases, the exponential term can be ignored since the steady state is reached. The straight line obtained arises from:

$$p = \frac{k_1 k_2 e_0 s_0 t}{K_m + s_0} - \frac{k_2 e_0 s_0}{k_1 (s_0 + K_m)^2}$$

And if s_0 is large compared to K_m (conditions of saturating substrate concentration), we have approximately:

$$p = k_2 e_0 t - (k_2 e_0 / k_1 s_0)$$

And for $p = 0$, i.e. where the linear part of the curve intersects the horizontal axis (see Fig. 5.3), then $t = 1/k_1 s_0$.

Therefore, the value of t gives k_1 directly. The validity of these expressions naturally implies the existence of a single intermediate complex. If several intermediate complexes are formed, these expressions become more complicated and depend on the rate-limiting steps.

5.2.2.2. REACHING THE STEADY STATE

Let us look again at the expression for p which comprises a first Michaelian term (I) followed by an exponential term (II). When time becomes very large, the exponential term cancels out. We rediscover the MICHAELIS equation by deriving dp/dt.

Figure 5.4 represents the changes in s, p and (ES) at different phases of the reaction. It shows clearly that a linear change in p over time is only obtained when (ES) is constant, i.e. when the steady state is reached.

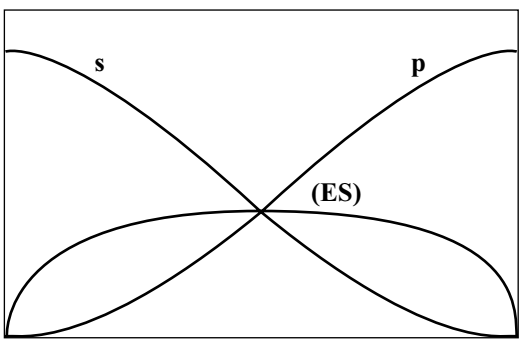


Fig. 5.4 Changes in s, (ES) and p over time

Under conditions approximating the steady state with the simplified scheme that we have already considered, the kinetic parameters have the following significance: the maximum rate, V_m, is equal to k_2e_0 **and the MICHAELIS constant, K_m, is equal to $(k_{-1} + k_2)/k_1$. It is a complex constant that does not reflect the inverse of the enzyme's affinity for the substrate, but depends on all the rate constants.**

5.2.2.3. APPROXIMATION TO A QUASI-EQUILIBRIUM

Whereas the steady state approximation involves no assumption as to the respective values of the specific rate constants, the approximation to a quasi-equilibrium is based on such an assumption. It assumes that $k_2 << k_{-1}$, i.e. that the equilibrium between E, S and ES is attained rapidly and that the chemical reaction is the limiting step. In these conditions, we always have:

$$v = k_2(ES)$$

but the concentration of ES is given by an equilibrium relationship:

$$(E)(S)/(ES) = k_{-1}/k_1 = K_s$$

Under these conditions, the experimental MICHAELIS constant, K_m, is equal to K_s, i.e. to the ratio k_{-1}/k_1, the dissociation constant of the enzyme-substrate complex. In this extreme case, it represents the inverse of the enzyme's affinity for its substrate. In other words, K_m for a given substrate can be likened to a constant of the substrate's dissociation from the ES complex. This approximation is only valid if the catalytic constant, k_2, is much smaller than k_{-1}, which is justified for certain enzymatic reactions, but it is always necessary to show it. *Never must* K_m *immediately be likened to a dissociation constant*.

5.2.2.4. ORDER OF ENZYMATIC REACTIONS

As for chemical reactions (see Chap. 4) we must consider, on the one hand, the order with respect to time, and on the other, the order with respect to concentration or the initial order.

Order with respect to time

It is possible to follow the kinetics of an enzymatic reaction until the amount of substrate is exhausted. Only the substrate is consumed, the enzyme concentration is the same at the end as at the start of the reaction. The kinetics are intermediate between zero and first order. Indeed, starting with an initial substrate concentration s_0, with s being the concentration at time t, we may write:

$$-\frac{ds}{dt} = \frac{k_2e_0s}{K_m + s}$$

By integrating and taking into account the initial conditions, we obtain:

$$k_2 e_0 t = K_m \ln \frac{s_0}{s} + (s_0 - s)$$

which contains a first-order term followed by a zero-order term.

This equation can be simplified in certain cases. If s_0 and consequently s are small compared to K_m, the reaction is first order:

$$\frac{k_2}{K_m} e_0 t = 2.3 \log \frac{s_0}{s}$$

Conversely, if s_0 is large compared to K_m and the product concentrations, i.e. at the beginning of the reaction when the substrate concentration is saturating ($s >> K_m$), we have:

$$k_2 e_0 t = s_0 - s$$

This corresponds to the portion of the curve that is zero-order and observed at the start of the reaction when the substrate concentration is sufficiently high.

Order with respect to concentration or initial order

As for chemical reactions, in order to determine the reaction order, we must vary the initial substrate or enzyme concentration and then determine *the initial reaction rate* at each concentration.

Study of the changes in the reaction rate as a function of the enzyme concentration

One of the experimenter's primary objectives is to determine the change in the initial reaction rate as a function of the enzyme concentration. The curve obtained generally contains a linear part, before curving to reach a plateau (Fig. 5.5).

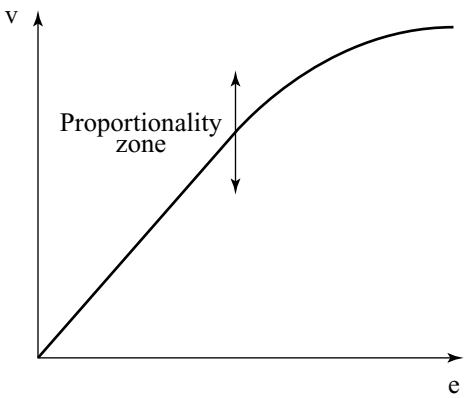


Fig. 5.5 Change in an enzymatic reaction as a function of the enzyme concentration

The simplest procedure whenever possible is to work in a range of enzyme concentrations that ensures proportionality between the reaction rate and the enzyme concentration. In this range, the reaction is first order with respect to the concentration.

It is important here to emphasise that care must be taken in determining the initial rate. Indeed, some enzymatic reactions cannot be followed by continuous-flow methods (see Chap. 6), and so only one method, "point by point", is at our disposal. This involves taking samples and then stopping the reaction after a certain time. It is advisable, therefore, to choose carefully the time period otherwise the initial rate risks being underestimated. Thus, if an experimenter titrates, relative to a control solution, an enzyme solution of unknown concentration for too long a time interval that no longer respects the conditions of linearity, an error ΔP will be made (Figs. 5.6 and 5.7a). It is essential, therefore, at least in an initial series of trials to obtain several points in order to check at what time there is a deviation from the initial rate (Fig. 5.6).

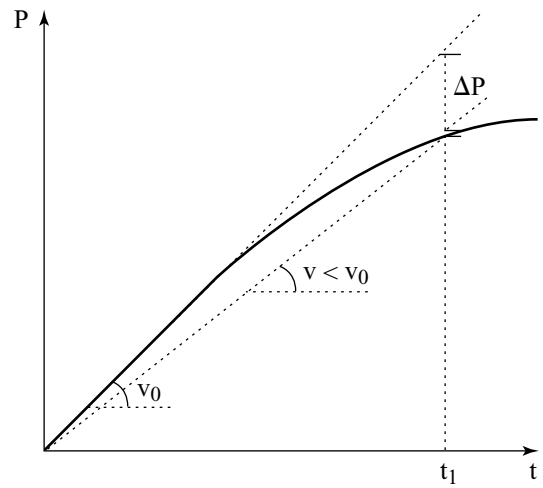


Fig. 5.6 Error in the determination of the rate by measuring p at time t_1

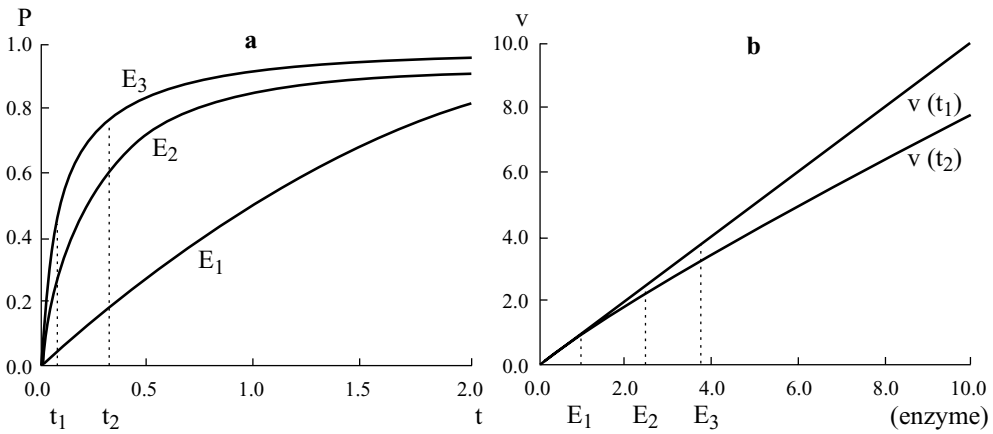


Fig. 5.7 (a) kinetics of the appearance of P over time for different enzyme concentrations: E_1, E_2 and E_3. Evaluation of the rate at times t_1 and t_2 (b) change in the reaction rate as a function of the enzyme concentration at times t_1 and t_2, taken from the data in the preceding curve

Indeed, when no longer working under the initial-rate conditions, the MICHAELIS equation is no longer satisfied and the change in reaction rate with respect to the enzyme concentration is no longer linear (Fig. 5.7a and b). The rate expression given previously is only applicable in these conditions. For the same reasons, it is wise to ensure that the pH of the reaction medium stays constant for the duration of the measurements (see Chap. 9).

Study of the changes in reaction rate as a function of the substrate concentration

For a given enzyme concentration, changes in the reaction rate as a function of the initial substrate concentration follow the hyperbolic law given by the MICHAELIS equation:

$$v = \frac{k_2 e_0 s_0}{K_m + s_0}$$

This relation also expresses a kinetic intermediate between zero and first order. Under conditions of extreme substrate concentrations, there is a tendency to lean either towards zero-order kinetics or towards first-order kinetics.

When s_0 is large relative to K_m ($s_0 >> K_m$), this expression simplifies to:

$$v = k_2 e_0 = V_m$$

The rate tends towards the maximal rate and Y_s becomes practically equal to 1. Therefore, the reaction rate becomes practically independent of the substrate concentration. These conditions are important from an experimental point of view.

When $s_0 << K_m$, the MICHAELIS equation becomes:

$$v = \frac{k_2}{K_m} e_0 s_0$$

The reaction rate is proportional to the ratio $k_2 e_0 / K_m$, which is the first-order rate constant for the reaction. The rate varies linearly as a function of s_0. Now, we can write:

$$\frac{k_2}{K_m} = \frac{k_2 k_1}{k_2 + k_{-1}}$$

The limit of the constant k_2/K_m is therefore determined by k_1, the rate constant for the formation of the ES complex. This rate is limited by molecular diffusion. It cannot be faster than the speed with which the molecules of E and S meet, which is controlled by diffusion. This is clearly a physical barrier that cannot be exceeded. At physiological temperatures, the magnitude of diffusion is between 10^8 and 10^9 $mol^{-1} . s^{-1}$, implying that:

$$\frac{k_2}{K_m} < 10^8\text{-}10^9 \ M^{-1} . s^{-1}$$

Certain enzymes such as catalase or carbonic anhydrase have k_{cat}/K_m values that are this order of magnitude. These enzymes have attained *catalytic perfection* and hence their activity rates are only limited by diffusion (see Chap. 11). Biological systems have found a way round this physical limit by creating **multi-enzyme complexes** in which the reaction product is transferred directly to the catalytic site of the enzyme catalysing the subsequent reaction. In Part V, we shall see examples of these multi-enzyme complexes, which catalyse several successive reactions on the same metabolic pathway.

In summary, under extreme conditions when the substrate concentration is low compared to the MICHAELIS constant, the kinetics are first order; conversely, if s is large relative to K_m, the kinetics are practically zero order. The reaction rate tends towards a maximum value $V_m = k_2 e_0$.

5.2.3. METHODS TO DETERMINE KINETIC PARAMETERS

In order to obtain with satisfactory precision the experimental parameters K_m and V_m, it is important to determine the initial reaction rate for a sufficient number of substrate concentrations situated either side of the K_m value. Several linear graphical plots are available for determining these parameters.

5.2.3.1. SEMI-LOGARITHMIC PLOT

The first graphical plot used by MICHAELIS involved plotting the reaction rate as a function of the logarithm of the substrate concentration, for a given enzyme concentration (Fig. 5.8). The inflexion point corresponds to $\log K_m$ on the horizontal axis. This method is very imprecise as a result of the difficulty in reliably determining V_m.

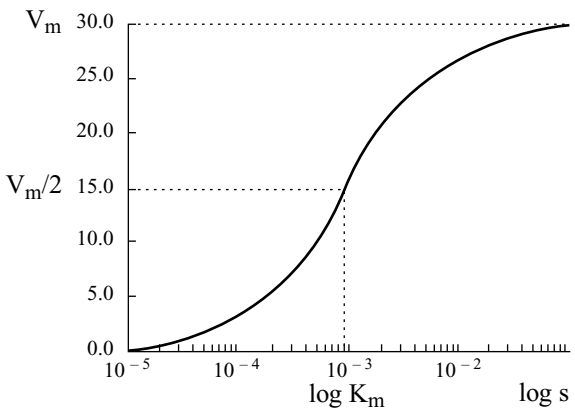


Fig. 5.8 Determination of the MICHAELIS constant and V_m by the semi-logarithmic plot

5.2.3.2. EADIE PLOT

The rate equation can be written in the following linear form:

$$v = -K_m v/s + V_m$$

If we plot v as a function of v/s, we obtain a straight line whose slope is equal to $-K_m$ and the vertical-axis intercept, V_m (Fig. 5.9).

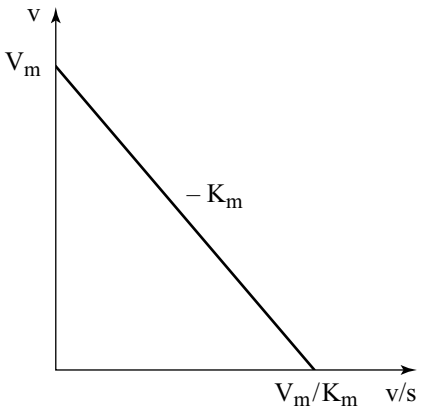


Fig. 5.9 EADIE plot

5.2.3.3. LINEWEAVER-BURK PLOT

A linear plot can also be obtained when using the inverse of the MICHAELIS equation:

$$\frac{1}{v} = \frac{K_m}{V_m s} + \frac{1}{V_m}$$

The slope of the straight line obtained by plotting 1/v versus 1/s is equal to K_m/V_m and the vertical-axis intercept, $1/V_m$. The intercept on the horizontal axis gives the value $-1/K_m$ (Fig. 5.10).

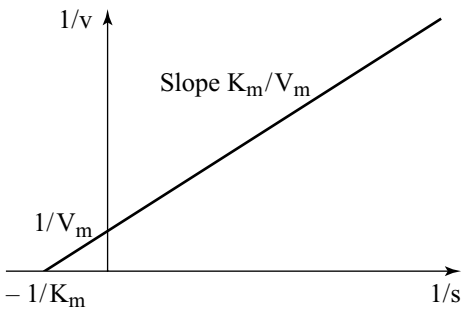


Fig. 5.10 LINEWEAVER-BURK plot

5.2.3.4. HANES-DIXON PLOT

Another equivalent form of the rate equation is:

$$s/v = s/V_m + K_m/V_m$$

The plot of s/v as a function of s leads to a straight line whose slope directly gives $1/V_m$ and the vertical-axis intercept is equal to K_m/V_m (Fig. 5.11 below).

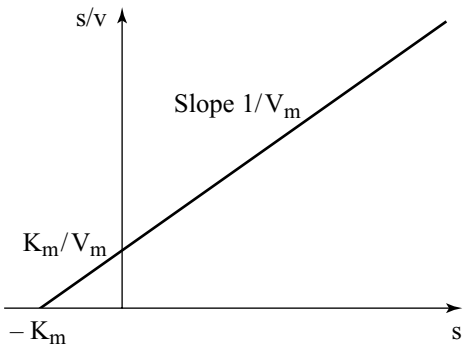


Fig. 5.11 HANES-DIXON plot

5.2.3.5. PLOT DERIVED FROM THE INTEGRATED RATE EQUATION

The integrated rate equation written above can be put in the following form:

$$\frac{2.3 \log s_0 / s}{t} = \frac{V_m}{K_m} - \frac{(s_0 - s)}{K_m t}$$

s and $(s_0 - s)$ are measured in a series of defined time-points and on a graph we plot $(2.3 \log s_0/s)/t$ as a function of $(s_0 - s)/t$; a straight line should be obtained if the kinetics follow the classical law well. This line has a slope of $-1/K_m$ and the horizontal-axis intercept is V_m (Fig. 5.12). This plot is therefore applicable in principle even for cases where the previous plots are not suitable as a result of the imprecision in the estimation of the initial rates. Furthermore, only a single experiment is required in principle, although it is necessary to check that the same straight line is obtained with several substrate concentrations.

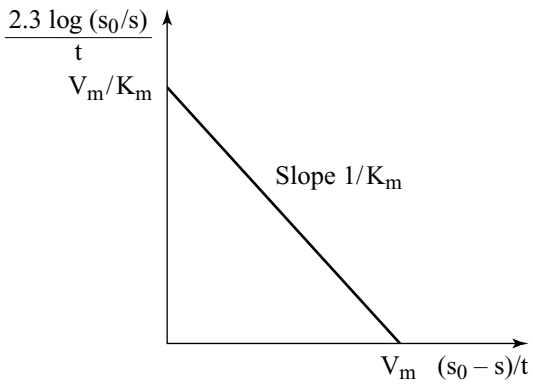


Fig. 5.12 Graphical plot from the integrated MICHAELIS equation

5.2.3.6. DIRECT PLOT OF EISENTHAL AND CORNISH-BOWDEN

This method involves plotting, for each experiment, v on the vertical axis versus –s on the horizontal, then drawing the corresponding lines. According to the MICHAELIS equation, these straight lines intersect at the same point having coordinates of V_m and K_m. Figure 5.13a illustrates the graphical procedure. However, the effect of experimental error means that the intersection point is not unique. The number of intersection points is $1/2n(n-1)$. For the five experimental curves

presented in Fig. 5.13b there are 10 intersection points. Each of these points gives an estimation of K_m and V_m, and then the average value of these is determined.

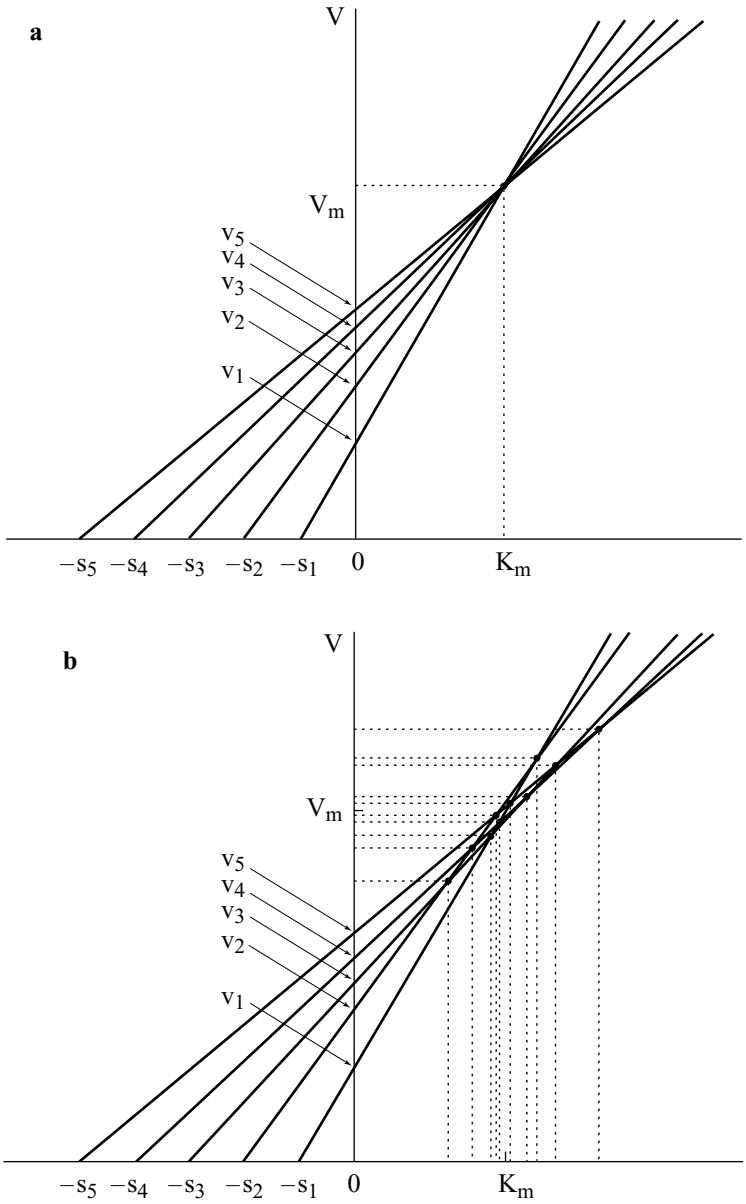


Fig. 5.13 Direct plot of EISENTHAL and CORNISH-BOWDEN

(a) *determination of V_m and K_m. Each* straight slope *represents one experiment for a substrate concentration s, giving an initial rate v. The intersection point gives V_m and K_m –* ***(b)*** *the intersection point can regress if there is an error in the straight line. Each intersection gives an estimation of K_m and V_m. For each group, the average value is considered to be the best value for the parameters*

5.2.3.7. VALIDITY OF THE DIFFERENT GRAPHICAL PLOTS

The problem of which method is the best to analyse the experimental data and evaluate the kinetic parameters from the MICHAELIS equation is a very old problem. The direct plot of v against s_0 is an equilateral hyperbola passing through the origin and having the asymptotes $v = V_m$ and $s_0 = -K_m$. But it is practically impossible to obtain V_m and K_m with precision from this plot because:

- only finite and positive values of s are measurable,
- it is rarely possible to use sufficiently high substrate concentrations to determine precisely the plateau corresponding to the maximal rate.

In the MICHAELIS-MENTEN plot (or semi-logarithmic plot) there is an inflexion point at $s = K_m$, and the maximum slope at this point is 0.576. This method is statistically correct, however, the maximum rate must be measured experimentally with great precision. Consequently, most enzymologists prefer to use one of the linear plots that have just been described. The validity of the kinetic constants estimated from these diverse plots has been widely discussed, in particular by WILKINSON (1961), JOHANSEN and LUMRY (1961), DOWD and RIGGS (1965), COLQUHOUN (1971), and CORNISH-BOWDEN and EISENTHAL (1974).

As analysed by DOWD and RIGGS, the frequency distribution of the K_m and V_m values derived from the three linear plots show clearly the great inferiority of the LINEWEAVER-BURK plot. This is evident by examining the diagrams in Fig. 5.14a opposite, which give the frequency distribution of V_m and K_m for these three plots, based on 500 identical experiments; the error in v is assumed to be constant and relatively high. 40 values estimated for V_m and 35 values for K_m were superior to 100 or inferior to 0 in the LINEWEAVER-BURK plot. An analogous conclusion can be drawn from the same diagram, but making the assumption that the error in v increases practically proportionally to v (Fig. 5.14b). Even allowing a small error in v, the LINEWEAVER-BURK plot leads to the least correct evaluation, since in this case the authors obtained a non-negligible number of negative values in estimating the parameters.

The two other linear plots: v against v/s_0 or s_0 against s_0/v, are therefore superior to the LINEWEAVER-BURK plot. Nevertheless, they do not permit a statistically significant analysis since the values plotted on both axes are not independent variables. For this reason, as shall be analysed in the last paragraph of this chapter, we prefer to employ statistical methods of analysis based on the non-linear MICHAELIS-MENTEN plot and then to evaluate the parameters by a non-linear regression method, especially since today's computer programs enable direct data processing.

Whatever the case, **the marked inferiority of the LINEWEAVER-BURK plot compels us to advise against its use for the estimation of kinetic parameters**.

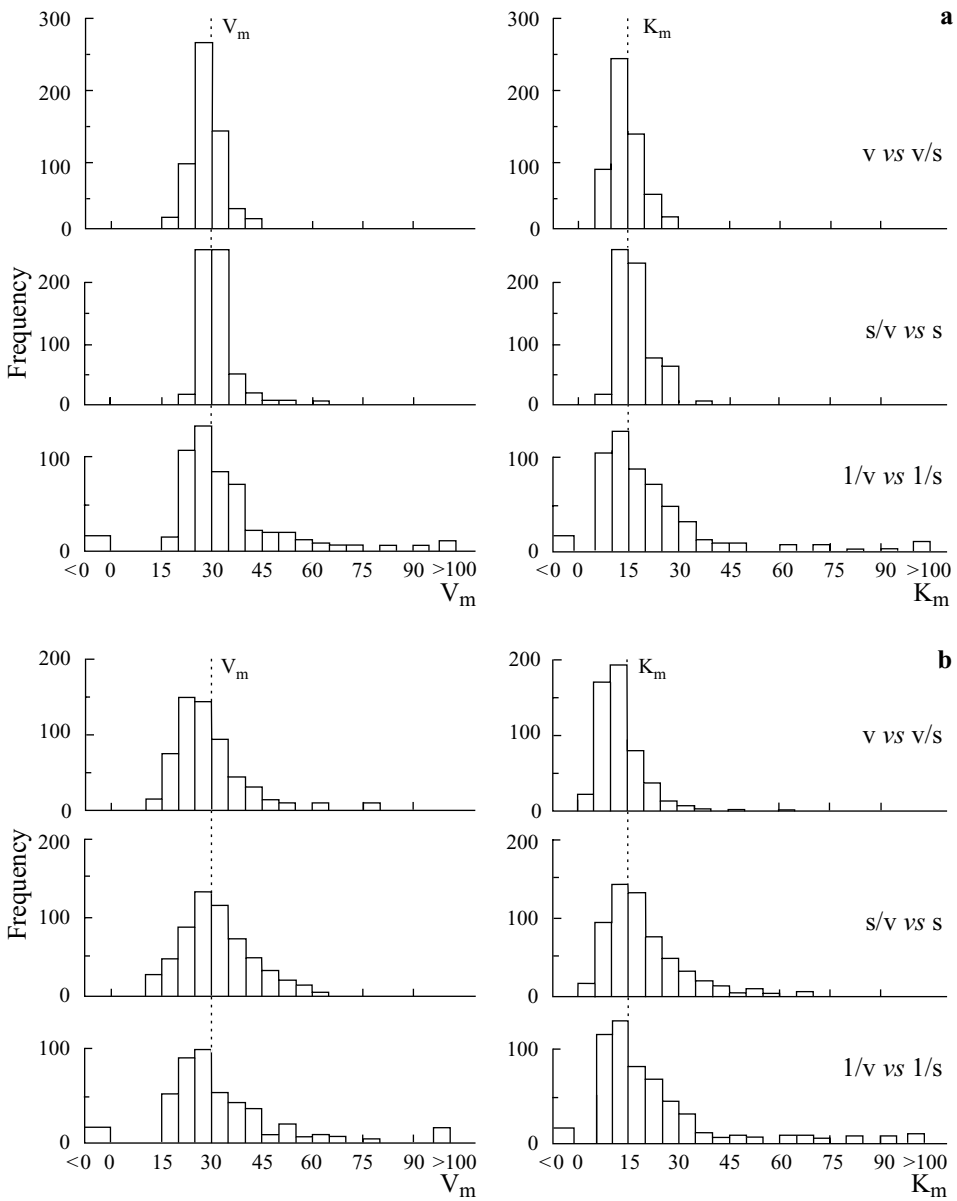


Fig. 5.14 Frequency distribution for K_m and V_m values for the three linear plots *(**a**) frequency distribution based on 500 identical experiments – (**b**) same diagram, but with the assumption that the error in determining the rate increases along with the rate value.* (From *J. Biol. Chem.*, **240**, DOWD J.E. & RIGGS D.S., 863. © (1965) with permission from The American Society for Biochemistry and Molecular Biology)

5.3. KINETICS OF ENZYMATIC REACTIONS IN THE PRESENCE OF EFFECTORS (INHIBITORS OR ACTIVATORS)

5.3.1. KINETICS OF ENZYMATIC REACTIONS IN THE PRESENCE OF INHIBITORS

The kinetics of enzymatic reactions can be considerably modified by the presence of inhibitors in the reaction medium. The phenomenon of inhibition is very frequent in enzymology and the most diverse chemical components are capable of inhibiting enzymatic reactions; this, of course, depends on the enzyme and reactant.

As we shall see in Part VI, the regulation of cellular metabolism relies for the most part on physiological mechanisms of inhibiting enzyme activity.

Furthermore, it is interesting to provoke inhibition of an enzymatic reaction using substances of known structure in order to obtain information about the mechanism of enzyme action. The inhibition of enzymes in vivo increasingly underpins the numerous chemotherapeutic procedures. The study of inhibitory phenomena is therefore of primary importance in enzymology.

Some types of inhibition result from the reversible association of an inhibitor to an enzyme. Others are the consequence of irreversible binding; the action of irreversible inhibitors will be considered in Parts III and V.

Different types of reversible inhibition exist. The most simple and the most classic are competitive inhibition, non-competitive inhibition and uncompetitive inhibition or inhibition by blocking the intermediate complex. In these types of inhibition, the presence of the inhibitor on the enzyme totally abolishes its activity, although partial inhibition also exists. We shall now examine successively total and partial inhibition.

5.3.1.1. TOTAL INHIBITION

Competitive inhibition

There is competitive inhibition when the binding of an inhibitor molecule, I, to the enzyme prevents substrate binding, and reciprocally, the inhibitor cannot practically bind to the ES complex; **there is exclusive binding of either the inhibitor or the substrate.** Competitive inhibition occurs in particular when the inhibitor and substrate are structurally analogous and bind to the same site on the enzyme, but this is not the only example. The scheme for an enzymatic reaction subjected to competitive inhibition is written thus:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

$$E + I \underset{k_{-i}}{\overset{k_i}{\rightleftharpoons}} EI$$

with the respective dissociation constants:

$$K_s = \frac{k_{-1}}{k_1} = \frac{(E)(S)}{(ES)} \quad \text{and} \quad K_i = \frac{k_{-i}}{k_i} = \frac{(E)(I)}{(EI)}$$

There is competition between substrate and inhibitor molecules with respect to the enzyme; an excess of substrate displaces the inhibitor. The rate equation at steady state is obtained by taking the equation for enzyme conservation:

$$e_0 = (E) + (ES) + (EI)$$

Furthermore, when $(I) >> e_0$, the free-inhibitor concentration is practically equal to the total-inhibitor concentration; this condition generally arises in experiments in vitro in which the inhibitor, like the substrate, is in excess relative to the enzyme concentration.

The rate equation is:

$$v = \frac{k_2 e_0 s_0}{K_m\left[1+\frac{(I)}{K_i}\right]+s_0} = V_m \frac{s_0}{K_m\left[1+\frac{(I)}{K_i}\right]+s_0}$$

It shows that, in this type of inhibition, only the apparent MICHAELIS constant for the reaction varies as a function of the inhibitor concentration:

$$K'_m = K_m\left[1+\frac{(I)}{K_i}\right]$$

The maximal reaction rate, V_m, remains unchanged whatever the inhibitor concentration. An excess of substrate displaces the inhibitor.

If we use the EADIE plot to determine the kinetic parameters, we obtain a series of straight lines whose slopes increase in absolute value as a function of the inhibitor concentration by a factor of $[1 + (I)/K_i]$, but which converge to the same point on the vertical axis, i.e. V_m (Fig. 5.15a below). The LINEWEAVER-BURK plot also leads to a beam of straight lines coinciding at a point on the vertical axis that has the value $1/V_m$ (Fig. 5.15b). From these K'_m values for different inhibitor concentrations, it is straightforward to obtain K_m and K_i with a **secondary plot** (Fig. 5.15c).

Another graphical plot, suggested by DIXON, to determine directly the constant K_i involves plotting 1/v as a function of (I) according to the equation:

$$\frac{1}{v} = \frac{K_m\left[1+\frac{(I)}{K_i}\right]+s_0}{V_m s_0}$$

For different concentrations of s_0, we obtain the graph depicted in Fig. 5.15d.

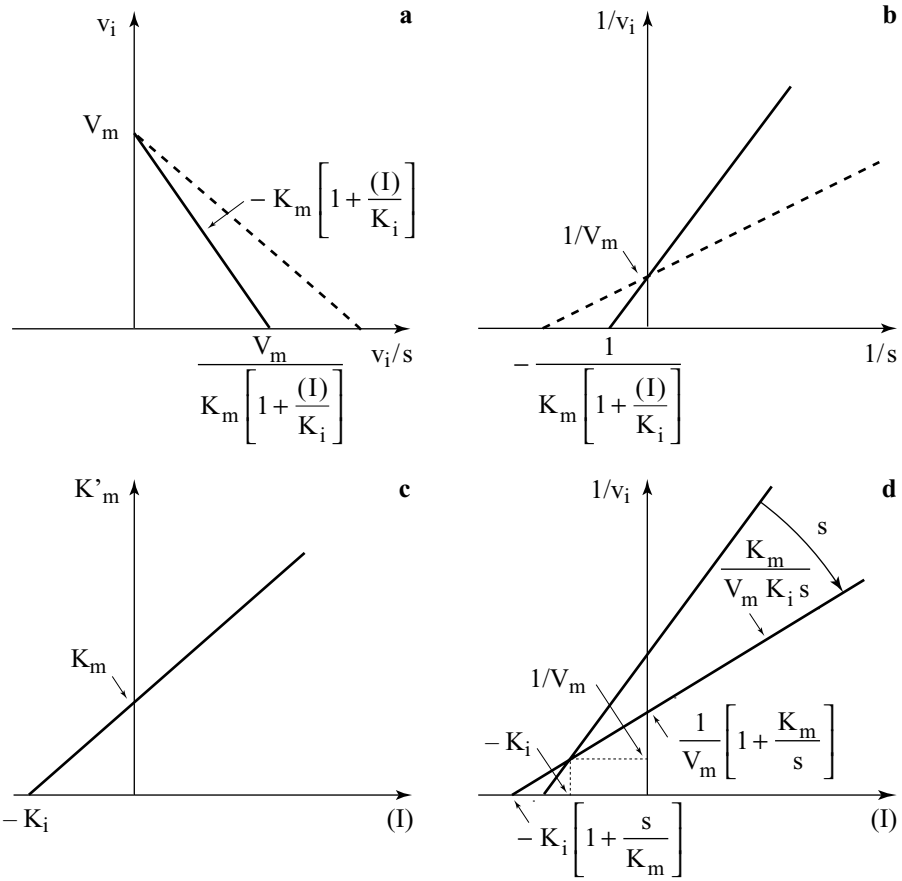


Fig. 5.15 Competitive inhibition

*(**a**) EADIE plot – (**b**) LINEWEAVER-BURK plot – (**c**) secondary plot of K'_m as a function of (I) permitting the determination of K_m and K_i – (**d**) DIXON plot of $1/v_i$ as function of (I) for two substrate concentrations. In (**a**) and (**b**) unbroken lines: reactions in the presence of inhibitor; dashed lines: reactions in the absence of inhibitor*

When K_i is very small relative to (I) and K_m, i.e. when the inhibitor has a high affinity for the enzyme, it can be difficult to determine the nature of the inhibition. Indeed, in the expression for v, we have $(I)/K_i >> 1$ and $K_m(I)/K_i >> s_0$.

The rate expression is thus simplified to the following:

$$v = \frac{k_2 e_0 s_0}{\frac{K_m (I)}{K_i}}$$

In this case, it is indistinguishable from non-competitive inhibition (see the next paragraph).

Non-competitive inhibition

In this type of inhibition, **substrate and inhibitor binding are not exclusive; they are independent.** The inhibitor binds without altering the affinity of the enzyme for its substrate. A ternary ESI complex can thus be formed, but it is inactive. The presence of the inhibitor alone may also lead to the formation of an inactive EI complex. This can be represented by the following scheme:

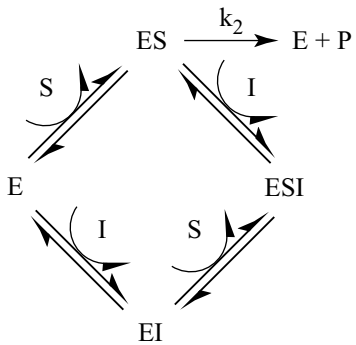
$$E + S \underset{}{\overset{K_s}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

$$E + I \overset{K_i}{\rightleftharpoons} EI$$

$$ES + I \overset{K_i}{\rightleftharpoons} ESI$$

$$EI + S \overset{K_s}{\rightleftharpoons} ESI$$

This scheme may also be written as follows:



The equation for enzyme conservation is:

$$e_0 = (E) + (EI) + (ES) + (ESI)$$

with $(I) >> e_0$ as before. Solving the system at steady state leads to the rate equation:

$$v = \frac{k_2 e_0 s_0}{K_m + s_0} \cdot \frac{1}{1 + \frac{(I)}{K_i}}$$

The apparent MICHAELIS constant does not change; only the maximum reaction rate as a function of the inhibitor concentration:

$$V'_m = \frac{V_m}{1 + \frac{(I)}{K_i}}$$

The EADIE plot gives rise to a series of parallel lines with decreasing values for the vertical-axis intercept as the inhibitor concentration increases (Fig. 5.16a).

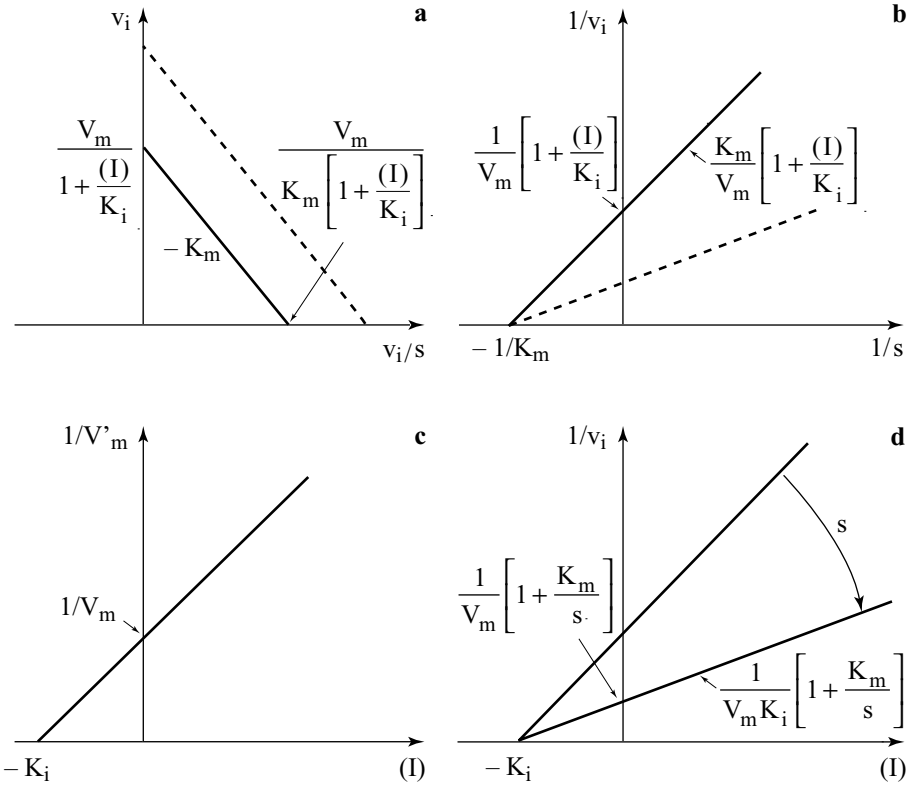


Fig. 5.16 Non-competitive inhibition

(a) EADIE plot – (b) LINEWEAVER-BURK plot – (c) secondary plot of 1/V'm as a function of (I) – (d) DIXON plot. The symbols are the same as in Fig. 5.15

The LINEWEAVER-BURK plot gives a series of straight lines which converge at a point on the horizontal axis whose values is $-1/K_m$ (Fig. 5.16b). The changes in apparent maximum rate as a function of the inhibitor concentration permits determination of the constant K_i with a secondary plot (Fig. 5.16c).

Based on the equation:

$$\frac{1}{v} = \frac{K_m + s_0}{V_m s_0}\left[1+\frac{(I)}{K_i}\right]$$

the DIXON plot, 1/v against (I), produces the graphs indicated in Fig. 5.16d. The straight lines thus obtained for different values of s converge at the point $-K_i$ on the horizontal axis.

"Uncompetitive" inhibition or inhibition by blocking the intermediate complex

A type of inhibition called "uncompetitive" inhibition exists which involves a mechanism of **inhibition by blocking the intermediate complex**. In this case actually, the inhibitor binds to the enzyme-substrate complex and not to the free enzyme, giving an inactive ternary complex. The reaction scheme can be represented as follows:

$$E + S \underset{}{\overset{K_s}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

$$ES + I \overset{K_i}{\rightleftharpoons} ESI$$

The equation for enzyme conservation is:

$$e_0 = (E) + (ES) + (ESI)$$

and the rate equation:

$$v = \frac{k_2 e_0 s_0}{\dfrac{K_m}{1+\dfrac{(I)}{K_i}} + s_0} \cdot \frac{1}{1+\dfrac{(I)}{K_i}}$$

The apparent MICHAELIS constant, K'_m, and the apparent maximum rate, V'_m, vary as a function of the inhibitor concentration; they become proportionately smaller:

$$V'_m = \frac{V_m}{1+\dfrac{(I)}{K_i}} \quad \text{and} \quad K'_m = \frac{K_m}{1+\dfrac{(I)}{K_i}}$$

It seems as though the presence of the inhibitor facilitates the formation of an intermediate complex, since K'_m decreases as (I) increases, yet at the same time the reaction is prevented from taking place.

In this type of inhibition the EADIE plot produces a series of straight lines whose slopes decrease while the inhibitor concentration increases (Fig. 5.17a below). These lines converge to a point on the horizontal axis whose value is equal to the ratio V_m/K_m and which does not alter. The LINEWEAVER-BURK plot gives a series of parallel lines since the slope is equal to K_m/V_m (Fig. 5.17b).

The DIXON plot also gives rise to a series of parallel lines according to the equation:

$$\frac{1}{v} = \frac{K_m + s_0\left[1+\dfrac{(I)}{K_i}\right]}{V_m s_0}$$

The intercept with the horizontal axis is $-K_i(1 + K_m/s_0)$ and with the vertical, $(K_m + s_0)/V_m s_0$; its slope is $1/V_m K_i$ (Fig. 5.17d).

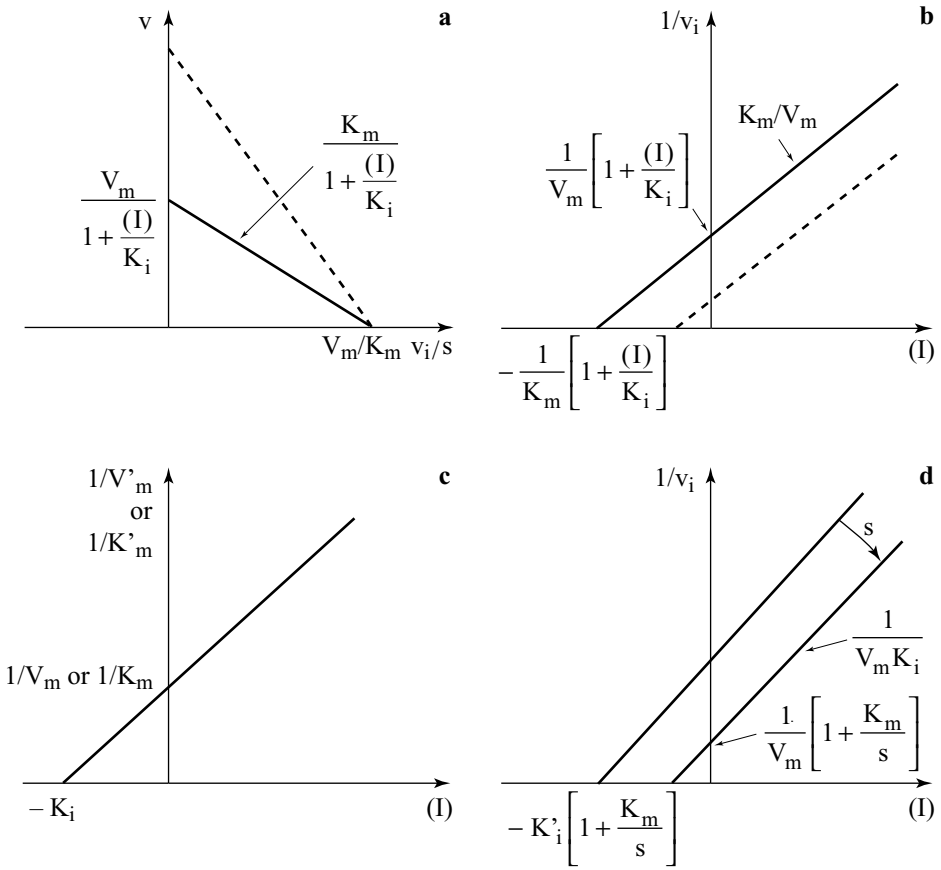


Fig. 5.17 "Uncompetitive" inhibition

(a) EADIE plot – (b) LINEWEAVER-BURK plot – (c) secondary plot of 1/V'm or 1/K'm as a function of (I) – (d) DIXON plot. The symbols are the same as in Fig. 5.15

Inhibition by the binding of the inhibitor to the substrate

In some enzymatic reactions, in particular in proteolysis reactions where the substrate can be a molecule with large dimensions, the inhibitor may bind to the substrate and not to the enzyme. This does not happen frequently, but ought to be borne in mind. The inhibition may either be competitive or non-competitive. For each of these types of inhibition, we end up with the same rate equation whether the inhibitor binds to the enzyme or the substrate, on the condition that free (S) stays large compared to e_0 (YON, 1961).

Consequently, it is not always enough to know the type of inhibition in order to determine the inhibitor's mode of action. Sometimes it is necessary to do a study with different substrates and possibly with other enzymes that recognise the same substrate.

Inhibition by high substrate concentrations

Some enzymatic reactions obey the law of HENRI-MICHAELIS at low substrate concentrations, but at high concentrations the rate, after having reached a maximum, diminishes. For such reactions, the plot of the initial reaction rate as a function of the substrate concentration gives rise to the curves shown in Fig. 5.18. This occurs when the enzyme is liable to bind several substrate molecules in the active site. Only the complex to which the substrate binds in a favourable orientation is active. We might imagine that, when the substrate concentration increases, two or more molecules may bind at each of the enzyme's sub-sites, and since none at this point is in a favourable orientation to ensure a reaction, the ternary (or higher order) complex remains inactive.

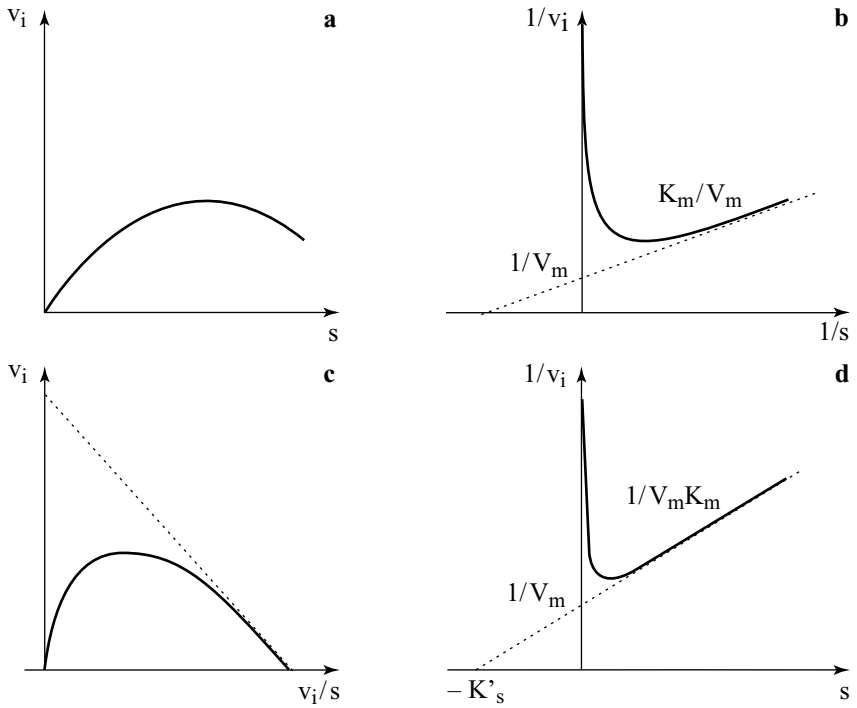


Fig. 5.18 Inhibition by excess substrate
*(**a**) change in the rate as a function of the substrate concentration – (**b**) LINEWEAVER-BURK plot – (**c**) EADIE plot – (**d**) change in 1/v as a function of s*

This situation, which is a true case of non-competitive inhibition by the substrate itself, has been theoretically studied by HALDANE, and corresponds to the following scheme:

$$E + S \underset{}{\overset{K_s}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

$$ES + S \overset{K'_s}{\rightleftharpoons} ES_2$$

K'_s represents the dissociation constant of the ternary complex.

The rate equation is:

$$v = \frac{k_2 e_0}{1 + \frac{K_m}{s_0} + \frac{s_0}{K'_s}}$$

In this case, however, it is possible to determine the MICHAELIS constant of the system by using sufficiently low substrate concentrations that the inhibition is negligible. Furthermore, we can determine the inhibition constant K'_s if we use conditions in which the substrate concentration is sufficiently high relative to K_m such that the term K_m/s_0 becomes negligible. In these conditions, the rate equation simplifies to:

$$v = \frac{V_m K'_s}{K'_s + s_0}$$

or:

$$\frac{1}{v} = \frac{s_0}{K'_s V_m} + \frac{1}{V_m}$$

The plot of 1/v as a function of s_0 leads to a straight line in the region corresponding to high substrate concentrations where this simplification is valid. The horizontal-axis intercept gives a value of $-K'_s$ (Fig. 5.18d). When the substrate concentration diminishes, the equation ceases to be valid and we observe a curvature leading to a minimum and 1/v increases. This sort of inhibition arises with urease or acetylcholinesterase, for example.

Inhibition by the reaction products

Another frequently observed case in enzymology is inhibition by the reaction products, **even when the reverse reaction cannot take place.** Indeed, the reaction products often have very similar structures to that of the substrate or a part of the substrate, and thus are capable of forming a specific complex with the enzyme. Again, this is competitive inhibition, but none of the plots described above are able to show this. In fact, these diagrams rely on the determination of initial rates, and under the initial reaction conditions the inhibition is not yet apparent, as the product concentration is negligible. However, if the reaction products were to be added at time zero (t = 0), then inhibition would be observed. Inhibition by the reaction products can be schematised as follows:

$$E + S \underset{}{\overset{K_s}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

$$E + P \overset{K_i}{\rightleftharpoons} EP$$

with $(P) = s_0 - s$ and $e_0 = (E) + (ES) + (EP)$, and K_i being the dissociation constant of the enzyme-product complex, EP. The solution of this system leads to the following rate equation:

$$v = \frac{k_2 e_0 s_0}{K_m\left[1+\frac{(s_0 - s)}{K_i}\right] + s_0}$$

The equation in its integrated form is:

$$k_2 e_0 t = 2.3\, K_m\left(1+\frac{s_0}{K_i}\right)\times \log\frac{s_0}{s} - \left(\frac{K_m}{K_i} - 1\right)(s_0 - s)$$

If $K_i = K_m$, the kinetics are first order throughout the reaction:

$$k_2 e_0 t = 2.3(K_m + s_0)\times \log\frac{s_0}{s}$$

and yet the first-order constant varies with the initial substrate concentration. If $K_m = K_i$, the plot obtained from the integrated rate equation gives a series of lines whose slopes vary as a function of the initial substrate concentration.

The integrated rate equation can be written:

$$\frac{2.3}{t}\log\frac{s_0}{s} = \frac{V_m}{K_m\left(1+\frac{s_0}{K_i}\right)} - \frac{1}{K_m\left(1+\frac{s_0}{K_i}\right)}\left(1-\frac{K_m}{K_i}\right)\left(\frac{s_0 - s}{t}\right)$$

If we plot $2.3/t \times \log (s_0/s)$ as a function of $(s_0 - s)/t$, we obtain a series of straight lines. The intercept of these on the horizontal axis decreases as s_0 increases and the absolute values of their gradients diminish as a function of s_0 (Fig. 5.19).

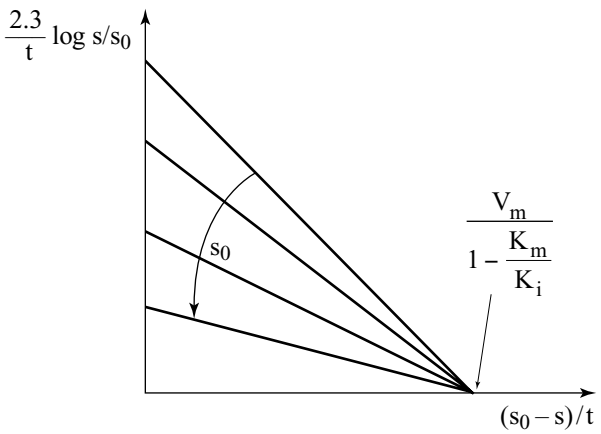


Fig. 5.19 Inhibition by the reaction product
Plot based on the integrated MICHAELIS equation for different substrate concentrations

The slope of these lines can, incidentally, be positive or negative depending on the respective values of K_m and K_i. If $K_m > K_i$, then the value of $[1 - (K_m/K_i)]$ is negative and the slope is positive. If $K_m < K_i$, the slope is negative, which represents the most general case – the enzyme more often than not has a greater affinity for its substrate than for the reaction products. The different curves obtained converge to the same point on the horizontal axis whose value corresponds to:

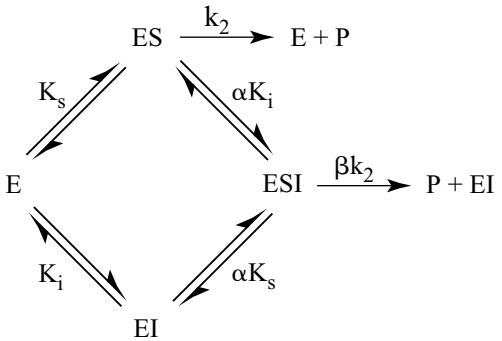
$$\frac{V_m}{\left(1-\frac{K_m}{K_i}\right)}$$

HARMON and NIEMANN demonstrated inhibition of this type for the tryptic hydrolysis of benzoyl-L-arginine amide. One of the hydrolysis products, benzoyl-L-arginine, inhibits the reaction. If this product is added to the reaction at time 0, it behaves like a classic competitive inhibitor, as would be the case in general, and benzoyl-L-arginine has a K_i value of 6.5×10^{-3} M (BECHET et al., 1956).

When we come to study two-substrate reactions, we shall analyse in detail inhibition by the reaction products, as it provides important information about the reaction scheme and often enables distinguishing between two possible schemes.

5.3.1.2. PARTIAL INHIBITION

◤ To address partial inhibition, we shall present first of all a generalised formulation for diverse types of reversible inhibition, in the case of an approximation to a quasi-equilibrium, i.e. when all complexes are in rapid equilibrium with their components. We always assume conditions where $e_0 << s_0$ and (I). The general scheme is written:



α represents the change in the dissociation constant of the enzyme-substrate complex under the effect of the inhibitor, or in the dissociation constant of the enzyme-inhibitor complex under the effect of substrate binding. The coefficient β is the factor of change in the rate constant when the inhibitor is present. The constants of dissociation equilibrium are as follows:

$$K_s = \frac{(E)(S)}{(ES)} \qquad \alpha K_s = \frac{(EI)(S)}{(EIS)}$$

$$K_i = \frac{(E)(I)}{(EI)} \qquad \alpha K_i = \frac{(ES)(I)}{(ESI)}$$

From the following relationships:

$$e_0 = (E) + (ES) + (EI) + (ESI)$$

$$v = k_2(ES) + \beta k_2(ESI)$$

we derive the rate equation:

$$v = V_m \frac{s_0[\alpha K_i + \beta(I)]}{s_0(I) + \alpha[K_i s_0 + K_s(I) + K_s K_i]}$$

with the apparent parameters:

$$V'_m = \frac{V_m[\alpha K_i + \beta(I)]}{K_i + \beta(I)}$$

$$K'_m = \frac{K_s[\alpha K_i + \beta(I)]}{K_i + \beta(I)}$$

For the diverse types of inhibition, generally $\alpha > 1$ and $\beta < 1$, although the first condition is not absolutely necessary; the rate may decrease even if $\alpha < 1$ on the condition that β is sufficiently small.

Partial competitive inhibition

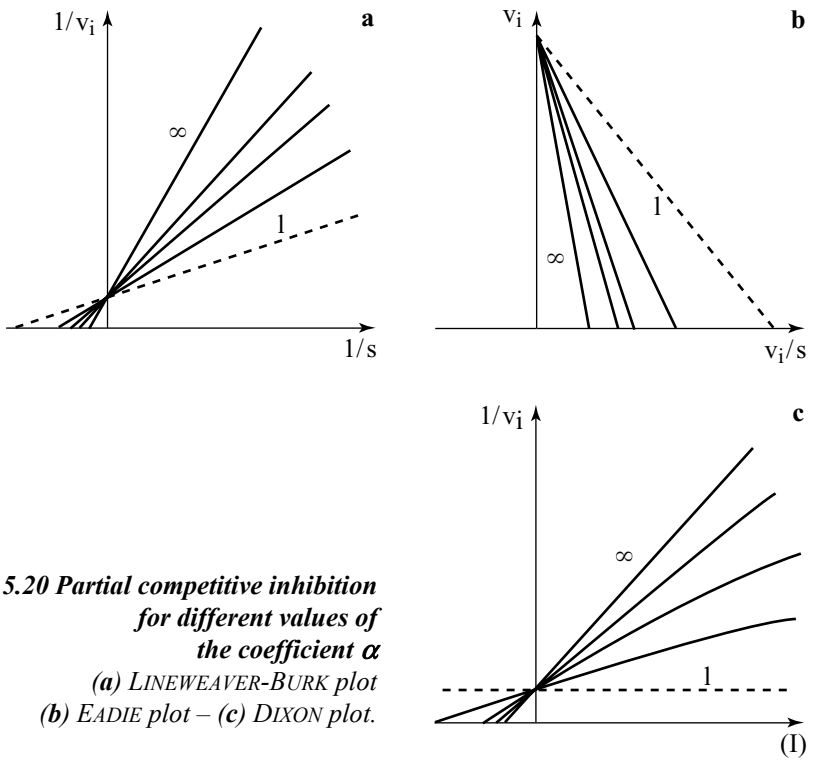


Fig. 5.20 Partial competitive inhibition for different values of the coefficient α
(a) *LINEWEAVER-BURK plot*
(b) *EADIE plot – (c) DIXON plot.*

In total competitive inhibition $\alpha \longrightarrow \infty$ and we have the expression given previously (see the paragraph on competitive inhibition). In partial competitive inhibition $\infty > \alpha > 1$ and $\beta = 1$. The inhibitor only partially prevents substrate binding and does not affect the degradation rate of the complex. The rate equation is:

$$v = V_m \frac{s_0}{s_0 + K_s \dfrac{\alpha(I) + \alpha K_i}{(I) + \alpha K_i}}$$

$$K'_m = K_s \frac{\alpha(I) + \alpha K_i}{(I) + \alpha K_i}$$

The maximum rate remains identical in the presence of inhibitor. Only the apparent MICHAELIS constant varies. The graphs obtained using the EADIE or LINEWEAVER-BURK plots do not differ from those when the inhibition is totally competitive. However, the DIXON plot of 1/v against (I) is no longer linear, but gives a series of curves (Fig. 5.20 above).

Partial non-competitive inhibition

In total non-competitive inhibition, $\alpha = 1$ and $\beta = 0$. The equations are as previously indicated. In partial non-competitive inhibition $\alpha = 1$ and $0 < \beta < 1$. The rate equation is:

$$v = V_m \frac{\beta(I) + K_i}{(I) + K_i} \quad \frac{s_0}{s_0 + K_s}$$

K'_m alone varies as the inhibitor concentration increases. As before, 1/v as a function of (I) is no longer linear. Figure 5.21 illustrates how 1/v changes as a function of (I) for different values of β.

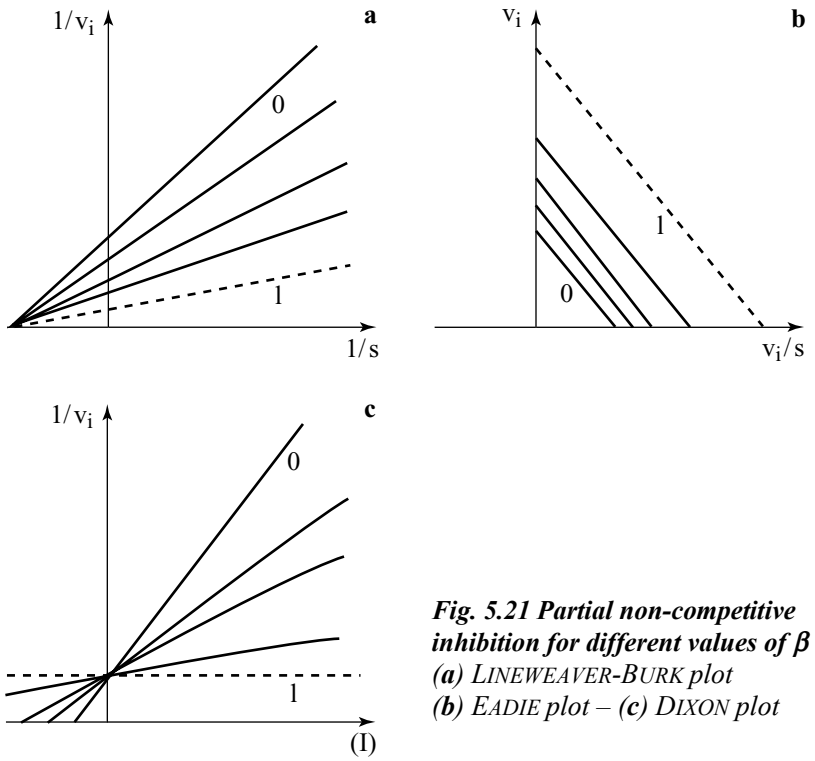


Fig. 5.21 Partial non-competitive inhibition for different values of β
(a) *LINEWEAVER-BURK plot*
(b) *EADIE plot – (c) DIXON plot*

Mixed inhibition

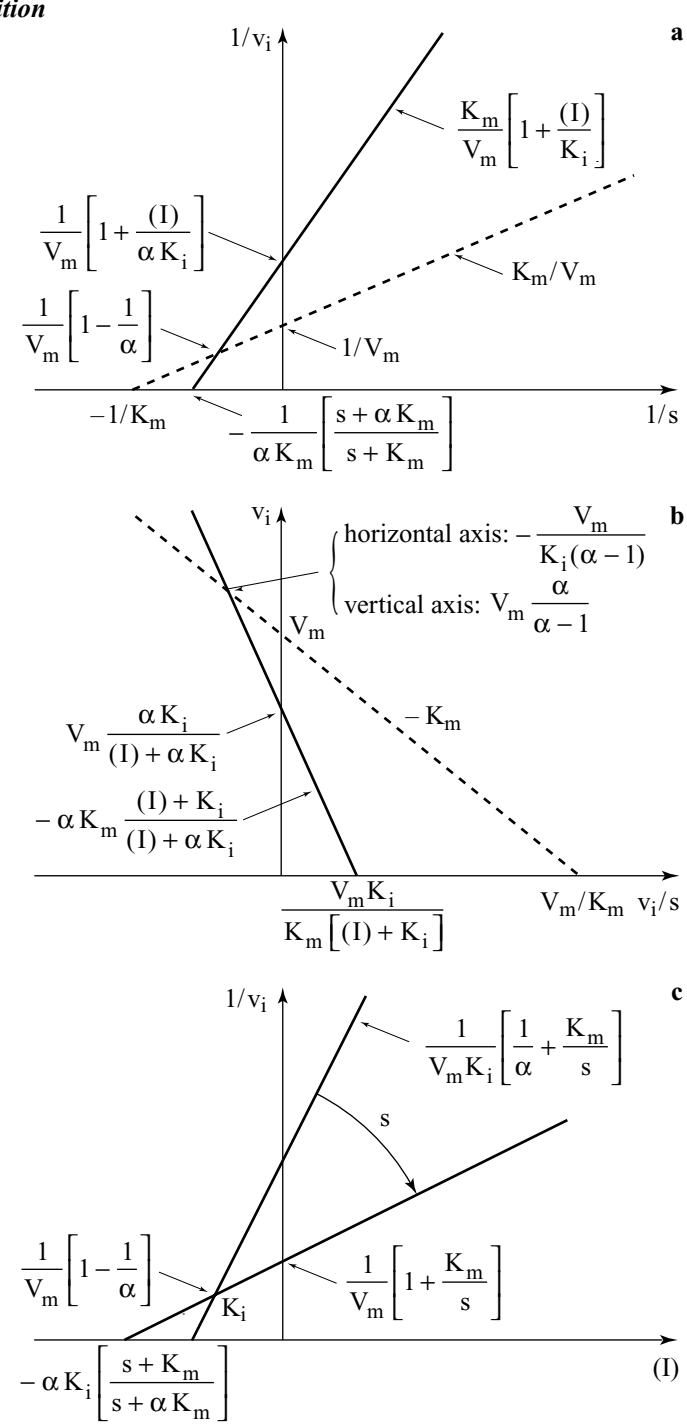


Fig. 5.22 Mixed inhibition
*(**a**) LINEWEAVER-BURK plot – (**b**) EADIE plot – (**c**) DIXON plot*

In general, there is mixed inhibition when $\alpha > 1$ and $\beta < 1$. The apparent affinity of the enzyme for its substrate is reduced in the presence of inhibitor and the rate of degradation of the ESI complex is less than that of the ES complex. At the same time, V'_m and K'_m are modified by the presence of inhibitor. The behaviour is intermediate between partial competitive and partial non-competitive inhibition.

If $0 < \beta < 1$, the inhibition is partially mixed. If $\beta = 0$, the inhibition is totally mixed. The ESI complex can no longer be converted. The reaction rate is given by the relationship:

$$v = V_m \frac{\alpha K_i}{(I) + \alpha K_i} \frac{s_0}{s_0 + K_s \dfrac{\alpha[(I) + K_i]}{(I) + \alpha K_i}}$$

The apparent parameters are as follows:

$$V'_m = V_m \frac{\alpha K_i}{(I) + \alpha K_i}$$

$$K'_m = K_m \frac{\alpha(I) + \alpha K_i}{(I) + \alpha K_i}$$

Figure 5.22 above indicates the determination of these parameters by the diverse graphical methods. ◢

We have outlined above the different types of simple inhibition, either total or partial. **In very general terms, we term competitive inhibition all those types that only modify the MICHAELIS constant, and non-competitive inhibition all those that modify the maximum reaction rate.** The importance of this generalisation will be revealed in the analysis of inhibition by the reaction products, for reactions with several substrates.

5.3.2. KINETICS OF ENZYMATIC REACTIONS IN THE PRESENCE OF AN ACTIVATOR

While certain effectors have an inhibitory effect on enzymatic reactions, conversely, others are able to activate them. Diverse types of activator exist: metal ions, anions, various natural molecules and the substrate itself can behave as an activator. These phenomena also arise in the regulation of cellular metabolism; in particular, they help to coordinate the regulation of several metabolic pathways. Coenzymes have sometimes been considered to be activators. Bearing in mind the kinetic mechanisms involved, we shall treat dissociable coenzymes as second substrates.

It is also important to consider the activators' modes of action. The mechanisms by which a substance is capable of activating an enzyme reaction can be as varied as the inhibitory mechanisms: either there is total activation (the enzyme has no activity in the absence of the activator), or there is partial activation (the enzyme has a low, but non-zero, activity in the absence of the activator); this only serves to increase the reaction rate.

5.3.2.1. TOTAL ACTIVATION

We shall distinguish several cases depending on whether the activator and the substrate bind to the enzyme randomly or in a sequential manner.

Independent binding of the substrate and activator

This situation is described by the following scheme:

$$E + S \overset{K_s}{\rightleftharpoons} ES$$

$$E + A \overset{K_a}{\rightleftharpoons} EA$$

$$EA + S \overset{K_s}{\rightleftharpoons} EAS \overset{k_2}{\longrightarrow} EA + P$$

$$ES + A \overset{K_a}{\rightleftharpoons} EAS \overset{k_2}{\longrightarrow} EA + P$$

The ES complex is inactive; only the ESA complex is capable of giving rise to the reaction products. In this case, there are no interactions between the substrate and activator binding sites. K_a represents the dissociation constant of the enzyme-activator complex or the enzyme-substrate-activator complex:

$$\frac{(E)(A)}{(EA)} = \frac{(ES)(A)}{(ESA)} = K_a$$

Similarly, the dissociation constant of the enzyme-substrate complex, K_s, is the same as the dissociation constant of the enzyme-activator-substrate complex:

$$\frac{(E)(S)}{(ES)} = \frac{(EA)(S)}{(EAS)} = K_s$$

The rate equation becomes: $$v = \frac{k_2 e s_0}{K_m + s_0} \frac{1}{1 + \frac{K_a}{a}}$$

where a is the activator concentration, having the condition that $a >> e_0$. Only V'_m increases as a function of the activator concentration.

Dependent binding of the activator and substrate

In this case, the binding of one of the ligands, activator or substrate, affects the affinity of the enzyme for the other. Assuming a quasi-equilibrium, we have the following relationship:

$$E + S \overset{K_s}{\rightleftharpoons} ES$$

$$E + A \overset{K_a}{\rightleftharpoons} EA$$

$$EA + S \underset{}{\overset{K'_s}{\rightleftharpoons}} EAS \xrightarrow{k_2} EA + P$$

$$ES + A \underset{}{\overset{K'_a}{\rightleftharpoons}} EAS \xrightarrow{k_2} EA + P$$

This system could be treated as a two-substrate reaction (see further on) approximating a quasi-equilibrium and with a single reaction product. It would therefore exhibit a random Bi-Uni mechanism. The rate equation is exactly the same:

$$v = \frac{V_m}{1 + \frac{K'_a}{a} + \frac{K'_s}{s_0} + \frac{K'_a K_s}{a s_0}}$$

but the activator is not converted during the reaction.

Sequential binding of the activator and substrate to the enzyme

The following reaction scheme describes this scenario:

$$E + A \underset{}{\overset{K_a}{\rightleftharpoons}} EA$$

$$EA + S \underset{}{\overset{K_s}{\rightleftharpoons}} EAS \xrightarrow{k_2} EA + P$$

The rate, in a state of quasi-equilibrium, is given by the relationship:

$$v = \frac{k_2 e s_0}{K_m\left(1 + \frac{K_a}{a}\right) + s_0}$$

Only the apparent MICHAELIS constant, K'_m, varies with the activator concentration.

Sequential binding of the substrate and activator to the enzyme

In this instance, the activator can only bind to the enzyme-substrate complex, which is described by the scheme:

$$E + S \underset{}{\overset{K_s}{\rightleftharpoons}} ES$$

$$ES + A \underset{}{\overset{K_a}{\rightleftharpoons}} EAS \xrightarrow{k_2} E + A + P$$

The rate, in a state of quasi-equilibrium, is:

$$v = k_2(ESA) \qquad \text{with } e_0 = (E) + (ES) + (ESA)$$

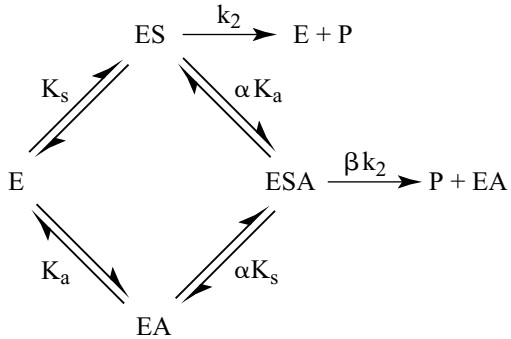
and so:

$$v = \frac{k_2 e_0}{\frac{K_s}{\frac{(K_a + a)}{K_a}} + s_0} \; \frac{1}{1 + \frac{K_a}{a}}$$

Both of the parameters K'_m and V'_m are affected by the presence of the activator.

5.3.2.2. PARTIAL ACTIVATION

There are a great number of enzyme reactions in which the ES complex reacts only weakly but whose turnover rates can be considerably enhanced by the presence of an activator. As before, for inhibition in which both the catalytic constant and the affinities are affected, the situation can be described in a general way using the following scheme:



As for partial inhibition, assuming that the system is in a state of quasi-equilibrium, the reaction rate is given by:

$$v = \frac{V_m s_0[\alpha K_a + \beta(A)]}{s_0(A) + \alpha[K_a s_0 + K_s(A) + K_s K_a]}$$

with $\beta > 1$; α can be > 1 if β is large enough for the rate to be increased in the presence of the activator. The apparent kinetic parameters are as follows:

$$V'_m = V_m \frac{\alpha K_a + \beta(A)}{\alpha K_a + (A)}$$

$$K'_s = K_s \frac{\alpha K_a + \alpha(A)}{K_a + \beta(A)}$$

If $\alpha = 1$, the kinetic parameters become:

$$V'_m = V_m \frac{K_a + \beta(A)}{K_a + (A)}$$

$$K'_s = K_s \frac{K_a + (A)}{K_a + \beta(A)}$$

and thus: $$V'_m \times K'_s = V_m \times K_s$$

Furthermore, if β is very large, then $\beta k_2 >> k_2$, which becomes again a case of independent substrate and activator binding.

5.3.2.3. EXAMPLES OF ENZYMATIC ACTIVATION

We shall give the example of the activation of β-galactosidase from *E. coli* by Mg^{++} ions although its properties are a little more complex than in the simple systems described in this section. The enzyme is a tetramer with a molecular weight of 540 000 daltons, formed from four identical protomers, and catalyses the hydrolysis of β-D-galactosides. β-galactosidase is sensitive to the action of diverse cations. Na^{+} ions are indispensable for its activity. Certain divalent cations, in particular Mg^{++} and Mn^{++}, are activators; on the contrary, others like Be^{++} and Ca^{++}, are inhibitors. The enzyme possesses a weak activity, yet significant in the absence of Mg^{++}. The properties of the enzyme deprived of Mg^{++} were determined by TENU et al. (1972).

This quantitative study of activation by Mg^{++} was carried out at the optimal reaction pH, in well-controlled conditions, in particular by keeping the Na^{+} concentration (0.145 M) and the ionic strength (0.17 ± 0.02) constant. Phenylgalactoside was chosen as the substrate; with this substrate, the first chemical step is rate-limiting, which facilitated the interpretation. Additionally, the conditions were chosen such that the substrate concentration was much above K_m. Figure 5.23 shows the activation by different concentrations of Mg^{++}. The latent phase observed when Mg^{++} is added to the reaction at time 0 shows that activation by magnesium is a slow process. After a variable time-period dependent on the Mg^{++} concentration, the reaction reaches a steady state, which corresponds to the linear kinetic phase; the rate becomes constant. If the enzyme is first incubated in the presence of magnesium, this latent phase is no longer observed. The slow activation by magnesium is independent of the substrate's nature. Therefore, the activation must result from the binding of Mg^{++} to the enzyme.

In this section, we shall only present the results obtained when the steady phase of the activation is reached, as represented by the linear parts of the curves in Fig. 5.23.

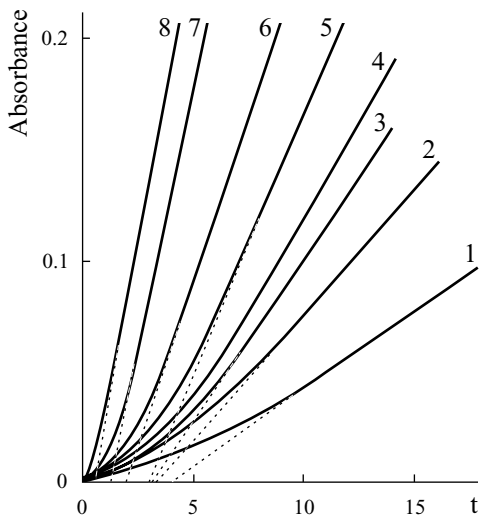


Fig. 5.23 Activation of β-galactosidase from E. coli by Mg^{++}
Enzyme concentration: 3.26 μM; substrate concentration: 10 mM; pH 7.0 at 5°C; Mg^{++} concentration: (1) 78 nM, (2) 202 nM, (3) 317 nM, (4) 420 nM, (5) 670 nM, (6) 1.23 μM, (7) 7.03 μM, (8) 50 μM.
(Reproduced from *Eur. J. Biochem.,* **26**, TENU J.P. *et al.*, Activation of β-galactosidase by Mg^{++}, 110. © (1972) with permission of Blackwell Publishing Ltd)

By studying the change in the catalytic constant as a function of the Mg^{++} concentration (M), we can determine an apparent dissociation constant for the enzyme-Mg^{++} complex, using this type of representation:

$$\Delta k_{cat} = (\Delta k_{cat})_{max} \frac{(M)}{(M + K)}$$

Δk_{cat} represents the difference in values of the catalytic constant k_{cat} for a given Mg^{++} concentration and for zero Mg^{++} concentration (k_{cat} being different from 0 under these conditions); $(\Delta k_{cat})_{max}$ is the corresponding difference at saturating Mg^{++}. K is the dissociation constant of the enzyme-Mg^{++} complex. Figure 5.24 reveals that the saturation curve is Michaelian. The constant K has a value of $(0.65 \pm 0.05) \times 10^{-6}$ M, which corresponds to a strong apparent affinity of the enzyme for the metal ion. Direct measurements of Mg^{++} binding to the enzyme showed that a single Mg^{++} ion binds per enzyme protomer.

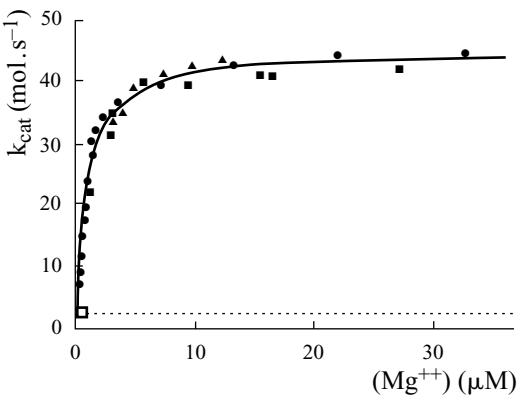


Fig. 5.24 β-galactosidase activation by Mg^{++} ions *Changes in the catalytic constant for the hydrolysis of phenylgalactoside by the enzyme as a function of the Mg^{++} concentration.* Dashed line*: residual activity in the absence of Mg^{++}.* (Reproduced from *Eur. J. Biochem.*, **26**, TENU J.P. *et al.*, Activation of β-galactosidase by Mg^{++}, 110. © (1972) with permission of Blackwell Publishing Ltd)

5.3.2.4. ACTIVATION BY THE SUBSTRATE

There are numerous examples of substrate inhibition described in the literature, much rarer are cases describing activation. However, this phenomenon has been observed for some enzymes. ALBERTY et al. (1954) noted the activation of fumarase by its substrate. BÉCHET and YON (1964) described a similar phenomenon for the hydrolysis of several substrates by trypsin (Fig. 5.25 below).

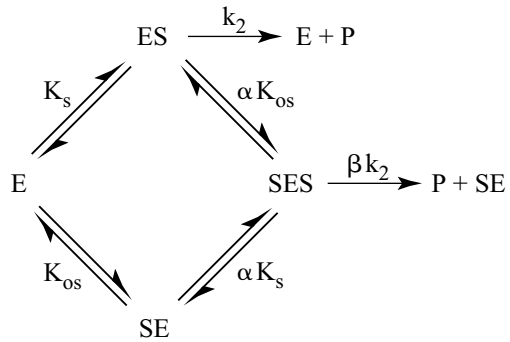
The activation by the substrate can be described by a phenomenological rate equation of the form:

$$v = \frac{e_0(as + bs^2)}{1 + cs + ds^2}$$

where a change in s to the second degree appears in the numerator and denominator; a, b, c, and d are constants that depend on the experimental kinetic parameters of the reaction.

If this phenomenological equation remains valid, diverse mechanisms may be involved in the apparent activation of an enzyme by high substrate concentrations. Clearly, the significance of these parameters depends on the mechanisms involved.

In this paragraph, we shall only explore the simplest mechanism, involving the existence of two substrate-binding sites. One is the normal site where the substrate is converted and the other is an activator site where the substrate undergoes no modification. The scheme is shown below:



If $\alpha = 1$, the sites are equivalent and independent. This scheme thus reminds us of the general scheme given for inhibition; but with activation, $\beta > 1$. The reaction rate is:

$$v = V_m \frac{s_0(\alpha K_{os} + \beta s_0)}{s_0^2 + \alpha(K_{os}s_0 + K_s s_0 + K_s K_{os})}$$

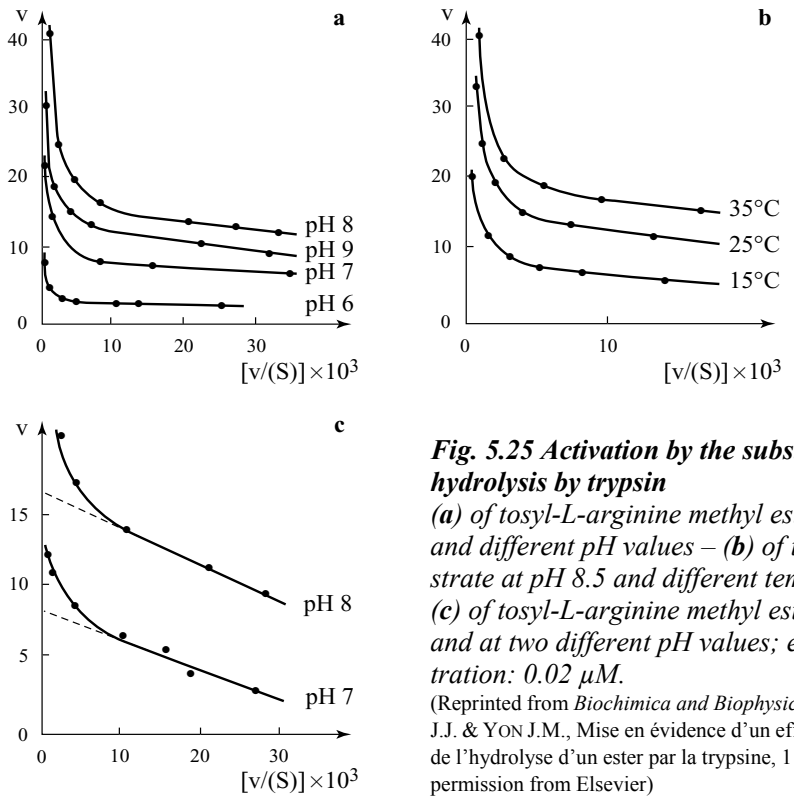


Fig. 5.25 Activation by the substrate during hydrolysis by trypsin
*(**a**) of tosyl-L-arginine methyl ester at 35°C and different pH values – (**b**) of the same substrate at pH 8.5 and different temperatures (**c**) of tosyl-L-arginine methyl ester at 35°C and at two different pH values; enzyme concentration: 0.02 μM.*
(Reprinted from *Biochimica and Biophysica Acta*, **89**, BECHET J.J. & YON J.M., Mise en évidence d'un effet allostérique lors de l'hydrolyse d'un ester par la trypsine, 117. © (1964) with permission from Elsevier)

Figure 5.25 illustrates the activation by an excess of substrate during the hydrolysis of benzoyl-L-arginine methyl ester by trypsin. The EADIE plot is no longer linear. It is possible to define two values for V_m and K_m in the zones of extreme substrate concentrations.

It is important to note that other types of mechanism lead to kinetic profiles of this sort as we shall see later on, and this is particularly true with, for example, mnemonic, hysteretic and allosteric enzymes.

5.4. *ENZYMATIC REACTIONS WITH ONE SUBSTRATE AND SEVERAL INTERMEDIATE COMPLEXES*

5.4.1. *KINETICS AT THE STEADY STATE*

The numerous data currently available, studied in detail for diverse enzymatic reactions, in many cases generally show that the enzymatic reaction pathway includes several intermediate complexes. However, the number of intermediate complexes remains limited.

By way of example, we shall discuss the following scheme, for the general case of a reversible reaction. We shall then turn to the reaction in a single direction when the reverse reaction is negligible:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} Y \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} E + P$$

Several methods exist for determining the rate equation at steady state, i.e. when $d(X)/dt = d(Y)/dt = 0$. The equations for the system under these conditions are:

$$\begin{aligned} d(X)/dt &= k_1(E)s - (X)(k_{-1} + k_2) + k_{-2}(Y) = 0 \\ d(Y)/dt &= k_{-3}(E)p + k_2(X) - (Y)(k_{-2} + k_3) = 0 \\ (E) &+ (X) + (Y) = 0 \\ s &+ p = s_0 \end{aligned}$$

The last two equations are, respectively, the equations for the conservation of enzyme and substrate.

Under conditions in which only the two intermediate complexes are present, these equations are readily solvable. However, when the number of intermediates increases, the equations become more complex but there are adequate methods to solve them. In particular some simple methods for kinetic analyses exist, such as:

- by determinants,
- by graphical methods, including that of KING and ALTMAN.

These methods allow the determination of the reaction rate which, in the direction of product appearance, is given by the relationship:

$$v = dp/dt = k_3(Y) - k_{-3}(E)p$$

It is useful to determine (X), (Y) and (E).

5.4.1.1. KINETIC ANALYSIS BY DETERMINANTS

The rules of CRAMER enable the determination of (X), (Y) and (E) by writing the determinants of $X(D_x)$, $Y(D_y)$ and $E(D_e)$ and the determinant of the coefficients (D).

We have: $(E) = D_e/D \quad ; \quad (X) = D_x/D \quad ; \quad (Y) = D_y/D$

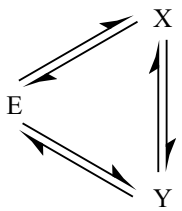
The determinant of the free enzyme is written below as an example:

$$D_e = \frac{\begin{vmatrix} 0 & -(k_1+k_2) & k_{-2} \\ 0 & k_2 & -(k_2+k_3) \\ e_0 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} k_1 s & -(k_{-1}+k_2) & k_{-2} \\ k_{-3}p & k_2 & -(k_{-2}+k_3) \\ 1 & 1 & 1 \end{vmatrix}}$$

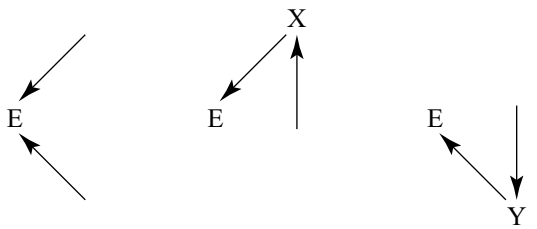
Solving this system leads to the following equation:

$$v = \frac{e_0(k_1k_2k_3s - k_{-1}k_{-2}k_{-3}p)}{k_{-1}k_3 + k_2k_3 + k_{-1}k_{-2} + k_1(k_2+k_{-2}+k_3)s + k_{-3}(k_{-1}+k_2+k_{-2})p}$$

5.4.1.2. ANALYSIS BY THE GRAPHICAL METHOD OF KING AND ALTMAN



To solve this system, we write that $(E)/e_0$, $(X)/e_0$ and $(Y)/e_0$ are the sums of all allowed combinations leading to E, X and Y respectively, divided by the total number of all combinations. A closed loop is a forbidden loop. Thus, for the free enzyme the allowed combinations are the following configurations:



Then, we have:

$$\frac{(E)}{e_0} = \frac{k_{-1}k_3 + k_2k_3 + k_{-2}k_{-1}}{\Sigma}$$

$$\frac{X}{e_0} = \frac{k_1k_{-2}s + k_{-3}k_{-2}p + k_1k_3s}{\Sigma}$$

$$\frac{(Y)}{e_0} = \frac{k_1k_2s + k_{-1}k_{-3}p + k_2k_{-3}p}{\Sigma}$$

Σ represents the sum of the numerators of the three expressions above. Solving this system, of course, leads to the same rate equation as before.

5.4.1.3. ANALYSIS BY OTHER GRAPHICAL METHODS

Other graphical methods have been suggested for deriving rate equations for complex kinetic schemes. The theory of graphs has been applied to solve enzyme kinetics at steady state and to problems involving simple inhibition by VOLKENSTEIN and GOLDSTEIN (1966). CHOU and FORSEN (1980), and CHOU (1980, 1981) applied the method to solve enzyme reactions composed of branched schemes. With all these methods, including that of KING and ALTMAN, there is a fundamental formula that expresses the concentration of the mth form of the enzyme in the reaction:

$$(E_m) = \frac{N_m}{\sum_i^n N_i} e_0$$

n being the total number of enzyme forms in the reaction scheme. The different methods involve obtaining N_i (i = 1, 2, 3 … m … n) in the simplest way.

Let us consider the scheme comprising the two intermediates X and Y. **In the first step,** we draw a graph of the enzyme's states in which each point represents one species of the enzyme; the arcs joining them up represent the paths for going from one species to another (Fig. 5.26a below). **In the second step,** each point on the graph is associated with a loop whose weight is equal to the sum of the arcs that leave that point. Next, the sign is changed for each arc on the graph. Graph D is thus transformed into graph D^+ (Fig. 5.26b). **In the third step,** we take a reference point, for example the species Y, and we trace every graph that comprises a path going from Y to E and all cycles which neither have an intersection with them nor with the paths (Fig. 5.26c). For each sub-graph, we take the weight of all the points multiplied by the sign $(-1)^{n+c+1}$, n being the number of points, and c, the number of cycles in the corresponding sub-graph.

Thus:

$$N_E = (-1)^{3+1}(-k_{-1})(-k_{-2}) + (-1)^{3+1+1}(-k_{-3})(k_{-1} + k_2)$$
$$= k_{-1}k_{-2} + k_3(k_{-1} + k_2)$$

In the same manner we have:

$$N_X = k_1k_3s + k_{-2}(k_1s + k_{-3}p)$$
$$N_Y = k_1k_2s + k_{-3}p(k_{-1} + k_2)$$

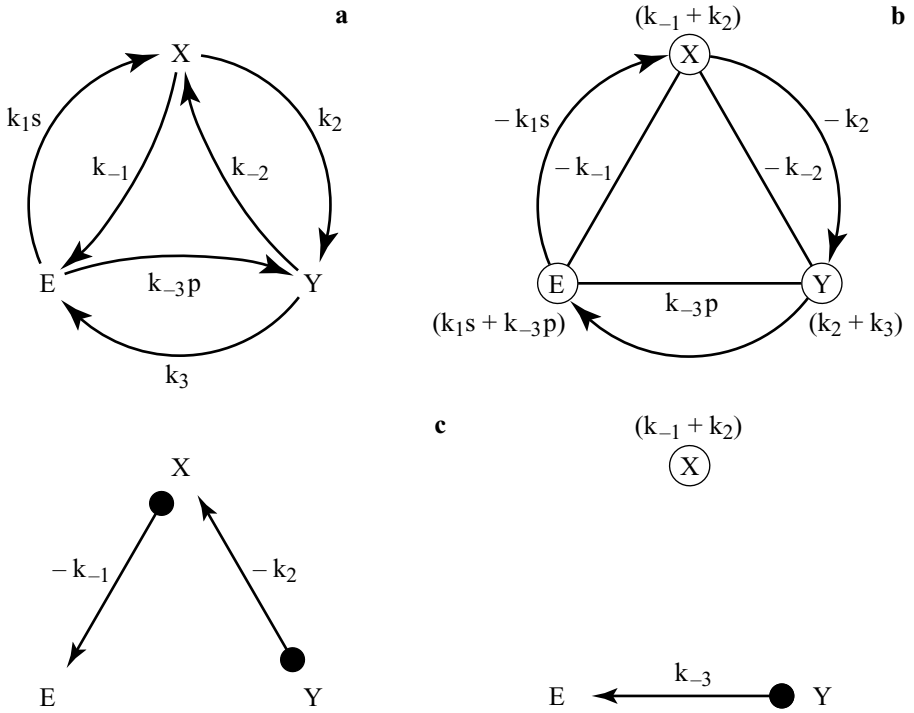


Fig. 5.26 Solving a system comprising two reaction intermediates by the graphical method

In the last step, to avoid omitting a sub-graph and to facilitate verification, the authors recommend the use of methods that involve construction of a matrix $A = a_{ij}$ from the transformed graph D^+:

$|a_{ij}| = 1$ if a path exists between E_i and E_j,

$|a_{ij}| = 0$ if a path does not exist between E_i and E_j.

Thus, for a point E_m, if we take E_q as a reference point, the number of sub-graphs is:

$$n^{q \to n} = A(D^+)_{mq}$$

$a(D^+)$ is the sub-matrix obtained by eliminating the nth row and the qth column from $A(D^+)$. Thus, in the scheme:

$$A(D^+) = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 2$$

it is clear that there are two sub-graphs. The rate expression is given by the relationship:

$$v = k_3 y - k_{-3}(E)p$$

$$v = \left(k_3 \frac{N_Y}{N_E + N_X + N_Y} - k_{-3} p \frac{N_E}{N_E + N_X + N_Y} \right) e_0$$

and we find once again the rate expression obtained by the previous methods.

The graphical method is useful because it enables a certain number of simplifications by direct operations on the graphs. Thus:

- parallel branches can be added. The value of the resulting branch is equal to the sum of the values of the branches;
- the graph can be simplified if some branches fuse together by using the graph's symmetry;
- the number of nodes can be reduced.

In order to illustrate these operations, let us look again at the example of the action of a modifier M (a partial inhibitor or activator) on the enzymatic reaction that we discussed in the previous section. Taking the scheme:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

$$E + M \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} EM$$

$$ES + M \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} EMS \xrightarrow{\beta k_2} EM + P$$

$$EM + S \underset{k_{-5}}{\overset{k_5}{\rightleftharpoons}} EMS \xrightarrow{\beta k_2} EM + P$$

Let E_1 be the free enzyme, E_2 the ES complex, E_3 the EM complex and E_4 the ternary complex EMS. We can sketch the schema D (Fig. 5.27a). The first property –addition of parallel branches– can be used to write a simplified form of the graph (Fig. 5.27b). The number of nodes can be reduced.

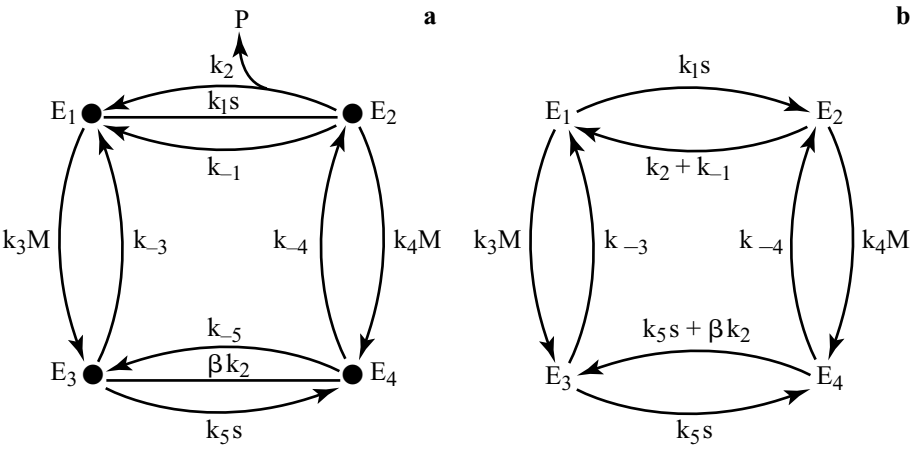


Fig. 5.27 Graph of an enzyme system containing a modifier, M

To calculate E_3 by taking E_2 as a reference point, two paths exist to go from one to the other: one is $E_2 \longrightarrow E_4 \longrightarrow E_3$, the other $E_2 \longrightarrow E_1 \longrightarrow E_3$, which, by

applying the principles outlined above from the transformed graph (Fig. 5.28a), would give the two sub-graphs indicated in Fig. 5.28b. However, each of these paths can be condensed into a single point as indicated in Fig. 5.28c. As a result:

$$N_{E3} = (k_1 s + k_3 M)(k_{-5} + k_2) + (k_{-5} + k_2 k_{-4})(k_2 + k_{-1})k_{-3} M$$

In the same way, we would determine N_{E1}, N_{E2} and N_{E4}.

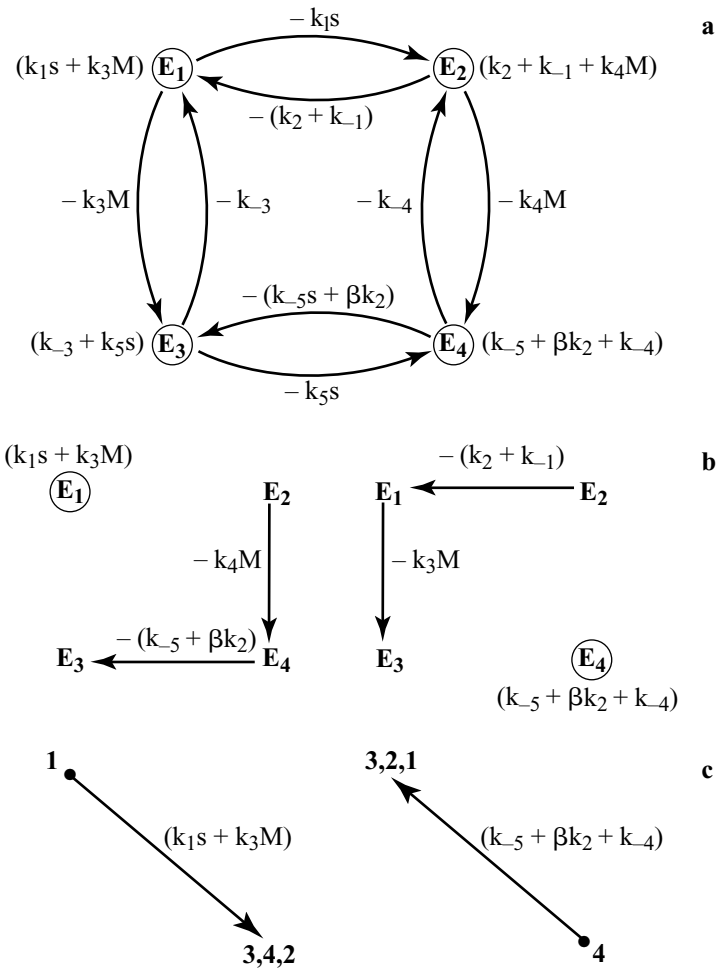


Fig. 5.28 Simplification of the graph in Fig. 5.27

5.4.1.4. RELATIONSHIP BETWEEN THE PARAMETERS OF THE RATE EQUATION IN REACTIONS WITH A SINGLE SUBSTRATE AND TWO INTERMEDIATE COMPLEXES

The general rate equation can be written in a more expressive form:

$$v = \frac{e_0(N_1 s - N_2 p)}{C_0 + C_1 s + C_2 p}$$

In this expression:

$$N_1 = k_1k_2k_3$$
$$N_2 = k_{-1}k_{-2}k_{-3}$$
$$C_0 = k_{-1}k_3 + k_{-1}k_{-2} + k_2k_3$$
$$C_1 = k_1(k_2 + k_{-2} + k_3)$$
$$C_2 = k_{-3}(k_{-1} + k_2 + k_{-2})$$

It is interesting to note that the reaction parameters are related to the complex constants. So, the equilibrium constant $K_{eq} = N_1/N_2$. The kinetic parameters of the forward reaction are related to the constants by the relationships:

$$k_{cat,A} = V_{m,A}/e_0 = N_1/C_1$$
$$K_{m,s} = C_0/C_1$$

and those for the reverse reaction:

$$k_{cat,R} = V_{m,R}/e_0 = N_2/C_2$$
$$K_{m,p} = C_0/C_2$$

The kinetic parameters of the forward reaction are obtained from the general rate equation when p is zero; those of the reverse reaction are obtained when s is zero.

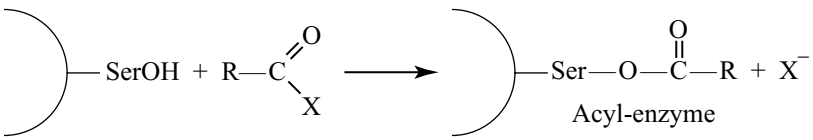
The expression for K_m clearly shows that this parameter has a complex value; thus the MICHAELIS constant with respect to the substrate S is:

$$K_{m,s} = \frac{k_{-1}k_3 + k_{-1}k_{-2} + k_2k_3}{k_1(k_2 + k_{-2} + k_3)}$$

It is clear that this parameter, which experimentally always represents the substrate concentration that leads to the half-maximal reaction rate, is a complex parameter; it does not reflect the inverse of the enzyme's affinity for its substrate.

5.4.2. EXAMPLE: ENZYMATIC REACTIONS CATALYSED BY SERINE PROTEASES

It has been demonstrated that hydrolytic enzymes possessing a serine in the active site form an acyl-enzyme intermediate with the substrate, an acyl-serine (see Chap. 12), according to the scheme:



This is a covalent complex between part of the substrate and a serine, the other part of the substrate or "leaving group" is liberated after this step in the reaction. In the following step, the acyl-enzyme is hydrolysed with the incorporation of a water molecule and the enzyme is regenerated.

The reaction scheme is written as follows:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES_1 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} ES_2 \xrightarrow{k_3} E + P_2$$
$$+ P_1$$

ES_1 corresponds to the classic MICHAELIS complex, ES_2 is the acyl-enzyme, P_1 is the leaving group and P_2 the second reaction product. The leaving group can be a non-specific part of the substrate as in the case of trypsin, chymotrypsin, elastase, or a specific part as seen in acetylcholinesterase.

Letting p and q be the respective concentrations of ES_1 and ES_2, at steady state we can write the following equations:

$$k_1(e_0 - p - q)s + k_{-2}\, q = (k_1 + k_2)p$$
$$k_2 p = (k_{-2} + k_3)q$$

The reaction rate is $v = k_3 q$. We end up with the rate equation below:

$$v = \frac{\dfrac{k_2 k_3 e_0}{k_2 + k_{-2} + k_3}}{1 + \dfrac{1}{s} \times \dfrac{k_{-1}k_{-2} + k_{-1}k_3 + k_2 k_3}{k_1(k_2 + k_{-2} + k_3)}}$$

5.4.3. SIGNIFICANCE OF THE KINETIC PARAMETERS

The kinetic parameters of the MICHAELIS equation, K_m and V_m, become the apparent parameters that we may determine experimentally and which correspond to these complex expressions:

$$V_{m,app} = \frac{k_2 k_3 e_0}{k_2 + k_{-2} + k_3}$$

$$K_{m,app} = \frac{k_{-1}k_{-2} + k_{-1}k_3 + k_2 k_3}{k_1(k_2 + k_{-2} + k_3)}$$

If k_{-2} is very small relative to k_3, i.e. if the acylation process is practically irreversible under the experimental reaction conditions, these expressions may be simplified to:

$$V_{m,app} = \frac{k_2 k_3 e_0}{k_2 + k_3}$$

$$K_{m,app} = \frac{k_3(k_{-1} + k_2)}{k_1(k_2 + k_3)}$$

with $V_{m,app}/e_0 = k_{cat}$ and $K_{m,app} = K_m\, k_3/(k_2 + k_3)$.

Two extreme cases can thus arise depending on which of the two processes, acylation or deacylation, is limiting.

5.4.3.1. ACYLATION IS LIMITING: $k_2 \ll k_3$

$$V_{m,app} = k_2 e_0 \quad \text{and} \quad k_{cat} = k_2$$
$$K_{m,app} = (k_{-1} + k_2)/k_1 = K_m$$

5.4.3.2. DEACYLATION IS LIMITING: $k_2 \gg k_3$

$$V_{m,app} = k_3 e_0 \quad \text{and} \quad k_{cat} = k_3$$
$$K_{m,app} = \frac{(k_{-1} + k_2)}{k_1} \frac{k_3}{k_2} = K_m \frac{k_3}{k_2}$$

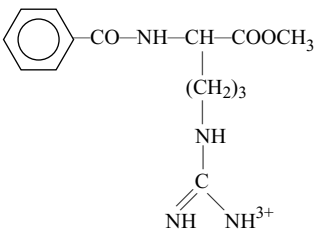
In the most general case, a simple relationship exists between the parameters k_{cat} and $K_{m,app}$:

$$\frac{k_{cat}}{K_{m,app}} = \frac{k_2}{K_m}$$

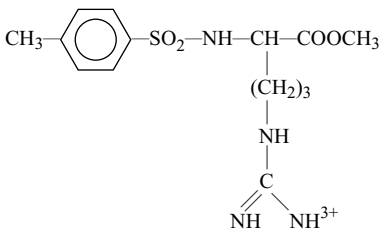
5.4.4. DETERMINATION OF THE ELEMENTARY KINETIC CONSTANTS

Generally, the studies carried out under steady state conditions do not permit the determination of experimental parameters with a complex significance. The parameters of individual steps can only be obtained in extreme cases. It is however possible for particular enzymatic systems, e.g. reactions catalysed by serine proteases, to determine all system parameters including the individual rate constants by using a range of carefully chosen substrates.

◤ Let us take the reactions catalysed by trypsin as examples. In the biological context, this enzyme hydrolyses peptide bonds where the amino acid at the position α to the carboxyl is L-lysine or L-arginine. But it can also hydrolyse ester and amide bonds requiring the same specificity. The tryptic hydrolysis of ester and amide derivatives of L-benzoyl arginine and tosyl-L-arginine has been studied:



Benzoyl-L-arginine methyl ester



Tosyl-L-arginine methyl ester

The corresponding ester and amide, e.g. benzoyl-L-arginine ethyl ester (BAEE) and benzoyl-L-arginine amide (BAA), give the same acyl-enzyme intermediate, benzoyl-L-arginine trypsin. These two substrates will, therefore, have the same rate constants for deacylation, k_3.

A kinetic study of the ester and amide hydrolyses gave the following kinetic parameters for the two substrates:

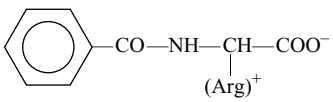
For BAEE: $k_{cat} = 23\ s^{-1}$ and $K_{m,app} = 2.6 \times 10^{-6}$ M

For BAA: $k_{cat} = 1.1\ s^{-1}$ and $K_{m,app} = 2.4 \times 10^{-3}$ M

It is of note, upon inspection of these values, that:

- the deacylation step cannot be rate-limiting for the hydrolysis of these two substrates because the values of k_{cat} are not identical;
- the acylation step does not seem to be rate-limiting for the two substrates. If this were the case, then $K_{m,app}$ would be equal to K_m. Now, there is a difference of 10^3 between the $K_{m,app}$ values of the two substrates whose structures differ extremely little. Nevertheless, no conclusions can be drawn and further experiments are necessary.

Complementary information was obtained by studying the inhibition by the reaction product, benzoyl-L-arginine:



Benzoyl-L-arginine behaves as a competitive inhibitor with respect to the enzyme, with an inhibition constant $K_i = 2.5 \times 10^{-3}$ M. Comparing the K_m of the amide to the K_i of the inhibitor analogue, which has the same value, it is reasonable to think that for amide hydrolysis $K_m = K_s$, and therefore that acylation is rate-limiting; consequently, $k_{cat} = k_2$.

On the other hand, the difference between K_s and $K_{m,app}$ for ester hydrolysis is a factor of 10^3, suggesting that deacylation is rate-limiting, and so:

$$\frac{K_{m,app}}{K_s} = \frac{k_3}{k_2} = 10^{-3}$$

Thus, it is possible to obtain the kinetic parameters for all the elementary steps; k_3 is identical for the two substrates and $k_2 = 1\ 000\ k_3$ for the ester. Table 5.1 summarises the results of this study.

Table 5.1 Values for the kinetic parameters corresponding to the tryptic hydrolysis of a benzoyl-L-arginine ester and amide

Substrate	*k_{cat} (s^{-1})*	*k_2 (s^{-1})*	*k_3 (s^{-1})*	*K_s or K_i (10^{-3} M)*	*$K_{m,app}$ (M)*
BAEE	23	23 000	23	2.4	2.6×10^{-6}
BAA	1.1	1.1	23	2.4	2.4×10^{-3}
BA				2.5	

5.4.5. STUDY OF NUCLEOPHILIC COMPETITION

A large number of hydrolytic enzymes do not possess a very narrow specificity for their second substrate, namely, water. In certain cases, it is possible to substitute other more potent nucleophilic agents for water. Analysis of the reaction products therefore provides a means to understand how these nucleophilic agents participate in an enzymatic reaction. This method suggested by KOSHLAND and HERR (1957) has been applied with success to a great number of systems such as ribonuclease (FINDLAY et al., 1960), diverse cholinesterases (WILSON et al., 1950), papain and subtilisin (GLAZER, 1966), chymotrypsin (BENDER et al., 1966), and to trypsin and β-galactosidase (YON et al., 1967; 1973).

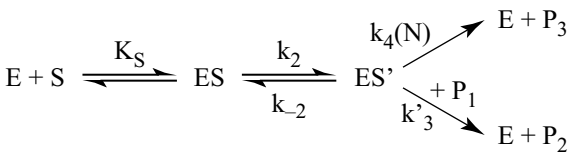
We have just seen that during reactions catalysed by trypsin and chymotrypsin a covalent acyl-enzyme intermediate is formed. This entity reacts with water to liberate the second reaction product, the first product having been liberated during the formation of the acyl-enzyme. Such a mechanism, involving a covalent intermediate, has been established for several enzymes, including serine proteases and esterases, cysteine proteases and glycosidases. YON and co-workers demonstrated this mechanism in reactions catalysed by β-galactosidase (1973).

For serine proteases, competition between water and other nucleophilic agents takes place at the acyl-enzyme step; for β-galactosidase, the competition arises with the galactosyl-enzyme. The nucleophile acts as an acceptor, the part of the substrate covalently linked to the enzyme can thus be transferred to this acceptor. This property is used industrially in peptide synthesis, in which the nucleophile is an amino acid or a peptide that can be linked to a part of a protein or to another peptide.

In kinetic studies of nucleophilic competition we must distinguish two situations according to whether or not the enzyme presents a significant affinity for water and the acceptor.

5.4.5.1. STUDY OF NUCLEOPHILIC COMPETITION IN THE CASE WHERE NO BINDING SITE EXISTS FOR WATER AND ITS ANALOGUES

BENDER et al. (1966) studied the kinetic consequences of nucleophilic competition in hydrolysis reactions of esters using chymotrypsin. The kinetic scheme that was proposed is shown below:



Scheme 5.1

where S is a substrate of the form: R–CO–X–R′ with X = O in the case of esters, or X = NH in the case of amides and peptides. If the substrate is a peptide bond and the nucleophile another amino acid, a transpeptidation reaction occurs. Such reactions, as well as transglycosylations, take place in a biological environment. P_1, the

first reaction product, is R′–XH; P_2 is the hydrolysis product of the acyl-enzyme, R–COOH; P_3, the deacylation product from the nucleophilic agent, N = R″YH, has the form: R–CO–Y–R″. K_s is the dissociation constant of the ES complex, k_2 is the acylation constant and k_4 is the deacylation constant from the nucleophile N. If P_1 is itself utilised as a nucleophile, k_4 is identical to k_{-2}, the reverse acylation constant. The constant k'_3 is assumed to be equal to $k_3(W)$, (W) being the water concentration and may be considered to be constant if, in the reaction medium, it changes negligibly.

General equations

Solving the system at steady state, for the initial rates of product appearance P_i (i = 1, 2, 3), leads to the following equations:

$$\frac{dP_i}{dt} = \frac{k_{cat,i}e_0}{1+\frac{K_{m,app}}{s}} \qquad [4]$$

The rate of appearance of P_1 is:

$$k_{cat,1} = \frac{k_2[k'_3 + k_4(N)]}{k_2 + k'_3 + k_4(N)} \qquad [5]$$

and of P_2:

$$k_{cat,2} = \frac{k_2 k'_3}{k_2 + k'_3 + k_4(N)} \qquad [6]$$

and of P_3:

$$k_{cat,3} = \frac{k_2 k_3(N)}{k_2 + k'_3 + k_4(N)} \qquad [7]$$

The apparent MICHAELIS constant, identical in the three cases, is written:

$$K_{m,app} = K_s \frac{k'_3 + k_4(N)}{k_2 + k'_3 + k_4(N)} \qquad [8]$$

We note that:

$$\frac{dP_1}{dt} = \frac{dP_2}{dt} + \frac{dP_3}{dt} \qquad [9]$$

From the previous equations, it is possible to distinguish three important cases depending on the relative values of the rate constants, k_2 and k'_3, and assuming that $k_4(N)$ is the same order of magnitude as k'_3, which is always achievable by choosing a suitable concentration of the added nucleophile N.

Case I – If $k_2 >> k'_3$, i.e. if the deacylation step is rate-limiting, expressions [5], [6], [7] and [8] can be simplified. The apparent MICHAELIS constant is written:

$$K_{m,app} = K_s \frac{k'_3 + k_4(N)}{k_2}$$

The catalytic constant for the appearance of the product P_1 (case I-P1) is:

$$k_{cat,1} = k'_3 + k_4(N)$$

with $k_{cat,1}$ and $K_{m,app}$ being linear functions of (N) and possibly (W). The rate of appearance of product P_1 increases with (N) when $s >> K_{m,app}$; when $s << K_{m,app}$, no further effect is observed.

The rate of appearance of product P_2 (case I-P2):

$$k_{cat,2} = k'_3$$

is independent of (N). This is strictly identical to competitive inhibition with:

$$K_I = (k_4/k'_3)^{-1}$$

The rate of product appearance P_3 (case I-P3):

$$k_{cat,3} = k_4(N)$$

increases linearly with (N).

Case II – If $k'_3 >> k_2$, i.e. if the acylation step is rate-limiting, the expressions [5], [6], [7] and [8] may also be simplified. The apparent MICHAELIS constant $K_{m,app}$ becomes equal to K_s, the dissociation constant of the first enzyme-substrate complex.

The rate of appearance of product P_1 (case II-P1) is:

$$k_{cat,1} = k_2$$

In this case, (N) has no effect on $k_{cat,1}$ nor on $K_{m,app}$.

The rate of appearance of product P_2 (case II-P2) is:

$$k_{cat,1} = \frac{k_2 k_3}{k'_3 + k_4(N)}$$

$K_{m,app}$ being independent of (N):

$$\left(\frac{dP_2}{dt}\right)^{-1} = \left(\frac{1}{k_2} + \frac{k_4(N)}{k_2 k'_3}\right)\left(1 + \frac{K_s}{s}\right)(e_0)^{-1}$$

The inverse of the appearance rate of product P_2 is a linear function of (N), irrespective of the concentration of s. This case is identical to non-competitive inhibition with:

$$K_I = (k_4/k'_3)^{-1}$$

The appearance rate of product P_3 (case II-P3) is:

$$k_{cat,3} = \frac{k_2 k_4(N)}{k'_3 + k_4(N)}$$

The appearance rate of product P_3 increases with (N), and tends towards the limit value k_2 when $k_4(N) >> k'_3$.

Case III – If $k'_3 \sim k_2$, i.e. if both acylation and deacylation are partially rate-limiting, the previous expressions can no longer be simplified, and (N) affects both $K_{m,app}$ and k_{cat}. $K_{m,app}$ is a complex function of (N) which tends towards a limit value equal to K_s when $k_4(N)$ becomes very large relative to k_2.

The rate of appearance of product P_1 (case III-P1), $k_{cat,1}$ increases with (N) and tends towards a limit value equal to k_2 when $k_4(N) >> k_2$. For $s << K_{m,app}$, dP_1/dt is independent of (N).

Regarding the appearance of product P_2 (case III-P2), $1/k_{cat,2}$ is a linear function of (N):

$$\frac{1}{k_{cat,2}} = \frac{1}{k_2} + \frac{1}{k'_3} + \frac{k_4(N)}{k_2 k_3}$$

The inhibition is a maximum when $s << K_{m,app}$, so we have:

$$\left[\frac{1}{se_0}\ \frac{dP_2}{dt}\right]^{-1} = \frac{K_s}{k_2}\left[1 + \frac{k_4(N)}{k'_3}\right]$$

The inhibition decreases when the concentration of s increases, but does not cancel out, even when the enzyme is saturated by the substrate, which distinguishes it from what is observed in case I-P2. This form of inhibition is formally identified as partial competitive and partial non-competitive inhibition.

The appearance rate of product P_3 (case III-P3), $k_{cat,3}$, increases with (N) and tends towards a limit value equal to k_2 when $k_4(N)$ becomes very large relative to k_3 and k'_3. Case III-P3 is interesting as it is possible to obtain all of the system's kinetic constants by studying the effect of the concentration of (N) on the kinetic parameters of the reaction. For example, in case III-P2, we can determine all of the system's kinetic constants from the experimental parameters.

Interpretation of experimental data

Thus, by studying the initial rate of product appearance P_2 in the presence of an added nucleophile as well as water, and depending on the kinetic characteristics of the hydrolysis reaction of the substrate studied, it is possible to observe various types of inhibition. These range from competitive pseudo-inhibition to non-competitive pseudo-inhibition. By following the release of product P_1 under the same conditions, we observe a more or less complex activation of the reaction in favourable cases; this activation disappears when $s << K_{m,app}$.

Figure 5.29 below summarises the effect of (N) on the kinetic parameters $k_{cat,1}$, $k_{cat,2}$ and $K_{m,app}$ for the different cases I, II and III. The corresponding expressions for these parameters are given in Table 5.2 opposite.

Table 5.2 Effect of nucleophilic competition on the kinetic parameters $k_{cat,1}$, $k_{cat,2}$, and K_m depending on the rate-limiting step(s) for the hydrolysis reaction involving a covalent intermediate

(Reproduced from Eur. *J. Biochem.*, **3**, SEYDOUX F. & YON J., Compétition nucléophile dans les réactions d'hydrolyse enzymatique: Analyse cinétique et application à l'hydrolyse trypsiqiue de quelques esters, 42. © (1967) with permission of Blackwell Publishing Ltd)

Limiting step(s)	*Measured rate*	*Relation between k_{cat} and (N)*	*Relation between K_m and (N)*	*Relations between $\frac{k_{cat}}{K_m}$ and (N)*	*Case*
Deacylation $k_3 << k_2$	$\frac{dP_1}{dt}$	$k_{cat,1} = k'_3 + k_4(N)$	$K_m = K_s \frac{k'_3 + k_4(N)}{k_2}$	$\frac{k_{cat,1}}{K_m}$ independent of (N)	IP_1
	$\frac{dP_2}{dt}$	$k_{cat,2}$ independent of (N)	$K_m = K_s \frac{k'_3 + k_4(N)}{k_2}$	$\frac{K_m}{k_{cat,2}} = K_s \frac{k'_3 + k_4(N)}{k_2 k'_3}$	IP_1
Acylation $k_2 << k_3$	$\frac{dP_1}{dt}$	$k_{cat,1}$ independent of (N)	$K_m = K_s$ independent of (N)	$\frac{k_{cat,1}}{K_m}$ independent of [N]	IIP_1
	$\frac{dP_2}{dt}$	$\frac{1}{k_{cat,2}} = \frac{1}{k_2} + \frac{k_4(N)}{k_2 k_3}$	$K_m = K_s$ independent of (N)	$\frac{K_m}{k_{cat,2}} = K_s \frac{k'_3 + k_4(N)}{k_2 k'_3}$	IIP_2
Acylation and deacylation $k_2 \sim k_3$	$\frac{dP_1}{dt}$	$k_{cat,1} = \frac{k_2[k'_3 + k_4(N)]}{k_2 + k'_3 + k_4(N)}$	$K_m = K_s \frac{k'_3 + k_4(N)}{k_2 + k'_3 + k_4(N)}$	$\frac{k_{cat,1}}{K_m}$ independent of [N]	$IIIP_1$
	$\frac{dP_2}{dt}$	$\frac{1}{k_{cat,2}} = \frac{1}{k_2} + \frac{1}{k'_3} + \frac{k_4(N)}{k_2 k'_3}$	$K_m = K_s \frac{k'_3 + k_4(N)}{k_2 + k'_3 + k_4(N)}$	$\frac{K_m}{k_{cat,2}} = K_s \frac{k'_3 + k_4(N)}{k_2 k'_3}$	$IIIP_2$

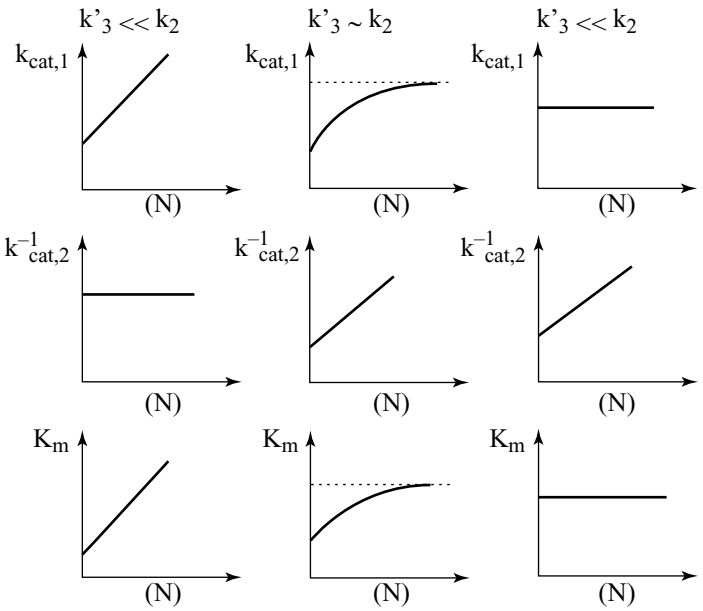
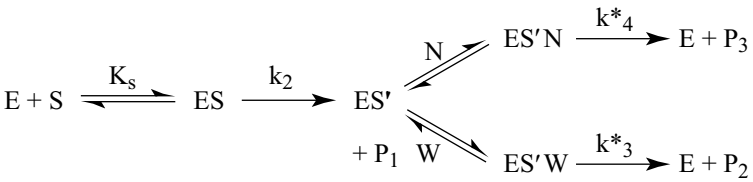


Fig. 5.29 Change in kinetic parameters as a function of the nucleophile concentration and depending on the rate-limiting step of the reaction

5.4.5.2 DETERMINATION OF THE KINETIC PARAMETERS IN THE CASE WHERE A BINDING SITE EXISTS FOR WATER AND ITS ANALOGUES

In Scheme 5.1 (pp. 153), it is assumed that the deacylation reactions are bimolecular, i.e. of the form $k_3(ES')(W)$ or $k_4(ES')(N)$. It may, however, be necessary to take into account non-covalent interactions between the enzyme and the water molecule or its analogues preceding the actual reaction event. This would imply the existence of a binding site for water and other nucleophiles in the active site of the enzyme.

It is reasonable to think that these sites are not independent. In this instance Scheme 5.1 becomes:



Scheme 5.2

In this scheme ES′W and ES′N are the complexes formed, respectively, between the acyl-enzyme and water, W, and the acyl-enzyme and the nucleophile N. k^*_3 and k^*_4 are the monomolecular deacylation constants relating to the degradation of the complexes ES′W and ES′N, respectively.

K_W and K_N are the dissociation constants for the complexes ES′W and ES′N, respectively. The kinetic analysis of this scheme at steady state leads to the following equations for $k_{cat,1}$, $k_{cat,2}$, $k_{cat,3}$ and $K_{m,app}$.

$$k_{cat,1} = \frac{k_2\left[k^*_3 + k^*_4 \frac{K_W(N)}{K_N(W)}\right]}{k_2\left[1 + \frac{K_W}{(W)} + \frac{K_W(N)}{K_N(W)}\right] + k^*_3 + k^*_4 \frac{K_W(N)}{K_N(W)}}$$

$$k_{cat,2} = \frac{k_2 k^*_3}{k_2\left[1 + \frac{K_W}{(W)} + \frac{K_W(N)}{K_N(W)}\right] + k^*_3 + k^*_4 \frac{K_W(N)}{K_N(W)}}$$

$$k_{cat,3} = \frac{k_2 k^*_4 \frac{K_W(N)}{K_N(W)}}{k_2\left[1 + \frac{K_W}{(W)} + \frac{K_W(N)}{K_N(W)}\right] + k^*_3 + k^*_4 \frac{K_W(N)}{K_N(W)}}$$

$K_{m,app}$ is the same in all three cases:

$$K_{m,app} = \frac{k^*_3 + k^*_4 \frac{K_W(N)}{K_N(W)}}{k_2\left[1 + \frac{K_W}{(W)} + \frac{K_W(N)}{K_N(W)}\right] + k^*_3 + k^*_4 \frac{K_W(N)}{K_N(W)}}$$

The expressions dP_1/dt, dP_2/dt and dP_3/dt can be explained as for the previous ones. In general, the forms of the functions $k_{cat,i} = f_i[(W),(N)]$ and $K_{m,app} = g[(W),(N)]$ depend on the order of magnitude of the quantities $K_W/(W)$ and $K_W(N)/K_N(W)$ compared to unity. One particularly interesting case arises where these quantities are negligible relative to 1, when $K_W << (N)$ and $(N) \sim K_N$, the molar H_2O concentration of an aqueous solution being about 55.5 M and the concentration of N rarely exceeding 5 M. If the structure of the nucleophile N resembles that of water, we have additionally, $K_N \sim K_W$. In this case, the equations are written:

$$k_{cat,1} = \frac{k_2\left[k^*_3 + k^*_4 \frac{K_W(N)}{K_N(W)}\right]}{k_2 + k^*_3 + k^*_4 \frac{K_W(N)}{K_N(W)}}$$

$$k_{cat,2} = \frac{k_2 k^*_3}{k_2 + k^*_3 + k^*_4 \frac{K_W(N)}{K_N(W)}}$$

$$k_{cat,3} = \frac{k_3 k^*_4 \frac{K_W(N)}{K_N(W)}}{k_2 + k^*_3 + k^*_4 \frac{K_W(N)}{K_N(W)}}$$

$$K_{m,app} = K_s \frac{k^*_3 + k^*_4 \frac{K_W(N)}{K_N(W)}}{k_2 + k^*_3 + k^*_4 \frac{K_W(N)}{K_N(W)}}$$

If the water concentration (W) can be considered as a constant, these expressions are identical in form to the expressions obtained in the Scheme 5.1. It is possible to pass from one scheme to the other by using the relations:

$$k_4 = k^*_4 \frac{K_W}{K_N(W)} \quad \text{and} \quad k_3 = k^*_3$$

If (W) is likely to vary, for example, when the nucleophile concentration becomes very high, the variable (N) must be replaced by the ratio (N)/(W). Thus, even if a binding site exists for water and its analogues in the enzyme's active site, it may not be possible to observe an effect of analogue saturation in a wide concentration range. This can be explained by the fact that the water concentration in the reaction medium is very high and the water and analogue sites are not independent.

The significance of the kinetic constants experimentally attained differs however between the reaction Schemes 5.1 and 5.2. With an equal intrinsic reactivity (identical k^*_4 values), two nucleophiles would be able to display different apparent reactivities if their affinities for the enzyme's receptor site are not the same (different K_N values); in this case, Schema 5.1 is insufficient. Thus, for example, the apparent reactivities of a series of normal primary aliphatic alcohols towards acyl-trypsins and chymotrypsins can vary considerably as a function of the hydrocarbon chain of the alcohol, whereas their **"intrinsic" reactivities,** evaluated by means of non-enzymatic alcoholysis reactions, remain appreciably constant.

If we use a nucleophile that reveals no saturation effect in a wide concentration range, we can apply the relations from Table 5.2, which give the graphs seen in Fig. 5.29. The analysis of these parameters as a function of nucleophile concentration enables the determination of the rate-limiting step of the reaction and allows us to obtain the values of the elementary rate constants for each reaction when no step is rate-limiting.

This method has been used successfully during the study of tryptic hydrolysis of some substrates (SEYDOUX & YON, 1967). Along with the previously described method this has made possible the determination of the elementary steps corresponding to the tryptic hydrolysis of various substrates as indicated in Table 5.3.

Table 5.3 Kinetic parameters corresponding to the tryptic hydrolysis of different substrates

Substrate	K_m *M*	k_{cat} s^{-1}	k_{cat}/K_m $s^{-1}.M^{-1}$	K_s *M*	k_2 s^{-1}	k_3 s^{-1}
(a) Bz-L-ArgNH$_2$	2.7×10^{-3}	0.54	2.0×10^{2}	2.7×10^{-3}	0.54	24
(a) To-L-ArgNH$_2$	7.5×10^{-3}	0.75	1.0×10^{2}	7.5×10^{-3}	0.75	95
(b) Bz-L-Arg OEt	2.6×10^{-6}	24	9.2×10^{6}	2.2×10^{-3}	2.0×10^{4}	24
(b) To-L-Arg OMe	1.5×10^{-5}	95	6.4×10^{6}	7.5×10^{-3}	4.8×10^{4}	95
(c) L-Lys OMe	2.8×10^{-4}	6.7	2.6×10^{4}	1.4×10^{-2}	3.4×10^{2}	6.7
(c) L-Lys OEt	4.9×10^{-4}	6.7	1.3×10^{4}	1.4×10^{-2}	1.9×10^{2}	6.7
(b) Ac-L-Tyr OEt	4.7×10^{-2}	36	7.6×10^{2}	47×10^{-3}	36	193
(b) Ac-L-Phe OMe	1.1×10^{-1}	55	5.0×10^{2}	1.1×10^{-1}	55	173
(b) To-L-Orn OMe	1.6×10^{-2}	5.4	3.4×10^{2}	1.6×10^{-2}	5.4	>5.4

Substrate	K_m *M*	k_{cat} s^{-1}	k_{cat}/K_m $s^{-1}.M^{-1}$	K_s *M*	k_2 s^{-1}	k_3 s^{-1}
(b) Bz-D-Arg OEt	7.4×10^{-4}	9.1×10^{-2}	1.2×10^{2}	2.2×10^{-3}	0.28	0.14
(c) L-Tyr-OMe	1.8×10^{-1}	0.7	3.8	1.8×10^{-1}	0.68	>0.68
(b) Ac-Gly OEt	8.8×10^{-1}	1.9×10^{-2}	3.6×10^{-2}	8.8×10^{-1}	3.2×10^{-2}	0.6

(a) parameters determined at the optimal pH and at 35°C and extrapolated to 25°C (from CHEVALLIER & YON, 1966); (b) kinetic parameters determined at the optimal pH and at 25°C (from SEYDOUX & YON, 1967); (c) (from SEYDOUX, 1967).

A comparable analysis undertaken by VIRATELLE and YON (1973) for reactions catalysed by the *E. coli* β-galactosidase demonstrated for the first time that these reactions also proceed *via* the formation of an intermediate chemical component which may degrade by reacting with water or another nucleophilic component. Figure 5.30 below indicates the change in the parameters $k_{cat,i}$ and $K_{m,app}$ for the hydrolysis of two substrates, o-phenyl galactoside and m-nitrophenyl galactoside as a function of the nucleophile methanol. This profile is characteristic of a reaction in which k_2 and k_3 are of the same order of magnitude. By an analogous study, it was possible to determine the kinetic parameters, including the rate constants for the elementary steps, for diverse substrates that were more or less specific to the enzyme. Table 5.4 lists these values.

Table 5.4 Kinetic parameters corresponding to the hydrolysis of different substrates of E. coli β-galactosidase at pH 7.0, 25°C, 10^{-3} M $MgSO_4$, 0.146 M NaCl *(from YON, 1976)*

Substrate	k_{cat} s^{-1}	K_m *M*	k_{cat}/K_m $s^{-1}.M^{-1}$	K_s *M*	k_2 s^{-1}	k_3 s^{-1}
β-D-galactosides						
Phenyl	45	10^{-4}	4.5×10^{5}	10^{-4}	45	
o-NO_2-phenyl	750	1.1×10^{-4}	6.8×10^{6}	3.1×10^{-4}	2 100	1 200
mNO_2-phenyl	800	1.5×10^{-4}	5.3×10^{6}	3.6×10^{-4}	1 900	1 400
pNO_2-phenyl	90	0.3×10^{-4}	3.0×10^{6}	0.3×10^{-4}	90	
oNH_2-phenyl	56	4.5×10^{-4}	1.2×10^{5}	4.5×10^{-4}	56	1 200
pNH_2-phenyl	90	3.3×10^{-4}	2.7×10^{5}	3.3×10^{-4}	90	1 135
2,4-diNO_2-phenyl	1 200	1.6×10^{-4}	7.5×10^{6}		>5 000	
3,5-diNO_2-phenyl	1 135	6.3×10^{-4}	1.8×10^{6}		>5 000	
Cinnamyl	230	1.2×10^{-2}	1.9×10^{4}	1.5×10^{-2}	285	1 200
Methyl	61	8.0×10^{-3}	7.6×10^{3}	8.0×10^{-3}	61	
Naphtyl	280	1.6×10^{-3}	1.75×10^{4}	1.6×10^{-4}	280	
α-L-arabinosides						
o-NO_2-phenyl	44	4.3×10^{-3}	1.0×10^{4}	4.3×10^{-3}	44	
β-D-fucoside						
o-NO_2-phenyl	4	3.0×10^{-3}	1.3×10^{3}	3.0×10^{-3}	4	
β-D-glucoside						
oNO_2-phenyl	0.14	6.2×10^{-3}	2.2×10^{1}	6.2×10^{-3}	0.14	

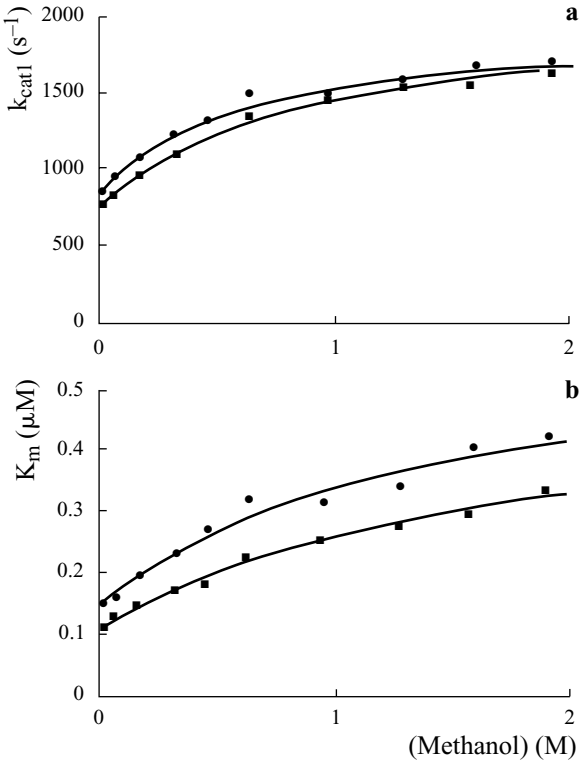


Fig. 5.30 Study of nucleophilic competition by methanol in the hydrolysis of m-nitrophenyl galactoside (●) and of o-nitrophenyl galactoside (■) by β-galactosidase from E. coli
(Reproduced from *Eur. J. Biochem.,* **33**, VIRATELLE O.M. & YON J.M., Nucleophilic competition in β-galactosidase catalyzed reactions, 110. © (1973) with permission of Blackwell Publishing Ltd)

5.4.6. KINETIC STUDY OF THE PRE-STEADY STATE: TITRATION OF ENZYME ACTIVE SITES

The analysis of the preceding kinetic scheme (Sect. 5.4.1) assumed a steady state, i.e. $d(ES_1)/dt = d(ES_2)/dt = 0$. In pre-steady state conditions, P_1 appears before P_2. When the steady state is reached, the rates of appearance of P_1 and P_2 are the same, as indicated in Fig. 5.31.

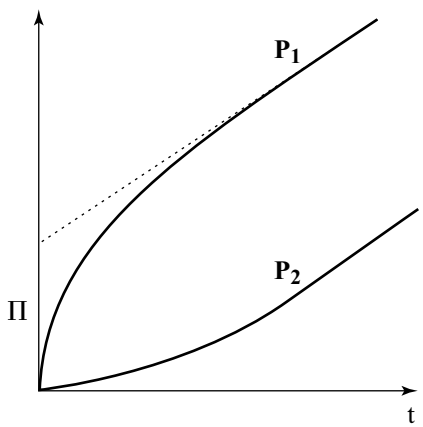


Fig. 5.31 Appearance of P_1 and P_2 over time in the pre-steady state and steady state phases

Let us recall the general scheme:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES_1 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} ES_2 + P_1 \xrightarrow{k_3} E + P_2$$

The kinetic treatment at the pre-steady state for the appearance of product P_1 in conditions where $s >> e_0$, i.e. $s = s_0$ with $k_2 < k_1$, leads to the following expression:

$$P_1 = \frac{k_{cat} e_0 s_0}{K_{m,app} + s_0} t + e_0 \left(\frac{\frac{k_{cat}}{k_3}}{1 + \frac{K_{m,app}}{s_0}} \right) (1 - e^{-kt})$$

with: $k_{cat} = \frac{k_2 k_3}{k_2 + k_3}$ and $K_{m,app} = K_s \frac{k_3}{k_2 + k_3}$

the constant of the exponential is:

$$k = k_3 + \frac{k_2}{1 + \frac{K_m}{s_0}} \qquad \text{with } K_m = \frac{k_{-1} + k_2}{k_1}$$

As t becomes large, the exponential term becomes negligible and the expression for P_1 can be written:

$$P_1 = \frac{k_{cat} e_0 s_0}{K_{m,app} + s_0} t + e_0 \left(\frac{\frac{k_{cat}}{k_3}}{1 + \frac{K_{m,app}}{s_0}} \right)^2 \quad \text{with } \Pi = e_0 \left(\frac{\frac{k_{cat}}{k_3}}{1 + \frac{K_{m,app}}{s_0}} \right)^2$$

Π represents the initial burst of the product extrapolated to time 0 of the reaction. When $s_0 >> K_{m,app}$ and $k_3 << k_2$, then: $\Pi = e_0$.

In these conditions, the initial burst of the product P_1 represents the quantity of enzyme that has effectively reacted with the substrate. This provides a means for stoichiometric titration of the active enzyme.

Such active-site titrants of trypsin exist: p-nitrophenyl acetate and p-nitrophenyl guanidinobenzoate. To titrate the active sites, the nitrophenol burst phase is followed by spectrophotometry.

When it is impossible to reach conditions of saturating substrate concentration, for example, when the substrate's solubility is low relative to $K_{m,app}$, it is still possible to evaluate the number of active sites by determining the initial burst of P_1 at several substrate concentrations. Indeed, the relation:

$$\Pi = e_0 \left(\frac{\frac{k_{cat}}{k_3}}{1 + \frac{K_{m,app}}{s_0}} \right)^2$$

simplifies, when k_3 is limiting, to:

$$\Pi = \left(\frac{e_0}{1 + \frac{K_{m,app}}{s_0}} \right)^2$$

and can be written: $$\frac{1}{\sqrt{\Pi}} = \frac{1}{\sqrt{e_0}} \left(1 + \frac{K_{m,app}}{s_0} \right)$$

The change in $1/\sqrt{\Pi}$ as a function of $1/s_0$ gives a linear plot, which, upon extrapolation to the vertical axis, yields the value $1/\sqrt{e_0}$. It is however desirable to obtain a sufficient number of points in order for the extrapolation to be precise. When k_3 is not really rate-limiting, analysis of the coefficient k from the exponential as a function of the substrate concentration allows the rate constants to be obtained for each individual step, by using the plot shown in Fig. 5.32.

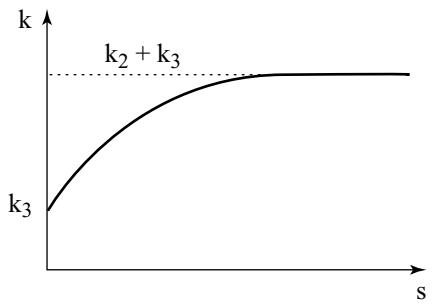


Fig. 5.32 Change in the exponential constant as a function of the substrate concentration

By way of example, the kinetics for the appearance of paranitrophenol during the titration of a β-galactosidase mutant are presented in Fig. 5.33a opposite. The initial burst is analysed as a function of the enzyme and substrate concentrations as shown, respectively, in Fig. 5.33b and c.

These few examples illustrate the diverse experimental approaches that enable the rate constants to be obtained for the elementary steps in a single-substrate reaction involving two intermediates. When the reaction products appear in a sequential order, the appearance of P_1 under pre-steady state conditions can incidentally lead to the stoichiometric titration of the active site. **This method is of the utmost importance as much in fundamental studies as in applied aspects of enzymology.**

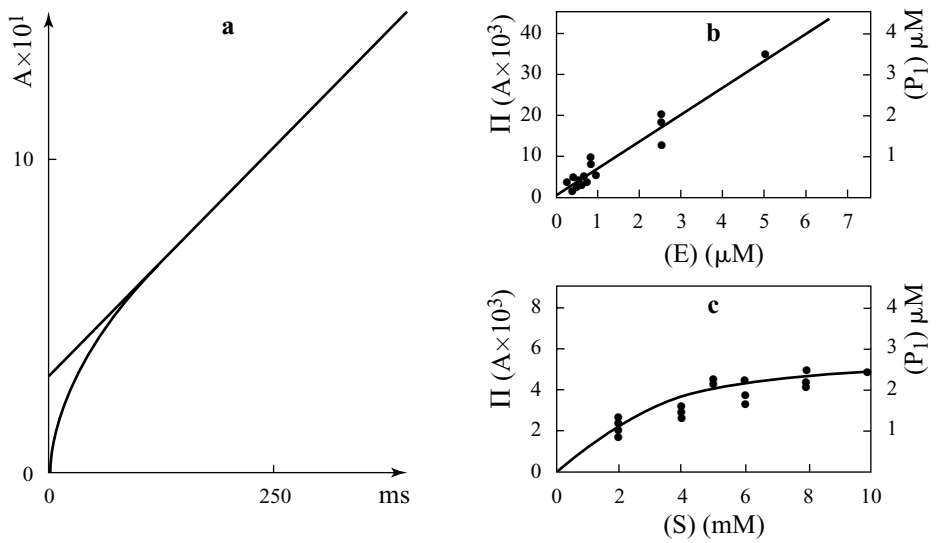


Fig. 5.33 Hydrolysis of p-nitrophenyl galactoside by a point mutant (13PO) of β-galactosidase from E. coli

*(**a**) kinetics of the appearance of p-nitrophenol – (**b**) analysis of the initial burst as a function of the enzyme concentration – (**c**) analysis of the initial burst as a function of the substrate concentration.* (Reprinted from *PNAS*, **75**, DESCHAVANNES P.J. *et al.*, 1892. © (1978) National Academy of Sciences, USA)

5.4.7. GENERALISATION FOR N INTERMEDIATES

The rate equation can be generalised for the case of n intermediate complexes. Let us take the following reaction scheme:

$$E + S \rightleftharpoons ES_1 \rightleftharpoons ES_2 \rightleftharpoons ES_3 \rightleftharpoons \ldots \rightleftharpoons ES_n \rightleftharpoons E + P$$

We can show by an iterative method that the kinetic parameters are given by the following relationships:

$$K_{m,app} = \frac{\Sigma \text{ of all allowed combinations of n constants not containing } k_1}{\Sigma \text{ of all allowed combinations of } (n-1) \text{ constants not containing } k_1, k_{-1} \text{ or } k_{-(n+1)}}$$

$$k_{cat} = \frac{\Pi \text{ of all allowed combinations of n constants, excluding } k_1}{\Sigma \text{ of all allowed combinations of } (n-1) \text{ constants not containing } k_1, k_{-1} \text{ or } k_{-(n+1)}}$$

Disallowed combinations are those that contain subscripts whose algebraic sum is equal to 0 or 1; i.e. constants having equal subscripts, opposite signs or consecutive subscripts (e.g. k_1k_{-1} or $k_{-1}k_2$). The number of intermediates, n, is never very large, but it can exceed 2. We see, therefore, the usefulness of the graphical method to solve the rate equation when the reaction scheme involves more than two intermediates.

5.5. ENZYMATIC REACTIONS WITH TWO SUBSTRATES

Enzymes nearly always catalyse reactions having several substrates, frequently two:

$$A + B \longleftrightarrow P + Q$$

Certain enzymes require the presence of a dissociable coenzyme. For kinetic analysis, the coenzyme can be formally considered as a second substrate. For example, alcohol dehydrogenase catalyses the reaction:

$$\text{ethanol} + NAD^+ \longleftrightarrow \text{acetaldehyde} + NADH + H^+$$

Creatine kinase uses creatine and ATP as substrates:

$$\text{creatine} + ATP \longleftrightarrow \text{creatine phosphate} + ADP$$

Commonly, the concentration of one of the substrates is in large excess and will not be significantly modified over the course of the reaction. In this case, when analysing the kinetics, only a single substrate need be taken into account. Enzymatic hydrolysis reactions use water as a second substrate. When they take place in aqueous solution, the second substrate does not contribute to the kinetics, as has been previously mentioned. In another solvent, the water concentration may be limiting and the kinetic analysis must take this into account.

Enzyme reactions that involve several substrates can follow very different reaction pathways. Their kinetics thus differ, and their analyses may reveal information about the reaction scheme.

5.5.1. NOMENCLATURE

Before commencing study of these reactions, we shall indicate some elements of the nomenclature employed. A **scheme** is first of all defined by the number of substrates and the number of products:

- a Uni Bi mechanism corresponds to one substrate and two products;
- a Bi Bi mechanism has two substrates and two products;
- a Ter Bi mechanism has three substrates and two products;
- a Ter Quad mechanism has three substrates and four products.

Sequential mechanisms are those in which the reaction substrates associate to the enzyme in a defined order; for **random mechanisms,** there is no defined order. A **ping-pong mechanism** refers to a reaction scheme in which the enzyme oscillates between two forms. Furthermore, **Iso** designates **a mechanism that involves enzyme isomerisation.** Here, the discussion is limited to mechanisms involving two substrates.

Before detailing these systems, it is appropriate to define the various kinetic parameters. There is one MICHAELIS constant with respect to substrate A and another with respect to substrate B; these are complex parameters. A dissociation constant

can also be defined for the first substrate, equal to (E)(A)/(EA). Different rules of nomenclature have been suggested to define these parameters and they are given in Table 5.5. Throughout the remainder of this book we shall adopt the nomenclature of CLELAND (1963).

Table 5.5 Nomenclature rules

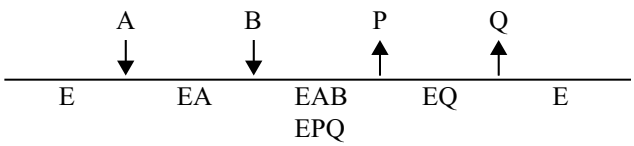
Author	*DALZIEL*	*ALBERTY*	*BLOOMFIELS*	*CLELAND*
MICHAELIS constant for A	Φ_1/Φ_0	K_A	K_{AB}/K_A	K_a
MICHAELIS constant for B	Φ_2/Φ_0	K_B	K_{AB}/K_B	K_b
Dissociation constant for A	Φ_{12}/Φ_2	K_{AB}/K_A	K_A	K_{ia}
V_{max}	E_0/Φ_0	V_f	V_{AB}	V_1
k_{cat}	$1/\Phi_0$	V_f/E_0	V_{AB}/E_0	V_1/E_0

In reaction schemes involving two substrates A and B and two products P and Q, it is useful to distinguish linear and branched schemes. For the sake of simplicity and because they are the most frequently encountered, we shall only look at a few examples of these types. The discussion is not exhaustive, but the general principles raised may be applied to solve other reaction schemes. According to the CLELAND plot, different states of the enzyme are represented by a horizontal line and the substrates and products by vertical arrows.

5.5.2. LINEAR SCHEMES

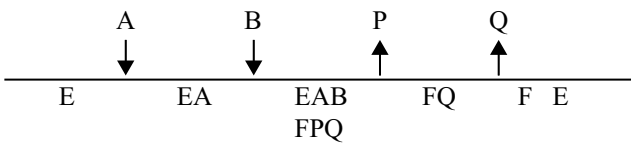
5.5.2.1. ORDERED BI BI MECHANISM

Substrates associate to the enzyme in a defined order; B can only bind to the EA complex and the reaction products are released in a defined sequence:



5.5.2.2. ISO-ORDERED BI BI MECHANISM

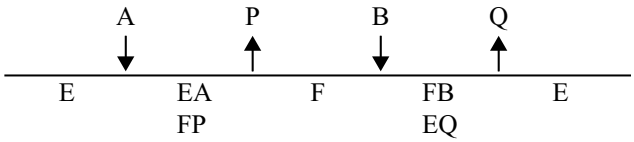
This is the same mechanism as before, but during the reaction the enzyme undergoes isomerisation. In the final step it once again adopts its initial, active conformation:



In these two types of mechanism the ternary complex is the reactive complex.

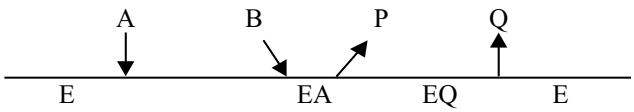
5.5.2.3. PING-PONG BI BI MECHANISM

In this mechanism the enzyme oscillates between two states, E and F:



5.5.2.4. THEORELL-CHANCE MECHANISM

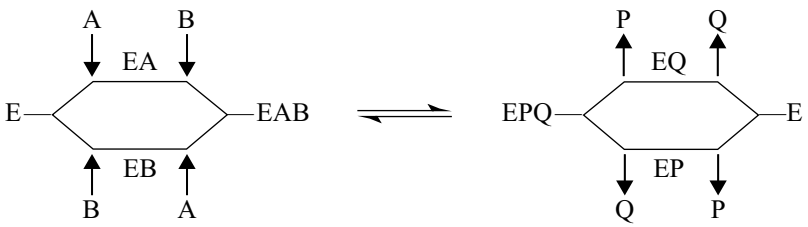
This mechanism was proposed by CHANCE and THEORELL to account for the kinetics of reactions catalysed by alcohol dehydrogenase. Substrate B, upon binding, ejects the first reaction product and converts it:



In the last two schemes the reaction complex is a binary complex between the enzyme and a single substrate.

5.5.3. BRANCHED SCHEMES: RANDOM BI BI MECHANISM

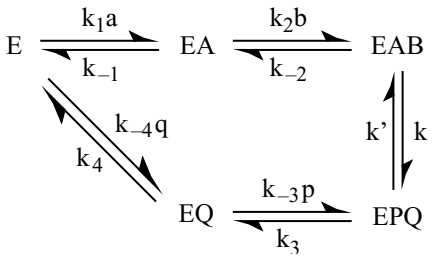
In this scheme, the two substrates can bind independently to the enzyme. The reactive complex is the ternary complex EAB:



5.5.4. KINETIC STUDY OF SOME TWO-SUBSTRATE REACTIONS

5.5.4.1. ORDERED BI BI MECHANISM

In order to solve this system under steady state conditions, we shall use the rules of KING and ALTMAN. The scheme can be written:



Considering the general case where the reverse reaction is not negligible, under steady state conditions we obtain the following rate equation:

$$\frac{v}{e_0} = \frac{N_{ab}ab - N_{pq}pq}{D_0 + D_a a + D_b b + D_{ab}ab + D_p p + D_q q + D_{qb}qb + D_{pq}pq + D_{abp}abp + D_{bpq}bpq}$$

The values of the coefficients are given in Table 5.6. The values of the rate parameters are:

$$V_1 = N_{ab}/D_{ab} \quad ; \quad V_2 = N_{pq}/D_{pq}$$

V_1 and V_2 are the maximal reaction rates for the forward and reverse reactions, respectively, using CLELAND's nomenclature.

Table 5.6 Parameter values for the ordered Bi Bi mechanism

N_{ab}	=	$k_1k_2kk_3k_4$
N_{pq}	=	$k_{-1}k_{-2}k'k_{-3}k_{-4}$
D_0	=	$k_{-1}kk_3k_4 + k_{-1}k_{-2}k_3k_4 + k_{-1}k_{-2}k'k_4$
D_a	=	$k_1kk_3k_4 + k_1k_{-2}k_3k_4 + k_1k_{-2}k'k_4$
D_b	=	$k_2kk_3k_4$
D_{ab}	=	$k_1k_2k_3k_4 + k_1k_2kk_4 + k_1k_2k'k_4 + k_1k_2kk_3$
D_p	=	$k_{-1}k_{-2}k'k_{-3}$
D_q	=	$k_{-1}kk_3k_{-4} + k_{-1}k_{-2}k_3k_{-4} + k_{-1}k_{-2}k'k_{-4}$
D_{pa}	=	$k_1k_{-2}k'k_{-3}$
D_{qb}	=	$k_2kk_3k_{-4}$
D_{pq}	=	$k_{-2}k'k_{-3}k_{-4} + k_{-1}kk_{-3}k_{-4} + k_{-1}k'k_{-3}k_{-4} + k_{-1}k_{-2}k_{-3}k_{-4}$
D_{abp}	=	$k_1k_2kk_{-3} + k_1k_2k'k_{-3}$
D_{bpq}	=	$k_2kk_{-3}k_{-4} + k_2k'k_{-3}k_{-4}$

The different MICHAELIS constants are given by the following relations:

- K_a, MICHAELIS constant with respect to A = D_b/D_{ab},
- K_b, MICHAELIS constant with respect to B = D_a/D_{ab},
- K_p, MICHAELIS constant with respect to P = D_q/D_{pq},
- K_q, MICHAELIS constant with respect to Q = D_p/D_{pq}.

The dissociation constant of the first substrate is:

$$K_{ia} = D_0/D_a = D_p/D_{pa} = K_{-1}/k_1$$

and that of the second product:

$$K_{iq} = D_0/D_q = D_b/D_{bq} = k_4/k_{-4}$$

The equilibrium constant is defined by:

$$K_{eq} = \frac{(P)_{eq}(Q)_{eq}}{(A)_{eq}(B)_{eq}} = \frac{N_{ab}}{N_{pq}}$$

From these diverse relationships, the HALDANE equations can be formulated:

$$K_{eq} = \frac{V_1 K_p K_{iq}}{V_2 K_{ia} K_{ib}} = \left(\frac{V_1}{V_2}\right)^2 \frac{K_{ip} K_q}{K_a K_{ib}}$$

with: $$K_{ib} = \frac{D_{pq}}{D_{bpq}} \quad \text{and} \quad K_{ip} = \frac{D_{ab}}{D_{abp}}$$

The reaction rate in the forward direction can be deduced from the rate equation by putting p = q = 0 in the initial conditions of the forward reaction. In its linear form:

$$\frac{1}{v} = \frac{1}{V_1}\left(1 + \frac{K_a}{a} + \frac{K_b}{b} + \frac{K_{ia}}{ab}\right)$$

Graphical plots

These include both primary and secondary plots.

Primary plots – The primary plots of LINEWEAVER-BURK (1/v versus 1/a or 1/b) yield concurrent straight lines; both substrates have symmetrical behaviour (Fig. 5.34). The intersection of the lines may be either above, below or even on the horizontal axis. The intersection of the linear curve 1/v as a function of 1/a has the horizontal-axis value of $-1/K_{ia}$; when the graph 1/v is plotted against 1/b, this value becomes $-K_a/K_{ia}K_b$.

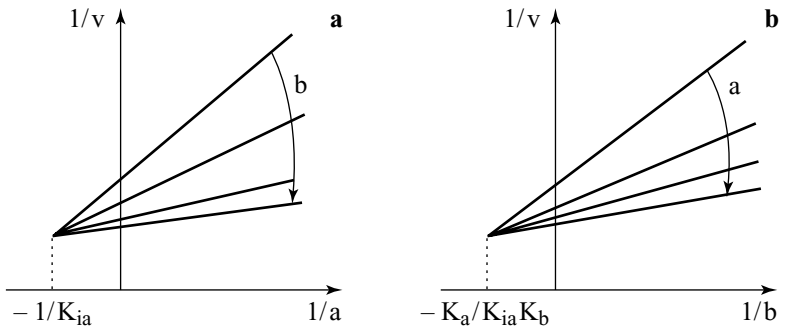


Fig. 5.34 Primary plots for an ordered Bi Bi mechanism at steady state

Secondary plots – From the primary plots, the slopes and ordinate (vertical-axis) intercepts as a function of the concentration of the other substrate provide secondary plots (Fig. 5.35 opposite), which enable the kinetic parameters of the reaction to be obtained. Thus, the slope of the curves in the plot 1/v versus 1/a is:

$$P_A = \frac{1}{V_1}\left(K_a + \frac{K_{ia} K_b}{b}\right)$$

and the ordinate intercepts: $$O_A = \frac{1}{V_1}\left(1 + \frac{K_b}{b}\right)$$

Similarly, we have:

$$P_B = \frac{1}{V_1}\left(1+\frac{K_{ia}}{a}\right)$$

$$O_B = \frac{1}{V_1}\left(1+\frac{K_a}{a}\right)$$

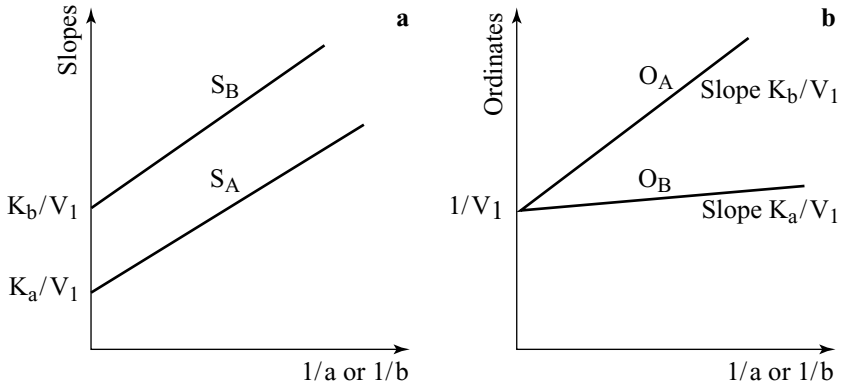


Fig. 5.35 Secondary plots corresponding to the mechanism in Fig. 5.34

Under conditions of quasi-equilibrium and no longer at steady state, the rate equation becomes:

$$\frac{1}{v} = \frac{1}{V_1}\left(1+\frac{K_{ib}}{b}+\frac{K_{ia}K_{ib}}{ab}\right)$$

The constants K_{ia} and K_{ib} are the true dissociation constants for the enzyme-substrate complexes. The primary plots differ slightly from those in the previous case (Fig. 5.36) and allow V_1 and K_{ia} to be obtained. The secondary plots give the value of K_{ib}.

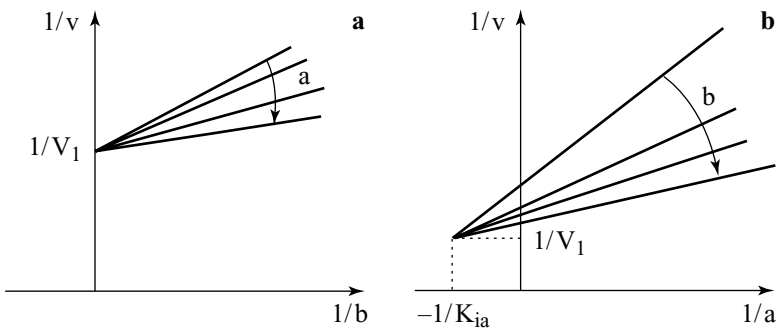
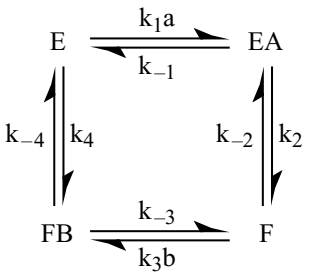


Fig. 5.36 Primary plots in an ordered Bi Bi mechanism when approximating a quasi-equilibrium

5.5.4.2. PING-PONG BI BI MECHANISM

Solving a system under steady state conditions is achieved using the rule of KING and ALTMAN. To this end, we write the scheme in the form shown below:



Under conditions that require taking the reverse reaction into account, the general rate equation is:

$$\frac{v}{e_0} = \frac{N_{ab}ab - N_{pq}pq}{D_a a + D_b b + D_{ab}ab + D_p p + D_q q + D_{pq}pq + D_{ap}ap + D_{bq}bq}$$

The values of the coefficients are provided in Table 5.7.

Table 5.7 Values for the parameters in a ping-pong Bi Bi mechanism

N_{ab}	=	$k_1k_2k_3k_4$
N_{pq}	=	$k_{-1}k_{-2}k_{-3}k_{-4}$
D_a	=	$k_1k_2k_4 + k_1k_2k_{-3}$
D_b	=	$k_2k_3k_4 + k_{-1}k_2k_3$
D_{ab}	=	$k_1k_3k_4 + k_1k_2k_3$
D_p	=	$k_{-1}k_{-2}k_{-3} + k_{-1}k_{-2}k_4$
D_q	=	$k_2k_{-3}k_{-4} + k_{-1}k_{-3}k_{-4}$
D_{pq}	=	$k_{-2}k_{-3}k_{-4} + k_{-1}k_{-2}k_{-4}$
D_{ap}	=	$k_1k_{-2}k_4 + k_1k_{-2}k_{-3}$
D_{bq}	=	$k_2k_3k_{-4} + k_{-1}k_3k_{-4}$

In the initial conditions of the forward reaction, the rate equation in LINEWEAVER-BURK form becomes:

$$\frac{1}{v} = \frac{1}{V_1}\left(1 + \frac{K_a}{a} + \frac{K_b}{b}\right)$$

It contains no term for 1/ab, which is due to the fact that the binary complexes are the active ones. In this expression, the MICHAELIS constants with respect to two substrates A and B have complex values, respectively equal to:

$$K_a = \frac{k_{-1} + k_2}{k_1} \quad \frac{k_4}{k_2 + k_4}$$

$$K_b = \frac{k_{-3} + k_{-4}}{k_3} \frac{k_2}{k_2 + k_4}$$

Graphical representations

As before, we shall consider the primary and secondary plots:

Primary plots – The primary plots of 1/v versus 1/a or 1/b both give a series of parallel lines typical for this type of mechanism (Fig. 5.37).

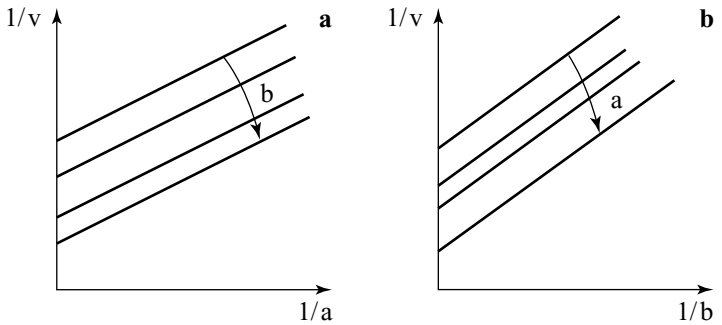


Fig. 5.37 Primary plots for a ping-pong Bi Bi mechanism

Secondary plots – The slopes of the lines in the primary plots are constant:

$$S_A = K_a/V_1 \quad ; \quad S_B = K_b/V_1$$

All of the parameters can be determined from the ordinate intercepts:

$$O_A = \frac{1}{V_1}\left(1 + \frac{K_b}{b}\right)$$

$$O_B = \frac{1}{V_1}\left(1 + \frac{K_a}{a}\right)$$

Figure 5.38 shows the secondary plots.

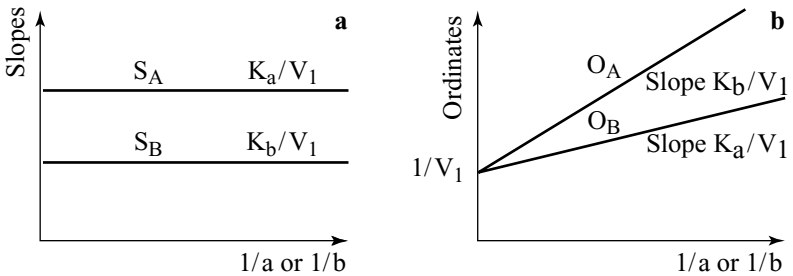
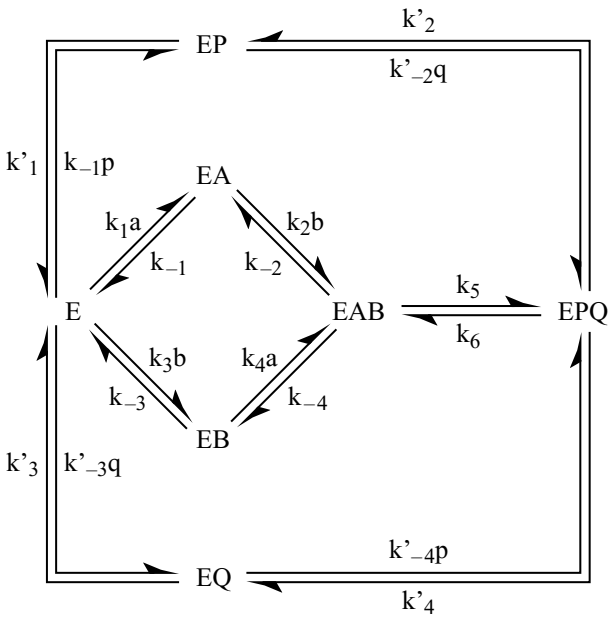


Fig. 5.38 Secondary plots for a ping-pong Bi Bi mechanism

5.5.4.3. A BRANCHED SCHEME: THE RANDOM BI BI MECHANISM

The scheme is represented as follows:



A scheme of this type represents the superposition of four ordered Bi Bi pathways. In other words, an ordered Bi Bi mechanism is a special case of this scheme in which the other pathways have a negligible probability of being followed.

Solving the system under steady state conditions

◤ By employing KING and ALTMAN's rules, we can solve the system under steady state conditions, i.e. when:

$$d(EA)/dt = d(EB)/dt = d(EAB)/dt = d(EP)/dt = d(EQ)/dt = 0$$

Even if we only consider the reaction in the direction of the formation of products P and Q, and assuming that k_6 is negligible, we end up – in the initial reaction conditions– with a non-linearisable rate equation having the general form:

$$\frac{v}{e_0} = \frac{N_1ab + N_2a^2b + N_3ab^2}{D_4ab + D_5b + D_6a + D_7a^2 + D_8b^2 + D_9a^2b + D_{10}ab^2 + D_{11}}$$

The values of the coefficients are listed in Table 5.8 opposite. This complex equation cannot be linearised. This system therefore leads to an equation that deviates from MICHAELIS' law.

Table 5.8 Parameter values for a random Bi Bi mechanism, approximating the steady state *(forward reaction where k_6 is negligible)*

N_1	=	$k_1k_{-3}k_2k_5 + k_{-1}k_3k_4k_5$
N_2	=	$k_1k_2k_4k_5$
N_3	=	$k_2k_3k_4k_5$
D_4	=	$k_1k_2k_{-3} + k_{-1}k_3k_4 + k_1k_2k_{-4} + k_{-2}k_3k_4$
D_5	=	$k_{-1}k_3k_{-4} + k_2k_3k_{-4} + k_{-1}k_{-2}k_3$
D_6	=	$k_{-3}k_{-2}k_1 + k_{-1}k_{-2}k_4 + k_1k_{-3}k_{-4}$
D_7	=	$k_4k_{-2}k_1$
D_8	=	$k_2k_3k_{-4}$
D_9	=	$k_1k_2k_4$
D_{10}	=	$k_2k_3k_4$
D_{11}	=	$k_{-1}k_{-3}k_2 + k_{-1}k_{-3}k_{-4}$

Solving a system under conditions of quasi-equilibrium

We can readily solve this system by assuming the existence of a quasi-equilibrium between the free enzyme and the different complexes. For this approximation we consider that the constants k_5 and k_6 are very small relative to all other rate constants in the system. This hypothesis is validated by the experimental results when the graphical plots (LINEWEAVER-BURK) are linear. Indeed, in this case, the rate equation takes on a Michaelian form. Two cases must be considered:

- either the binding of A and B to the enzyme is dependent, i.e. the binding of A modifies the affinity of the enzyme for B and vice versa;
- or it is independent; the binding of one substrate occurs in the same way in the presence or absence of the second.

Dependent binding

Dependent binding is the most general case. If the binding is dependent, we must define for the system in equilibrium the following four equilibrium constants:

- K_a for the dissociation of the complex EA: $K_a = (E)(A)/(EA)$,
- K_b for the dissociation of the complex EB: $K_b = (E)(B)/(EB)$,
- K'_a for the dissociation of A from the complex EAB: $K'_a = (EB)(A)/(EAB)$,
- K'_b for the dissociation of B from the complex EAB: $K'_b = (EA)(B)/(EAB)$.

Since the system is at equilibrium, the four constants are related by the thermodynamic relationship:

$$K_aK'_b = K'_aK_b$$

The system is therefore entirely defined by three of them.

The rate equation is thus linear and can be put in the form:

$$\frac{1}{v} = \frac{1}{V_1}\left(1 + \frac{K'_a}{a} + \frac{K'_b}{b} + \frac{K_aK'_b}{ab}\right)$$

which is strictly analogous to that describing the ordered Bi Bi mechanism assuming a steady state. ***These two schemes, which can be represented by the same phenomenological equation, are said to be "homeomorphic".***

Primary plots – The primary plots, as for sequential mechanisms, are represented by rays of lines converging to a point whose intercepts with the horizontal axis are $-1/K_b$ and $-1/K_a$; and with the vertical axis, $(K_b - K'_b)/V_1K_b$ and $(K_a - K'_a)/V_1K_a$, for the graphs $1/v = f(1/b)$ and $1/v = f(1/a)$, respectively. Now:

$$\frac{K_a - K'_a}{K_a} = \frac{K_b - K'_b}{K_b}$$

This shows clearly that when the dependence between the substrate-binding sites is positive (i.e. when the binding of the first substrate increases the enzyme's affinity for the second), the linear plots intersect at a point above the horizontal axis (Fig. 5.39a). On the contrary, if the interactions between the sites are negative, the lines intersect at a point below the horizontal axis (Fig. 5.39b).

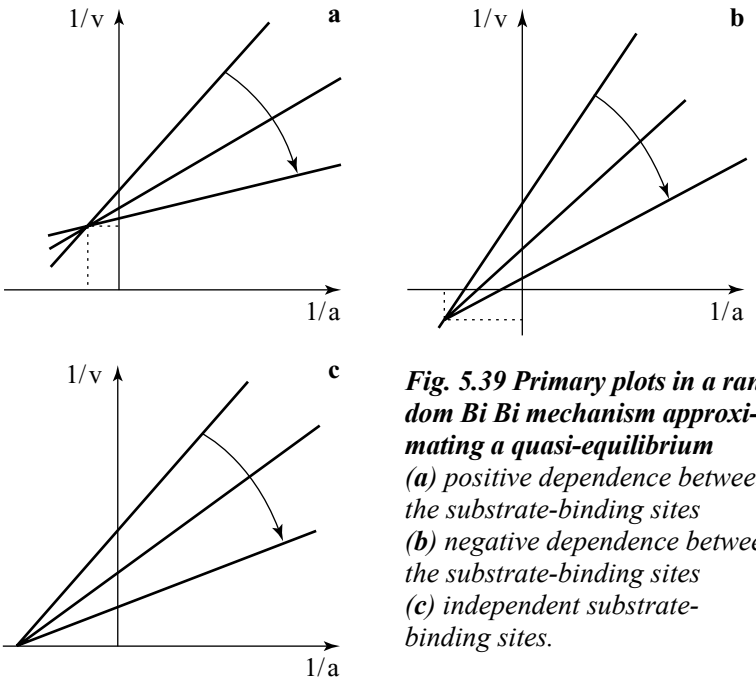


Fig. 5.39 Primary plots in a random Bi Bi mechanism approximating a quasi-equilibrium
*(**a**) positive dependence between the substrate-binding sites*
*(**b**) negative dependence between the substrate-binding sites*
*(**c**) independent substrate-binding sites.*

Independent binding

If the binding sites of the two substrates are sufficiently distinct and if there is no interaction between them, the binding of one substrate will not affect the binding of the other. In this instance, the reaction scheme contains no more than two equilibrium constants, $K_a = K'_a$ and $K_b = K'_b$. The rate equation is simplified to:

$$\frac{1}{v} = \frac{1}{V_1}\left(1 + \frac{K_a}{a} + \frac{K_b}{b} + \frac{K_a K_b}{ab}\right)$$

The linear plots intersect at points on the horizontal axis whose values are equal to $-1/K_a$ and $-1/K_b$, respectively, in the graphs $1/v = f(1/a)$ and $1/v = f(1/b)$, as indicated in Fig. 5.39c.

5.5.5. *HOMEOMORPHIC SCHEMES: HOW IS THE AMBIGUITY OF THE KINETIC RESPONSE REMOVED?*

In the few schemes just described, it clearly appears that a kinetic study, though necessary and indicative, is not in itself sufficient to determine unequivocally the type of kinetic scheme. We noted, for example, that the random Bi Bi mechanism treated under quasi-equilibrium conditions and the ordered Bi Bi mechanism treated at steady state lead to the same phenomenological expression for the reaction rate and constitute *homeomorphic schemes.*

It is useful therefore to choose alternative criteria in order to remove the ambiguity of the kinetic response. Among the methods used, one relies on analysing the inhibition kinetics of the reaction products or "dead-end" inhibitors. Another involves studying the binding equilibria of each substrate; lastly, fast kinetic methods may be used to study the reaction. The latter will be examined in the following chapter, which deals with experimental methods for studying kinetics.

5.5.5.1. STUDYING INHIBITION BY THE REACTION PRODUCTS: CLELAND'S RULES

The products that are formed during a reaction are inhibitors of their own formation by virtue of the microscopic reversibility principle. CLELAND determined several practical rules that follow directly from the analysis of the general expression for the reaction rate in the presence of two reaction products:

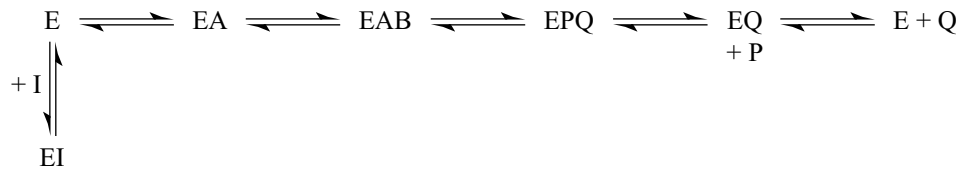
- if a reaction product and a substrate bind to the same form of the enzyme, the inhibition is competitive (change in the slope of the linear plot $v = f(v/s)$ as a function of the inhibitor concentration);
- if a reaction product and a substrate bind to two different forms of the enzyme separated by reversible steps, the inhibition is non-competitive (change in the value of the vertical-axis intercept).

Let us consider the preceding different schemes. We shall examine successively the inhibition by the products P and Q with respect to each substrate, A and B. In an ordered Bi Bi mechanism, inhibition by the product P with respect to A is non-competitive (P binds to EQ; A to E); it is also non-competitive with respect to B. Inhibition by the product Q is competitive with respect to A, yet non-competitive with respect to B.

In a ping-pong Bi Bi mechanism, inhibition by the product P is non-competitive with respect to A and competitive with respect to B.

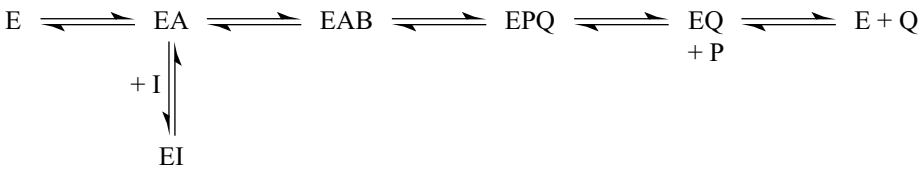
In a random Bi Bi mechanism, in a state of quasi-equilibrium, each reaction product behaves as a competitive inhibitor with respect to each substrate. This is valid if the diverse states of the enzyme are separated by reversible steps. If one of the steps separating two forms of the enzyme – one of which binds the substrate, the other the product – is irreversible, the situation is more complex.

In addition to the reaction products, other terminal or dead-end inhibitors exist that produce an abortive or dead-end complex with one of the enzyme forms. One part of the enzyme is thus trapped in an inactive form. If such an inhibitor binds to a form of the enzyme that has bound one of the substrates, the inhibition is uncompetitive with respect to this substrate. If it binds to the same enzyme form as the substrate, the inhibition is competitive. For example, in a sequential Bi Bi mech-anism, if there is a dead-end inhibitor I such that:



the inhibition by I with respect to A is competitive.

In the following case:



the inhibition by I with respect to A is uncompetitive.

Thus, based on this theoretical analysis, it is possible experimentally to obtain information about the kinetic mechanism of an enzymatic reaction by studying the influence of the reaction products and the dead-end inhibitors. In practice, the latter are generally substrate analogues. Table 5.9 indicates the types of inhibition by the reaction products for the reaction schemes that have been analysed above.

Table 5.9 Inhibition by the reaction products for the principal types of kinetic mechanism

Reaction scheme	*Inhibition by P with respect to A*	*Inhibition by P with respect to B*	*Inhibition by Q with respect to A*	*Inhibition by Q with respect to B*
Ordered Bi Bi	N-C	N-C	C	N-C
Ping-pong Bi Bi	N-C	C	C	N-C
Random Bi Bi	C	C	C	C

N-C: non-competitive; C: competitive

5.5.5.2. STUDY OF SUBSTRATES BINDING TO AN ENZYME

A complementary study to that discussed above involves analysing the equilibrium of each substrate binding separately to the enzyme. This requires the use of either equilibrium dialysis or dynamic dialysis developed by COLOWICK and KAPLAN (see Chap. 2).

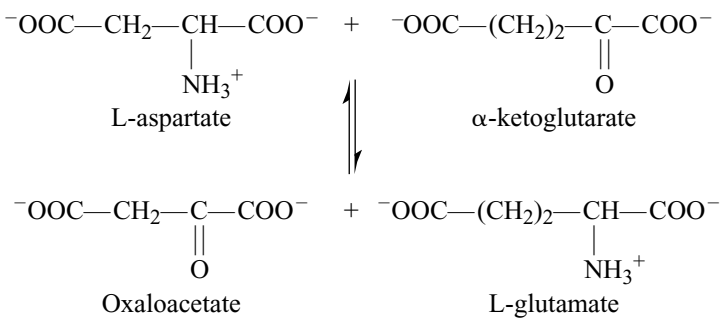
5.5.5.3. STUDY OF TRANSITORY STEPS BY RAPID KINETICS

It is clear that the precise determination of a reaction path requires the use of fast kinetic methods: flow and (or) relaxation. These methods are presented in the following chapter.

5.5.6. SOME EXAMPLES

5.5.6.1. L-ASPARTATE-2-OXOGLUTARATE AMINO TRANSFERASE

◤ This enzyme, still sometimes called glutamate oxaloacetate transaminase, catalyses the following reaction:



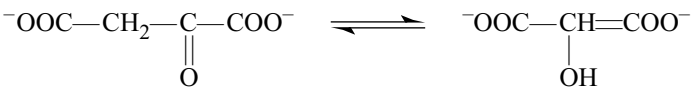
The enzyme catalyses the transfer of the amino group from L-aspartate to α-ketoglutarate. As with all transaminases, this enzyme uses pyridoxal phosphate as a coenzyme, which is transformed into pyridoxamine phosphate during the reaction. The majority of these studies initially used the enzyme from pig heart; it was in fact the first transaminase to be purified. Its sequence and three-dimensional structure are known, as well as those of transaminases from other organisms. The kinetic studies were carried out around the 1960s by the teams of BANKS and VERNON in the UK, VELICK and VAVRA in the USA, BRAUSTEIN in the former USSR, FASELLA in Italy (Rome), and YON in France (Paris). The total reaction equilibrium was studied and the value of the equilibrium constant:

$$K_{eq} = (Asp)(\alpha\text{–}KG)/(Glu)(OAA)$$

is of the order of 3–5. (α–KG) represents the concentration of α-ketoglutarate and (OAA) that of oxaloacetate.

Different reaction mechanisms could be invoked *a priori* to describe the reaction; they imply the formation of either a ternary complex, or a binary complex, with the associations being sequential or non-sequential. The kinetic studies have permitted the derivation of a reaction scheme for both the forward and reverse reactions.

For the forward reaction, the experiments were performed in the initial reaction conditions by following the appearance of oxaloacetate by spectrophotometry. This compound is in equilibrium with its enolic form which absorbs strongly:



In the experimental conditions used, the last reaction was not rate-limiting. Fig. 5.40 shows the change in 1/v as a function of the inverse of the α-ketoglutarate concentration for different aspartate concentrations.

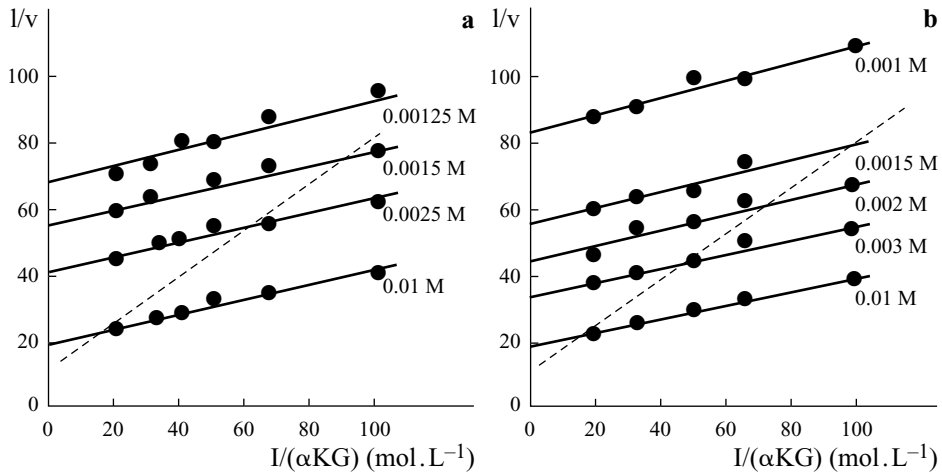
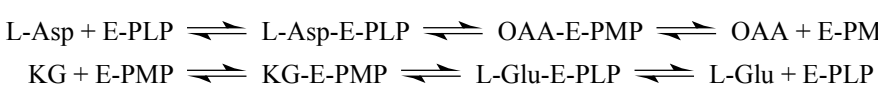


Fig. 5.40 Kinetic study of the forward reaction catalysed by aspartate amino transferase

(a) *0.1 M imidazole/HCl buffer, at pH 8.0 and 25°C –* ***(b)*** *at pH 7.1 and 25°C, for different aspartate concentrations.* (Reprinted from *Chemical and Biological Aspects of Pyridoxal Catalysis*, BANKS B.E.C. *et al.*, Kinetic studies of glutamic-aspartic transaminase, 205, 1963, with authors' permission)

Figure 5.41 opposite depicts the change in 1/v as a function of the inverse of the glutamate concentration for different concentrations of oxaloacetate. The parallel nature of the straight lines obtained for the two series of experiments indicates that the reaction takes place according to a ping-pong mechanism; the reaction is described by the corresponding equation and can be split into two half-reactions:



PLP and PMP are the pyridoxal and pyridoxamine forms of the coenzyme, respectively.

This reaction scheme has been confirmed by inhibition studies with glutarate which binds to the pyridoxal form of the enzyme and not to the pyridoxamine form.

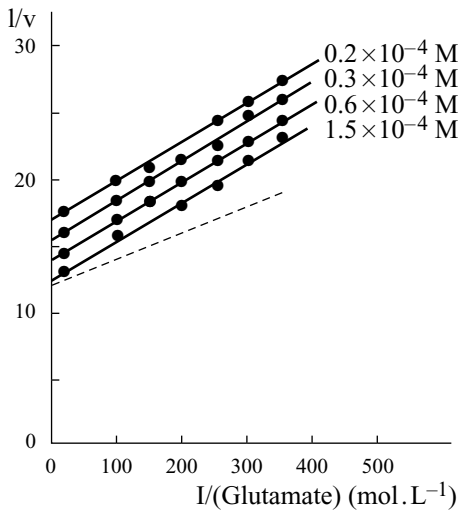


Fig. 5.41 Kinetic study of the reverse reaction catalysed by aspartate amino transferase at pH 7.8 and 25°C, for different oxaloacetate concentrations
(Reprinted from *Chemical and Biological Aspects of Pyridoxal Catalysis*, BANKS B.E.C. *et al.*, Kinetic studies of glutamic-aspartic transaminase, 206, 1963, with authors' permission)

The results are presented in Fig. 5.42a and b. In reality this scheme is further complicated by the existence of other reaction intermediates. The kinetic constants for the system were determined:

$$K_{asp} = 9\times10^{-4}\ M \quad ; \quad K_{KG} = 10^{-4}\ M \quad ; \quad K_{glu} = 4\times10^{-3}\ M \quad ; \quad K_{OAA} = 4\times10^{-5}\ M$$

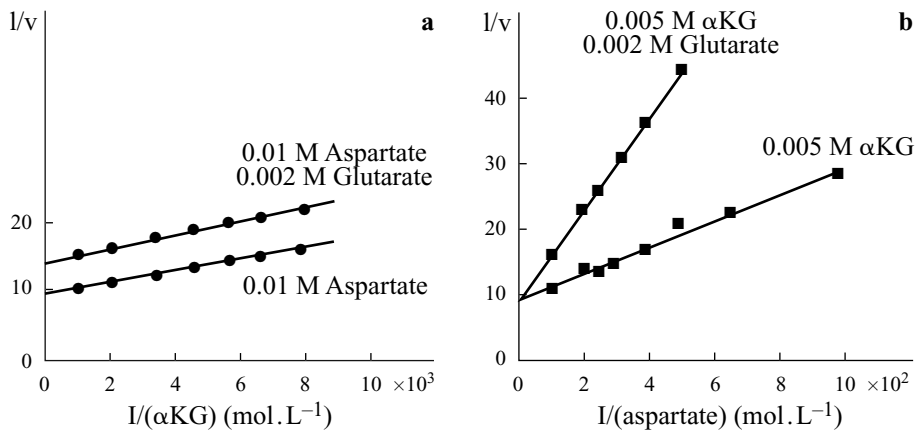


Fig. 5.42 (a) glutarate inhibition of the forward reaction catalysed by aspartate amino transferase, with respect to α-ketoglutarate (b) inhibition of the reaction by α-ketoglutarate with respect to aspartate
(Reprinted from *Chemical and Biological Aspects of Pyridoxal Catalysis*, BANKS B.E.C. *et al.*, Kinetic studies of glutamic-aspartic transaminase, 207, 1963, with authors' permission)

5.5.6.2. YEAST HEXOKINASE

Yeast hexokinase catalyses the following reaction:

$$\text{glucose} + \text{MgATP} \rightleftharpoons \text{glucose-6-phosphate} + \text{MgADP}$$

It has been particularly studied in J. RICARD's laboratory in Marseille (France). The authors showed that the reaction kinetics can be interpreted according to an ordered Bi Bi mechanism.

The primary plots for the forward reaction are given in Fig. 5.43. The secondary plots are given in Fig. 5.44a and b.

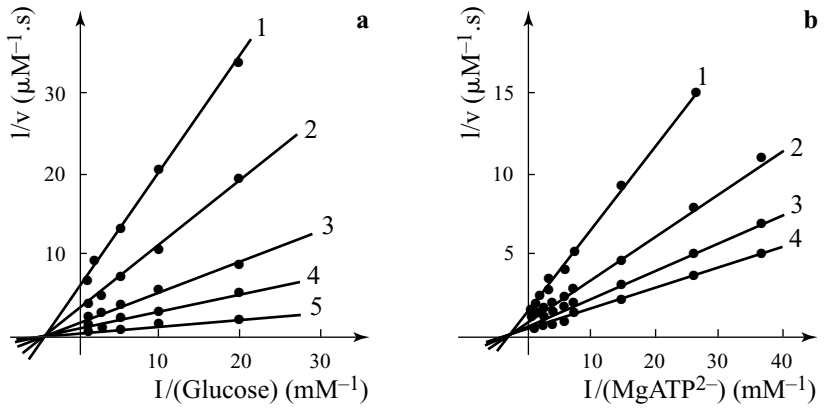


Fig. 5.43 Primary plots of the phosphorylation reaction of glucose by MgATP catalysed by yeast hexokinase

(Reproduced from *Eur. J. Biochem.,* 5, NOAT T.G. *et al.*, Kinetic study of yeast hexokinase 1 steady state kinetics, 59. © (1968) with permission of Blackwell Publishing Ltd)

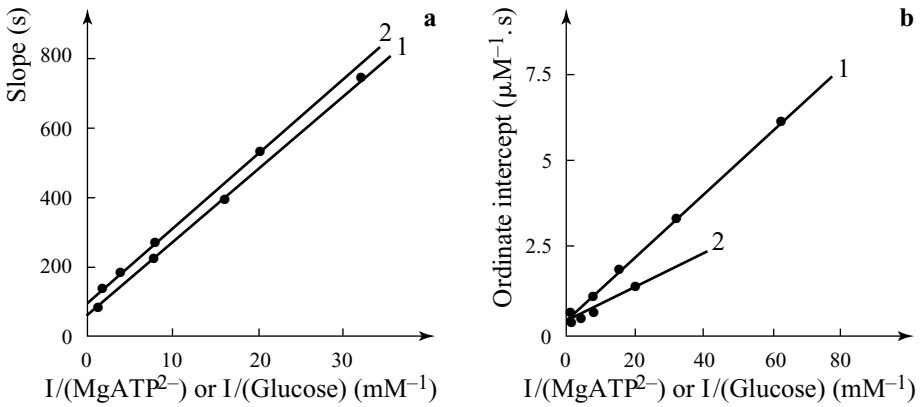


Fig. 5.44 Secondary plots obtained from the data in Fig. 5.43

(Reproduced from *Eur. J. Biochem.,* 5, NOAT T.G. *et al.*, Kinetic study of yeast hexokinase 1 steady state kinetics, 60. © (1968) with permission from Blackwell Publishing Ltd)

These results are compatible with several reaction schemes: either ordered Bi Bi or random Bi Bi when approximating a quasi-equilibrium. On the one hand, the study of the inhibition by MgADP indicates non-competitive inhibition with respect to two substrates (Fig. 5.45 opposite). On the other hand, inhibition by glucose-6-phosphate is competitive with respect to glucose and non-competitive with respect to MgATP

(Fig. 5.46a and b). The results of these studies of product inhibition indicate an ordered Bi Bi mechanism.

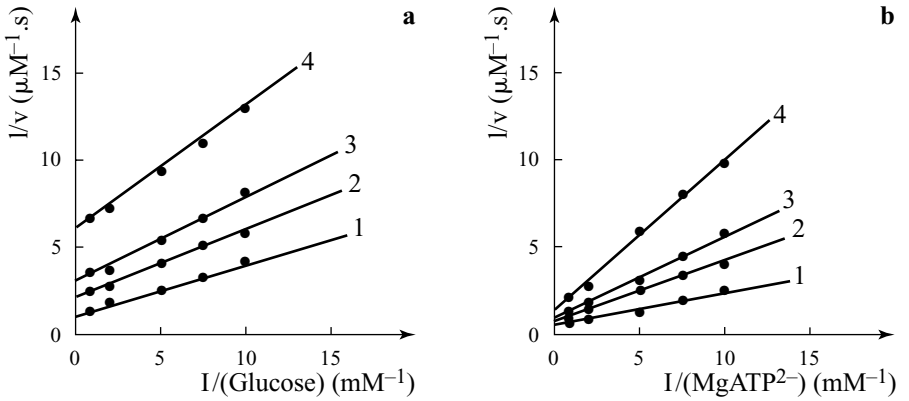


Fig. 5.45 Non-competitive inhibition by the reaction product during glucose phosphorylation catalysed by hexokinase

(Reproduced from *Eur. J. Biochem.*, 5, NOAT T.G. *et al.*, Kinetic study of yeast hexokinase 1 steady state kinetics, 66. © (1968) with permission from Blackwell Publishing Ltd)

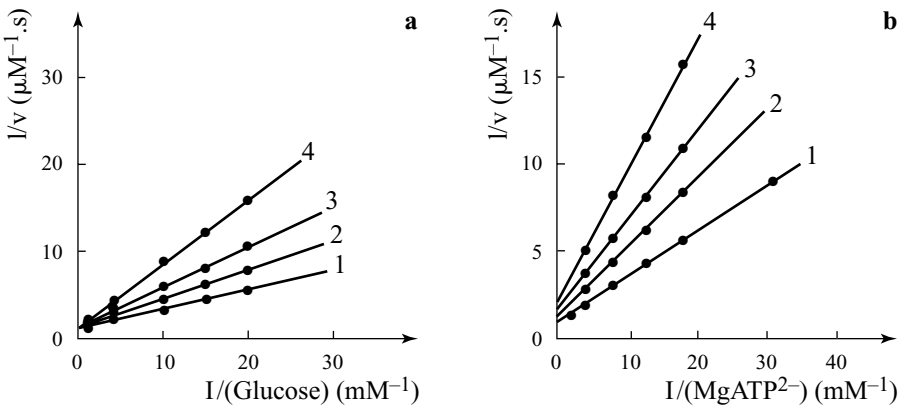


Fig. 5.46 Inhibition by glucose-6-phosphate of the hexokinase-catalysed reaction

(Reproduced from *Eur. J. Biochem.,* 5, NOAT T.G. *et al.*, Kinetic study of yeast hexokinase 1 steady state kinetics, 66. © (1968) with permission from Blackwell Publishing Ltd)

5.6. STATISTICAL ANALYSIS OF EXPERIMENTAL DATA

The aim of experimental data processing, in particular with kinetic data in enzymology, is to estimate the degree of reliability of the constants determined from the experimental results by using statistical techniques. These are generally based on least-squares methods. This facilitates the evaluation of the quantitative values in the interpretation and verification of the hypotheses. This is the case for the analysis

of the MICHAELIS equation by the method of non-linear regression which was employed from 1961 (JOHANSEN & LUMRY, 1961; WILKINSON, 1961; CLELAND, 1963). Such an analysis is largely facilitated today by the use of computers and the existence of appropriate software. We shall recall a few definitions, before succinctly outlining the principles of linear, multilinear and non-linear regression. For the reader who wishes to deepen his or her understanding of these aspects, we recommend the book by M. DESMADRIL: *L'Analyse des données en C* (*Data analysis in C*).

5.6.1. A FEW DEFINITIONS

◤ The arithmetic average of a variable is defined by the relation:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad [10]$$

The variance of the variable x is:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2 \qquad [11]$$

the standard deviation being defined by σ. The relation [11] can be written:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i^2 - 2x_i\bar{x} + x^2)$$

$$\sigma^2 = \frac{1}{n}\left(\sum_{i=1}^{n} x_i^2 - 2x\sum_{i=1}^{n} x_i + \bar{x}^2\right)$$

i.e.:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

$\overline{x^2}$ is the average of the x^2 values and $\bar{x}^2$ is the square of the average of the x values.

The covariance of the two variables x and y is given by the expression:

$$\sigma_{x,y} = \mathrm{cov}(x,y) = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \qquad [12]$$

which can also be written:

$$\sigma_{x,y} = \overline{xy} - \bar{x}.\bar{y}$$

The covariance is nil if the variables x and y are independent. It differs from 0 when they are linked.

5.6.2. SIMPLE LINEAR REGRESSION

Simple linear regression is applicable to all linear relationships of the form $y = ax + b$. For a given value of x, the observable is equal to the quantity $Y = \beta_0 + \beta_1 x$, plus a quantity ε, which represents the error, i.e. the deviation of the experimental value from the regression line:

$$y_i = \beta_0 + \beta_1 x + \varepsilon \qquad [13]$$

The method of linear regression consists of estimating the values of β_0 and β_1 that best account for the experimental results, i.e. the values (x_i, y_i). The value S, which represents the spread of the experimental points from the line $\beta_0 + \beta_1 x$, is defined by the relationship:

$$S = \sum_{i=1}^{n} w_i (y_i - \beta_0 - \beta_1 x_i)^2 = \sum_{i=1}^{n} w_i (y_i - Y)^2 \qquad [14]$$

where w_i is the statistical weight associated with each value of y_i, an experimental value of the variable Y. It is distributed with a variance σ^2, independent of the x_i values.

$$\sigma^2(Y) = \frac{\sum_{i=1}^{n} w_i (y_i - Y)^2}{n - p} = \frac{\sum_{i=1}^{n} w_i (y_i - Y)^2}{n - 2} \qquad [15]$$

p is the number of parameters to be determined, which here is equal to 2.

The estimates b_0 and b_1 of the parameters β_0 and β_1 which give the best Y values associated with (x_i, y_i) are those which lead to the smallest value of S. If we consider relationship [14] as a function of both β_0 and β_1, it is a minimum at the point $S_0(b_0,b_1)$ such that:

$$\begin{aligned} \frac{\partial S}{\partial \beta_0} &= -2\sum_{i=0}^{n} w_i (y_i - b_0 - b_1 x_i) = 0 \\ \frac{\partial S}{\partial \beta_1} &= -2\sum_{i=0}^{n} w_i x_i (y_i - b_0 - b_1 x_i) = 0 \end{aligned} \qquad [16]$$

This is again solved as a system of two linear equations with two unknowns:

$$\begin{aligned} \sum_{i=0}^{n} w_i y_i &= b_0 \sum_{i=0}^{n} w_i + b_1 \sum_{i=0}^{n} w_i x_i \\ \sum_{i=0}^{n} w_i y_i x_i &= b_0 \sum_{i=0}^{n} w_i x_i + b_1 \sum_{i=0}^{n} w_i x_i^2 \end{aligned} \qquad [17]$$

It can be solved by calculating matrices and applying the rules of matrix multiplication:

$$\begin{vmatrix} \sum_{i=0}^{n} w_i y_i \\ \sum_{i=0}^{n} w_i y_i x_i \end{vmatrix} = \begin{vmatrix} \sum_{i=0}^{n} w_i & \sum_{i=0}^{n} w_i x_i \\ \sum_{i=0}^{n} w_i x_i & \sum_{i=0}^{n} w_i x_i^2 \end{vmatrix} \times \begin{vmatrix} \beta_0 \\ \beta_1 \end{vmatrix}$$

By putting:

$$Y = \begin{vmatrix} \sqrt{w_1} y_1 \\ \sqrt{w_2} y_2 \\ \cdots \\ \sqrt{w_n} y_n \end{vmatrix} \qquad X = \begin{vmatrix} \sqrt{w_1} & \sqrt{w_1} x_1 \\ \sqrt{w_2} & \sqrt{w_2} x_2 \\ & \cdots \\ \sqrt{w_n} & \sqrt{w_n} x_n \end{vmatrix}$$

the preceding relationship gives:

$$b = |X'X|^{-1} \times |X'Y|$$

where X′ represents the matrix transposed from X.

Solving this system yields the following expressions for the parameters b_0 and b_1, respectively:

$$b_0 = \frac{\left(\sum_{i=0}^{n} w_i x_i^2 \times \sum_{i=0}^{n} w_i y_i - \sum_{i=0}^{n} w_i x_i \times \sum_{i=0}^{n} w_i x_i y_i\right)}{D}$$

$$b_1 = \frac{\left(\sum_{i=0}^{n} w_i \times \sum_{i=0}^{n} w_i x_i y_i - \sum_{i=0}^{n} w_i y_i \times \sum_{i=0}^{n} w_i x_i\right)}{D}$$

with:
$$D = \sum_{i=0}^{n} w_i \times \sum_{i=0}^{n} w_i x_i^2 - \left(\sum_{i=0}^{n} w_i x_i\right)^2$$

5.6.3. *MULTILINEAR REGRESSION*

The previous analytical method based on the principle of least-squares can be generalised to all linear relationships of the form:

$$y = \beta_0 + \sum_{k=1}^{n} \beta_k X_k \qquad [18]$$

Linear regression with a constant term of the order n + 1 can always be adapted to regression of the nth order with the constant term by an appropriate transformation of the variables. We suppose that the observed or transformed variable y_i follows the relationship:

$$y_i = \beta_0 + \sum_{k=1}^{n} \beta_k X_k + \varepsilon_i \qquad [19]$$

By applying the principle of least squares, the minimum is:

$$S = \sum_{i=1}^{n} w_i \left(y_i - \sum_{k=1}^{n} \beta_k X_k\right)^2 \qquad [20]$$

such that $\partial S / \partial \beta_k = 0$ for k = 1, 2, 3 … n. Once more we are solving a system of n linear equations with n unknowns. As before we put:

$$Y = \begin{vmatrix} \sqrt{w_1}\, y_1 \\ \sqrt{w_2}\, y_2 \\ \cdots \\ \sqrt{w_n}\, y_n \end{vmatrix} \qquad X = \begin{vmatrix} \sqrt{w_1}\, x_{1,1} & \sqrt{w_1}\, x_{1,2} & \cdots & \sqrt{w_1}\, x_{1,n} \\ \sqrt{w_2}\, x_{2,1} & \sqrt{w_2}\, x_{2,2} & \cdots & \sqrt{w_2}\, x_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ \sqrt{w_n}\, y_n & \sqrt{w_n}\, y_{n,2} & \cdots & \sqrt{w_n}\, y_{n,n} \end{vmatrix}$$

5.6.4. *NON-LINEAR REGRESSION ANALYSIS*

The principle of least squares is only applicable to a linear relationship having between two and several variables. More often in practice, however, we have to analyse non-linear relationships, in particular the MICHAELIS equation. In this situation two approaches are possible. Either, with the help of an appropriate transformation

of several variables, we obtain a linear relationship, or when it is impossible to linearise, an iterative method must be used. One of the conditions for using the least-squares method is that, in the linear relationship, the variables must be independent.

For example, let us take the MICHAELIS equation:

$$v = \frac{V_m s}{K_m + s} = y$$

Let $\Delta K_m = K_{m2} - K_{m1}$, the difference between two estimates of K_m, which gives a value of v such that $\sum w_i (y_i - y)$ is a minimum. We have:

$$y_i = y_0 + \sum_{i=1}^{n} \Delta K_m \times \frac{\partial y}{\partial K_m}\Bigg)_{K_{m0}}$$

The expansion as a series can be limited to the first two terms:

$$y_i = \frac{V_{m1} s}{K_{m1} + s} - K_{m1} \frac{V_{m1} s}{(K_{m1} + s)^2}$$

If we write:

$$\alpha_1 = V_{m1} \qquad x_1 = s/(K_{m1} + s)$$
$$\alpha_2 = K_{m1} V_{m1} \qquad x_2 = s/(K_{m1} + s)^2$$

the expression then becomes: $y_i = \alpha_{1i} x_{1i} + \alpha_{2i} x_{2i}$

and we return to the former case. We determine α_1 and α_2 by minimising the function $\sum w_i (y_i - y)^2$, which is again multilinear regression. The values of α are given by the relationship:

$$|\alpha| = |X'X|^{-1} \times |X'Y|$$

From these α values, the new values of K_m and V_m are thus used in subsequent multilinear regression. This iterative process is continued until there is negligible change in the parameter values between two successive iterations.

5.6.5. CHECKING THE ADEQUACY OF THE FIT

5.6.5.1. EXAMINATION OF THE RESIDUAL VALUES

Analysis of the results requires an initial model, for example, that described by the MICHAELIS equation, and thereafter those parameter values are determined that account most appropriately for the experimental results. Before determining the confidence intervals associated with each parameter, it is necessary to test whether or not the chosen model is correct. This is achieved by inspection of the residual values, and is a useful test whatever the regression method employed. The residual values comprise n differences:

$$\varepsilon_i = y_i - Y$$

y_i being the experimental value and Y the calculated value of the function. The ε_i values correspond to the observed errors if the model is adequate. If so, the residuals must be independent and distributed randomly, their mean value must be nil and have a constant variance.

Let us look at the MICHAELIS-MENTEN equation. Figure 5.47a displays an example of fitting using a non-linear regression method; Fig. 5.47b gives the distribution of residues. First, there is good apparent agreement between the experimental and calculated values; second, it is clear that the residuals are distributed randomly.

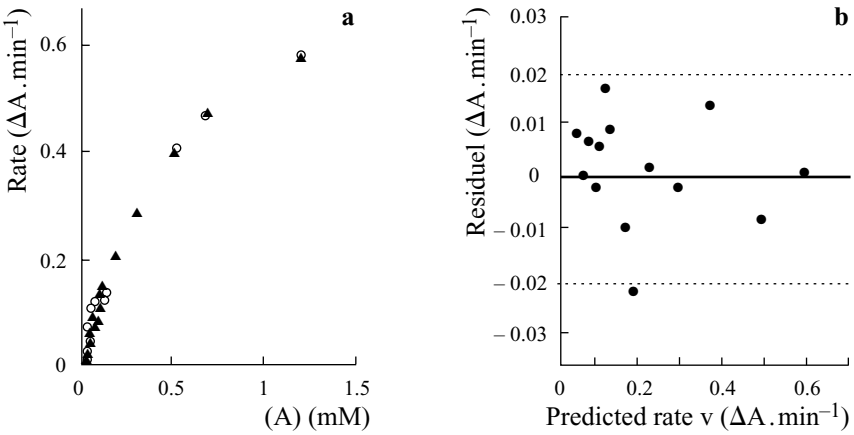


Fig. 5.47 (a) fit to the MICHAELIS-MENTEN equation by a non-linear regression method. Convergence of the regression function as proved by the correspondence between the calculated (▲) and experimental (○) values of the reaction rate for different substrate concentrations – (b) distribution of the residual values

Certain anomalies may arise as depicted in Fig. 5.48. They correspond either to (**a**) insufficient parameters, (**b**) an inadequate model, or (**c**) to a variance that depends on Y. It is thus useful to weight the experimental values. Another important criterion is that, in linear regression, the covariance of x and y must be nil.

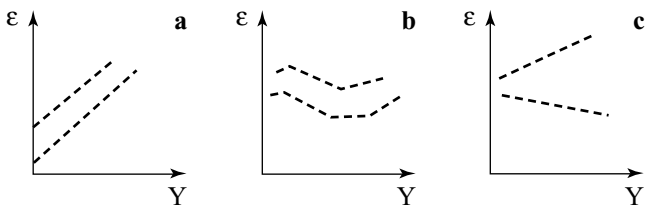


Fig. 5.48 Anomalies in the distribution of the residuals

5.6.5.2. THE WEIGHTING FACTOR, w_I

The variance is not always constant as can be seen by studying the residual values. To account for this, a statistical weight w_i must be applied to each value of y_i. It is inversely proportional to the variance of y_i and even to the square of the absolute error: $w_i = 1/(\Delta y_i)^2$. If Δy_i is constant, w_i is constant and it cancels out in the various expressions. This is equivalent to putting it equal to 1. If the relative error $\Delta y_i/y_i$ is constant, $w_i \neq K/y_i^2$. If the law that governs the change in error is not known in advance, it is helpful to carry out a polynomial regression with the pairs (ε_i, y_i) in order

to obtain the polynomial $f(y_i)$ which represents the error change in y_i. We then let $w_i = 1/f(y_i)$. ◢

5.6.5.3. GENERAL STRATEGY

As a general rule to analyse the kinetics of an enzyme reaction, we must first carry out a certain number of logical operations, and in a certain order. These are:

- correctly define the problem, for example, the dependence of the rate on the reactant concentrations;
- perform adequate experiments under optimal conditions for the intended purpose, which might be either the determination of parameters, or the choice of a better kinetic model;
- collect the experimental data in the order that they are acquired as well as any information that might reveal systematic errors;
- formulate the problem using a mathematical model, for example, the rate equation;
- analyse the data while considering the agreement between the model and the results, either in the estimation of parameters, or the choice of a kinetic scheme;
- interpretation of the results must lead to an understanding of the system, and to the possibility of generalising the conclusions. **An important test is the experimental checking of predictions that result from the model.**

BIBLIOGRAPHY

GENERAL WORKS

BERTRANDIAS F. & BERTRANDIAS J.P. –1997– *Mathématiques pour les sciences de la vie, de la nature et de la santé*, Grenoble Sciences Collection, EDP Sciences, Paris.

CORNISH-BOWDEN E., SAKS V. & JAMIN M. –2004– *Principes fondamentaux de la catalyse enzymatique*, Grenoble Sciences Collection, EDP Sciences, Paris.

DESMADRIL M. –1989– *L'Analyse des données en C*, Eyrolles, Paris.

DIXON M. & WEBB E. –1964– *Enzymes*, 2nd ed., Longmans, London.

HALDANE J.B.S. –1930– *Enzymes*, Longmans, London.

PROTASSOV K. –2002– *Analyse statistique des données expérimentales*, Grenoble Sciences Collection, EDP Sciences, Paris.

RICARD J. –1973– *Cinétique et mécanismes d'action des enzymes. I- Cinétique enzymatique phénoménologique*, Doin, Paris.

SEGEL I.M. –1975– *Enzyme kinetics*, John Wiley & Sons, New York.

YON J. –1961– *Mécanismes des réactions enzymatiques*, Publication of the Instituto de Biofisica, Rio de Janeiro, Brasil.

REVIEWS

BRAUNSTEIN A.E. –1947– Transamination and the interactive function of dicarboxylic acids in nitrogen metabolism, *Adv. Prot. Chem.* **3**, 1–52b.

BRAUNSTEIN A.E. –1957– Principal ways of assimilation and dissimilation in animals, *Adv. Enzymol.* **19**, 335–389.

CLELAND W.W. –1970– *The Enzymes*, 3rd ed., Vol. 2, 105, Acad. Press, New York.

CLELAND W.W. –1977– Determining the chemical mechanisms of enzyme-catalyzed reactions by kinetic studies, *Adv. Enzymol.* **5**, 273–387.

SPECIALISED ARTICLES

ALBERTY R.A. –1953– *J. Am. Chem. Soc.* **75**, 1928.

ALBERTY R.A., MASSEY V., FRIEDEN C. & FUHLBRIGGE A.R. –1954– *J. Am. Chem. Soc.* **76**, 2485.

BANKS B.E.C., LAURENCE A.J., VERNON C.A. & WOOTON J.F. –1963– Kinetic studies of glutamic-aspartic transaminase (from heart muscle), in *Chemical and biological aspects of pyridoxal catalysis*, E.E. SNELL, P.M. FASELLA, A.E. BRAUNSTEIN & A. ROSSI-FANELLI eds, 197–215, Pergamon Press, London.

BECHET J.J., GARDIENNET M.C. & YON J.M. –1966– *Biochim. Biophys. Acta* **122**, 101–115.

BECHET J.J. & YON J.M. –1964– *Biochim. Biophys. Acta* **89**, 117–126.

BENDER M.L., BEGUE-CANTON M.L., BLAKELEY B.L., BRUBACHER L.J., FEDER J., GUNTER C.R., KEZDY F.J., KILLHEFER J.U., MARSHALL T.H., MILLER C.J., ROEKE R.W. & STOOPS J.K. –1966– *J. Am. Chem. Soc.* **88**, 5890.

BLOOMFIELD V., PELLER L. & ALBERTY R.A. –1962– *J. Am. Chem. Soc.* **84**, 4367.

BRIGGS G.E. & HALDANE J.B.S. –1925– *Biochem. J.* **19**, 338.

CHEVALLIER J. & YON J. –1966– *Biochim. Biophys. Acta* **112**, 116–126.

CHOU K.C. –1980– *Eur. J. Biochem.* **113**, 195–198.

CHOU K.C. –1981– *Can. J. Biochem.* **59**, 757–761.

CHOU K.C. & FORSEN S. –1980– *Biochem. J.* **187**, 829–835.

CLELAND W.W. –1963– *Biochim. Biophys. Acta* **67**, 104–137.

COLQUHOUN D. –1971–*Lectures in Biostatistics,* 257–272, Clarendon Press, Oxford.

CORNISH-BOWDEN A. & EISENTHAL R. –1974– *Biochem. J.* **139**, 715-720.

DALZIEL K. –1957– *Acta Chem. Scand.* **11**, 1706.

DESCHAVANNES P.J., VIRATELLE O.M. & YON J.M. –1978– Functional properties of beta-galactosidase from mutant strain 13 PO of *Escherichia coli*, in *Proc. Natl Acad. Sci. USA* **75**, 1892–1896.

DIXON M. –1953– *Biochem. J.* **55**, 170.

DOWD J.E. & RIGGS D.S. –1965– *J. Biol. Chem.* **240**, 863–869.

EADIE G.S. –1942– *J. Biol. Chem.* **146**, 85–93.

EISENTHAL R. & CORNISH-BOWDEN A. –1974– *Biochem. J.* **139**, 721–730.

FINDLAY D., MATHIAS A.P. & RABIN B.R. –1960– *Nature* **187**, 601.

GLAZER A.N. –1966– *J. Biol. Chem.* **241**, 635.

HAMMES G.G. & FASELLA P.M. –1963– in *Chemical and biological aspects of pyridoxal catalysis*, E.E. SNELL, P.M. FASELLA, A.E. BRAUNSTEIN & A. ROSSI-FANELLI eds, 185-195, Pergamon Press, London.

HANES C.S. –1932– *Biochem. J.* **26**, 1406-1421.

HENRI V. –1902– *C.R. Acad. Sci. Paris* **135**, 916.

JOHANSEN G. & LUMRY R. –1961– *C.R. Trav. Lab. Carlsberg* **32**, 185–214.

KING E.L. & ALTMANN C. –1956– *J. Phys. Chem.* **60**, 1375–1378.

KOSHLAND D.E. & HERR E.B. –1957– *J. Biol. Chem.* **228**, 1021.

LINEWEAVER H. & BURK D. –1934– *J. Am. Chem. Soc.* **56**, 658–666.

MICHAELIS L. & MENTEN M.L. –1913– *Biochem. Z.* **49**, 333–369.

NOAT T.G., RICARD J., BOREL M. & GOT C. –1968– *Eur. J. Biochem.* **5**, 55.

ONSAGER L. –1931– *Phys. Rev.* **37**, 405.

SEYDOUX F. & YON J.M. –1967– *Eur. J. Biochem.* **3**, 42–56.

STRAUSS O.H. & GOLDSTEIN A. –1943– *J. Gen. Physiol.* **26**, 559.

TENU J.P., VIRATELLE O.M. & YON J.M. –1972– *Eur. J. Biochem.* **26**, 112–118.

VAVRA I. & VELICK S.F. –1962– *J. Biol. Chem.* **237**, 2109.

VIRATELLE O.M. & YON J.M. –1973– *Eur. J. Biochem.* **33**, 110–116.

VOLKENSTEIN M.V. & GOLDSTEIN B.N. –1966– *Biochem. Biophys. Acta* **115**, 471–477.

WILKINSON G.N. –1961– *Biochem. J.* **80**, 324–332.

WILSON I.B., BERGMANN F. & NACHMANSOHN D. –1950– *J. Biol. Chem.* **186**, 781.

YON J.M. –1976– *Biochimie* **58**, 61–69.

6 – EXPERIMENTAL METHODS TO STUDY ENZYMATIC REACTIONS

The study of enzymatic reaction mechanisms involves researching the elementary steps that intervene during the course of a reaction, and determining the rates of appearance and disappearance of reaction intermediates, as well as their relative energetic levels. The majority of intermediates that appear have very short lifetimes and are not generally detectable by classical techniques. Increasingly rapid methods have been progressively introduced for studying enzyme reactions as is the case more generally for all kinetic studies. The discovery of techniques that discern ever lower time constants marked great progress in the approaches taken to study enzymatic reaction mechanisms.

Kinetic studies aim to measure the progress of a reaction over time. As far as the time-scale is concerned, frequently used methods detect times greater than or equal to a second. Between 1 s and 10^{-10} s lies a zone of fast reactions still accessible to kinetic experimentation. Times of less than 10^{-10} s (between 10^{-10} and 10^{-15} s) represent the spectroscopic range, for example, the time taken for an electron to pass from the ground state to the excited state.

Routine methods	Fast methods		Spectroscopic zone	
1	10^{-3}	10^{-8}	10^{-10}	10^{-15}
Seconds				

Kinetic studies of enzyme reactions consist of various aspects. First of all, there is the global reaction for which kinetic parameters are determined, on a time-scale of the order of one second or a minute. Among the methods often employed to follow enzyme reactions, it is useful to distinguish between discontinuous methods (or "point by point"), continuous methods and coupled enzyme assays.

The detection and study of elementary steps require the use of fast kinetic techniques. Two types of method are at our disposal: flow and relaxation, which are based on very different principles. Flow methods permit study at the pre-steady or non-steady state. Relaxation methods apply to reactions at equilibrium or at the steady state.

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_7, © Springer-Verlag Berlin Heidelberg 2010

6.1. DISCONTINUOUS METHODS

Discontinuous methods consist of removing aliquots from the reaction mixture at given time-points and then stopping their reactions, either by changing the pH, or by adding a suitable reagent or an inhibitor specific to the enzyme. Dilute acids are often used, in particular trichloracetic acid, which not only stop the reaction by lowering the pH, but also precipitate the enzyme. So the aliquot simply requires centrifugation in order to recover the supernatant in which the reaction products remain plus any unused substrate, except if this is a protein. In the sample, we either measure the product that has appeared or the unconsumed substrate. The principal methods of quantification are essentially chemical methods or those involving radioactive isotopes when it is possible to separate the product from the substrate. These "point by point" methods present several disadvantages, however. Sometimes it is difficult to estimate precisely the initial rates; besides, such methods are often laborious. In general, continuous methods are preferable whenever possible.

Let us cite the example of the method used to follow the activity of aspartate transcarbamylase, an enzyme that catalyses the carbamylation of the amino group of aspartate according to the reaction:

$$[^{14}C]\text{-aspartate} + \text{carbamyl phosphate} \rightleftharpoons \text{carbamyl}[^{14}C]\text{-Asp} + P_i$$

Using ion exchange chromatography the radioactive aspartate [^{14}C], which has not reacted, and the reaction product, carbamyl [^{14}C]-aspartate, can be separated. Firstly, the reaction is stopped by adding acetic acid, which at the same time ensures the protonation of the aspartate's amino group. The reaction mix is then fractionated on a DOWEX 50 column, which is an anionic resin that retains any aspartate that has not reacted and allows the carbamyl aspartate to flow through as it is now devoid of positive charge. This is then directly collected in vials used specially for measuring radioactivity. In practice, the reaction in the aliquot is stopped at the moment of acetic acid addition before being loaded onto the DOWEX column. Thus, a series of samples can be taken, each spaced apart by time intervals of only a few seconds.

6.2. CONTINUOUS METHODS

In continuous methods, the reaction progress is recorded directly as a function of time. If the substrate or the product has characteristic absorption spectra, each shifted relative to the other, the detection can be achieved by **absorbance spectrophotometry** at the wavelength where the difference in absorbance between the substrate and product is the greatest. Figure 6.1 opposite shows the measurement at 373 nm of orthonitrophenol liberated during the hydrolysis of o-nitrophenyl galactoside by β-galactosidase from *E. coli*. The continous trace obtained is linear for a while and its precision allows v_0 to be well determined.

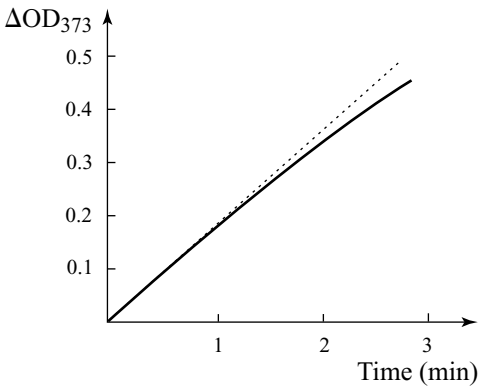


Fig. 6.1 Kinetics of the appearance of orthonitrophenol during the hydrolysis of ortho-nitrophenyl galactoside (ONPG) by E. coli β-galactosidase

Detection by **fluorescence** is a much more sensitive method to use when possible. For example, it is used to follow the appearance of NADH in dehydrogenation reactions catalysed by NAD^+ dehydrogenases.

Spectropolarimetry is used, for instance, to follow the mutarotation of glucose catalysed by invertase.

The **potentiometric method at a constant pH** using a pH-stat comprising an automatic titrant burette and a recorder enables reactions that release or capture protons to be followed continuously. In all other methods, it is important to work with a well-buffered solution in order to maintain a constant pH throughout the reaction. On the contrary, the potentiometric method requires the reaction medium to be unbuffered. The pH is maintained constant by adding, depending on the reaction, NaOH or HCl whose added quantities enable the reaction progress to be monitored. It is essential to maintain an ionic strength that is sufficiently high (in general, 0.1 M NaCl) for it to remain constant throughout the entire duration of the measurement. It is recommended to work under a nitrogen atmosphere to avoid acidification of the medium by atmospheric CO_2.

This method can be applied quantitatively to the enzymatic hydrolysis of ester substrates:

$$\mathrm{R{-}COOR' + H_2O \longrightarrow R{-}COOH + R'OH}$$

$$\mathrm{R{-}COOH \rightleftharpoons R{-}COO^- + H^+}$$

H^+ ions are neutralised by NaOH; hydrolases are generally active at neutral pH, i.e. well above the pK of the acid, which is entirely dissociated and the number of protons liberated is equal to the number of product molecules appearing. This method can also be applied to the study of protein hydrolysis by proteases. But depending on the pH of the reaction, most often only a fraction of the hydrolysed peptide bonds are titrated; this fraction can be determined. Indeed, the breaking of

a peptide bond at neutral pH causes a completely deprotonated COO^- group to appear and an amino group such as:

$$\underset{1-\alpha}{NH_3^+} \longrightarrow \underset{\alpha}{NH_2} + \underset{\alpha}{H^+}$$

with the dissociation constant: $K_{diss} = \dfrac{(NH_2)(H^+)}{(NH^+_3)} = \dfrac{\alpha(H^+)}{(1-\alpha)}$

giving: $$\alpha = \frac{1}{\dfrac{H^+}{K_{diss}} + 1}$$

If the pK of the amino group is much lower than the measurement pH, $\alpha \longrightarrow 1$; if the pK is equal to the measurement pH, only half of the amine groups liberated by hydrolysis are titrated. Therefore, firstly we have to determine α by titrating at different pH values. Once α is known, the measurement is quantifiable.

Continuous methods are particularly well adapted to the determination of kinetic parameters from graphical plots from the integrated MICHAELIS equation.

6.3. COUPLED ENZYME ASSAYS

When direct continuous methods are not available to record the appearance of product during an enzyme reaction, we can in certain cases and **working in suitable conditions** use a helpful trick. If the reaction product is itself a substrate for an enzyme that catalyses a reaction one of whose products is readily measurable by continuous methods, this second "***coupling***" enzyme can be used to follow the first reaction. We have therefore two consecutive reactions, the kinetics of which have been the subject of theoretical study (e.g. BERGMEYER, 1953; GUTFREUND, 1965; MC LURE, 1969; BARWELL & HESS, 1970; HART, 1970; GOLDMAN & KATCHALSKI, 1971; EASTERBY, 1973; STORER & CORNISH-BOWDEN, 1974).

The simplest example of enzyme coupling (EASTERBY, 1973) can be represented by the following scheme:

$$S \xrightarrow[E]{v_0} I \xrightarrow[E_1]{V_1/K_1} P$$

S is the substrate of the enzyme E that we wish to study; I, the reaction product, is the substrate for the coupled enzyme E_1; it is therefore a simple intermediate of two consecutives reactions. P is the final product, which we measure by a continuous method. The rate of the enzyme reaction under study is given by v_0; V_1 and K_1 are respectively the maximal rate and the MICHAELIS constant of the coupling enzyme. **For the enzyme assays to be valid, it is important that the second reac-**

tion is not rate-limiting, therefore V_1 must be much higher than v_0. The concentration of the intermediate I must be very low and much lower than K_1. Thus, the appearance rate of P is given by the relation:

$$\frac{dP}{dt} = \frac{V_1}{K_1}(I)$$

in which V_1/K_1 is a *first-order rate constant*.

The reaction rate is given by the following expression:

$$\frac{dI}{dt} + \frac{V_1}{K_1}(I) = v_0$$

with the following conditions: (I) = 0 and (P) = 0 at time t = 0, giving:

$$(I) = v_0\left(1 - e^{-t/\tau_1}\right)$$

$$(P) = v_0\left(t + \tau_1 e^{-t/\tau_1} - \tau_1\right)$$

with: $\tau_1 = K_1/V_1$

At steady state: $(I)_{ss} = v_0\tau_1 = v_0\dfrac{K_1}{V_1}$

Figure 6.2 shows the effect of the ratio V_1/K_1 on the time needed to reach the steady state.

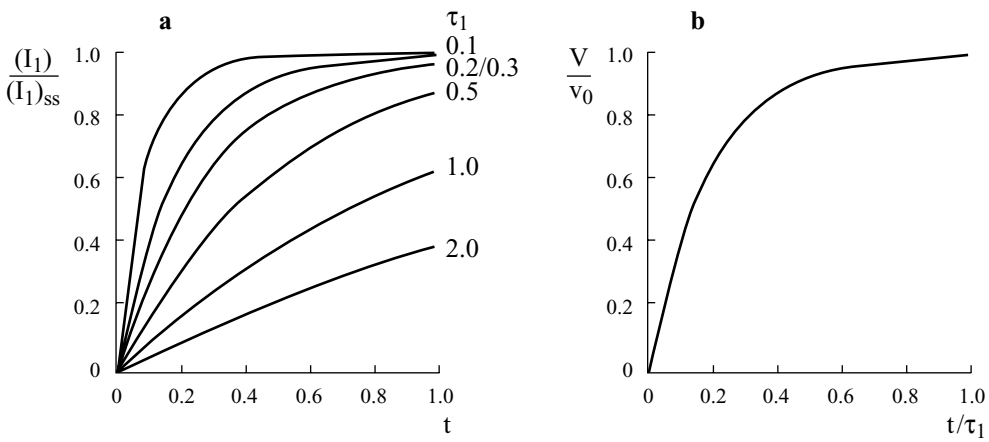


Fig. 6.2 Effect of the transient time $\tau_1 = K_1/V_1$ on the time needed to reach the steady state

***(a)** relative concentration of the intermediate $(I)/(I)_{ss}$ – **(b)** rate normalised with respect to v_0 as a function of t/τ_1.* (Reprinted from *Biochim. Biophys. Acta*, **293**, EASTERBY J.S., Coupled enzyme assays: a general expression for the transient, 552. © (1973) with permission from Elsevier)

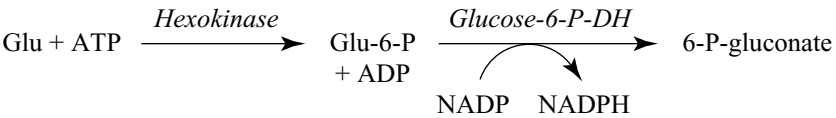
When time becomes sufficiently long, the exponential term becomes negligible in the expression for P, which takes on a simplified form:

$$(P) \to \infty = v_0(t - \tau_1)$$

giving:
$$\frac{dP}{dt} = v_0$$

The appearance rate of P represents well the rate of the reaction catalysed by the first enzyme.

By way of example, let us consider the coupled enzyme assays used to follow the reaction catalysed by yeast hexokinase. It utilises glucose-6-phosphate dehydrogenase as the coupling enzyme, which requires $NADP^+$. We have the following chain of reactions:



The change in absorbance due to the formation of NADPH is recorded at 340 nm. Figure 6.3 shows product appearance over time. The pre-steady state phase appears very clearly; it is followed by the steady state and corresponds to the linear part of the graph. This straight line may be extrapolated to a value equal to τ_1 on the horizontal axis and on the vertical axis, equal to I_{ss} or $v_0\tau_1$, from the expression for the appearance rate of P.

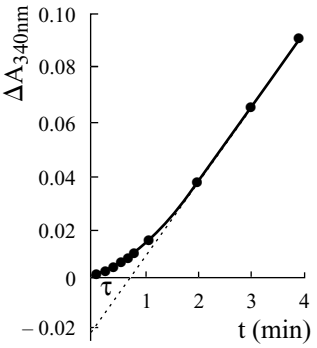


Fig. 6.3 Enzyme assay with hexokinase Kinetics of NADH appearance over time

It is sometimes necessary to use several coupled reactions. The kinetics of several consecutive coupled reactions was developed by BARWELL and HESS (1970) and by EASTERBY (1973). The reaction sequence can be written:

$$S \xrightarrow{v_0} I_1 \xrightarrow{V_1/K_1} I_2 \xrightarrow{V_2/K_2} I_3 \ \ldots \ I_n \xrightarrow{V_n/K_n} P$$

with V_1, $V_2 \ldots V_n$ and K_1, $K_2 \ldots K_n$ being, respectively, the values for the maximal rates and the MICHAELIS constants of these consecutive enzymatic reactions. **It is important to work in conditions in which every coupled reaction is first order.**

The system can be described by the following series of differential equations:

$$\frac{dI_1}{dt} + \frac{I_i}{i} = \frac{I_{i-1}}{i-1} \qquad i = 2, 3 \ldots n$$

with the initial conditions $I_i = 0$ at $t = 0$.

The solution to this equation for I_1 has already been given; for the successive intermediates, the solution is obtained by progressive substitution into the equation. The general solution is:

$$I_i = v_0\tau_i\left(1-\sum_{j=1}^{i}A_j e^{-t/\tau_j}\right) \qquad i = 1, 2, 3\ldots n$$

with:
$$A_j = \tau_j^{i-1}\prod_{k=j}^{i}\frac{1}{\tau_i-\tau_j}$$

The concentration of the intermediate I_i at steady state is given by the relation:

$$I_{i,ss} = v_0 - \tau_i \qquad \text{with } \tau_i = K_i/V_i$$

The final product concentration measured is:

$$\frac{dP}{dt} = \frac{V_n}{K_n}I_n$$

Taking into account the initial conditions: $P = 0$ at $t = 0$, we have:

$$P = v_0\left(t+\sum_{j=1}^{n}C_i e^{-t/\tau_i} - \sum_{i=1}^{n}\tau_i\right)$$

with:
$$C_i = \tau_i^k\prod_{j=1}^{n}\frac{1}{\tau_i-\tau_j} \qquad j \neq i$$

At steady state:
$$P_{t\to\infty} = v_0\left(t-\sum_{i=1}^{n}\tau_i\right)$$

the corresponding curve, analogous to that in the preceding figure, intersects the horizontal axis at the point:

$$t = \sum_{i=1}^{n}\tau_i$$

and the vertical axis at:
$$-v_0 = \sum_{i-1}^{n}\tau_i$$

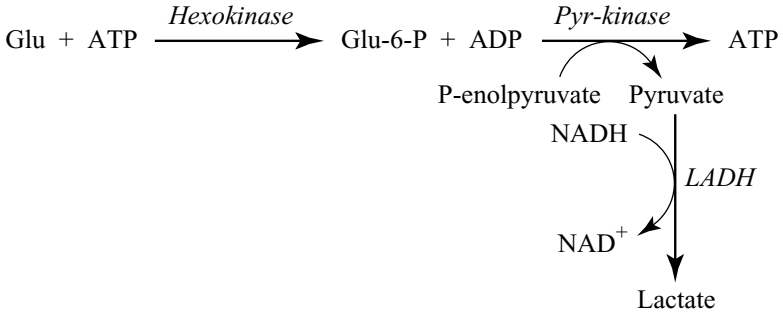
which is the sum (bearing a minus sign) of the concentrations of intermediate species at steady state.

Thus, when two coupled enzymes are used, the concentration of P is given by the expression:

$$P = v_0\left(t+\frac{\tau_1^2}{\tau_1+\tau_2}e^{-t/\tau_1}+\frac{\tau_2^2}{\tau_1+\tau_2}e^{-t/\tau_2}-\tau_1-\tau_2\right)$$

For example, the coupled enzymatic assay used for yeast hexokinase is not applicable to animal hexokinases because an accumulation of 6-phosphogluconate leads

to inhibition of these enzymes. In this case, we can employ a method using two coupling enzymes: pyruvate kinase and lactate dehydrogenase. Below are the corresponding reactions:



STORER and CORNISH-BOWDEN (1974) proposed a method for the kinetic analysis of coupled enzyme reactions that avoids simplification by considering that the second reaction is first order.

$$S \xrightarrow{v_1} I \xrightarrow{v_2} P$$

The reaction catalysed by the second enzyme is governed by MICHAELIS' law:

$$v_2 = \frac{V_2(I)}{K_2 + (I)} \qquad [1]$$

We have:

$$\frac{d(I)}{dt} = v_1 + v_2 \qquad [2]$$

From the MICHAELIS equation, we deduce:

$$(I) = \frac{K_2 v_2}{V_2 - v_2} \qquad [3]$$

and:

$$\frac{d(I)}{dv_2} = \frac{V_2 K_2}{(V_2 - v_2)^2} \qquad [4]$$

giving:

$$\frac{d(I)}{dt} = \frac{V_2 K_2}{(V_2 - v_2)^2} \frac{dv_2}{dt} \qquad [5]$$

By substituting into equation [2], we obtain:

$$\frac{V_2 K_2}{(v_1 - v_2)(V_2 - v_2)^2} dv_2 = dt \qquad [6]$$

and after integration:

$$t + Ct = -\frac{V_2 K_2}{V_2 - v_1}\left(\frac{1}{V_2 - v_2} + \frac{1}{V_2 - v_1} \ln \frac{v_1 - v_2}{V_2 - v_2}\right) \qquad [7]$$

The constant of integration, C, is determined by the initial conditions $v_2 = 0$ at $t = 0$. Thus:

$$t = \frac{V_2 K_2}{(V_2 - v_1)^2} \ln \frac{v_1(V_2 - v_2)}{V_2(v_1 - v_2)} - \frac{K_2 v_2}{(V_2 - v_2)(V_2 - v_1)} \qquad [8]$$

Equation [2] shows that at steady state $v_1 = v_2$ since $d(I)/dt = 0$. The appearance rate of P measured experimentally is equal to the appearance rate of (I). Figure 6.4 shows the correspondence between the theoretical and experimental curves for the reaction kinetics with glucokinase as followed by assay and using glucose-6-phosphate dehydrogenase as the coupling enzyme, for different values of its concentration. The difficulty resides in evaluating the rate's limit. It is therefore preferable to calculate the time necessary for v_2 to reach 0.99 v_1 rather than relying on the apparent linearity of the experimental curves. The calculation method thus involves at least a partial estimation of v_1.

With a first-order approximation for the coupling enzyme, the kinetic constants V_2 and K_2 are not separated, but the ratio V_2/K_2 represents the first-order constant. Also, it is not possible to separate the factors that alter each of the apparent values of these constants. On the contrary, in the general rate expression represented by equation [8], it is possible to separate them; the time needed for v_2 to reach a given percentage of v_1 is proportional to K_2. Consequently, the presence of a competitive inhibitor of the coupling enzyme diminishes the method's efficiency by delaying the time necessary to reach the steady state. Thus, during the assay of glucokinase by glucose-6-phosphate dehydrogenase, the latency period in the presence of ATP, a competitive inhibitor of the coupling enzyme, is notably extended.

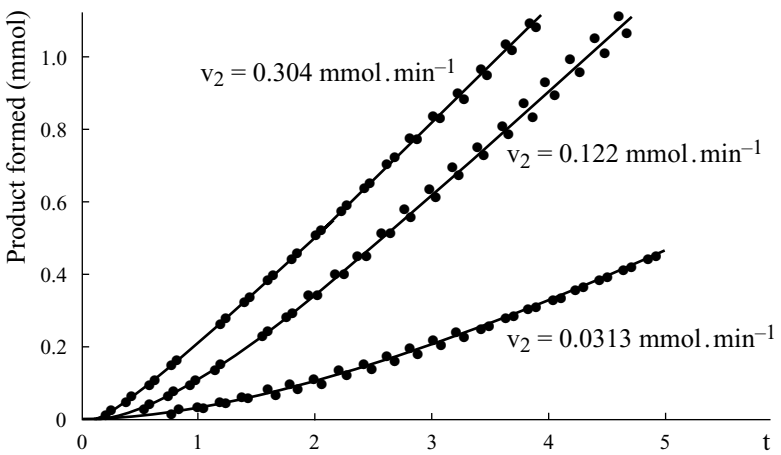


Fig. 6.4 Comparison of experimental and theoretical curves for the enzyme assay of glucokinase with glucose-6-phosphate dehydrogenase as the coupling enzyme

(Reproduced with permission from STORER A.C. & CORNISH-BOWDEN A. (1974), *Biochemical Journal*, **141**, 205. © The Biochemical Society)

A coupled enzyme assay requires sufficient coupling enzyme, but this must not exceed the necessary concentration, not merely in the interests of economy, rather because excessively high concentrations of coupling enzyme are sometimes the source of artefacts. For example, glucose-6-phosphate dehydrogenase can use glucose instead of glucose-6-phosphate, however, this parasitic reaction is negligible at the coupling enzyme concentrations necessary for the assay. The use of a coupled enzyme assay applies equally when the product formed is either unstable or difficult to detect. Now, carbamyl phosphate synthetase catalyses the formation of carbamyl phosphate from glutamine, bicarbonate and ATP. Being an acid anhydride, carbamyl phosphate is unstable. We must therefore use the coupling enzyme, aspartate transcarbamylase, which, in the presence of saturating aspartate concentrations, immediately converts the carbamyl phosphate into stable carbamyl aspartate that is assayed by the method described above.

6.4. FLOW METHODS

Flow methods are applicable to the study of enzymatic reactions **in pre-steady state conditions or more generally at the steady state.** They require an experimental device enabling the determinations to be carried out in very short time periods, of the order of a few milliseconds. The problems encountered in fast kinetic studies are the rapidity of mixing the reactants and the detection of modifications appearing in the mixture. The principle of the method consists of quickly injecting the two reactants, e.g. the enzyme and substrate, into a mixing chamber or an observation tube.

6.4.1. GENERAL PRINCIPLE OF FLOW METHODS

Whatever the flow type, the apparatus contains a device in which the solutions of the two reactants R_1 and R_2 are introduced by forced mixing at the entrance of an observation tube. Beyond this point, when the two solutions are instantaneously combined to homogeneity, the destiny of the reaction mixture differs depending on whether or not there is a continous flow.

6.4.2. CONTINUOUS-FLOW APPARATUS

In a continuous-flow apparatus (Fig. 6.5 opposite), the reaction mix progresses towards the interior of the observation tube at a constant rate v, such that in any section of the tube S_d, situated at a distance d from the point of mixing, a time t (equal to the ratio d/v) can be unequivocally associated to a well-defined point in the reaction. When the mixture travels a certain length of the tube at a distance d, it is in the same state that it would be in at time t if it remained immobile. While the rate, v, remains constant, a steady distribution of the reactants, products and possibly the

intermediates participating in the reaction studied is established all along the observation tube. If there is a physical property known to correspond to the changes in concentrations of the chemical species present in the reaction medium, for example spectral properties, then by placing an appropriate detector along the length of the observation tube we can obtain an electrical response that is a function of the distance d. Then it suffices to convert these experimental data into concentration and time, respectively.

The first apparatus of this type was constructed in 1922 by HARTRIDGE and ROUGHTON. It comprised an observation tube 5 mm in diameter and consumed 3–6 l of reactant per experiment! It was first used to study the combination of oxygen with carboxyhaemoglobin under the influence of flash photolysis:

$$COHb + O_2 \underset{}{\overset{h\nu}{\rightleftharpoons}} HbO_2 + CO_2$$

The authors succeeded in measuring the rate of the reverse reaction which was carried out in the dark. This reaction is fast and could not be followed at temperatures higher than 15°C. With the device it was possible to measure at 37°C the rate of combination of oxygen with haemoglobin, which occurs in 0.01 s.

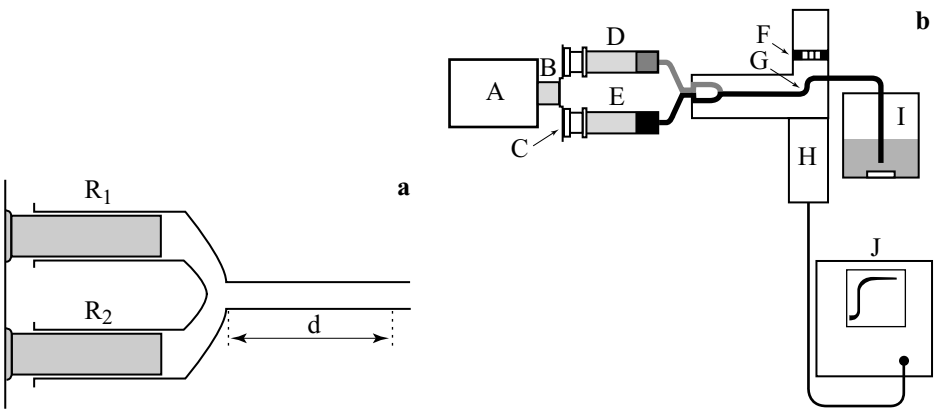


Fig. 6.5 Continuous-flow apparatus
*(**a**) scheme showing the principle of the method – (**b**) layout of a continuous-flow apparatus. A: constant rate motor; B: drive piston; C: end-stops for the syringe plungers; D and E: syringes; F: filter or monochromator; G: observation point; H: photomultiplier; I: mixing chamber; J: recorder*

The advantage with these continuous-flow apparatus is that systems for the detection of relatively long global response times can be adapted to them. This imposes a minimum duration with each stoppage of the detector at a given point in the observation tube and affects the total flow time and consequently the consumption of the reactants. However, in 1935 ROUGHTON and MILLIKAN constructed an improved apparatus consuming less reactant.

A continuous-flow apparatus with a resolution time of 100 μs was developed by REGENFUSS and co-workers in 1985. The system consisted of mixing two solutions by passing them through a tiny aperture of 10 μm. The resulting turbulence provoked a rapid mixing by diffusion in the tiny whirlpools created.

6.4.3. STOPPED-FLOW APPARATUS

The stopped-flow system is designed to provide fast and completely homogenous mixing of two reactants and this mixture, as quickly as possible, must fill the observation tube that is placed in the path of the detector. In this type of apparatus, the detector must be positioned as closely as possible to the point of mixing. Its response time as well as that of its associated circuits must also be as short as possible. Stopping the flow may be passive or induced. It was passive in the first apparatus used by CHANCE when, after the impulse given to the pistons, the system came to a halt by inertia. Thus, with these instruments the flow rate reaches a maximum value and then decreases, tending to zero in a time-interval that can be of the order of the preliminary period of flow. In the apparatus developed by GIBSON (Fig. 6.6), the stoppage is induced and this time it is infinitesimally small. The flow is blocked abruptly at the moment that the rate reaches its maximum. After having passed through the observation tube, the reaction mixture is ejected into a third syringe whose piston may be halted by an adjustable stop.

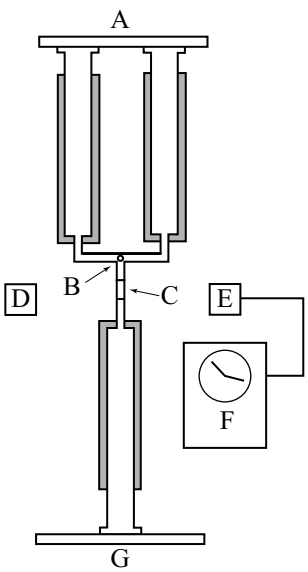


Fig. 6.6 Stopped-flow apparatus based on the GIBSON principle
A: drive pistons in the mixing syringes;
B: mixing; C: observation chamber;
D: monochromatic light source;
E: detector; F: recorder;
G: stop block for the stopper syringe.

The resolution time of these apparatus is generally of the order of several milliseconds (4–5 ms); in the best cases, they have resolution times of one millisecond and even half millisecond times have been reported. For this, a certain number of difficulties, which we shall soon address, have to be overcome. The detection system

most often used with stopped-flow equipment is absorption or emission spectrophotometry. It must be sensitive enough to permit detection of minor variations.

6.4.4. QUENCHED-FLOW APPARATUS

When no detection method is readily usable for measuring reaction rates, we can make use of fast mixing methods and then stop the reaction. The function of these multi-mixing or quenched-flow apparatus, whose diagrams are given in Fig. 6.7, relies on the following principles. The reactants are mixed quickly by means of a two-syringe system and left to incubate for the desired time; next, the reaction mixture is again rapidly mixed with a reagent to stop the reaction and the amount of product formed can be determined. The time during which the reaction is left to evolve is fixed according to the length of the incubation tube between the first and second mixer. The resolution time of these apparatus is of the order of a few tens of milliseconds (typically 20–30 ms).

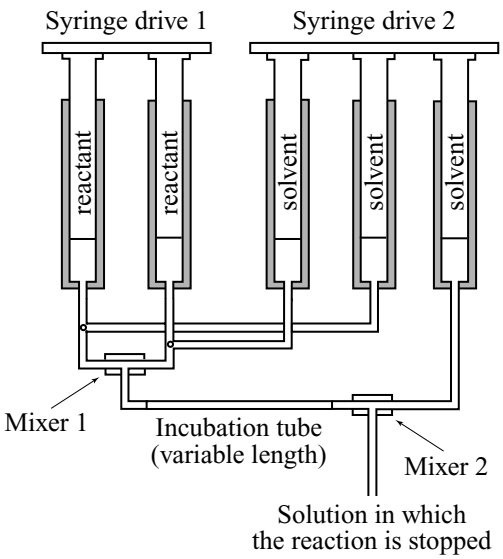


Fig. 6.7 Quenched-flow apparatus

6.4.5. CRITERION FOR HOMOGENEITY OF THE MIXTURE

Obviously, it is of the greatest importance that the time required to obtain an homogenous mix of the two reactants is short relative to the half-life of the reaction studied. From the very beginning of the development of flow methods, the mixing efficiency was investigated using either chemical or physical tests. The chemical tests consisted of measuring the progress of bimolecular reactions that were known to terminate in a time shorter than that measurable by the system. In this way, the mixing efficiency and the minimum response of the apparatus can be evaluated.

The most commonly used reaction is acid neutralisation by a base in the presence of a coloured indicator; the heat of neutralisation can also be measured. The physical tests include, for example, mixing two substances having different optical properties and then using an optical criterion as a means to test the homogeneity of the mixture. This principle was adopted by DUBOIS and TROWSE in the method of SCHLIEREN.

6.4.6. SOME TECHNICAL PROBLEMS

The resolution time of these apparatus depends on the mixing rate of the reactant solutions. However, this rate cannot pass certain limits. To obtain quickly good mixing by using only minimal quantities of reactants, it is important that the form of the mixing chamber is suitably adapted. In the first systems used, the chamber contained eight jets that delivered the liquid at a tangent, creating a rotational movement in the flow. CZERLINSKI developed a system of mixers with infinite jets involving two thin concentric layers facing one another at the top of a cyclindrical observation tube. Other systems have been described in which the many thin layers of liquid come into contact with each other. One such system comprised 25 layers and the flow rate was able to reach 10 $m . s^{-1}$. However, it is not beneficial to use too high a flow rate. Indeed, above a certain threshold value slightly higher than 10 $m . s^{-1}$, although this depends on the geometry of the mixing chamber, cavitation phenomena may occur. We note also that among the technical difficulties encountered while using this type of equipment were interfering vibrations and liquid leaks.

The methods for the titration of enzyme active sites using the burst technique that we presented in Section 5.4.6 often require the use of the stopped-flow technique. Similarly, the example given concerning the identification of an intermediate in reactions catalysed by β-galactosidase from a mutant *E. coli* strain by the appearance of a paranitrophenol burst was realised by means of a stopped-flow apparatus.

6.5. RELAXATION METHODS

Flow techniques enable rate constants between milliseconds and seconds to be achieved. However, they are not fast enough to allow the detection of certain reaction intermediates during enzymatic reactions. In particular, the formation of the first enzyme-substrate complex is far too fast to be detected by even the best stopped-flow apparatus. The relaxation methods introduced by EIGEN and DE MAEYER covered time-scales one order of magnitude smaller than flow methods. The time-scales associated with the different methods are displayed on page 193. Relaxation methods cover a time-scale of between seconds to 10^{-10} s.

6.5.1. PRINCIPLE OF RELAXATION METHODS

Relaxation methods are applicable to systems at equilibrium or in conditions of steady state. In physics, the perturbation of a system at equilibrium leads to its *relaxation* towards a new equilibrium state. Such is the case, for example, for the reorientation of dipoles after a change in the electric field. The analogy between these processes and those arising in chemical systems led EIGEN and DE MAEYER to use the term ***chemical relaxation***. We understand by chemical relaxation, any readjustment of a chemical equilibrium affected by a perturbation of the reactant concentrations, either following a change in the parameters related to that state, or affected by an external force influencing the equilibrium. The method consists of applying extremely rapidly to the system at equilibrium an external perturbation, such as an abrupt change in temperature, pressure or an electrical impulse, which can shift the equilibrium position. This definition also includes the processes studied by special flash techniques, where the equilibrium in the fundamental state is disturbed after transformation of the excited molecules, which become disactivated after reacting. If the perturbation is applied much more quickly than the time needed for readjustment of the chemical equilibrium, the concentrations of the different constituents will reach the new equilibrium position at a rate that is solely dependent on the reaction mechanism i.e. on the individual concentrations and the specific rate constants of each elementary step.

The technology has developed to such a degree that it is possible in practice to change variables such as the temperature or density of the electric field in times as short as 10^{-8} to 10^{-6} s. The discontinuity of the change relative to the rate of relaxation is such that the suffix *jump* is employed to designate this type of perturbation, as in for example the "temperature jump" method, often shortened to **T-jump.**

Two types of perturbation can be used depending on the reaction type to be studied: either a transient perturbation, or an alternative perturbation.

6.5.1.1. TRANSIENT PERTURBATION

A transient perturbation is applied rapidly to the system, which then attains its new state of equilibrium in a relatively long time compared to the perturbation time. Figure 6.8 below illustrates this sort of response for a simple equilibrium:

$$A \rightleftharpoons A'$$

In the thermal relaxation method, for example, a temperature rise of 10°C can be obtained in less than a microsecond by an impulse of 10^4–10^5 v through the solution. The rate equation is a linear differential equation and the relaxation curve is given by:

$$\Delta c = \Delta c_0 \, e^{-t/\tau}$$

Δc_0 and Δc are the changes in concentration at time 0 and t, respectively; τ represents the relaxation time. In the case of simple equilibrium, it is related in a simple way to the specific rate constants for the reaction as shown further on (§ 6.5.3.2). If the reaction comprises several steps, there is a complete spectrum of relaxation times and we obtain quite a complex expression.

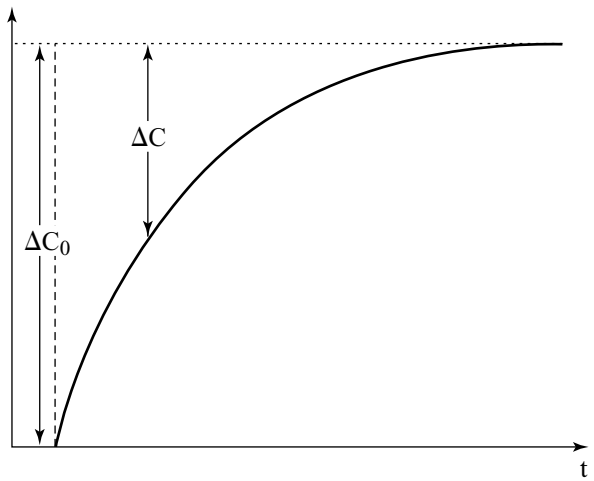


Fig. 6.8 Relaxation of a system subjected to a transient perturbation
The transient perturbation is shown as a dashed line*; the relaxation of the system, an* unbroken line

6.5.1.2. ALTERNATIVE PERTURBATION

In the alternative perturbation method, a periodic perturbation is applied to the system at equilibrium with a period of the order of time necessary to restore equilibrium. Due to the finite rate of chemical reactions, the response is periodic, but remains shorter than the perturbation frequency. The phase difference is a measure of the chemical reaction rate. The attenuation of the perturbation represents the energy dissipated by the system and can be measured. It is related to the phase difference by a FOURIER transform. Ultrasonic waves and electric fields are used to generate alternative perturbations. For a periodic perturbation provoked, for example, by an ultrasonic wave, we have the equation:

$$\overline{\Delta C} = ae^{i}\omega^{t}$$

a being a constant, ω the perturbation frequency and $i^2 = -1$. The concentration varies as a function of time according to the expression:

$$\Delta C = a^{ei\omega t}/(1 + i\omega t)$$

Thus ΔC oscillates with the same frequency as $\overline{\Delta C}$, but out of phase. Energy absorption per wavelength due to relaxation depends on the frequency, according to the following relation:

$$B\omega\tau/(1 + \omega^2\tau^2)$$

This function has a maximum when $\omega = 1/\tau$. The time limit of each direct method is given by the finite propagation rate of the signal through the system. Such methods were developed and used successfully for the perturbation of systems involving dipolar molecules in an alternative electric field.

6.5.2. PRINCIPAL RELAXATION METHODS

Different relaxation methods exist that differ in the nature of the perturbation and the detection method. Their use depends on the type of reaction under study.

6.5.2.1. THERMAL RELAXATION

Thermal relaxation (***temperature jump*** or ***T-jump***) is the appropriate method for studying transient phenomena in biochemical reactions such as the formation of enzyme-metal complexes, antibody-hapten reactions, enzymatic reactions and conformational changes in proteins and nucleic acids. It has been used with success in each of these different cases. Of course, its usage requires an obvious thermodynamic condition, namely, that **the reaction is accompanied by a non-zero enthalpy change,** ΔH. Indeed, the thermal dependence of an equilibrium constant is given by the classic equation of VAN T'HOFF:

$$\frac{d \ln K_{eq}}{dT} = \frac{\Delta H}{RT^2}$$

When ΔH is nil, the equilibrium constant does not change with temperature and consequently nor do the concentrations, therefore the equilibrium cannot be relaxed.

Thermal relaxation allows a time-scale to be covered that ranges from a second to about 10^{-7} s. In this method the perturbation is provoked by a heat shock generated by the discharge from a condenser at high voltage (100 kV), the capacity of the condenser being 0.02 μf. This technique produces a temperature jump of 6°C in less than 10^{-7} s in a volume of 1 mL, and it is homogenous throughout the solution. Such apparatus, schematised in Fig. 6.9 below, can be equipped for detection by absorption (single or double beam) or emission spectrophotometry, or even polarimetry, although this third detection method is generally too slow. The homogeneity of the temperature in the part of the cuvette in which the measurement takes place results from the particularly elaborate form of the cuvette, which, in principle, should also avoid any electrolysis caused by the discharge.

Thermal relaxation methods were very widespread a few years ago. Their use has led to the discovery of important intermediates in enzymatic reactions and helped to define the mechanisms for the functions of biological reactions. The most recent technological advances have given rise to the development of non-conventional T-jump apparatus. The temperature jump is provoked by a very fast impulse. This has permitted ever shorter resolution times to be achieved. Thus, in the apparatus constructed by PHILLIPS et al. (1995), the solutions are heated by a laser pulse, the rays are then absorbed by the dye molecules homogenously dispersed in the solution. A 10°C rise in temperature is obtained in 70 ps. With such methods, it is important to check that the dye does not interfere with the reaction being studied.

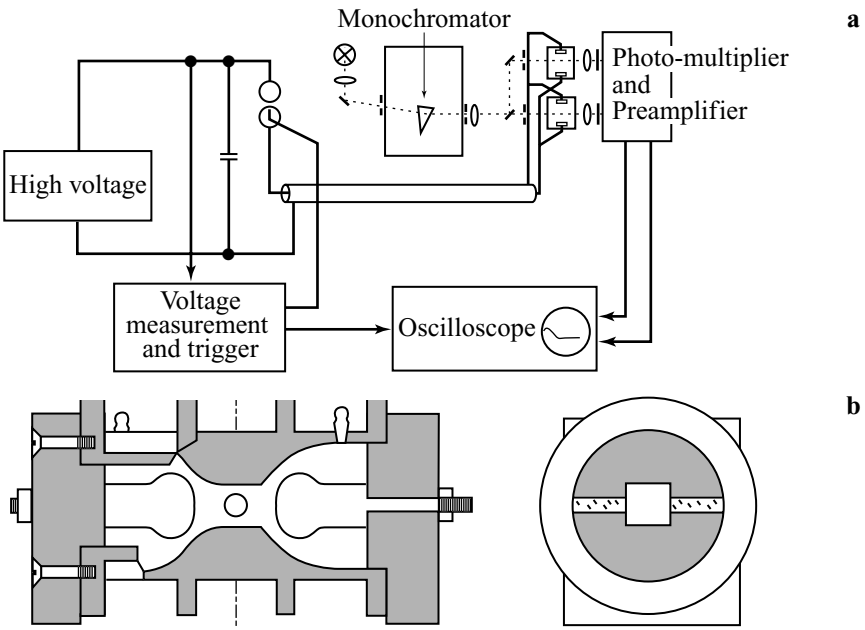


Fig. 6.9 Thermal relaxation apparatus
*(**a**) general scheme – (**b**) cell*

6.5.2.2. PRESSURE SHOCK

In the method of pressure shock (or **P-shock**), the perturbation is provoked by an abrupt change – either positive or negative – in pressure. The detection methods used are conductimetry or spectrophotometry. This technique gives time constants ranging from 10 to 5×10^{-5} s.

6.5.2.3. OTHER RELAXATION METHODS

The impulse from an electric field also provides a method of perturbation. In general, conductimetry is the detection technique used. These methods allow time-scales to be covered that range from 10^{-4} to 10^{-9} s and even lower.

The dielectric loss in intense electric fields constitutes another relaxation method which can achieve time constants of between 10^{-5} and 10^{-8} s.

The absorption and dispersion of sound can be used for studying certain systems. The time constants in these methods vary between 10^{-5} and 5×10^{-10} s. The detection methods make use of either the phenomenon of resonance, the reverber-ation of sound, or the diffraction of light at the sound gate.

6.5.3. ANALYSIS OF KINETIC DATA

If we consider the chemical processes, the previously described methods allow analysis of the kinetic data.

6.5.3.1. REACTIONS WITH N CONSECUTIVE STEPS

Let us first consider a reaction with n consecutives steps, which comprises n independent rate equations:

$$A \rightleftharpoons X_1 \rightleftharpoons X_2 \rightleftharpoons X_3 \rightleftharpoons \dots \rightleftharpoons X_n$$

The rate equations can be linearised to give equations of the form:

$$\frac{-\Delta C_i}{dt} = \sum_{j=1}^{n} a_{ij} \Delta C_j$$

in which the a_{ij} terms are functions of the specific constants for rate and concentration at equilibrium.

These coupled rate equations can be converted into a series of linear first-order equations having the form:

$$\frac{d\Delta y_i}{dt} = \frac{\Delta y_i}{\tau_i}$$

Δy_i being a linear combination of ΔC_i; the τ_i terms are relaxation times, functions of a_{ij}. The law for the change in concentration of individual species is written:

$$\Delta C_i = \sum_{j=1}^{n} A_{ij} e^{-t/\tau_i}$$

The n relaxation times are solutions of the determinant:

$$\begin{vmatrix} a_{11} - 1/\tau & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - 1/\tau & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - 1/\tau \end{vmatrix}$$

Thus the changes in concentration that result from the perturbation of equilibrium can be represented as the sum of the exponential terms:

$$\Delta C = A_1 e^{-t/\tau_1} + A_2 e^{-t/\tau_2} + \dots \quad 6$$

The set of relaxation times and their corresponding amplitudes form the relaxation spectrum. These values depend in a known way on the real concentrations and rate constants that determine the dynamics of the system. In principle, the equations could be derived for each kinetic mechanism. Conversely, the measurement of $1/\tau$ as a function of the concentration enables the determination of the individual rate constants and consequently leads to the elucidation of the mechanisms. The relaxation amplitudes are functions of the constants associated with the detection methods, for example, the molar extinction coefficients when absorption spectrophotometry is used for detection or the quantum yields if the detection method is fluorescence. It is not always straightforward to obtain values for the associated relaxation times and amplitudes from relaxation spectra. This is only simple when they are sufficiently separated in time.

One point needs to be emphasised, after the application of a transient perturbation, **the rate equations can only be solved if the perturbation is very weak** in such a way that the real concentration change introduced by this does not exceed a few percent of the absolute concentration of the reacting species. Although a weak perturbation leads to relatively small effects requiring a greater sensitivity in the detecting equipment, the resulting simplification of the mathematical analysis largely compensates for this disadvantage. In these conditions only, the rate equations corresponding to the kinetic laws for each individual step can be simplified by ignoring the products of the concentration change. This linearisation allows the use of matrix calculations to solve all the linear differential equations simultaneously. Thus for a step having a given order, the rate equation can be represented by the single relationship:

$$\frac{d\Delta C_i}{dt} + \frac{1}{\tau_i}\overline{\Delta C_i} = \frac{1}{\tau_i}\Delta C_i$$

For single reactions comprising only one elementary step, a single exponential relaxation is obtained. The relationship between $1/\tau$, the specific rate constants and the concentrations can be readily known.

6.5.3.2. FIRST-ORDER REACTION WITH A SINGLE STEP: ISOMERISATION

Let us now consider an isomerisation reaction such as:

$$A \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} B$$

The rate equation is:

$$\frac{dA}{dt} = -k_{12}A + k_{21}B$$

When the perturbation of the equilibrium is weak, the concentrations of A and B can be replaced by the final concentrations at the new equilibrium and the shift ΔA and ΔB from this final equilibrium, such that:

$$A = \bar{A} + \Delta A$$

$$B = \bar{B} + \Delta B$$

Thus:

$$\frac{d\left(\bar{A} + \Delta A\right)}{dt} = \frac{d(\Delta A)}{dt} = -k_{12}\left(\bar{A} + \Delta A\right) + k_{21}\left(\bar{B} + \Delta B\right)$$

From the law of mass action we may write:

$$\Delta A = -\Delta B \quad \text{or} \quad \Delta A + \Delta B = 0$$

The equilibrium relationship gives:

$$k_{12}\bar{A} = k_{21}\bar{B}$$

giving: $$\frac{d(\Delta A)}{dt} = -(k_{12} + k_{21})\Delta A$$

and by integrating: $$\Delta A_t = \Delta A_0 e^{-(k_{12}+k_{21})t}$$

ΔA_t is the change in the concentration of A at time t, ΔA_0 the total change in A between the initial and final equilibrium. Therefore, **the relaxation time:**

$$1/\tau = k_{12} + k_{21}$$

for a monomolecular, single-step process is independent of the concentrations of the species. It is equal to the sum of the individual rate constants of the forward and reverse reactions. The equilibrium constant $K_{eq} = k_{12}/k_{21}$ must be known in order to determine each of them. Consequently, relaxation methods do not dispense with the use of classical methods, they require and complete them.

6.5.3.3. BIMOLECULAR REACTION WITH ONE STEP

Let us consider the following bimolecular equilibrium:

$$A + B \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} C$$

The rate equation is written: $$\frac{dA}{dt} = -k_{12}AB + k_{21}C$$

As before, we have the relations:

$$A = \bar{A} + \Delta A$$

$$B = \bar{B} + \Delta B$$

$$C = \bar{C} + \Delta C$$

The equations for mass conservation are written:

$$\Delta A = \Delta B = -\Delta C$$

$$\Delta A = \Delta B$$

$$\Delta A + \Delta C = 0$$

which makes the existence of two equations of mass conservation appear more clearly. The equilibrium is written:

$$k_{12}\bar{A} \cdot \bar{B} = k_{21}\bar{C}$$

and by substituting in the rate equation and ignoring the terms in $\Delta A \, . \, \Delta B$, we obtain the equation:

$$\frac{d(\Delta A)}{dt} = -k_{21} + k_{12}\left(\bar{A} + \bar{B}\right)\Delta A$$

After integrating, this expression becomes:

$$\Delta A_t = \Delta A_0 e^{-t/[k_{21}+k_{12}(\bar{A}+\bar{B})]}$$

The time constant of a simple bimolecular reaction with one step depends on the sum and not the product of the concentrations of the reacting species at the final equilibrium; this results from the weak perturbation and the linearisation of the equation. Thus by plotting $1/\tau$ as a function of the sum of the concentrations $\left(\bar{A}+\bar{B}\right)$, we obtain k_{21} from the intercept of the linear plot with the vertical axis and k_{12} from its slope. Again the concentrations of A and B at the final equilibrium must be known. Here, yet again, the use of relaxation methods requires having studied the equilibrium beforehand.

6.5.3.4. BIMOLECULAR REACTION FOLLOWED BY ISOMERISATION

$$A + B \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} C \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} D$$

This type of reaction is interesting in biochemistry since it represents the first steps of an enzyme reaction, namely, the formation of the MICHAELIS complex followed by its isomerisation. The reacting components could be an enzyme and a poor substrate or a competitive inhibitor.

These two reactions are coupled *via* a common intermediate C; therefore the dependence of the concentrations cannot be given by the previously derived expressions. In principle, two relaxation times could be observed after perturbation of the equilibrium. If the rate constants have the same order of magnitude, the expressions obtained are complex functions of the rate constants and the concentrations. When the two elementary steps have very different time constants, the expression can be simplified.

Let us consider a bimolecular reaction that is much faster than a monomolecular reaction. After the perturbation we observe a relaxation corresponding to the bimolecular step which is not coupled to the subsequent step. The first fast relaxation time is therefore that of a simple bimolecular reaction:

$$1/\tau_1 = k_{21} + k_{12}\left(\bar{A}+\bar{B}\right)$$

The first reaction has already reached its new state of equilibrium whereas the relaxation of the second has not yet occurred. However, the relaxation of this second phase is coupled to the readjustment of the first equilibrium; the second step therefore corresponds to a change in the concentrations coupled to that of the first step. We have the rate equation:

$$\frac{dD}{dt} = k_{23}C - k_{32}D$$

as well as the relations:

$$A = \bar{A} + \Delta A$$
$$B = \bar{B} + \Delta B$$
$$C = \bar{C} + \Delta C$$
$$D = \bar{D} + \Delta D$$

and the conservation equation:

$$\Delta A + \Delta C + \Delta D = 0$$

The rate equation may be written:

$$\frac{d\left(\bar{D} + \Delta D\right)}{dt} = k_{23}\left(\bar{C} + \Delta C\right) - k_{32}\left(\bar{D} + \Delta D\right)$$

The equilibrium relationship is: $k_{23}\bar{C} = k_{32}\bar{D}$

The first fast equilibrium is always established during the course of the relaxation of the second. Thus we have:

$$\frac{(A)(B)}{(C)} = \frac{k_{21}}{k_{12}}$$

or:

$$k_{12}\left(\bar{A} + \Delta A\right)\left(\bar{B} + \Delta B\right) = k_{21}\left(\bar{C} + \Delta C\right)$$

By ignoring the product $\Delta A \,.\, \Delta B$ and taking into account the mass conservation expression: $\Delta A = \Delta B$, we obtain ΔA:

$$\Delta A = \frac{k_{21}\Delta C}{k_{12}\left(\bar{A} + \bar{B}\right)}$$

By substituting this expression in the equation for mass conservation, we obtain:

$$\Delta C\left[1 + \frac{k_{21}}{k_{12}\left(\bar{A} + \bar{B}\right)}\right] + \Delta D = 0$$

giving:

$$\frac{d\Delta D}{dt} = -k_{32}\Delta D - k_{23}\frac{k_{12}\left(\bar{A} + \bar{B}\right)}{k_{21} + k_{12}\left(\bar{A} + \bar{B}\right)}$$

by integrating:

$$\Delta D = \Delta D_0\, e^{-t/\tau_2}$$

with:

$$\frac{1}{\tau_2} = k_{32} + k_{23}\frac{k_{12}\left(\bar{A} + \bar{B}\right)}{k_{21} + k_{12}\left(\bar{A} + \bar{B}\right)}$$

Figure 6.10 below shows the changes in relaxation times with the concentration of the species concerned. It clearly shows that only for high concentrations of $\bar{A} + \bar{B}$ does the relaxation time become constant. From this hyperbolic curve we can obtain the constant k_{32} from its intercept with the vertical axis (extrapolated to zero concentration). The plateau corresponds to the sum of the constants $k_{23} + k_{32}$.

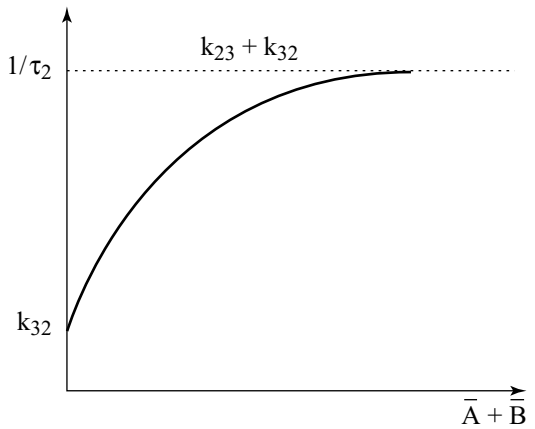


Fig. 6.10 Changes in $1/\tau_2$ as a function of $\bar{A} + \bar{B}$ in a second-order reaction followed by isomerisation

6.5.3.5. DIMERISATION EQUILIBRIUM

Dimerisation equilibrium is another example that applies to proteins, which in certain conditions can self-associate and in others, conversely, undergo dissociation into subunits. Let us consider the following equilibrium:

$$2\,M \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} D$$

The rate equation has the form:

$$-\frac{1}{2}\frac{d(M)}{dt} = \frac{d(D)}{dt} = -k_{12}(M)^2 + k_{21}(D)$$

As before, we have the relations:

$$M = \bar{M} + \Delta M$$
$$D = \bar{D} + \Delta D$$

and by substituting into the rate equation:

$$-\frac{1}{2}\frac{d\left(\bar{M} + \Delta M\right)}{dt} = -k_{12}\left(\bar{M} + \Delta M\right)^2 + k_{21}\left(\bar{D} + \Delta D\right)$$

The equation for mass conservation is:

$$\frac{\Delta M}{2} = -\Delta D$$

At equilibrium, we have: $k_{21}\bar{M}^2 = k_{21}\bar{D}$

and:
$$-\frac{1}{2}\frac{d\Delta M}{dt} = -2k_{12}\bar{M}\cdot\Delta M - k_{21}\frac{\Delta M}{2}$$

giving:
$$\frac{d\Delta M}{dt} = -\left(4k_{12}\bar{M} + k_{21}\right)\Delta M$$

and the relaxation time is given by the relationship:

$$1/\tau = 4k_{12}\bar{M} + k_{21}$$

6.5.3.6. ANALYSIS OF RELAXATION DATA

From these few simple examples, which illustrate well the use of relaxation methods, it is helpful to give some practical guidelines particularly concerning thermal relaxation, which is the most commonly used technique for studying biochemical reactions.

The number of relaxation times observed is a first indication about the chemical reactions studied, in particular about the number of reacting species and their related equilibria. This number is at most equal to the number of reacting species minus the number of conservation equations. Thus in the isomerisation equilibrium, we can count two species and one equation of conservation, giving a single relaxation time. Similarly, in a single-step bimolecular reaction, the number of species is three, with two equations of conservation and therefore a single relaxation time. In the example cited in Sect. 6.5.3.4, there are four species and two equations of conservation, and therefore two relaxation times in theory. However, these two times are only observable if they are sufficiently separated in time.

The determination of relaxation times and the study of their dependence on the concentrations of the system's constituents lead to defining the reaction mechanism itself. Thus it is easy to distinguish between an isomerisation, where $1/\tau$ is independent of the reactant concentrations, and a bimolecular process where this parameter varies linearly with the sum of the concentrations of the two components at equilibrium.

The analysis of the relaxation amplitudes can yield different types of information. Firstly, the change in amplitudes as a function of the concentration provides complementary information about the reaction scheme and its stoichiometry. The relaxation amplitudes are functions of the reaction enthalpy and characteristics related to the observed physical parameter. Thus in a bimolecular reaction:

$$A + B \rightleftharpoons C$$

the total amplitude of the phenomenon is:

$$\Delta P = (\Phi_A A + \Phi_B B + \Phi_C C)e^{-t/\tau}$$

Φ_A, Φ_B and Φ_C being the physical characteristics of the species. If the physical parameter is absorbance, we have:

$$\Delta P = (\varepsilon_A A + \varepsilon_B B + \varepsilon_C C)e^{-t/\tau}$$

ε_A, ε_B and ε_C are the molar absorbances of the species A, B and C, respectively. Thus knowledge of either the physical characteristics or the enthalpy of the reaction enables the determination of the other parameters from the values of the amplitudes.

6.5.3.7. EXAMPLE OF STUDYING AN ENZYMATIC REACTION BY MEANS OF THERMAL RELAXATION

In order to determine if the interaction of trypsin with a substrate, or a competitive inhibitor analogous to a substrate, is followed by an isomerisation step such as a conformational change in the protein, the trypsin-benzamidine interaction was studied. However, this competitive enzyme inhibitor absorbs light in a region of the spectrum poorly measurable by experiment, because the enzyme absorbs at the same wavelengths, making detection of the enzyme-inhibitor complex very difficult. We can however use another trick: proflavin, which is an acridine dye and also a competitive inhibitor of the enzyme, can be used as an indicator. In this study, two types of experiment were carried out. First, the study of the simple trypsin-proflavin equilibrium and then the binding of benzamidine in the presence of the indicator, proflavin.

Binding of proflavin to trypsin

For the equilibrium: $E + P \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} EP$, the reaction occurs in a single step.

We obtain a single relaxation time characteristic of a bimolecular reaction. It is a function of the concentrations of the reacting species:

$$1/\tau_1 = k_{12}\left(\bar{E} + \bar{P}\right) + k_{21}$$

Figure 6.11a indicates the changes in $1/\tau_1$ as a function of $\left(\bar{E} + \bar{P}\right)$; it shows a satisfactory linearity. We deduce from this the kinetic parameters of the reaction:

$$k_{12} = 7.6 \times 10^7 \text{ mol}^{-1} . \text{s}^{-1} \quad ; \quad k_{21} = 9\,500 \text{ s}^{-1}$$

with $K_p = 1.1 \times 10^{-4}$ M.

Binding of benzamidine to trypsin

We have the following group of reactions:

$$\begin{array}{ccc} P + E & \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} & EP \\ + & & \\ I & & \\ \Big\| k_{31} \quad k_{13} \Big\| & & \\ EI & & \end{array}$$

If each equilibrium is a simple bimolecular reaction, we should expect two relaxation times. Indeed, there are 5 molecular species and 3 equations of conservation as follows:

$$\Delta P + \Delta EP = 0$$

$$\Delta I + \Delta EI = 0$$

$$\Delta E + \Delta EP + \Delta EI = 0$$

However, as there are two coupled equilibria, we can only detect them if they are adequately separated in time. If the first equilibrium (EP) is fast relative to the second, the first relaxation time will be independent of the second equilibrium. Having independently studied this first equilibrium, it was noticeable that this is indeed the case. Taking into account the fact that the re-equilibration of proflavin and the trypsin-proflavin complex is quite fast so that the first system re-equilibrates independently of the second equilibrium, we have:

$$\frac{1}{\tau_2} = k_{31} + k_{13}\left(\bar{E} + \alpha\bar{I}\right)$$

with:
$$\alpha = \left(\bar{E} + K_p\right)\left(\bar{E} + \bar{P} + K_p\right)$$

and:
$$K_p = k_{21}/k_{12}$$

The plot of $1/\tau_2$ as a function of $\left(\bar{E} + \alpha\bar{I}\right)$ gives a straight line that is satisfactory enough to justify the hypothesis of a simple scheme (Fig. 6.11b) and permits determination of the system's constants:

$$k_{13} = 2.9\times10^{7}\ \text{mol}^{-1}.\text{s}^{-1} \quad ; \quad k_{31} = 600\ \text{s}^{-1} \quad ; \quad K_I = 1.2\times10^{5}\ \text{M}^{-1}$$

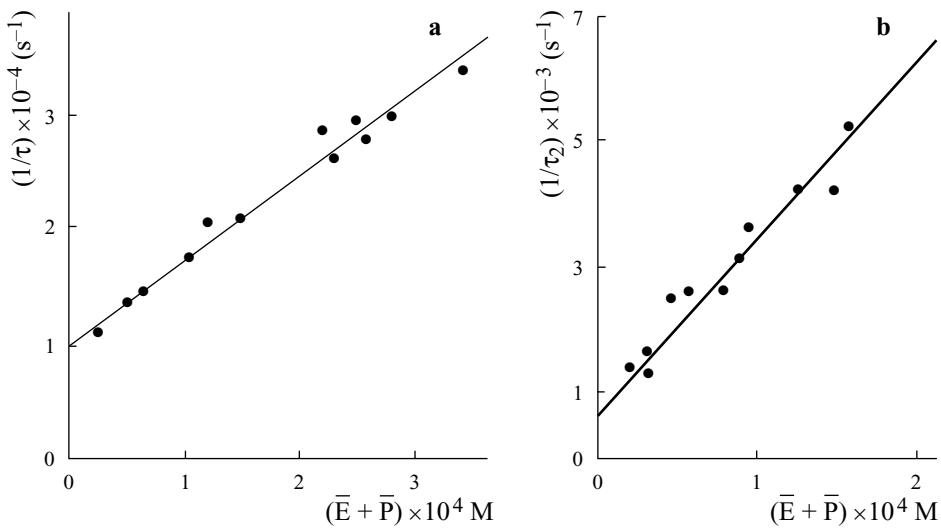


Fig. 6.11 Study of the trypsin-benzamidine interaction with the aid of an indicator, proflavin

*(**a**) changes in the inverse of the relaxation time as a function of the sum of the trypsin and proflavin concentrations at the new equilibrium $(\bar{E}+\bar{P})$, in the absence of benzamidine – (**b**) change in the inverse of the relaxation time as a function of $(\bar{E}+\bar{P})$ in the presence of benzamidine, with $\alpha = (\bar{E}+K_p)(\bar{E}+\bar{P}+K_p)$. These representations are characteristic of second-order kinetics*

6.6. STUDY OF ENZYMATIC REACTIONS AT LOW TEMPERATURES: CRYOENZYMOLOGY

As the different steps in an enzyme reaction are dependent on temperature, it is conceivable to slow them down to low temperatures in order to facilitate their study. The use of solvent mixtures that lower the freezing point of water has enabled examination of these reactions at temperatures as low as –80°C. These methods have led to the isolation, in stable form, of reaction intermediates that had never been amenable to study until that time, due to their very short lifetimes. While employing techniques at low temperatures, it is important to remember that the addition to the solvent of substances that lower the freezing point (e.g. glycerol, ethylene glycol and ethanol) may cause other effects. In particular, the influence of these substances on the viscosity or the dielectric constant may change considerably the structural and functional properties of the proteins studied (e.g. interactions between subunits, pK of the groups involved in interactions of the protein with its ligands or with itself even, surface exposed to solvent, etc.). In this way it was shown that the presence of such co-solvents stabilised one or other of the conformations involved in the allosteric properties of haemoglobin (BULONE et al., 1983), glycogen phosphorylase (DREYFUS et al., 1978) and of aspartate transcarbamylase (DREYFUS et al., 1984). Several cases of oligomeric enzyme dissociation at low temperature have in fact been reported.

Table 6.1 below indicates the change in dielectric constants of various water-ethylene glycol mixtures at different temperatures. It shows that water-solvent mixtures can be obtained for which the dielectric constant practically reaches 80, i.e. that of pure water at 20°C. The dielectric constant varies with temperature according to the relation:

$$\log D = a - bT$$

a and b are empirical coefficients and T the absolute temperature.

Table 6.1 Dielectric constants of ethylene glycol-water mixtures

(Reprinted from *Cryobiochemistry: an introduction,* DOUZOU P., 1977, Academic Press, London)

Solvent (%)	*Temperature (°C)*													*a*	*b*
	+20	+10	0	–10	–20	–30	–40	–50	–60	–70	–80	–90	–100	*(10^3)*	
0	80.4	84.2	88.1											1.945	2.00
10	77.7	81.4	85.3											1.931	2.02
20	75.1	78.4	82.5	86.9										1.916	2.15
30	72.0	75.7	79.5	84.0										1.912	2.20
40	68.1	72.1	76.3	80.2	84.4									1.880	2.30
50	64.5	68.4	72.4	76.5	80.7	85.0	89.3							1.860	2.35
60	61.1	64.6	67.9	72.0	76.3	80.8	85.3	90.1	95.7					1.832	2.41

Solvent (%)	*Temperature (°C)*													*a*	*b*
	+20	+10	0	–10	–20	–30	–40	–50	–60	–70	–80	–90	–100	*(10^3)*	
70	56.9	60.0	63.4	67.5	71.3	75.3	79.8	84.5	89.5	94.6	100.0	106.1	112.1	1.803	2.47
80	53.0	55.6	58.8	62.3	66.2	70.0	74.2	78.5	83.2	88.1	93.0			1.770	2.50
90	47.5	50.5	53.5	56.8	60.2	63.8	67.8	72.0						1.728	2.52
100	41.9	44.7	47.6	50.3										1.675	2.54

A certain number of systems have thus been studied with success, among which, horseradish peroxidase (DOUZOU et al., 1970), several serine proteases (FINK, 1973, 1974, 1976; FINK & ANGELIDES, 1975), β-galactosidase from *E. coli* (FINK & ANGELIDES, 1975) and several dehydrogenases (FINK & GEEVES, 1979).

6.7. STUDY OF ENZYMATIC REACTIONS UNDER HIGH PRESSURE

6.7.1. PRINCIPLE

During the last few years, numerous works based on the influence of pressure variations on the structure and properties of proteins have been undertaken. Initially destined for the study of structure-function relationships in enzymes from living organisms at atmospheric pressure, these methods today arouse additional interest after the discovery in 1978 of organisms living in deep-sea volcanic environments under pressures of up to 1 000 bars.

The influence of pressure on the structure and the catalytic and regulatory properties of enzymes is linked to LE CHATELIER's law (1888): "Any modification of one of the factors determining the state of a system at equilibrium provokes a shift of this equilibrium in the direction which tends to oppose the change in the factor considered." Thus, if a chemical reaction leads to a decrease in volume, it will be enhanced by an increase in pressure. If the reaction:

$$A + B \underset{}{\overset{K_{eq}}{\rightleftharpoons}} C + D$$

is accompanied by a change in volume, the pressure favours the direction of this reaction that leads to a reactant mixture occupying the smallest volume.

The volume change ΔV is linked to the equilibrium constant K_{eq} by the relationship:

$$\frac{d \ln K_{eq}}{dP} = -\frac{\Delta V}{RT}$$

at a constant temperature.

Thus, at a given temperature, by plotting the variation of 2.3 log K_{eq} as a function of the pressure P, we derive ΔV from the slope of the linear plot obtained. This slope can be positive or negative.

Similarly, for protonation equilibria:

$$RH \rightleftharpoons R + H^+$$

In most cases, the solution of the deprotonated form occupies a smaller volume, which results in particular from the phenomenon of electrostriction. This goes for ionic groups on proteins:

$$R—COO^- + H^+ \quad \text{and} \quad RNH_2 + H^+$$

whose ΔV values for the deprotonation are, respectively, 21.3 and 20 $cm^3 \cdot mol^{-1}$. Thus, the pressure will tend to diminish the ionisation pK of a protein's amino and carboxylic groups with any consequences that this may entail for its conformation. However, the protonation of the imidazole moiety of histidine residues is characterised by a very small increase in volume (1.1 $cm^3 \cdot mol^{-1}$).

The same considerations apply to the conformational changes in enzymes and more particularly, to allosteric transitions (see Chap. 13):

$$T \overset{L_0}{\rightleftharpoons} R$$

If the conformations R and T occupy different volumes, the pressure will affect the allosteric constant L_0 and will tend to stabilise the conformation occupying the smallest volume in solution. This has been established in the case of aspartate transcarbamylase (HERVÉ et al., 2004).

6.7.2. ACTIVATION VOLUME

Just as a chemical reaction is characterised by its free energy of activation, $\Delta G^{\neq}$, (see Chap. 1), so it also is by its activation volume, $\Delta V^{\neq}$. Thus, if the conversion of a substrate A into B involves the formation of a transition state $A^{\neq}$ whose formation is governed by the rate constant k_1 such that:

$$A \overset{k_1}{\rightleftharpoons} A^{\neq} \rightleftharpoons B$$

the activation volume $\Delta V^{\neq}$ is given by the relationship:

$$\Delta V^{\neq} = V_{A^{\neq}} - V_A$$

and is related to the rate constant k_1 by the relationship:

$$\frac{d \ln k_1}{dP} = -\frac{\Delta V^{\neq}}{RT} \qquad \text{at a constant temperature}$$

Thus, at a given temperature, the graphical plot of 2.3 log k_1 as a function of the pressure P provides the value of $\Delta V^{\neq}$.

6.7.3. EQUIPMENT

Figure 6.12 shows a diagram of an apparatus constructed to study enzymatic reactions at steady state. With this apparatus, the reactants can be injected, stirred and samples taken at the desired time intervals, without the pressure of the mixture being modified. It is composed of a cell in which the reaction mixture is subjected to the pressure exerted by piston P. Temperature and pressure sensors allow permanent control of these two parameters.

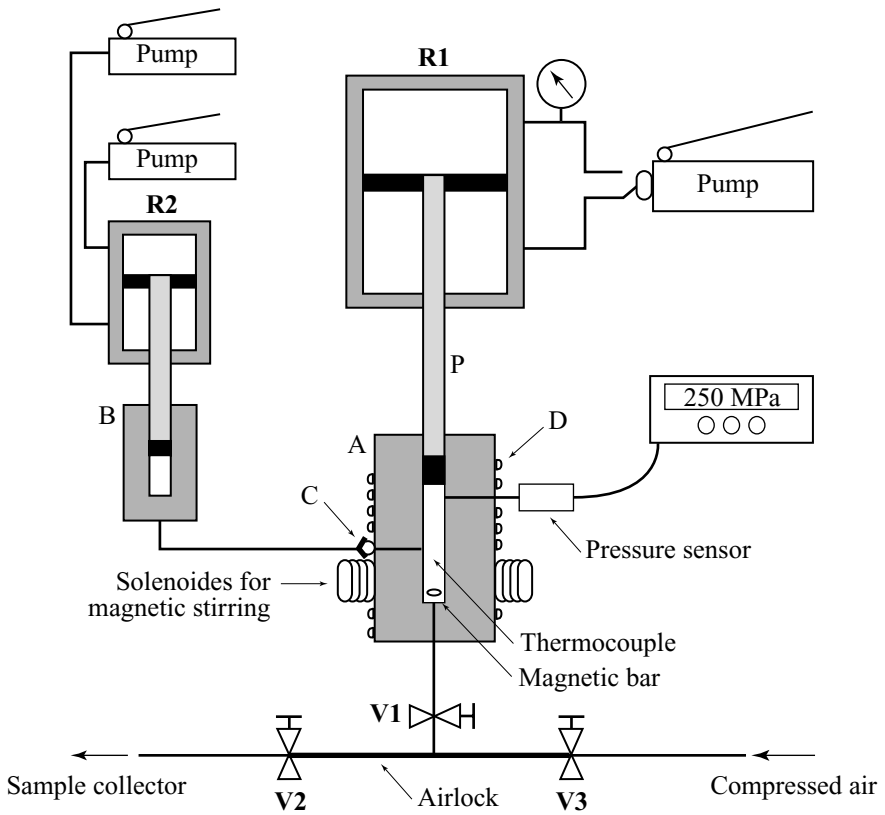


Fig. 6.12 Diagram of the apparatus developed to study enzyme kinetics at high pressure

A: incubation chamber; B: injection system; C: one-way valve; D: system for temperature regulation; P: main piston; R1 and R2: hydraulic presses; V1, V2, V3: valves.

(Reprinted from *Anal. Biochem.*, **187**, HUI BON HOA G. *et al.*, A reactor permitting injection and sampling for steady state studies of enzymatic reactions at high pressure: teste with aspartate transcarbamylase, 258. © (1990) with permission from Elsevier)

This cell is constructed in a non-magnetic alloy in such a way as to allow magnetic stirring of the sample. A one-way injection system is used to trigger the reaction. The samples are taken with the help of a small-volume locked chamber which,

contrary to what is schematically represented in the figure, is localised in the body of the cell (HUI BON HOA et al., 1990).

For kinetic studies in the pre-steady state, another type of device, enabling the use of the stopped-flow technique under pressure, was developed by HEREMANS et al. (1980) and BALNY et al. (1984). A diagram of this apparatus is displayed in Fig. 6.13. The stopped-flow module is inserted in the central volume of the high-pressure apparatus. After hermetic closing, the pressure is exerted by a pump and a transmission liquid injected by the inlet PI. When the system has reached the desired pressure, the device DM is used to trigger the stopped-flow system. The incident light beam OB is transmitted into the observation chamber by windows made from saphire and the signal obtained is then recorded.

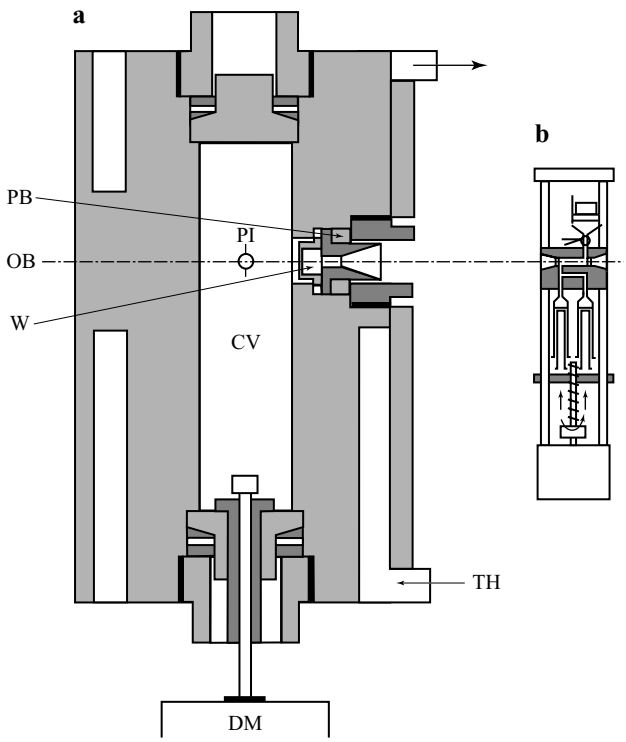


Fig. 6.13 Diagram of a stopped-flow apparatus under high pressure
(a) *schematic view of the high pressure cell in the central volume (CV) in which the stop-ped-flow device is inserted –* ***(b)*** *OB: incident light beam; W: saphire windows; TH: thermoregulation system; PI: inlet for the pressure transmission liquid; DM: trigger for the stopped-flow system.* (Reprinted from *Anal. Biochem.*, **139**, BALNY C. *et al.*, High-pressure stopped-flow spectrometry at low temperatures, 182. © (1984) with permission from Elsevier)

BIBLIOGRAPHY

BOOKS

BERNHARD S. –1968– *The structure and function of enzymes*, Benjamin Inc., New York.

CHANCE B., EISENHARD R.H., GIBSON Q.H. & LONGSBERG-HOLM K.K. –1965– *Rapid mixing and sampling techniques in biochemistry*, Acad. Press, New York.

DOUZOU P. –1977– *Cryobiochemistry: an introduction*, Acad. Press, London.

GUTFREUND H. –1965– *An introduction to the study of enzymes*, Blackwell Scientific Pub., Oxford.

HAMMES G.G. –1982– *Enzyme catalysis and regulation*, Acad. Press, New York.

LE MAIRE M., CHABAUD L. & HERVE G. –1989– *Un modèle d'étude: l'aspartate transcarbamylase. Théorie et guide d'expériences*, Masson, Paris.

ROUGHTON F.J.W. –1963– *Rates and mechanisms of reactions*, J. Wiley, New York.

GENERAL REVIEWS

BERGMEYER H.U. –1953– *Methods of enzyme analysis*, 10–13, Acad.Press, New York.

EIGEN M. & HAMMES G.G. –1963– Elementary steps in enzyme reactions (as studied by relaxation spectrometry), *Adv. in Enzym.* **25**, 1–38.

EIGEN M. & DE MAEYER L. –1963– Relaxation methods, in *Techniques in Organic chemistry*, Vol. VIII, 2nd ed., FRIESS S.L., LEWIS E.S. & WEISSBERGER A. eds, 806–1054, Interscience, New York.

ROUGHTON F.J.W. & CHANCE B. –1963– Rapid reactions, in *Techniques in Organic chemistry*, Vol. VIII, 2nd ed., FRIESS S.L., LEWIS E.S. & WEISSBERGER A. eds, 703–727, Interscience, New York.

YON J. –1963– Etude des réactions enzymatiques, in *Techniques de Laboratoire*, LOISELEUR éd., Tome II, 2e Partie, 936–959, Masson, Paris.

FINK A.L. –1977– *Accounts for Chem. Res.* **10**, 233–239.

FINK A.L. & GEEVES M.A. –1979– in *Methods Enzymol.* **63**, 336–370.

SPECIALISED ARTICLES

BALNY C., SALDANA J.L. & DAHAN N. –1984– *Anal. Biochem.* **139**, 178–189.

BARWELL C.J. & HESS B. –1970– *Hoppe Seyler's Physiol. Chem.* **351**, 1531–1536.

BULONE D., CAPANE A. & CORDONE L. –1983– *Biopolymers* **22**, 119–122.

DREYFUS M., FRIES J. & BUC H. –1978– *FEBS Lett.* **95**, 185–189.

DREYFUS M., VENDENBUNDER B., TAUC P. & HERVÉ G. –1984– *Biochemistry* **23**, 4852–4859.

EASTERBY J.S. –1973– *Biochim. Biophys. Acta* **293**, 552–558.

FINK A.L. –1973– *Biochemistry* **12**, 1736.

FINK A.L. –1974– *J. Biol. Chem.* **249**, 5027.

FINK A.L. –1976– *Biochemistry* **15**, 1580.

FINK A.L. & ANGELIDES K.J. –1975– *Biochem. Biophys. Res. Commun.* **64**, 701.

FINK A.L., & WILDI –1974– *J. Biol. Chem.* **249**, 6089.

GEKKO K. & TIMASHEFF S.N. –1981– *Biochemistry* **20**, 4667-4686.

GOLDMAN R. & KATCHALSKI E. –1971– *J. Theor. Biol.* **32**, 243–257.

HART W.M. –1970– *Mol. Pharmacol.* **6**, 31–40.

HEREMANS K., JUANWAERT J. & RIJKENBERG J. –1980– *Rev. Sci. Instr.* **51**, 806–808.

HERVÉ G., SCHMITT B. & SERRE V. –2004– *Cell. Mol. Biol.* **4**, 347–352.

HUI BON HOA G., HAMEL G., ELSE A., WEILL G. & HERVÉ G. –1990– *Anal. Biochem.* **187**, 258–261.

LEE B. & TIMASHEFF S.N. –1981– *J. Biol. Chem.* **256**, 7193–7202.

MC LURE W.R. –1969– *Biochemistry* **8**, 2782–2786.

PHILLIPS C.M., MIZUTAMI Y. & HOCHSTRASSER R.M. –1995– *Proc. Natl Acad. Sci. USA* **92**, 7292–7296.

REGENFUSS P., CLEGG R.M., FULWYLER M.J., BARRANTES F.J. & JOVIN T.M. –1985– *Rev. Sci. Instr.* **56**, 283–290.

STORER A.C. & CORNISH-BOWDEN A. –1974– *Biochem. J.* **141**, 205–209.

PART III

FORMATION AND STRUCTURE OF THE ACTIVE CENTRE OF ENZYMES

7 – ENZYME ORIGIN AND EVOLUTION

The resolution of the three-dimensional structures of proteins has led to an ever more precise understanding of the topology of active sites in a great number of enzymes and in some even, the molecular motion associated with it. The architectural complexity of currently known enzyme molecules prompts the following two questions:

- How have such structures, endowed with catalytic activity, progressively formed and been selected? A corollary to this question would be: why one particular structure and not another?
- How are they formed today in vivo from the information contained within a cell?

The origin of enzyme function practically overlaps with the origin of life. For a long time, it was thought that discussions relating to the origin of life arose more in lounge conversation than in serious science. It is true that this problem will always be of a speculative nature, as none of us was present at the beginning of time to witness it. However, experimentation has enabled simulation of some aspects of the formation of biological molecular structure; furthermore, much progress has been made in dating fossils with the help of isotopes allowing information to be obtained about the emergence of enzymes. We are therefore in a position to suggest plausible hypotheses about the appearance of these molecules, enzymatic activity and even the first cells. We are faced with diverse hypotheses, however, which we shall consider and discuss.

7.1. TIME-SCALE OF EVOLUTION

The Earth is estimated to be 4.6×10^9 years old. For a long time, authors interested in the origin of life assumed that the primitive atmosphere was most likely a very reducing atmosphere containing H_2O, CH_4, NH_3, H_2, N_2 and CO. Progressively, some of these molecules, in particular H_2, NH_3 and CH_4, disappeared and the atmosphere became less reducing. However, this hypothesis has had to be reconsidered. Current geochemical models increasingly indicate that the principal source of carbon was CO_2 and that there was very little or no methane. Ferrous ions would have played a key role in the photoreduction of CO_2 (BOROWSKA & MAUZERALL, 1988). However, all hypotheses about the origin of life assumed the absence of oxygen and the presence of SH_2 in the primitive atmosphere.

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_8, © Springer-Verlag Berlin Heidelberg 2010

Numerous observations suggested that the first organic molecules were formed from inorganic molecules in the atmosphere under the influence of UV radiation, electrical discharges or heat. The first organic molecules were dissolved in the sea, which at the time covered nearly all of the Earth's surface. This period of chemical evolution is estimated to have taken about 10^9 years, which represents a fifth of the Earth's history. The oldest organic material was discovered in schistose deposits in Fig Tree, South Africa. This material was composed of carbohydrates, porphyrins, purines and pyrimidines. The use of isotopes enabled the substance to be dated back 3.1×10^9 years. The same isoprenoid components that were found are present in cells today. Additionally, in these sediments analogues of cells were revealed, 0.6 mm in length. It is probable that the first living cells were heterotrophic anaerobic organisms capable of using dissolved organic substances from the sea. When these substances became scarce the only organisms to survive were those able to use simple carbon components. It is believed that the first photosynthetic organisms, blue algae, appeared a little later. Fossils of blue algae have been discovered in the same schistose sediments from South Africa and dated to 3.1×10^9 years ago. Current understanding suggests that until their appearance, there was little or no oxygen in the atmosphere. A further 100 thousand million years were needed before the development of anaerobic vertebrates and only in the last two thousand million years did *Homo sapiens* make its appearance on Earth. Table 7.1 illustrates the time-scale of evolution.

Table 7.1 Time-scale of evolution

Time	*Geological period*	*Approximate origin*	
-2×10^6	Phanerozoic	Man Mammals, birds Terrestrial plants Fish	
-5.6×10^6	Proterozoic	Invertebrates Multi-cellular organisms Eukaryotic cells	*Biological evolution*
-10^9	Archaea	Aerobic bacteria First fossils First photosynthetic prokaryotes	
-4×10^9		First cells	*Chemical evolution*
-4.4 to -4.5×10^9	Formation of the Earth		
-15×10^9		*Big Bang*	

The first theories about the origin of life were based on the hypothesis of a highly reducing primitive atmosphere, and from this an entire prebiotic chemistry was developed aiming to simulate the composition of the **'*original soup*'** of precursor

molecules consisting of proteins and nucleic acids. The experiments began with methane, which produces cyanide, this in turn being the precursor of biomolecules. The **"original soup"** theory was strongly criticised using thermodynamic, chemical and geochemical arguments. An alternative was proposed in 1988 by WÄCHTERSHÄUSER, the **theory of surface metabolism,** which favours the notion that triose phosphates are at the origin of metabolism.

7.2. PREBIOTIC CHEMISTRY ACCORDING TO THE "ORIGINAL SOUP" HYPOTHESIS

7.2.1. FORMATION OF SOME SIMPLE ORGANIC MOLECULES

Laboratory experiments were destined to simulate prebiotic conditions. In particular, they were carried out in a reducing atmosphere. The energy sources used were either ionising or ultraviolet radiation, electrical discharges or indeed heat (temperatures were probably higher near to volcanic sites). To these experiments we associate the names OPARIN, MILLER, FOX and ORGEL. Generally, in this type of experiment an electrical discharge is applied to a mixture of NH_3, CH_4, H_2 and H_2O. As early as 1920, OPARIN had suggested that the first organic molecules likely to generate more complex biological molecules were formed under these conditions. MILLER, in 1953, performed experiments of this type and observed the formation of biochemical components; in particular he obtained: glycine, alanine, sarcosine, β-alanine, γ-amino butyric acid, N-methylamine, and aspartic, glutamic, iminodiacetic, formic, acetic, glycolic, lactic, β-hydroxybutyric and succinic acids, urea and methylurea. A series of reactions generated these molecules; among the intermediates, cyanide and nitriles played a critical role. Here are the principal chemical reactions leading to the formation of alanine when an electrical discharge is applied:

$$CH_4 + NH_3 \longrightarrow HCN + 3H_2$$

$$C_2H_4 + HCN \longrightarrow CH_3CH_2CN \quad \text{(nitrile)}$$

$$CH_3CH_2CN \longrightarrow CH_3CH_2COOH \quad \text{(propionic acid)}$$

$$CH_3CHOHCN \longrightarrow CH_3CHNH_2CN \quad \text{(aminopropionitrile)}$$

$$CH_3CHNH_2CN \longrightarrow CH_3CHNH_2COOH \quad \text{(alanine)}$$

Some studies have shown that cyanide is at the basis of the formation of pyrimidines and purines, but also amino acids. The following scheme (from LEHNINGER) summarises the principal reactions that cyanide may undergo. It is interesting to notice that the elementary molecules obtained during experiments simulating prebiotic conditions are the same as those found in ancient rocks and sediments and also in certain meteorites.

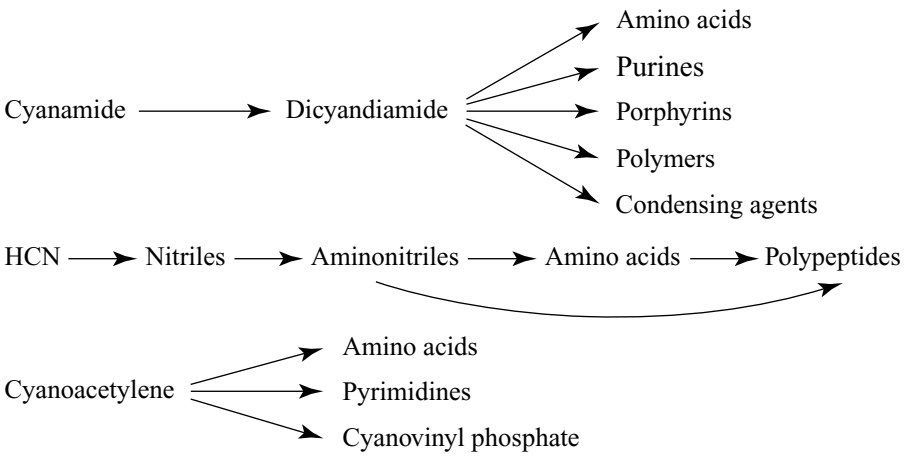


Table 7.2 below gives an idea of the sources of energy existing on the Earth's surface and their respective values.

Table 7.2 Sources of energy, average values at the Earth's surface

(Reprinted from *Biochemistry*, LEHNINGER A.L., chapter 24.
© (1972) with permission from W.H. Freeman and Company/Worth Publishers)

Source	*Calories/cm^2/year*
Solar radiation	260 000
Ultraviolet light	660
Electrical discharges	4
Radioactivity	0.8
Volcanic activity	0.13
Cosmic radiation	0.0015

The simple molecules thus formed can be the dominant products because their formation takes place along favourable energy pathways or because they are more stable than other products under defined conditions.

7.2.2. FORMATION OF MACROMOLECULES

The following step is the condensation reaction that leads to the formation of oligomers and polymers. During condensation water molecules are liberated. However, the bonds formed are hydrolysable; the peptide and glycosidic bonds are unstable in aqueous solution under prebiotic conditions. They tend to be hydrolysed, therefore, only a small number of molecules would succeed in being formed. These molecules can appear in anhydrous conditions, for example, at temperatures higher than the boiling point of water. FOX thus carried out the polymerisation of amino acids. The role played by ATP in the formation of peptide bonds during protein biosynthesis led to the suggestion that polyphosphates might be involved in the process of condensation in prebiotic conditions. Substances such as carbodiimides could have played an analogous role.

7.2.2.1. ABIOTIC FORMATION OF POLYPEPTIDES

FOX and other investigators obtained proteinoids from a mixture of amino acids either by heating to 170°C for several hours, or by a more gentle heat treatment (50–60°C) in the presence of polyphosphates for a longer time.

Starting from a mixture of all amino acids he obtained polymers of high molecular weight able to exceed 10 000 D and which did not form by chance. There was a mix of acidic, basic and neutral polymers produced which could be separated by chromatography. Besides, these polymers possessed properties that are common to all proteins such as precipitation at the isoelectric point, "salting in" and "salting out". Similarly, the incubation of a mixture of aminoacyl adenylates at pH 9 led to the formation of proteinoids with a high molecular weight; this reaction was carried out at mild temperatures. It is possible to synthesise proteinoids successfully by the polymerisation of aminoacetonitriles, suggesting that such a reaction could be produced under prebiotic conditions.

Some proteinoids thus appearing in abiotic conditions manifest a weak though significant biological activity. For example, certain proteinoids possess a **catalytic activity.** Their presence accelerates the hydrolysis of paranitrophenyl acetate, or increases the rate of decarboxylation of pyruvic acid. It is true that even imidazole is able to catalyse the hydrolysis of paranitrophenyl acetate. Furthermore, the activities of these proteinoids present pH profiles comparable to those observed in enzymatic reactions and obey MICHAELIS' law. Of course, these activities are much weaker than those of the corresponding enzyme (proteases or pyruvate decarboxylase). The proteinoids having an activity have been pointed out, for instance, an MSH (melanocyte-stimulating hormone) activity associated with a proteinoid has been observed.

7.2.2.2. ABIOTIC FORMATION OF NUCLEOTIDES AND NUCLEIC ACIDS

The experiments destined to recreate prebiotic conditions also ended with the formation of nucleotides from their corresponding bases (adenosine, deoxyadenosine). PONNAMPERUMA and co-workers managed to simulate the formation of AMP, ADP and ATP by heating or irradiating adenine and ribose in the presence of a phosphorylating agent. The condensation of mononucleotides was achieved by heating to 50–65°C in the presence of polyphosphoric acid. In abiotic conditions the 2′-5′ bond is formed predominantly; the 3′-5′ bond seems to be more difficult to create. Possibly the most interesting experiments were those carried out in the presence of template molecules, such as the polynucleotides from WATSON and CRICK's model of the double helix. In this way, the specificity in the pairing of bases, cytosine-adenine and uracil-guanine, was reproduced. MILES and TS'O showed that mononucleotides have a tendency to associate preferentially according to this complementarity and that a helical structure is formed along the template strand in which the nucleotides are covalently linked. ORGEL continued with the same type of experiment, but added a carbodiimide as a condensing agent and, using polyuridylic acid (poly U) as a template and AMP, he obtained poly A. Even under these conditions the polynucleotides form with 2′-5′ bonds.

Some amino acids are liable to associate to nucleotides with a certain selectivity. Poly-L-lysine and poly-L-arginine readily associate to guanylic and adenylic acids.

Due to this property, it was suggested that an α helix composed of a poly-amino acid could bind to adenylic acid and form helical structures in which a residue R would correspond to three nucleotides, thus foreshadowing the genetic code.

7.2.3. DISCUSSION OF THE NATURE OF THE FIRST BIOLOGICAL MOLECULES

Due to the highly speculative nature of the subject, it is not surprising that diverse theories exist all seeking to explain how life appeared on Earth. There is considerable discrepancy in the various suggested explanations, covering the nature of the first molecules and even the mechanisms of life's appearance. Two contrasting theories propose that life began either with proteins or with nucleic acids.

The first hypothesis, attributable to OPARIN and FOX, suggested that life started without nucleic acids, with the formation of droplets called ***protobionts*****:** microspheres containing proteins and possessing an elementary metabolism sometimes limited to a single reaction. According to OPARIN, the first cells existed when a membrane formed around a few macromolecules including catalytically active proteins. In OPARIN's view, the development of the genetic machinery was a late event in evolution. A cellular phase would have been formed from the primitive soup by coacervation. In the protobiont model, OPARIN and co-workers attempted an experiment with systems forming coacervates. Concentrated polymer solutions (5–50%) of polypeptides and polysaccharides in aqueous media were able to form visible droplets under the microscope; they had volumes of 10^{-8} to 10^{-6} cm^3. One of these coacervates containing glycogen phosphorylase was immersed in a solution of glucose-1-phosphate and was able to produce starch. When amylase was added to the system, maltose was formed. Similarly, a coacervate containing NADH dehydrogenase could reduce a solution of an oxidised dye (Fig. 7.1).

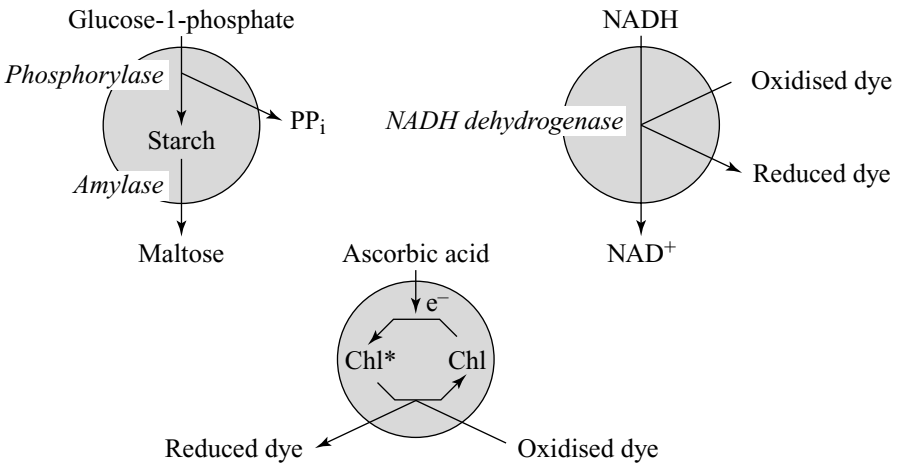


Fig. 7.1 Enzymatic activities within coacervates

FOX obtained microspheres by the progressive cooling of a hot and concentrated solution of proteinoids. When the pH was suitably adjusted, a bilayered structure analogous to a membrane was formed in the absence of lipids. These microspheres were

about 2HM thick and capable of dividing both in the presence of $MgCl_2$ and following a change in pH. The phenomenon of budding, similar to that observed with yeast, was sometimes produced. It is also thought that the action of wind at the sea surface helped to bring about membrane formation (lipid bilayers) which, upon closing up again, encapsulated proteins inside the vesicles (Fig. 7.2).

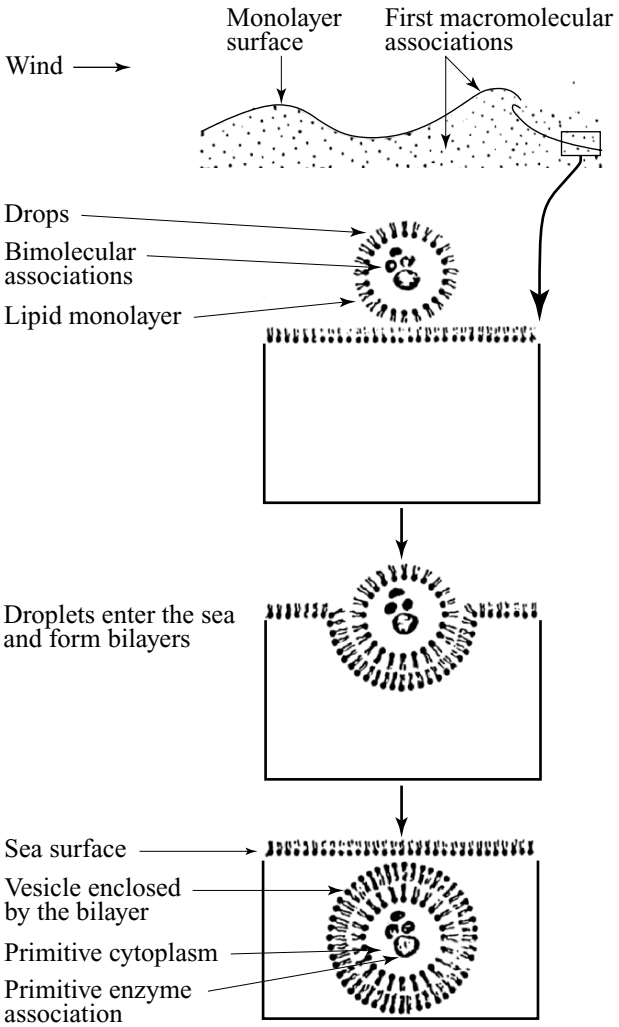


Fig. 7.2 Hypothetical mechanism for the formation of vesicles that encapsulate proteins from the sea surface

The stability of these vesicles is maximal around a particular size. The difference between the changes in internal volume and those at the external surface led them to divide like living cells by a purely physical phenomenon resulting from the variation in surface tension. It is possible to imagine that late in evolution the vesicles containing active molecules (enzymes and catalysts) and those containing information molecules (nucleic acids) fused, giving rise to structures similar to living cells and

capable of dividing. Whatever the case, without a genetic machinery to conserve the information, these first molecules would not have been able to replicate themselves.

Over time, the genetic hypothesis gave way to the notion of genetic machinery. MILLER, in 1929, proposed that life began with the abiotic formation of one or more genes. He considered that the minimum properties of a living organism, metabolism and reproduction, are potentially present within genes; when these are placed in an adequate molecular environment at the heart of a cell, they can give rise to a daughter cell. The genetic theory, which remained for a long time incompletely developed, was taken up again with the advent of genetic knowledge by various authors including CRICK, ORGEL and HOROWITZ. Membrane formation and the appearance of catalysts are considered to be late events in this paradigm. The genetic hypothesis was supported by the capacity of viruses to self-replicate and by modern developments in molecular genetics. CRICK and ORGEL suggested that the origin of life is allied to the origin of ribosomes, tRNA and the genetic code (tRNA anticodon triplets). The partisans of the genetic hypothesis postulated that before the appearance of activating enzymes the first forms of tRNA and rRNA filled one of the functions of enzyme molecules, namely, they provided specific binding sites for amino acids. The association of amino acids to tRNA in the initial conditions would have been non-covalent. Only later would ribosomes have acquired the ability to synthesise an enzyme capable of catalysing the formation of peptide bonds. In line with this hypothesis, it is noteworthy that RNA (and not DNA) possesses a tertiary structure; DNA supposedly appeared much later in evolution. As we shall see further on, the discovery in 1986 by ZAUG and CHECH that certain RNAs have a catalytic activity has reinforced the genetic hypothesis. **The capacity to self-organise into a three-dimensional structure seems to be vital for the origin of life**. Whatever the nature of the first biological molecules that appeared in the origin, proteins or nucleic acids (and here we find ourselves faced with the chicken-and-egg problem), the systems have subsequently evolved in response to environmental changes.

It is clear that, as interesting as these simulations of prebiotic chemistry are, they all build on the preconceived idea of an "original soup" created under strongly reducing conditions. This theory has been widely criticised by a number of authors, in particular CAIRNS, SMITH, WÄCHTERSHÄUSER and DANCHIN, who considered that the formation of biological macromolecules from an "original soup" is highly improbable from a thermodynamic, chemical and geochemical viewpoint. Furthermore, among the many molecules thus formed, as many "poisons" as "viable molecules" would have been present. The problem that seemed important for these authors is that of selectivity and selection during the formation of the first biological molecules. Also, as an alternative to the preceding hypotheses, WÄCHTERSHÄUSER proposed a theory of surface metabolism at life's origin. ◢

7.3. THEORY OF SURFACE METABOLISM

The theory of surface metabolism suggested by WÄCHTERSHÄUSER (1988) arises from the hypothesis that life in the beginning was autotrophic and consisted of autocatalytic metabolism confined to an organic bilayer. Anionic "metabolists"

were linked to positively charged surfaces such as pyrites, at the interface with hot water. Their association at the surface had to have been strong; the surface must have possessed positive charges due, for example, to polyvalent metals (Mg^{++}, Ca^{++}, Fe^{++}, Mn^{++} and Zn^{++}) and capable of forming insoluble salts. The organic constituents had to have been anionic possessing at least two negative charges in order to bind strongly enough to the surface. Thus, WÄCHTERSHÄUSER suggested that all polyanionic constituents from ancient metabolic pathways are ancient surface metabolites; for example, surface binding may have been an ancient function of phosphate groups. From a thermodynamic point of view, the reactions taking place on a surface are favoured by a positive entropy value which can compensate for the negative ΔH values that are otherwise unfavourable for these reactions. Surface metabolism therefore favours the formation of large molecular structures from less reactive groups with respect to the system in solution. For example, phosphotriose molecules cannot form intermolecular hemiacetal bonds in aqueous solution, but theory predicts that this would be possible on a surface. Figure 7.3 shows the condensation of glyceraldehyde phosphate and dihydroxyacetone phosphate to form structures named "phosphotriboses" by the author.

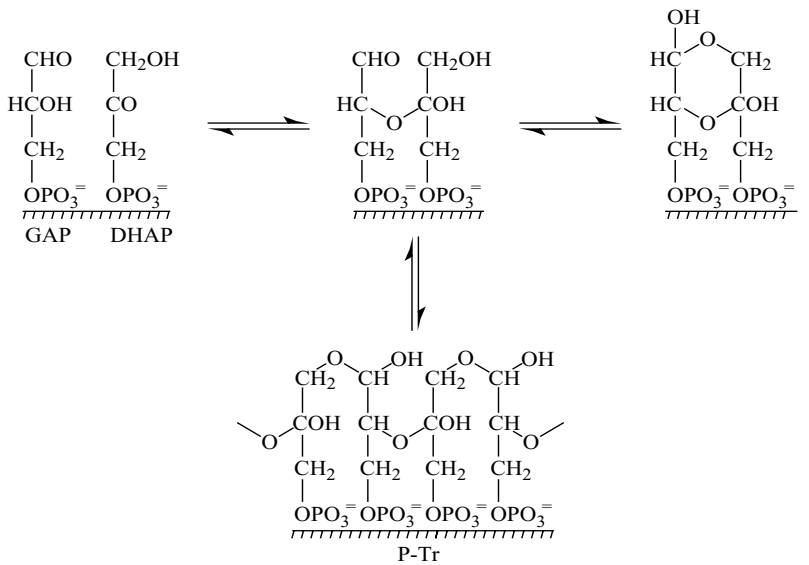


Fig. 7.3 Formation of phosphotriose (P-Tr) from glyceraldehyde phosphate (GAP) and dihydroxyacetone phosphate (DHAP) linked to a positively charged surface

(From *Microbiol. Mol. Biol. R.*, 1988, **52**, WÄCHTERHAÜSER G., 452, reproduced with permission from American Society for Microbiology)

These structures tightly bound to the surface are highly stable. The author suggested that phosphorylated sugars are the precursors to nucleic acids and coenzymes comprising a purine base. With similar reasoning, polypeptide formation from amino acids such as aspartic and glutamic acids or phosphoserine is less unfavourable on a surface than in solution. Surface metabolism offers several advantages,

not only thermodynamic, but also kinetic since it is a quasi-intramolecular process and therefore accelerates reaction rates; furthermore, it enables a high selectivity. It requires high temperatures, which were likely to have prevailed in primitive conditions.

7.3.1. AUTOTROPHIC SURFACE "METABOLISTS"

Surface metabolism gives rise to several types of transfer reaction, which are not observed in solution and which have been preserved in part in existing metabolic pathways. Small molecules like CO_2, HCOOH, CH_2O, CH_3COOH, CH_3CHO, $CH_2OH–CHO$, NH_3^+ and H^+ can be transferred from one molecule to another linked at the surface. This explains the formation of coenzymes such as pyridoxal phosphate, thiamine pyrophosphate, NAD^+ and haems. A sort of self-sufficient "life" would have established itself on a single-layered surface supporting the absorption of nutrients such as CO_2 and autocatalytic production of constituents, which would have spread out over the available surface.

This theory assumes that the first reactions were oxidoreduction reactions, in which a plausible electron source would have come from the formation of iron pyrite:

$$Fe^{++} + 2H_2S \longrightarrow FeS_2 + 4H^+ + 2e^-$$

Iron pyrite is able to reduce CO_2, CO or COO^- as shown in Fig. 7.4, via the formation of a thio acid.

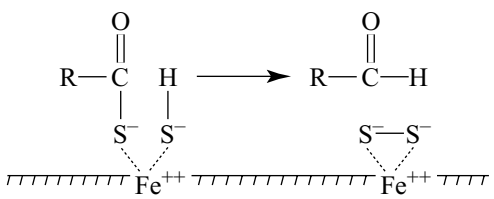


Fig. 7.4 Mechanism for the reduction of a carboxylic acid by iron pyrite via the formation of a thio acid
(From *Microbiol. Mol. Biol. R.*, 1988, **52**, WÄCHTERHAÜSER G., 466, reproduced with permission from American Society for Microbiology)

WÄCHTERSHÄUSER considered that this suggestion was in agreement with geochemistry. Indeed SH_2 was, and still is, abundant in the emissions of the Earth; besides, ferrous irons are ubiquitous and have been found in ancient sediments. Pyrite formation is anaerobic, which is consistent with the absence of oxygen in prebiotic conditions. Thereafter, nitrogen and phosphorous would have participated in this surface metabolism; C–N bonds would have been formed and ammonia, being insoluble in water, would have been instantly eliminated. The universal implication of ferredoxins and other iron-sulphur proteins in the electron transport chain and additionally the central role of cysteine residues and coenzyme A in numerous metabolic processes are probably relics from this primitive metabolism. It is remarkable that, even today, ferredoxins are short anionic polypeptides and their ancestral sequences contain only 11 of the 20 amino acids and none of the more complex amino acids that appeared later in evolution, namely: Met, His, Trp, Tyr, Phe, Leu, Thr, Lys and Arg. This signifies that the ancestral sequence of ferredoxins would

have pre-existed in the translation process. SH_2 would have played a double role, that of electron source and of nucleophilic agent, functions that would have become subsequently separated. The nucleophilic function would have been ensured by the cysteinyl groups on peptides linked to the surface, the activation of the carboxylate would have taken place via the formation of a thio acid generating a variety of peptides. Afterwards, a family of components would have been formed with acetyl-CoA.

Evolution would have progressed by **innovation and selection.** New metabolic cycles would have emerged by leaps. It has often been suggested that certain current coenzymes might be the vestiges of a pre-enzymatic metabolism. Many coenzymes are polyanionic, others like vitamin B12 and the quinones are derivatives of polyanionic biosynthetic precursors. The theory of surface metabolism suggests that coenzymes bound to the surface are not only catalytic for a class of reactions, but also **autocatalytic for their own synthesis from components bound to the surface**. Thus, the synthetic pathway for nicotinamide coenzymes would have involved the anaerobic formation of quinolinic acid from dihydroxyacetone phosphate and aspartic acid in the presence of an oxidant. Later, quinolinic acid would have been ribosylated by phosphoribosyl pyrophosphate (PRPP) to form $NADP^+$, as illustrated in Fig. 7.5.

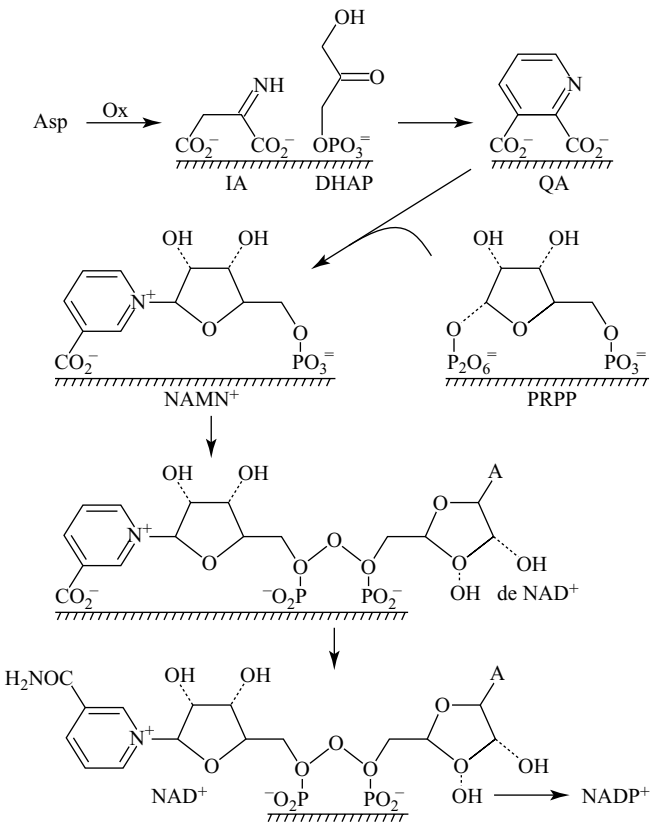


Fig. 7.5 Example of the anaerobic formation of coenzymes by surface metabolism *Quinolinic acid (QA) is formed from iminoaspartate (IA) and dihydroxyacetone phosphate (DHAP) and, in the presence of PRPP, gives rise to nicotinic acid mononucleotide ($NAMN^+$) and then to deamido-NAD^+ (from NAD^+) which can be phosphorylated to give $NADP^+$.*
(From *Microbiol. Mol. Biol. R.*, 1988, **52**, WÄCHTERHAÜSER G., 458, reproduced with permission from American Society for Microbiology)

Thiamine pyrophosphate would have been formed in a similar manner. It is clear that no surface metabolism has been found in nature. Furthermore, WÄCHTERSHÄUSER considered that the process has disappeared and proposed a phylogenetic model, which relates step by step the hypothetical precursor to systems currently existing.

7.3.2. THE CHANGE TOWARDS CELLULAR METABOLISM

A change in the environment could cause a sudden change in one organism, which may then acquire transmissible characteristics, whereas others may perish. The cellular revolution and genetic control profoundly modified evolutionary mechanisms. On an open surface, metabolism and evolution are identical. The change towards cellular metabolism marked a separation between *ontogenesis* and *phylogenesis*. On an open surface, all processes involving a detachment are eliminations. The appearance of closed cells represents an entirely different means of selection with their metabolic control systems. At this stage, evolution brought about a dichotomy between the organism and the environment.

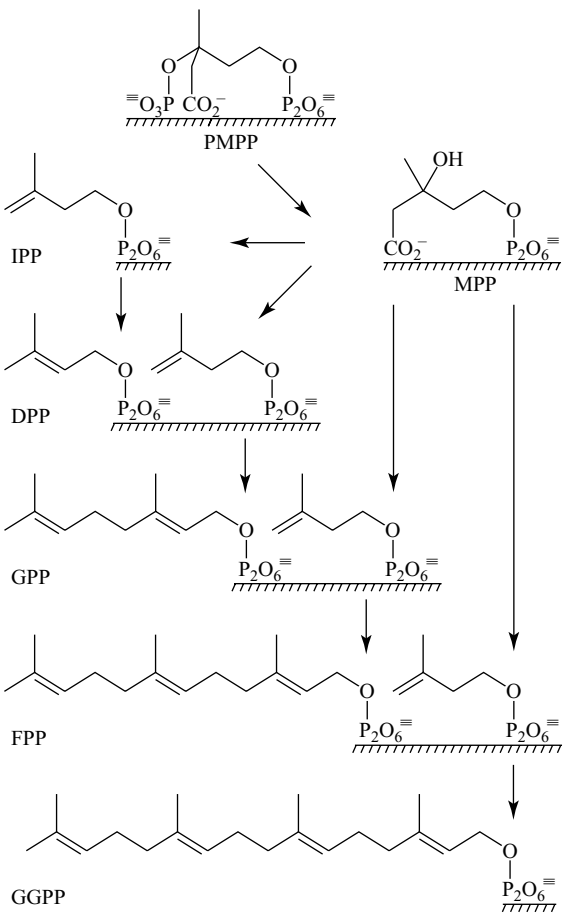


Fig. 7.6 Mechanism of stepwise extension of an isoprenoid lipid chain fixed to a surface, from 3-phosphomevalonate 5-pyrophosphate (PMPP), via mevalonic acid pyrophosphate (MPP), isopentenyl pyrophosphate (IPP), dimethylallyl pyrophosphate (DMPP), geranyl pyrophosphate (GPP), farnesyl pyrophosphate (FPP) and geranyl geranyl pyrophosphate (GGPP)
(From *Microbiol. Mol. Biol. R.*, 1988, **52**, WÄCHTERHAÜSER G., 463, reproduced with permission from American Society for Microbiology)

The appearance of lipids and membranes therefore constitutes the next step in primitive metabolism. WÄCHTERSHÄUSER assumed that surface metabolites produced isoprenoid lipids (Fig. 7.6 opposite) that remained attached to the surface and then accumulated as regular membranes.

The hydrophobic environment created by the accumulation of lipids had the effect of shifting the equilibria towards the formation of large molecules (polypeptides and nucleic acids). Lipid membranes fixed to the surface became detached to form *semi-cellular structures* still retaining a mineral support and leading to a cytosolic metabolism (Fig. 7.7). The change would have been helped by the emergence of reactions enabling the elimination of charged groups bound to the surface.

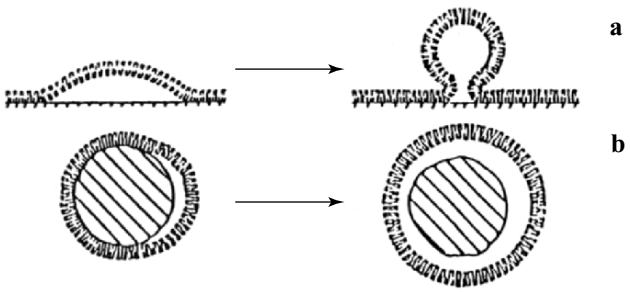


Fig. 7.7 Formation of primitive cells
(a) by constriction – (b) by circular detachment and mineral inclusion
(From *Microbiol. Mol. Biol. R.*, 1988, **52**, WÄCHTERHAÜSER G., 458,
reproduced with permission from American Society for Microbiology)

Evolution might have proceeded therefore by surface metabolism, semi-cellular metabolism and finally cellular metabolism. The genetic apparatus, enzymes, membrane pumps and electron transporters would have been established before the appearance of actual cells lacking a mineral support. WÄCHTERSHÄUSER believed that his theory was consistent with the emergence of three kingdoms in cellular life as proposed by WOESE: archaebacteria, eubacteria and eukaryotes. Isoprenoid lipids are still found in archaebacterial membranes. Other biological membranes arose from the metabolic pathway of fatty acids and would have been formed by enzymatic action.

With cellularisation, certain peptide precursors would have been incorporated into enzyme primary structures and then refolded into a three-dimensional structure. Coenzyme A would have become a true coenzyme by forming a covalent bond with a protein serine residue and thus becoming acyl carrier protein.

During evolution electron transporters bound to the membrane, tetrapyroles, seem to have played an essential role early on with the formation of haems, whereas bacteriochlorophyll would only have appeared much later on; it only exists in eubacteria and chloroplasts.

7.3.3. EVOLUTION OF THE GENETIC APPARATUS

The proposed model hypothesised an increasingly complex evolution of enzymes having a three-dimensional structure and replication machinery, as well as capable of translating nucleic acids – a process also suggested to have evolved in stages. Phosphotrioses were to play an important role; they are able to form polymeric structures, phosphotriboses, by binding carbon after activation by a thio acid. Purines would have been formed first, in the initial phase of surface metabolism, and pyrimidines last during the semi-cellular phase of evolution. Figure 7.8 opposite shows the hypothetical pathway of purine formation on a surface, which implicates the participation of imidazole. The structures formed from phosphotrioses and imidazole bound to the surface have been termed *tribonucleic acids* (TNA). Thus, a variety of purine-containing TNA structures would have been produced with novel structural properties and ultimately leading to complementary base-pairing.

Figure 7.9a below gives an example of a structure with two anti-parallel strands linked to the same surface. The bases would have appeared in the following order: U, A, G, C. The phosphoribose moieties produced by transfer of C2 and then a rearrangement would have formed glycosidic bonds with the purine bases by the same mechanism as the phosphotriboses.

The theory of surface metabolism assumed that translation would have preceded replication. Ribonucleosides would have associated with the TNA by base-pairing. These constituents, coupling to amino acids via ester linkages to the ribose moiety, could catalyse the formation of peptide bonds. From the outset, TNA molecules would therefore have catalysed the formation of peptide bonds without there being translation. The first translation process would likely have been autocatalytic, with TNA and ribonucleotides ensuring the synthesis of their own bases. TNA would have grown by extension and sequence changes brought about by modification of the purines. With the appearance of phosphoribosides, a new type of nucleic acid would have been formed by polymerisation, the TNA serving to position the elements in an orientation favourable to the formation of phosphodiester bonds. The TNA-TNA ribbon structure would have been replaced by a TNA-RNA hybrid (Fig. 7.9b). RNA would have thus inherited stereoregularity from the TNA. Linked in this way to TNA, it would have blocked the TNA and hence acted as a repressor. The situation would have changed during folding and as a result of RNA detachment; transcription would then have become differentiated. Later evolution would have been marked by the appearance of new amino acids and new bases. The first replication process would thus have involved RNA.

The appearance of DNA, considered to be a late event in the "original soup" hypothesis due to the instability of deoxyribose, is accepted as a relatively early event in the theory of surface metabolism. DNA would have been formed at the same time as the accumulation of surface lipids which would have influenced the formation of phosphoanhydrides. The deoxyribonucleotide pyrophosphates thus formed are capable of oligomerising on an RNA or TNA template (the former being reverse transcription).

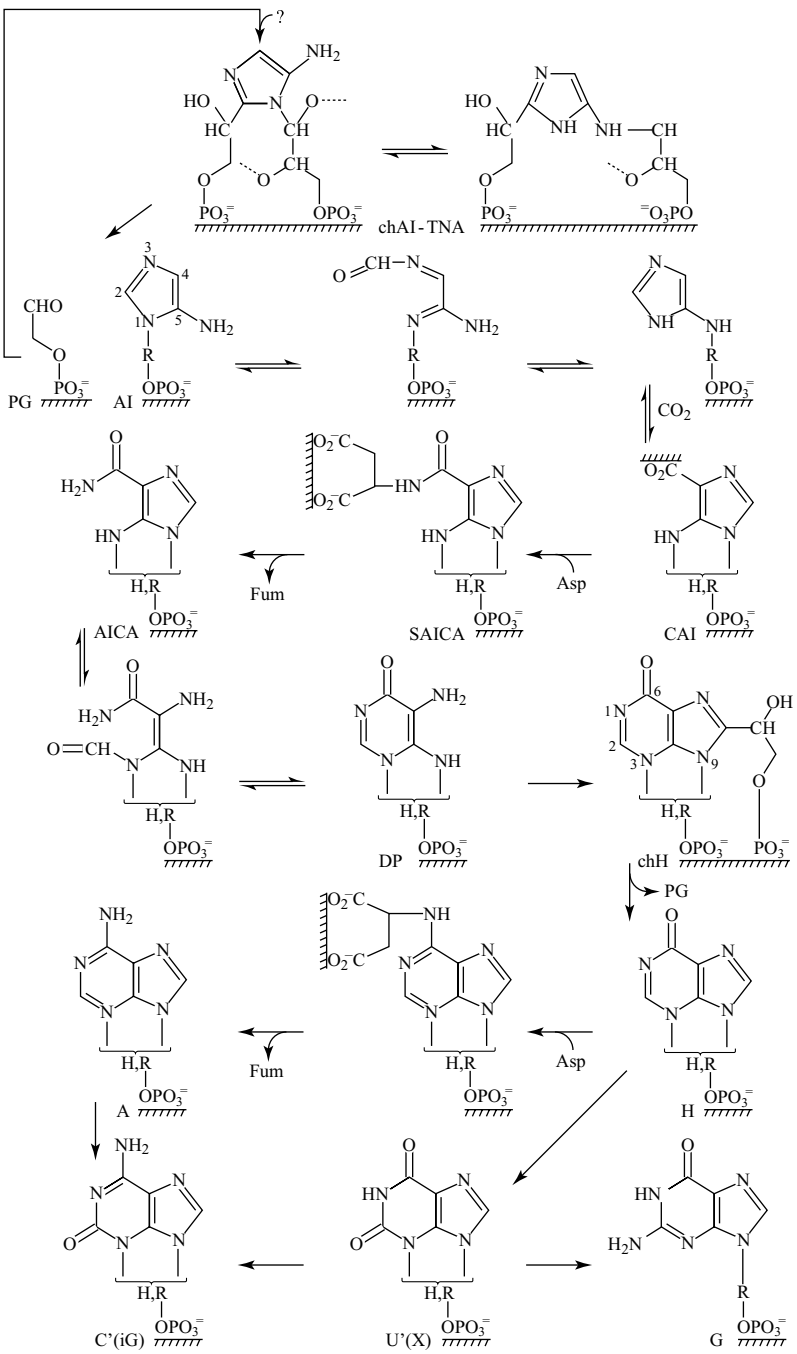


Fig. 7.8 Hypothetical pathway for the formation of purines on a surface
chAI: charged aminoimidazole; PG: phosphoglyceric aldehyde; CAI: carboxyaminoimide; SAICA: succino-aminoimidazole carboxamide; AICA: carboxamide derivative; Fum: fumaric acid; iG: isoguanine. (From *Microbiol. Mol. Biol. R.*, 1988, **52**, WÄCHTERHAÜSER G., 470, reproduced with permission from American Society for Microbiology)

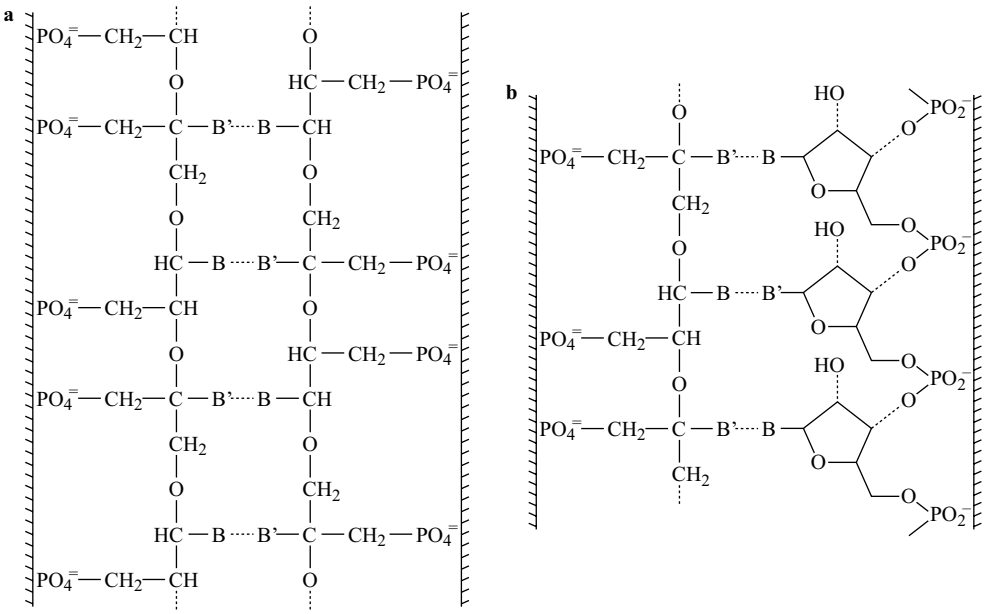


Fig. 7.9 (a) ribbon structure of two TNA strands linked to the same surface (b) hybrid structure of TNA-RNA linked to the same surface

(From *Microbiol. Mol. Biol. R.*, 1988, **52**, WÄCHTERHAÜSER G., 472 and 474, reproduced with permission from American Society for Microbiology)

In parallel to the evolution of transcription and replication, the nature of the bases changed, increasing their structural stability and with an improvement in folding. In this theory, an important role is assigned to phosphoribosyl pyrophosphate in pyrimidine synthesis which is thought to have been produced after the formation of semi-cellular structures. Finally, DNA would have folded itself, encircling the grains of ferrite (Fig. 7.7b) to form a closed, circular DNA molecule. The disappearance of ferrite grains in the cellular stage would therefore have led to the formation of the double helix.

The first peptides would have been formed by translation from building blocks having anionic side-chains (Asp, Glu, P-Ser). Certain amino acids would have undergone conversion e.g. phosphoserine to serine or cysteine. The sequences that ensured a strong bond from each anionic group with the surface would have been selected. Their configuration and conformation would have been determined by their surface association; only the amino acids with the L configuration would have been retained. Over time, more and more amino acids would have been incorporated into polypeptides, first of all hydrophilic amino acids, then smaller hydrophobic ones and lastly aromatic and cationic amino acids. WÄCHTERSHÄUSER favoured the presence of β and β-turn structures in early protein folding. The last amino acids would have entered into the genetic code after the appearance of synthetases. Specifically, these were histidine (derived from imidazole), tryptophan and methionine, which was perhaps the very last. The most recent experiments

done by WÄCHTERSHÄUSER and co-workers demonstrated the plausibility of this theory. The investigators showed in particular that a co-precipitate of FeS and NiS reduced carbon monoxide (HUBER & WÄCHTERSHÄUSER, 1997). Furthermore, they obtained peptides by the activation of amino acids with CO on co-precipitated (Fe,Ni)S (HUBER & WÄCHTERSHÄUSER, 1998; HUBER et al., 2003). These experiments took place in conditions that reproduced primordial temperatures. The existence of volcanic faults in sea depths, where spurts of very hot water rich in transition metals and sulphur mixed with cold sea water, makes this a very attractive theory. These zones are indeed regions of highly abundant biological activity.

Thus, according to WÄCHTERSHÄUSER's theory, evolution progressed in parallel co-evolution to amino acids and the genetic code, and RNA appeared before DNA. The central role of RNA has been previously emphasised in the theories of CRICK and ORGEL. The most recent data on the catalytic properties of RNA molecules seem to support the hypothesis of their particular role since the origin, but do not enable us to decide if they arose before proteins.

7.3.4. CATALYTIC PROPERTIES OF RNA

The discovery by ZAUG and CHECH, in 1986, indeed indicates that RNA possesses a catalytic activity. In their article entitled "The intervening sequence of Tetrahymena is an enzyme", they showed that splicing of this particular RNA sequence takes place without protein involvement, in the presence of micromolecular quantities of GTP or even simply guanosine which attacks RNA at position 413, at its 5′ extremity. This is a transesterification reaction in which the guanosine hydroxyl group attacks the phosphodiester bond between the 3′ end of the first exon and the first nucleotide of the intron. The guanosine remains attached to the 5′ end of the intron. Then, the 3′ end of the liberated exon attacks the extremity of base 413 from the intron to bring together the exon's two ends (Fig. 7.10 below). Thus, the mechanism involves a series of transesterifications. The intron itself undergoes a cascade of spontaneous reactions, which are rapid and relatively specific. This was remarked upon by WESTHEIMER (1986), as many organic molecules are able to undergo internal rearrangements without the aid of a catalyst. WESTHEIMER cites, for example, the cyclisation of 2-hydroxyphenyl propionic acid into dihydrocoumarin. The pentamethyl derivative cyclises 10^{10} times faster than the non-substituted derivative and hence with a high efficiency, and without considering the phenolic acids to be enzymes. However, in the reactions discovered by CHECH, firstly, the distance between the catalytic groups and the hydrolysable bond is great, which imposes a particular folding constraint on these RNAs. Secondly, the reaction is produced at a precise site even though other possible sites exist on the molecule. In the presence of this "truncated sequence", different polynucleotides are hydrolysed at a rate of 2 min^{-1}, which is slow, yet much quicker than the spontaneous hydrolysis of RNA and roughly corresponds to the rate of DNA hydrolysis by restriction enzymes or by non-specific phosphodiesterases. This reaction not only constitutes

autocatalysis, but also true catalysis since the intron can act on some other RNAs. These introns are classed as group I introns, and since their discovery a second class, group II introns, has been found. Their mechanism of excision and splicing proceeds slightly differently, in particular, the presence of guanosine is not required; the attack is carried out by the 2′OH group of a specific adenylate in the intron.

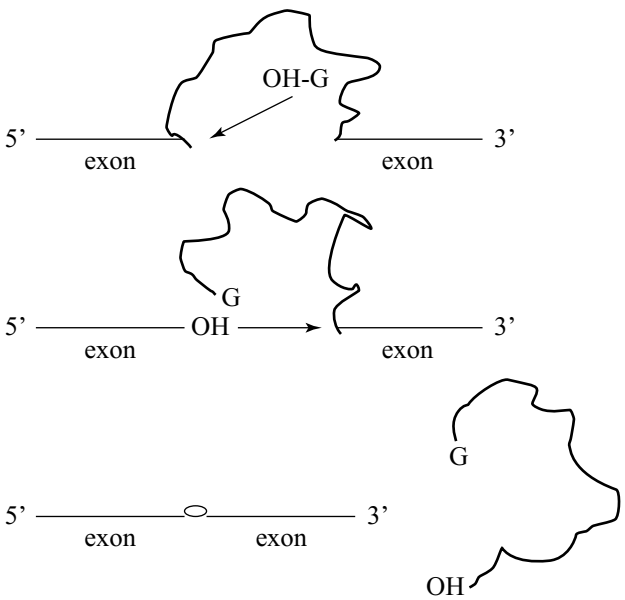


Fig. 7.10 Schematic representation of the catalytic activity of RNA

Over the last few years, numerous RNAs possessing catalytic activities have been discovered in diverse organisms. In addition, many synthetic RNAs capable of binding metabolites specifically and cofactors have been produced in the laboratory. Besides, from a catalytic RNA, M.C. MAUREL and co-workers obtained a variant whose activity is dependent on the presence of an adenine cofactor, thus showing that ribozymes can, like protein enzymes, function by means of a prosthetic group (MELI et al., 2003). The great plasticity and adaptability of RNA, which is demonstrated by these various observations, brings supportive arguments to the implication that an “RNA world” would have preceded the nucleoprotein world in the processes governing the appearance of life (see *La Naissance de la vie* [The birth of life], M.C. MAUREL, 2003).

Once more it is important to stress that the molecules harbouring a catalytic activity require a three-dimensional structure, whereas information storage is one-dimensional, which facilitates reading the code and copying the message. The fact that RNA can have catalytic properties argues for a primordial role of RNA in the origin, but does not necessarily signify that it appeared before proteins. **Whatever the scenario chosen to attempt to explain the origin of life and the first macromolecules, there is unanimous agreement that the existence**

of three-dimensional structures was vital for the expression of biological activity such as enzyme catalysis.

Returning to the theory of WÄCHTERSHÄUSER, life would have begun with metabolism on surfaces in contact with water at neutral pH and high salinity, at high temperatures and probably high pressure in an environment containing SH_2, CO_2, N as well as ferrous ions and other catalytic metals. Whereas the "original soup" favoured an initial appearance of proteins or nucleic acids according to its proponents, in the theory of surface metabolism, metabolic pathways, which resulted in protein and nucleic-acid formation, evolved in parallel. The existence of submarine, volcanic, hydrothermal sources has revived this theory, which today interests many experts investigating the origin of life.

7.4. CHIRALITY OF BIOLOGICAL MOLECULES

Biological molecules are asymmetric. Natural amino acids have the L configuration. Figure 7.11 shows the configuration of amino acids. In order to recall it, the rule is simple: if we look from the hydrogen along its bond linking the asymmetric carbon, the three other groups R, N, C′ appear clockwise in that order in the L configuration; in the D configuration they appear anti-clockwise. Trioses, such as glyceraldehyde, which only possess one asymmetric carbon, have the D configuration; the same goes for pentoses and hexoses.

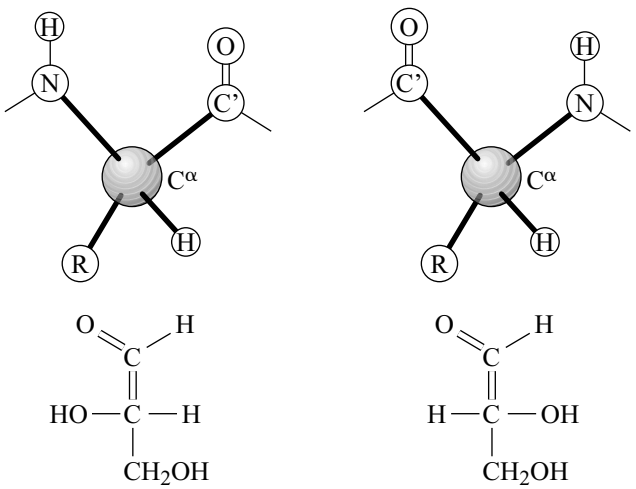


Fig. 7.11 Chirality of biological molecules

Diverse hypotheses have been proposed to explain the chirality of biological molecules. According to the "original soup" idea, these molecules supposedly existed

as racemic mixtures. The three-dimensional structure of polymers is only stable, however, if these mixtures are formed exclusively from L or D elements. Both polymer types must have existed initially. It has been suggested that the L form of amino acids was selected by chance. This might imply, consequently, that all cells derive from the same primitive cell or very similar cells. Molecular asymmetry would result either from the polarisation of light at sea surfaces, or from the chirality of the Universe. Using surface metabolic theory, the chirality of biomolecules (the first selection process) can be explained by considering the nature of the surface itself.

7.5. OTHER THEORIES ON THE ORIGIN OF LIFE

While discrepancies exist about the nature of the first molecules present in the beginning as we have just described, the mechanisms by which they are formed from original chaos are also the subject of controversy. In addition to the previously described models, various others based on different theories have been proposed.

For PRIGOGINE and his school, the origin of life can be described in terms of the spontaneous formation of structures in solution far from equilibrium (see Chap. 3). This is produced when, in a series of reactions, a flux of reactants exists along with a regulatory feedback loop (retroinhibition, retroactivation), the products disappearing when a steady state develops. Following instability at any given moment, the system can oscillate in time as well as in space, and dissipative structures may form. Currently, in many fields in physics and for the past few years in astronomy, in celestial mechanics and even in biology, chaos theories have aroused a growing interest. Indeed, environmental fluctuations may constitute an important aspect of a system's dynamics.

EIGEN, in his hypercycle theory, adopts quite a similar position. He considered that the spontaneous formation of a tRNA molecule is plausible; with mRNA serving as the adaptor, the spontaneous synthesis of a polynucleotide from individual nucleotides is conceivable. He thought that the initial apparatus was thus formed before a mechanism capable of removing the errors could be put in place. According to EIGEN, the filtering mechanism appeared later by coupling two or more cycles, introducing cooperation. The resulting hypercycle dominated and allowed the selection of information from all possible errors. For EIGEN, the first appearance of cooperative mechanisms was the mutual interaction of cycles that produced replicases. He suggested that three conditions were nevertheless necessary to overcome the loss of information across all possible replication errors:

- each replicative unit must selectively maintain its information in competition with its own error distribution;
- the competition between replicative units belonging to the same cooperativity must cease to be operational;
- the functional unit must be capable of competition between alternative units.

Other authors, in particular KUHN and WASER, refused the notion of such a spontaneous appearance of structures resulting from instability. They proposed an appearance

in stages, each stage having a reasonable probability and averaging quite long time periods. They suggested that, first of all, short segments would be formed, which are then condensed into longer segments, the first molecules being RNAs. Erroneous replication would have been rejected during the formation of aggregates, which are favoured by cooling. ◢

In summary, different scenarios have been proposed to attempt to explain the appearance of biological macromolecules in the beginning, and they have given rise to several controversies. The first assumed a very reducing atmosphere and through the effects of electrical discharges, the molecules present, including methane, would have generated an ensemble of biological molecules forming the "original soup". Then the macromolecules would have appeared, certain authors giving precedence to proteins, others to nucleic acids. The "primitive soup" hypothesis has been much criticised on chemical, geochemical and thermodynamic grounds. The most recent theory of surface metabolism developed by WÄCHTERSHÄUSER offers a more appealing alternative. It stems from the hypothesis that life's origin was autotrophic and consisted of an autocatalytic mechanism confined to a bilayer, the anionic components being linked to positively charged surfaces such as pyrites, and thus ensuring molecular selection. Diverse coenzymes would have been formed in this period along with polymeric components, the *phosphotriboses*. With the formation of isoprenoid lipids, changes would have proceeded in stages, firstly towards semi-cellular structures, then towards cellular structures. The pathways for the formation of proteins and nucleic acids would have evolved in parallel, and RNA appeared prior to DNA.

It is difficult to come up with proof for the theories of life's origin, which can only suggest more or less plausible scenarios. These theories evolve with progress in the understanding of currently existing biological systems. In this sense, it is reasonable to wait for useful information arising from studies into archaebacteria living in conditions of extreme temperature and pressure.

BIBLIOGRAPHY

GENERAL WORKS

ALLEGRE C. –1992– *Introduction à une histoire naturelle*, Fayard, Paris.

DANCHIN A. –1990– *Une aurore de pierre: aux origines de la vie*, Editions du Seuil, Paris.

MAUREL M.C. –2003– *La naissance de la vie*, Dunod, Paris.

OPARIN A.I. –1964– *Life, its origin, nature and development*, Acad. press, New York.

ORGEL L.E. –1973– *The origin of life*, Chapman & Hall Ltd, London.

ORO J., MILLER S.L., PONNAMPERUMA C. & YOUNG R.S. eds –1973– *Cosmochemical evolution and the origins of life. Proceedings of the IVth International Conference on the Origin of Life*, Dordrecht Holland Pub., Dordrecht.

REVIEWS

LEHNINGER A.L. –1972– in *Biochemistry*, Chap. 24, Worth Pub., Englewood clitts NJ.

WÄCHTERSHAÜSER G. –1988– Before enzymes and templates: theory of surface metabolism, in *Microbiol. Mol. Biol. R.* **52**, 452-484.

SPECIALISED ARTICLES

BOROWSKA Z. & MAUZERALL D. –1988– *Proc. Natl Acad. Sci. USA* **85**, 6577–6580.

CHECH T.R. –1986– *Cell* **44**, 207–210.

CRICK F.H.C. –1968– *J. Mol. Biol.* **38**, 367–379.

EIGEN M. –1971– *Naturwissenschaften* **58**, 465–522.

FOX S.W. –1969– *Naturwissenschaften* **56**, 1–9.

GILBERT W. –1986– *Nature* **319**, 618.

HUBER C. & WÄCHTERSHÄUSER G. –1997– *Science* **276**, 245.

HUBER C. & WÄCHTERSHÄUSER G. –1998– *Science* **281**, 670.

HUBER C., EISENREICH W., HECHT S. & WÄCHTERSHÄUSER G. –2003– *Science* **301**, 938.

LEWIN R. –1986– *Science* **231**, 545–546.

MELI M., VERGNE J. & MAUREL M.C. –2003– *J. Biol. Chem.* **278**, 9835–9842.

OPARIN A.I. –1965– *Adv. Enzymol.* **27**, 347–380.

ORGEL L.E. –1968– *J. Mol. Biol.* **38**, 381–393.

WESTHEIMER F.H. –1986– *Nature* **319**, 534–536.

ZAUG A.J. & CHECH T.R. –1986– *Science* **231**, 470–475.

8 – FORMATION OF THE FUNCTIONAL STRUCTURE OF ENZYMES: CO- AND POST-TRANSLATIONAL EVENTS

The acquisition of a three-dimensional structure is a defining event for the appearance of a functional active site in a protein. For this reason we feel it is important to describe succinctly here the mechanisms that ultimately generate a defined spatial architecture and thus enable the expression of activity. In view of the spectacular progress that has been made in understanding the mechanisms of genome expression, we might be tempted to believe that the events leading to the formation of an active protein are entirely elucidated. **The information contained in DNA, the genetic message, is one-dimensional information** which, during the diverse processes of biosynthesis – including transcription and translation – leads to the formation of a polypeptide chain having a well-defined sequence.

A number of key events also occurs in order that **the biological function, which requires a three-dimensional structure,** can be expressed. These events take place either throughout biosynthesis (co-translational events), or after the termination of the polypeptide chain (post-translational events). There are two groups: first, the covalent processes such as limited proteolysis and chemical modifications; second, the non-covalent processes, for instance, folding of the polypeptide chain and in certain cases the self-assembly of subunits to form a quaternary structure. However, all these events are still incompletely understood and poorly controlled, yet they are the subject of in-depth studies. Their importance is paramount since they generate the functional properties of proteins required for recognition, transport, catalysis and regulation. These functions generally only appear in very specific places within cells or organisms, permitting very precise and fine regulation to ensure the harmonious functioning of living beings.

Covalent and non-covalent processes will be considered in succession, although during protein formation they may arise chronologically in a different order.

8.1. COVALENT PROCESSES

8.1.1. LIMITED PROTEOLYSIS

During or after termination of the biosynthesis of polypeptide chains, or during their transport to particular regions of the cell or even the organism, some proteins

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_9, © Springer-Verlag Berlin Heidelberg 2010

undergo limited proteolysis. Essentially two types of proteolytic processing exist, which are important events for the acquisition of the functional structure of proteins; they are:

- cleavage of the signal peptide, and
- activation of precursors, zymogens or prohormones.

In 1975, BLOBEL and DOBBERSTEIN discovered that secreted proteins, particularly those that are synthesised on ribosomes linked to the rough endoplasmic reticulum (RER), when synthesised by cell-free systems not containing the RER membrane, possess a supplementary sequence of about twenty amino acids at their N-terminus, called the **signal peptide.** This sequence is largely hydrophobic and always begins at the N-terminus with a methionine residue. The addition of RER membranes to a cell-free preparation causes this sequence to disappear. The signal peptide, due to its hydrophobic character, enables the assisted deposition of the polypeptide chain into the intercisternal space. After traversing the membrane, it is digested by a membrane signal peptidase (Fig. 8.1a).

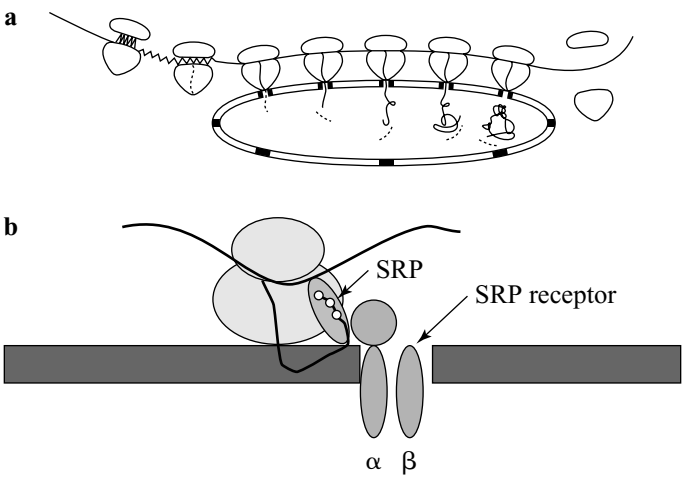


Fig. 8.1 Mechanism of enzymatic cleavage of the signal peptide during the transfer of the nascent polypeptide chain across the RER membrane

*(**a**) general scheme for the assisted transfer of the nascent protein across the membrane (**b**) interaction between the ribosome and signal recognition protein (SRP) associated to the anchoring protein.* (8.1a – © BLOBEL G. & DOBBERSTEIN B., 1975. Originally published in the *J. Cell Biol.*, **67**, 835–851)

- It is assumed that the signal peptide carries information for the association of the ribosome with the membrane via its interaction with particular proteins. These proteins were called ribophorins by KREIBICH et al. (1978). The mechanism was detailed later by MEYER et al. (1982) who identified the protein that recognises the signal sequence, signal recognition protein (SRP). BLOBEL and WALTER showed that SRP is composed of an assembly of 300 nucleotides and 6 different polypeptide chains forming a complex of 325 kDa. This complex is associated to another

integral membrane protein formed from one α subunit of 68 kDa and one β subunit of 30 kDa (Fig. 8.1b). The sequence of events optimises the co-translational process for secreted proteins. This process happens even when synthesis begins on the free ribosomes. Translation is blocked by the binding of SRP after a polypeptide chain comprising about 70–80 amino acids has been synthesised, when the signal sequence emerges from the large ribosomal subunit.

The pause in translation persists until contact is established with the anchoring protein. Translation then continues and translocation ensues. This safety-lock mechanism ensures that the protein is not terminated in the cytoplasm. The GTP-GDP cycle regulates the dissociation of the signal sequence from its receptor (RAPPOPORT, 1992). The anchoring protein is the first site of interaction between the nascent secreted protein and the endoplasmic reticulum membrane. However, the signal peptide is not always cleaved from membrane proteins, which are therefore not delivered to the intercisternal space, but are instead inserted into the membrane. The existence of signal peptides is not specific to proteins from eukaryotic cells; they have also been observed in proteins from prokaryotic organisms. Signal peptides also exist in proteins synthesised by free polysomes, such as mitochondrial proteins. ◢

The sequence of the signal peptide has been determined for a significant number of proteins. It is compatible with a helical structure, which would seem to be the preferential structure. The size of the helical segment formed would correspond approximately to the thickness of the RER membrane.

◤ To what extent is the proteolytic removal of the signal peptide necessary for the correct folding of the polypeptide chain and consequently for the emergence of a protein's function? There is no general answer to this question; indeed, pre-ribonuclease has an enzymatic activity whereas pre-amylase is inactive. Many proteins are synthesised as preproteins; the cleavage of the signal peptide leads, in some cases, to the gain of functional properties, but this situation really depends on the protein. ◢

As well as the removal of signal peptides, other controlled proteolytic processing is necessary in order for biological activity to manifest itself in certain proteins. Proteolysis of the precursor proteins triggers their activation and thus provides a means to regulate the formation of these active proteins in specific compartments within an organism. Many biological systems are regulated by proteases, for example, the formation of hormones and active peptides, the activation of zymogens to produce active enzymes, the blood coagulation cascade and the activation of complement, protein assembly in the heads of phages and the transformation of fibrinogen into fibrin. The role of proteases is also essential in fertilisation processes, development, inflammatory reactions or malignant cell migration and has been the subject of numerous studies.

◤ Highly specific proteolytic cleavage is used to generate certain peptide hormones. For instance, proinsulin is converted into insulin by the separation of a peptide of 33 amino acids, which gives rise to two protein chains, A and B, and is accompanied by a large change in conformation (Fig. 8.2 below).

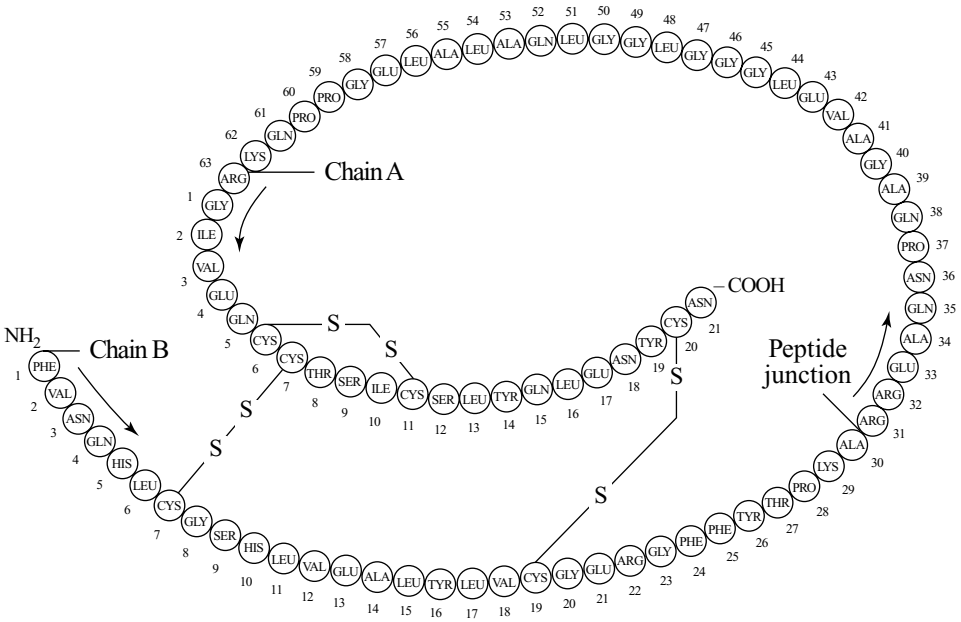


Fig. 8.2 Conversion of proinsulin, formed from a single chain, into insulin
The start sites of chains A and B are indicated by arrows

Multifunctional precursors also exist, such as pro-ACTH, which undergoes several types of proteolytic cleavage in a variety of places. It is not only the precursor of ACTH (adrenocorticotropic hormone), but also of α- and β-MSH (melanocyte-stimulating hormone), α- and β-LPH (lipotropic hormone), endorphins and enkephalins (Fig. 8.3).

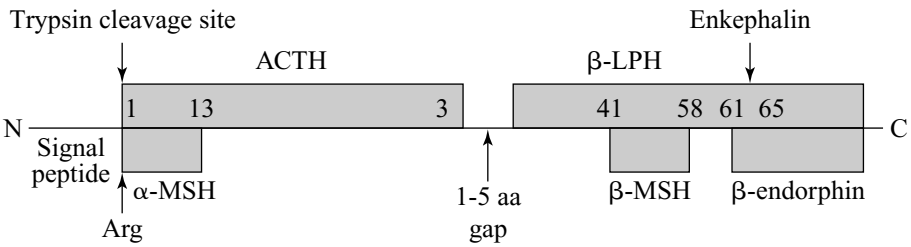


Fig. 8.3 Pro-ACTH: a multifunctional precursor of adrenocorticotropic hormone and other peptide hormones

The activation of zymogens to enzymes by limited proteolytic cleavage generally only leads to small variations in the polypeptide chain, but is sufficient for their ultimate conformational maturation. Thus with the serine proteases limited proteolysis of the zymogen induces, in appropriate conditions, the formation of a salt bridge between the amino group liberated by proteolytic digestion and a carboxylate neighbouring the reactive serine. This salt bridge stabilises the enzyme's functional structure.

Figure 8.4 illustrates the different activation pathways of chymotrypsinogen.

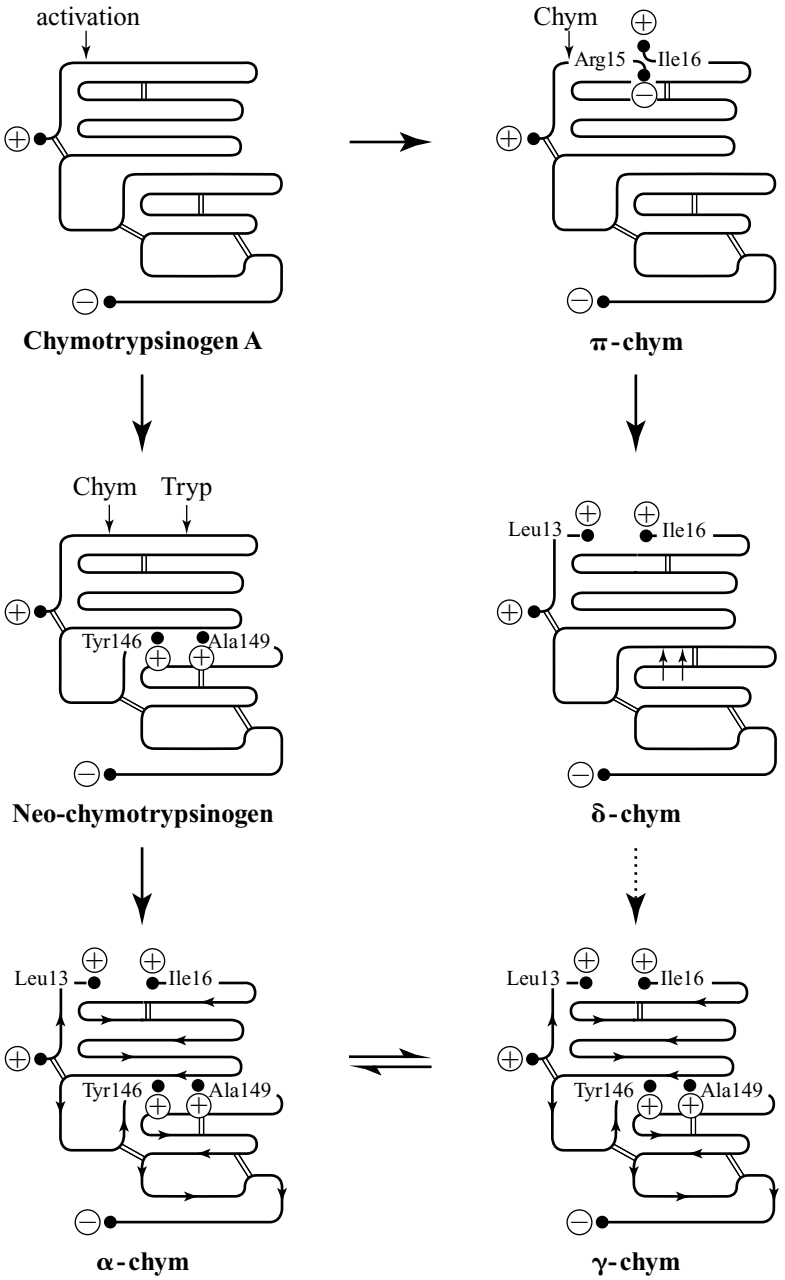


Fig. 8.4 The activation pathways of chymotrypsinogen

(Reprinted from *The Enzymes*, 3rd ed., Vol. III, BLOW D., The structure of chymotrypsin, 187.

Figure 8.5 shows, based on crystallographic data, the salt bridge formed between Ile 16 and Asp 194 adjacent to serine 195 in the active site of chymotrypsin. It is interesting to note that the active site is partially functional in the zymogen. Indeed, chymotrypsinogen has a very weak but significant catalytic activity; it is 10^6 to 10^7 times weaker than that of the enzyme (GERTLER et al., 1974). This system is studied in detail in Chap. 12. After limited proteolysis and creation of the salt bridge, the conformational rearrangements are only minor, yet decisive for augmenting the activity. Experimental evidence suggests that the zymogen is in a folded state that is close to the native state and yet different, probably a state that precedes the native one on the enzyme's folding pathway. The formation of the salt bridge must then promote the conformational coupling of the two structural domains that make up the enzyme: this coupling is necessary for the emergence of its activity (GHÉLIS & YON, 1979).

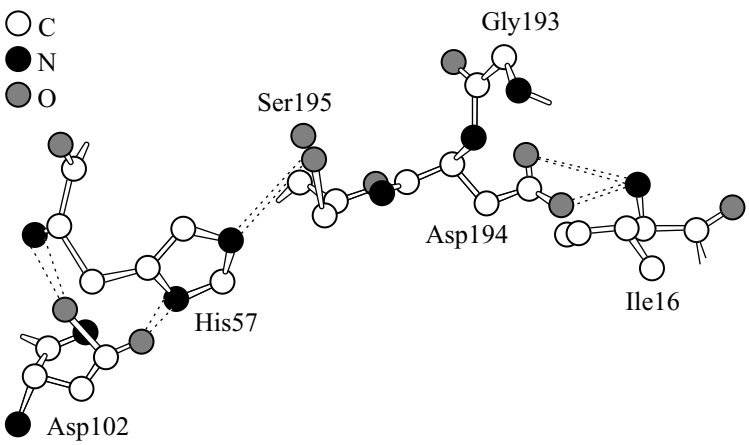


Fig. 8.5 The salt bridge between Asp 194 and Ile 16, which promotes chymotrypsin's activity (taken from crystallographic data)

(Reprinted from *The Enzymes*, 3rd ed., Vol. III, BLOW D., The structure of chymotrypsin, 187. © (1971) Academic Press, with permission from Elsevier)

The zymogen-activation cascades responsible for blood coagulation also come about from highly specific limited proteolyses. In this system, illustrated in Figure 8.6 opposite, each enzyme formed at a given step in the sequence activates a zymogen in the subsequent step, which in turn continues the cascade of events that culminates in the conversion of fibrinogen into fibrin.

Two different pathways exist to initiate the blood coagulation cascade involving either a contact factor, or a tissue factor. These processes engender a considerable amplification, which enables organisms to respond quickly to injury. Furthermore, they are modulated by non-covalent inhibitors at each step of the cascade. The cascade system represents an extremely efficient mechanism of covalent regulation, which we return to in Part V.

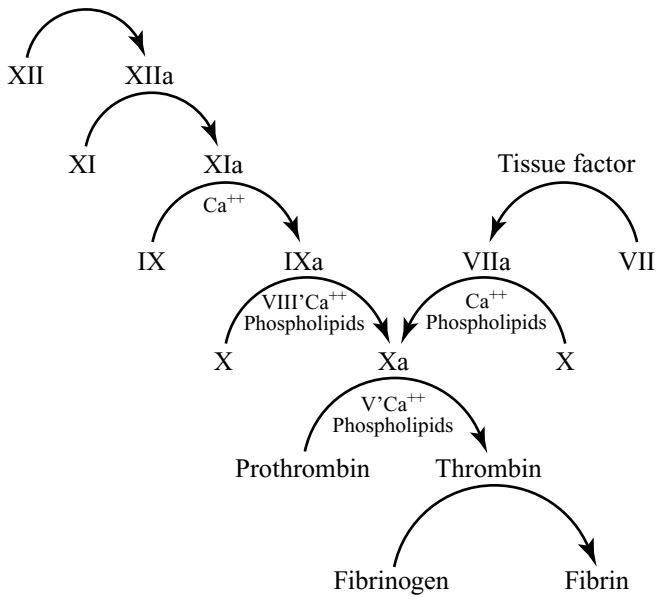


Fig. 8.6 Blood coagulation cascade

Several types of protein exist in circulation systems that are capable of generating, by limited proteolysis and in specific places, particular well-regulated enzymes in high concentrations. The fibronectins are an example. Apparently inert matrix proteins can also be activated in this way.

8.1.2. CHEMICAL MODIFICATIONS

Aside from proteolyses, diverse chemical events contribute to the genesis or the stabilisation of functional proteins. Their co- and post-translational chemical modifications are very varied. It is beyond the scope to detail them all here, but we mention briefly those which seem the most crucial from a structural and functional viewpoint. Many reviews deal with this problem, though we refer in particular to those that feature in *The enzymology of post-translational modifications of proteins*, edited by R.B. FREEMAN and H.C. HAWKINS (1980).

Let us firstly note that the formation of **disulphide bonds** results from the oxidation and condensation of sulphydryl groups. The formation of disulphide bonds is a major co- or post-translational modification in protein biosynthesis. Many free thiols are found in intracellular proteins that do not generally have disulphide bonds, yet on the other hand are frequently oligomeric. Disulphide bonds predominate in extracellular proteins. In fact the intracellular environment is more reducing than the circulating fluids in organisms. Furthermore, disulphide bonds confer greater stability on proteins helping them to maintain their functional structure despite environmental fluctuations. If the chemistry of sulphydryl groups and disulphide bonds is well known, disulphide-bond formation in vivo has not been completely elucidated.

◤ Thiol oxidation to give disulphide bonds is a spontaneous process that happens when oxidoreduction conditions are favourable. The existence of a microsomal enzyme, disulphide isomerase, capable of catalysing the conversion of thiols into disulphide bonds suggests that an enzymatic process occurs in vivo. For a few proteins, there is evidence that seems to indicate that disulphide bonds are formed on the nascent polypeptide chain after the completion of each structural domain (FREEDMAN & HILLSON, 1980). ◢

Glycosylation is a very widespread chemical modification, found in particular in membrane proteins, but also in many soluble proteins (PHELS, 1980). Only some amino acids are capable of binding to a sugar group, these include: serine, threonine, asparagine and hydroxylysine. The glycosidic motifs that link to these different residues are not the same, however. Glycosylation frequently arises on Asn residues in loops (β-turns) and therefore at the protein's surface in contact with solvent. This event confers on the protein precise structural properties in a given environment. The presence of sugar groups has the general effect of increasing the protein's stability and making it more resistant to proteases. Indeed, a glycosidic chain on the protein's surface creates an important screen between the protein and the solvent. In addition, the role of this carbohydrate extension in the specificity of cell-surface recognition is fundamental for cell adhesion, the immune response, hormone binding and differentiation.

Among other modifications, the **hydroxylation of prolines and lysines** leads to the formation, respectively, of 4-hydroxyproline, 3-hydroxyproline and hydroxylysine (Fig. 8.7).

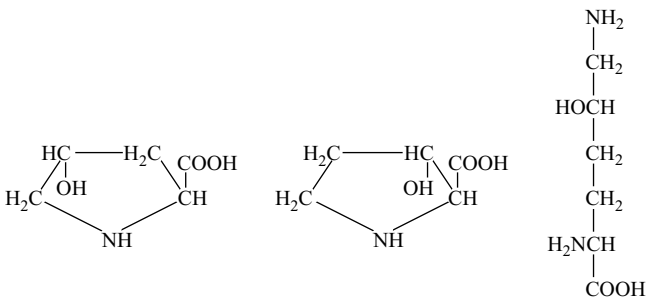


Fig. 8.7 Structures of 4-hydroxyproline, 3-hydroxyproline and 5-hydroxylysine

◤ These modifications are present in collagen and a few proteins having similar sequences such as acetyl cholinesterase, protein Clq from complement, elastin and some plasma proteins (KIVIRIKKO & MILLYLÄ, 1980). They are catalysed by three different enzymes: prolyl-4-hydroxylase, prolyl-3-hydroxylase and lysyl hydroxylase, enzymes which are found in higher organisms, and also in lower organisms and plants. These enzymes have been identified and characterised. ◢

Phosphorylations are important modifications for the regulation and control of a great number of cellular processes (HUNTER, 1987). They require the action of specific kinases, which transfer a phosphate group to the protein substrate. Nearly a hundred protein kinases have been identified to date. The reversibility of this covalent modification by the activity of phosphatases is a means to control certain cellular processes. Phosphorylation generally affects serine, threonine or tyrosine residues and can considerably modify a protein's properties. For enzymes, this may mean a change in activity either causing inhibition or, conversely, activation. Many enzymes are activated by phosphorylation, for instance, glycogen phosphorylase, phosphofructokinase, triacyl glycerol lipase, tyrosine hydroxylase and DNA-dependent RNA polymerase to name but a few. Others are inhibited by phosphorylation; a few examples include glycogen synthase, pyruvate dehydrogenase and glycerophosphate acyltransferase. A certain number of proteins having no enzymatic activity are phosphorylated in vivo, among which, troponin I, ribosomal proteins and perilipin, a protein from adipose tissue. The activity of certain kinases is dependent on cyclic AMP (cAMP) or cyclic GMP (cGMP). The cAMP-dependent kinases are in general composed of two types of subunit: catalytic (C) subunits, which catalyse the phosphorylation reaction and regulatory (R) subunits, which bind to cAMP. They are frequently tetramers having the composition R_2C_2. These enzymes are very specific for their protein substrates, more specifically the phosphorylation sites. Specific sequences are found proximal to the phosphorylated serine, namely, Lys-Arg-X-X-Ser(P) or Arg-Arg-X-Ser(P), where X denotes any amino acid. The phosphorylations catalysed by this type of enzyme are integral to metabolic control, muscle contraction, as well as transcription and translation. These aspects are discussed in Part V, since phosphorylation participates in the covalent regulation of the activity of several enzymes.

The **carboxylation** of glutamyl residues to form γ-carboxyglutamyl residues (γ-carboxyglutamic acid or Gla) plays an important role in the activity of proteins required for blood coagulation (SUTTIE, 1980). This modification has been observed in several plasma proteins, in particular prothrombin where the exact Gla residues have been identified; they are localised to the molecule's N-terminus. This modification is also present in the clotting factors IX and X and probably in factor VII, which possesses a similar N-terminal sequence to that of prothrombin; the Gla residues are found in this region. Other plasma proteins such as bovine protein C and human protein S contain γ-carboxyglutamic acid. These also have close homology to the sequences of the above-mentioned proteins. Let us also briefly mention protein Z from bovine plasma and osteocalcin from chicken bones. The carboxylase that introduces the additional carboxylate group on these proteins requires vitamin K.

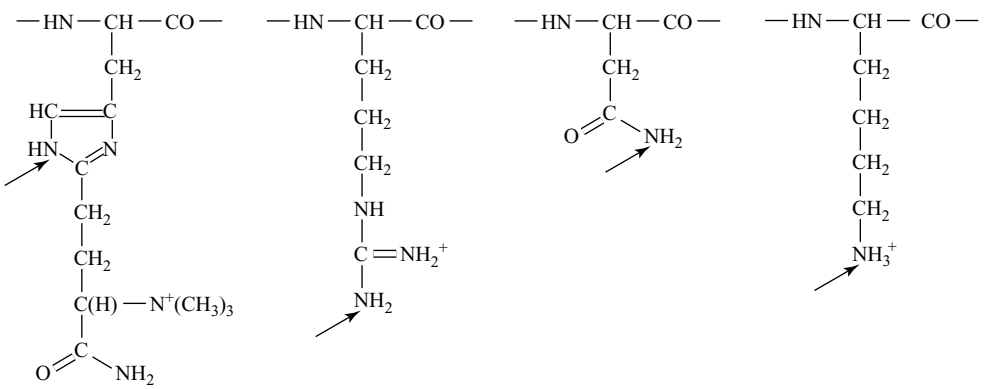
Among the other post-translational chemical modifications, **methylations** take part in various processes. O-methyl aspartate, O-methyl glutamate and N-methyl glutamine are present in a variety of proteins, e.g. N-methyl histidine in actin, α-N-trimethyl alanine in ribosomal protein from *E. coli* and myosin light chain, and α-N-dimethyl proline in bacterial cytochrome c. The ε-amino group of lysine

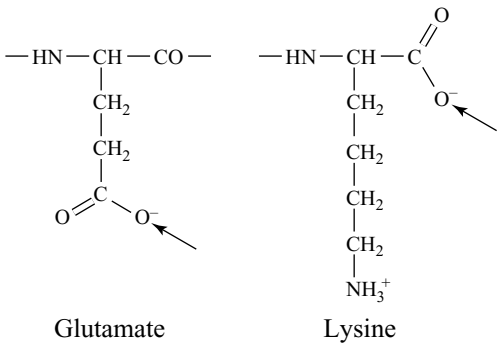
can be methylated giving rise to three derivatives: ε-N-methyl lysine, ε-N-dimethyl lysine, ε-N-trimethyl lysine. These derivatives have been found in a great variety of eukaryotic and prokaryotic proteins. The enzymes catalysing this reaction are protein-lysine methylases which have been isolated from different organisms and characterised.

Acetylation of the N-terminus (N-α-acetylation) of polypeptide chains is amino acid-dependent, occurring preferentially with Ala or Ser, and sometimes Met, Gly and Asp. N-α-acetylation has also been observed for Asn, Ile, Thr and Val, but none of the 11 remaining amino acids undergo this modification. It would seem that N-α-acetylation follows a phylogenetic scheme. For instance, cytochrome c is always N-α-acetylated in vertebrates and higher plants, but never in yeast or bacteria. The same is true for other enzymes such as enolase and several pancreatic enzymes. However, α-amylase is found to be acetylated in all species studied. N-α-acetylation of proteins is catalysed by the enzyme N-α-acetylase, which should be distinguished from N-ε-acetylase, the enzyme that acetylates the ε-amino group of lysines. Acetylation seems to take place on the nascent polypeptide chain after cleavage of the signal peptide.

The **iodination of tyrosines** in thyroglobulin yielding 3-iodotyrosine and 3,5-diiodotyrosine is an important chemical modification in the biosynthesis of thyroxine.

ADP ribosylation refers to a whole group of post-translational modifications on proteins. The ADP-ribosylation reactions are classed into two main groups: mono-ADP ribosylations and poly-ADP ribosylations. These two groups are distinct not only because of the ADP-ribose chain, but also due to the chemical nature of the bond between the ADP ribosyl and the protein, which is an N-glycoside bond in the first group and an O-glycoside bond in the second. The amino acid acceptors in the first group are: lysine, arginine, asparagine and diphtamide (a histidine derivative); and in the second group: glutamate and lysine (COO^- at the C-terminus) (Fig. 8.8).





Glutamate Lysine

Fig. 8.8 Amino acid acceptors in ADP-ribosylation reactions
The arrows *indicate the atoms involved in covalent-bond formation*

Many diverse proteins undergo these post-translational modifications, which are catalysed by specific enzymes. For instance, elongation factor EF2 is ADP-ribosylated, with diphatamide acting as the acceptor. Other proteins such as transducin, protein Ns from the adenylate cyclase complex, microtubule proteins and some *E. coli* and eukaryotic proteins are adenylated on their arginine or asparagine residues. The role of ADP-ribosylation in enzyme regulation is elaborated upon in Part V.

◤ Among other post-translational modifications, we also point out modification by the addition of fatty acids, namely, **N-myristylation and S-palmitoylation.** The myristate is linked by an amide bond on the N-terminal side of a glycine. The palmitate typically forms a thioester bond with a cysteine. Palmitoylated proteins are synthesised on free ribosomes and transported to the plasma membrane. **Isoprenylation,** well-known for G proteins, also occurs on Ras proteins. This modification is produced at a protein's C-terminus having the sequence CAAX by modifying the cysteine residue. The **glycolipid bond** is found in proteins anchored on the external side of cell membranes. ◢

8.2. NON-COVALENT PROCESSES

Although covalent modifications are not common to all proteins, the first non-covalent event, i.e. folding of the polypeptide chain, does apply to all proteins. For oligomeric proteins, the self-assembly of subunits constitutes an additional process.

8.2.1. PROTEIN FOLDING

The polypeptide chain has to fold in order to adopt a compact, globular three-dimensional structure. Protein folding is a vital process as it transforms one-dimensional information into three-dimensional information and only then can biological activity emerge, as has been previously emphasised. In this sense, protein folding can be

considered as the first act of morphogenesis. Studies into protein folding are however not detailed in this book; we refer the interested reader instead to other general books and reviews on the topic including: *Protein folding* by GHÉLIS and YON (1982), and the reviews by DOBSON and KARPLUS (1999), DILL and CHAN (1997), WOLYNES et al., (1995) and YON (2001). We shall nonetheless briefly outline the principal aspects of this research field.

It has been assumed for a long time that all the information necessary for a protein to acquire its three-dimensional structure is contained in its sequence, and thus in the genome. In 1958, F. CRICK wrote: "[Protein] folding is simply a function of the order of the amino acids." In 1973, following his remarkable work on the folding of ribonuclease, ANFINSEN clarified: "The three-dimensional structure of [...] the native [protein] conformation is determined by the totality of the inter-atomic interactions and hence by the amino acid sequence, in a given environment." The role of the environment, especially the solvent, is fundamental for the formation and stabilisation of the functional structure of proteins. A corollary of this well-known concept, designated the ANFINSEN postulate, was the assumption that the native structure of a protein represents the most energetically stable structure, corresponding to a minimum in GIBBS energy in defined conditions. The thermodynamic control of protein folding was called into question in 1968 by LEVINTHAL, and then by WETLAUFER in 1973 regarding temporal considerations. These authors pointed out that a random search for the native conformation amongst all possible conformations of a polypeptide chain would require an astronomical amount of time (10^{26} years for a chain only 150 amino acids in length!), which is of course incompatible with the folding times observed both in vitro and in vivo. LEVINTHAL proposed a kinetic rather than a thermodynamic control of the folding process. This familiar problem, termed the LEVINTHAL paradox, had dominated discussions for nearly thirty years. It is clear that a random search for the most stable structure of a protein along the entire length of its polypeptide chain is not a plausible hypothesis. It became obvious that evolution had found an efficient solution to solve this combinatorial problem.

Diverse models have been proposed to solve the LEVINTHAL paradox, which were experimented with in order to find and characterise the folding intermediates. A sequential and hierarchical model, in which elementary structures would form and interact with each other to generate the native protein, was supported by numerous authors until relatively recently (KIM & BALDWIN, 1990). The diffusion-collision model developed by KARPLUS and WEAVER (1976, 1994) implicated the formation of nuclei of elementary structures in different parts of the polypeptide chain. These nuclei would diffuse and self-associate. Several centres would be formed at the same time and self-assemble if they had the correct structure, thus enabling a polypeptide chain of 100–200 amino acids to acquire its native structure in less than a second. In 1985, HARRISON and DURBIN introduced the "jigsaw puzzle" model. Using the metaphor of puzzle building, they assumed that folding followed multiple

routes before reaching a unique solution. This model, which implied the existence of heterogenous species in the course of folding, became quite controversial at the time.

With the convergence of theoretical and experimental studies in 1995, a new vision of protein folding, the *new view*, emerged. Presented by WOLYNES et al. in terms of an energy landscape using the metaphor of a folding funnel, this model described the kinetic and thermodynamic behaviour of an ensemble of unfolded molecules. The number of conformations to be explored gradually diminishing as the process advanced, until an energy minimum representing the native structure was reached. The model implied the existence of several folding pathways and the heterogeneity of its intermediate species, and was in accordance with the "jigsaw puzzle" model. This vision of the folding process has now been confirmed experimentally. The advantage of this model is that it takes into account the possibility of incorrect folding which might result from the existence of local energetic minima in which the molecules could become "trapped", thus delaying their folding or forcing them towards aggregration. Two types of aggregate exist: amorphic aggregates, which appear as inclusion bodies when the gene coding for a protein is overexpressed in a foreign host, and aggregates organised into amyloid structures. Amyloid structures are the basis of serious animal and human pathologies among which bovine spongiform encephalopathy (BSE, or mad cow disease), CREUTZFELD-JACOB disease and ALZHEIMER's disease (see the review by YON, 2004).

Another question is raised: how does folding proceed in a cellular context? The assumption is that the same mechanisms are involved in folding processes *in vivo* and in vitro. The discovery of molecular chaperones (ELLIS, 1987) has led to a reconsideration of this question. The rapid development of research in this field, the resolution of the structures of chaperonins and molecular chaperones, as well as the characterisation of their interactions with partially folded proteins has permitted the elucidation of the roles of these molecules. Molecular chaperones, through their transient association with a nascent or unstable protein when under conditions of stress (heat shock, for example), prevent its potentially incorrect folding and subsequent aggregation. They also interact with proteins that undergo translocation. Their association is predominantly mediated by hydrophobic interactions. Molecular chaperones do not interact with native proteins. They do not have any information capable of directing the protein towards a conformation different from that determined by the sequence. Furthermore, they increase the folding yield but do not alter the folding rate; **molecular chaperones are not folding catalysts.** Lastly, molecular chaperones assist the folding of a limited number of proteins. The incorrect folding of a protein and its subsequent aggregation arises from kinetic competition between the correct folding pathway and an alternative side pathway. When the formation of the correct structure is kinetically favoured, the presence of a chaperone is unnecessary.

Taken together, the most recent data enable us to conclude that the principles governing protein folding, established by studies *in vitro* and *in silico*, also govern the folding of the nascent polypeptide chain in a cellular context.

Currently, protein folding represents a very active research field comprising different aspects of biology, physics, chemistry and computing. The fundamental principles have applications for the use of the information contained in genomic sequences, in the development of new therapeutics, understanding certain pathologies and the design of non-natural proteins harbouring specific functions.

8.2.2. ASSEMBLY OF SUBUNITS

Oligomeric proteins result from the association of identical (α_n homo-oligomers) or non-identical ($\alpha_n\beta_n$ hetero-oligomers) subunits. In the nomenclature proposed by MONOD, WYMAN and CHANGEUX, the term oligomer signified that the number of subunits is limited. It is in contrast to the term polymer for which the association consists of a large, indeed, an unlimited number of subunits. When the subunits making up the oligomer are identical, the authors employed the term protomer, a monomer being a molecule in which a single subunit exists. Thus, a haemoglobin molecule ($\alpha_2\beta_2$) has four subunits and two protomers.

The self-assembly of subunits takes place in the final stage of the folding process and is typically accompanied by specific conformational rearrangements. These may only be of a very subtle nature, but ensure the correct positioning of the functional sites (e.g. active site) in an enzyme. There are few known cases of oligomeric enzymes whose isolated subunits are functional. This is linked to the fact that, in many oligomeric enzymes, the active site is localised to the interface between two subunits and includes amino acid side chains belonging to each of them. Only the aldolases and transaldolases provide examples of oligomeric enzymes whose isolated subunits have a significant enzymatic activity. It is however important to emphasise that, even in these cases, the activity of the isolated protomer is lower than that of the oligomeric structure.

Thus, *the different levels of protein structure are essential for the formation of an enzyme's active site and for modulating its properties*.

BIBLIOGRAPHY

BOOKS

KIVIRIKKO K.I. & MILLYLA R. –1980– in *The Enzymology of post translational modifications of proteins* R.B. FREDMAN & H.C. HAWKINS eds, Academic Press, London p.53.

PHELPS C.F. –1980– in *The Enzymology of post translational modifications of proteins* R.B. FREDMAN & H.C. HAWKINS eds, Academic Press, London, p.112.

GENERAL REVIEWS

ANFINSEN C.B. –1973– Principles that govern protein folding, *Science* **181**, 223–230

CROSS G.A.M. –1990– Glycophorin anchoring of plasma membrane proteins, *Annu. Rev. Cell Biol.* **6**, 1–39.

DILL K.A. & CHAN H.S. –1997– From Levinthal paradox to funnel, *Nat. Struct. Biol.* **4**, 10–19.

DOBSON C.M. & KARPLUS M. –1999– The fundamentals of protein folding: bringing together theory and experiments, *Curr. Opin. Struct. Biol.* **9**, 92–101.

ELLIS R.J. & HARTL F.U. –1999– Principles of protein folding in the cellular environment, in *Curr. Opin. Struct. Biol.* **9**, 102–110.

GIBBS J.M. –1991– Ras C-terminal processing enzymes – New drug targets? *Cell* **65**, 1–4.

HUNTER T. –1987– A thousand and one protein kinases, *Cell* **50**, 823–829.

JAENICKE R. –1987– Folding and association of proteins, *Prog. Biophys. Biol. Mol.* **49**, 117–237.

KIM P.S. & BALDWIN R.L. –1990– Intermediates in protein folding reactions of small proteins, *Annu. Rev. Biochem.* **59**, 631–660.

TANFORD C. –1968– Protein denaturation, *Adv. Prot. Chem.* **23**, 121–182.

UEDA K. & HAYAISHI O. –1985– ADP ribosylation, *Annu. Rev. Biochem.* **54**, 73–100.

YON J.M. & BETTON J.M. –1991– Protein folding in vitro and in the cellular environment, *Biol. Cell* **71**, 17–23.

YON J.M. –2001– Protein folding: a perspective for biology, medicine and biotechnology, *Braz. J. Med. Biol. Res.* **34**, 419–435.

YON J.M. –2002– Protein folding in the post-geneomic era, *J. Cell. Mol. Biol.* **6**, 307–327.

YON J.M. –2004– Protein aggregation, in *Encyclopedia of Molecular Biology and Molecular Medicine*, Vol. I, 23–52, ed. by R.A. Meyers, Willey VCH, New York.

SPECIALISED ARTICLES

ANFINSEN C.B., HABER E., SELA M. & WHITE F.H. –1961– *Proc. Natl Acad. Sci. USA* **47**, 1309.

BLOBEL G. & DOBBERSTEIN B. –1975– *J. Cell Biol.* **67**, 835–862.

BLOW D.W. –1971– in *The Enzymes*, Vol. III, 3rd ed., P.D. BOYER ed., Acad. Press, New York.

CRICK F.H. –1958– *Symp. Soc. Exp. Biol.* **13**, 138.

ELLIS J.R. –1987– *Nature* **328**, 378.

GERTLER A., WALSH & NEURATH H. –1974– *Biochemistry* **13**, 1302.

GHÉLIS C. & YON J.M. –1979– *C. R. Acd. Sci.* **282**, 197.

HILLSON D.A. & FREEDMAN R.B. –1980– *Biochem. J.* **191**, 373.

JAMES G. & OLSEN E.N. –1990– *Biochemistry* **29**, 2623.

KARPLUS M. & WEAVER D.L. –1976– *Nature* **260**, 404.

KARPLUS M. & WEAVER D.L. –1994– *Protein Sci.* **3**, 650.

KREIBICH G., FREINSTEIN C.M., PEREYRA B.N., ULRECH B.N. & SABATINI D.D. –1978– *J. Cell Biol.* **77**, 464 and 488.

LEVINTHAL C. –1968– *J. Chim. Phys.* **65**, 44.

MEYER D.L., KRAUS E. & DOBBERSTEIN B. –1982– *Nature* **297**, 647.

RAPOPORT T.A. –1992– *Science* **258**, 931.

SUTTIE, J.W. –1980– *CRC Crit. Rev. Biochem.* **8**, 191.

WETLAUFER D.B. & RISTOW S. –1973– *Annu. Rev. Biochem.* **42**, 135.

WOLYNES P.G., ONUCHIC J.N. & THIRUMALAI D. –1995– *Science* **267**, 1619.

9 – TOPOLOGY OF THE ACTIVE CENTRE OF ENZYMES

For most enzymes, there is a disproportion between the size of the enzymatic molecule and the size of the substrate. Such an observation brought about very early the notion of the **active centre,** the fact that a very small proportion of the surface of the enzyme enters into contact with the substrate. However, the concept of the active centre remained for a long time rather poorly defined. For KOSHLAND, the active centre was constituted by all the enzyme atoms that are in contact fixed in the boundaries of the VAN DER WAALS radius with the substrate atoms, meaning those that stay at a minimal distance such that the electron clouds are not perturbed. KOSHLAND distinguishes thus the contact residues and the auxiliary residues, the latter being able to play a role in enzymatic activity (Fig. 9.1).

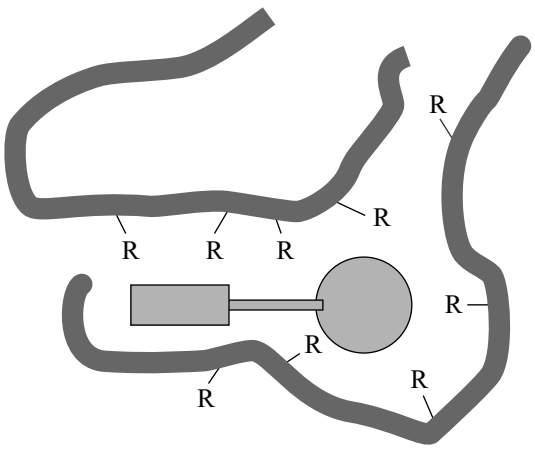


Fig. 9.1 Schematic representation of the active centre of an enzyme with the different residues R that interact with the substrate

Meanwhile, such an idea remains insufficient. It is purely spatial and does not take into account the functional aspect. Indeed, the residues of the enzyme that come into contact with the substrate can have very diverse roles, either interfering with the binding of the substrate or participating directly in the catalysis. Some do not play any role; they are present at the active centre by consequence of the polypeptide sequence and have direct influence neither on the enzyme-substrate association nor on the catalysis. Their modification or replacement by other amino acids in crossing from one species to another, for example, or replacement by site-directed mutagenesis, does not change their activity. On the contrary, certain amino acids which are not bound by VAN DER WAALS forces to the substrate can play an essen-

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_10, © Springer-Verlag Berlin Heidelberg 2010

tial role, favoring the binding of the substrate either by creating a microenvironment favorable to the functioning of the active centre or by maintaining its functional conformation.

Upon enzyme-substrate association, diverse events that bring about the displacement of certain enzyme atom groups can occur. For example, in carboxypeptidase, tyrosine 248 undergoes a rotation of 12 Å upon substrate binding; in the free enzyme this residue is situated near the molecule exterior far from the other residues of the active centre. For other enzymes, including hexokinase and phosphoglycerate kinase, there occurs a bringing together of domains upon substrate binding. It is important to specify these notions somewhat further. One must distinguish:

- the catalytic site that concerns the groups directly involved in the chemical act;
- the binding site that corresponds with enzyme groups establishing non-covalent interactions with the substrate;
- the conformational site that corresponds with specific residues important for maintaining an active conformation of the enzyme.

The active centre is constituted by the ensemble; it includes altogether the catalytic site, the binding site and the conformational site.

However, it is not always easy to know with precision the topology of the active centre. Several methods of investigation have been used for this aim. The kinetic approach consists of analysing variations in kinetic parameters as a function of pH; it has historically represented a first approach to the problem. The dependence of kinetic constants on pH is attributed to ionisation of enzymatic groups important for the activity. Nevertheless, as we will see, this method lets stand the most often an ambiguity concerning even the nature of these groups. It remains important for specifying their role but permits only to hint at their nature and not to specify it. The chemical methods to specifically label protein groups constitute precious instruments for determining the nature of essential groups. Genetic methods, in particular site-directed mutagenesis permitting the selective replacement of one amino acid by another, offer a very powerful tool complementary to chemical methods. The most precise topological information without any doubt is supplied by radiocristallographic studies and nuclear magnetic resonance. Structural determinations by X-ray diffraction require obtaining crystals of complexes that form an enzyme with substrate analogs. But even when the structure of the complexes has been solved, the functional studies are necessary to define the role of different residues at the active centre. Nuclear magnetic resonance studies have the advantage of being applicable to proteins in solution, but they are limited by the strong concentrations they require. Therefore, these different methods do not exclude each other but are complementary. These diverse approaches and the information that they bring are analysed successively.

9.1. KINETIC APPROACH – ANALYSIS OF pH PROFILES

This approach consists of analysing the influence of pH on enzymatic reactions. For a long time one had remarked that enzymes are active in a more or less limited pH range. In most cases, one defines even rather precisely an optimal pH (Fig. 9.2). The effect of pH on enzymatic reactions, like all pH effects, issues changes in the ionisation state of the system components as a function of pH. The existence of an optimum can result in different effects. Some of them are of a conformational order and are not connected to the ionisation of the active centre groups. One situation fairly often encountered results from the decrease in enzyme stability at extreme pH where a denaturation occurs that can be irreversible. This is relatively common with globular proteins. In other cases, reversible denaturations due to a variation of some charges of the molecule happen at extreme pH. In short, the more specific conformational effects are at times responsible for the variations in activity as a function of pH. In some cases, the ionisation of a single group or of a pair of groups plays a key role in the active conformation of an enzyme. Some examples will be given later.

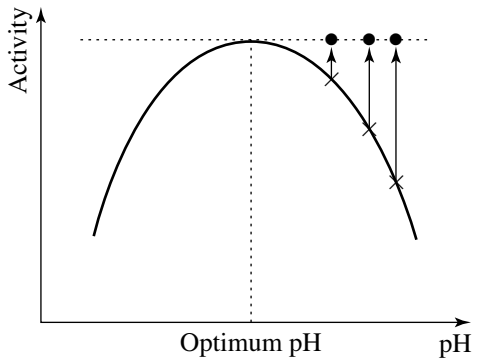


Fig. 9.2 Variation in enzymatic activity as a function of pH
The crosses represent the pH values of enzyme incubation; the points *indicate the activity after the return to optimal pH*

The action of pH on the simple enzyme-substrate association does not result in a variation in V_m, but modifies the parameter K_m. It can be assessed more directly by using a competitive inhibitor. The activity curve as a function of pH results also, at least in part, from the ionisation of catalytic groups. In these conditions, the effect of pH affects both parameters V_m and K_m.

Finally, if the substrate itself is susceptible to being ionised, and if its ionisation state moderates its binding, the enzyme affinity for the substrate will vary as a function of pH. It is convenient to distinguish experimentally these diverse effects. Some of them are easily analysed and must be approached at the beginning; first of all the conformational state of the protein, then the ionisation states of the substrate that are easy to distinguish separately. Moreover, it is important to analyse the role of ionisable enzyme groups in the association of the substrate and in the catalytic act itself.

9.1.1. EFFECT OF pH ON THE CONFORMATIONAL STATE OF THE PROTEIN

9.1.1.1. IRREVERSIBLE DENATURATION

Once the curve of global activity variation as a function of pH is established, it is easy to determine if the decrease in activity, either in alkaline or in acidic conditions, or in the two conditions, is due to an irreversible denaturation of the enzyme. The very simple experiment consists of pre-incubating the enzyme at these extreme pH values for a determined time, then placing it at the optimal pH to measure its activity. If, under these conditions, the activity stays constant whatever the pre-incubation pH and equal to the previously measured activity at optimal pH, an irreversible denaturation cannot be the cause of the decrease in activity at extreme pH. It is the same if the activity returns to a value superior to the value observed at the pH of the pre-incubation although inferior to the previously observed value at the optimal pH; one can be situated then in a case of slowly reversible denaturation (see Fig. 9.2).

9.1.1.2. REVERSIBLE DENATURATION

If the decrease in enzymatic activity at extreme pH is due to a reversible denaturation, the return of the activity to its initial value measured at optimal pH can be more or less slow. In many cases this phenomena is not immediate and it is possible to follow the kinetics. As renaturation is on the contrary a rapid phenomenon, the existence of a denaturation at extreme pH risks happening unnoticed. **In order to be exempt from this phenomenon, it is important to measure, as a function of pH, a physical parameter revealing the conformational state of the protein.**

9.1.1.3. IONISATION OF GROUPS THAT INTERFERE SPECIFICALLY WITH THE ACTIVE CONFORMATION OF THE ENZYME

As previously mentioned, in order to obtain a clear response on the role of ionisable groups that interfere with the activity of an enzyme, the conformational studies as a function of pH must be carried out in parallel with kinetic studies. Indeed, the active conformation of certain enzymes is maintained by a specific interaction between two charged groups. The changes in pH involving charge modification of one or another of these groups delete the interaction. It follows that either the substrate binding or the catalytic constant depends on the ionisation of these groups, although they interfere only in the conformation of the active site of the enzyme, and not in the catalytic act.

For several serine proteases, it has been shown that an ionic interaction secures the optimal conformation of the active centre. For chymotrypsin, this interaction is formed between the ⟨-amino group of isoleucine 16 at the N-terminal position of the enzyme in the NH_3^+ state and the carboxylate of aspartate 194 in the proximity of serine 195 of the catalytic site. For trypsin and elastase that are proteases of the same family, the same type of interaction exists. Each of these groups undergoes an

ionisation that induces a structural transition on both sides of the optimal pH. The amino group is ionised with an apparent pK of 10 for trypsin (CHEVALLIER & YON, 1966), of 8.9 for chymotrypsin (OPPENHEIMER et al., 1966); the carboxylate is ionised with a pK of 4 for trypsin and 3 for chymotrypsin. In alkaline medium, these enzymes exhibit therefore an equilibrium between a neutral form comprising the existence of a salt bridge that maintains the active conformation of the catalytic site and an alkaline form in which the salt bridge is broken. In acidic medium they are in equilibrium between the neutral form and the acidic form in which the interaction is equally broken (Fig. 9.3).

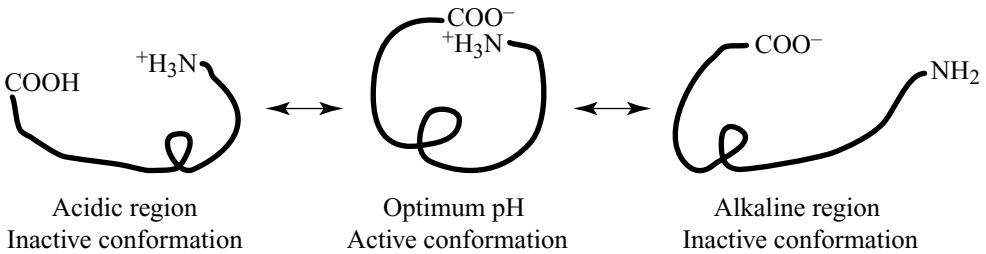


Fig. 9.3 Conformational equilibrium of serine proteases as a function of pH

The variations of diverse physical parameters such as rotatory power, fluorescence, and spectrum difference, as a function of pH reveal these two pK (Fig. 9.4). The results of crystallography studies have confirmed the existence of salt bridge in serine proteases. Figure 8.5 of the preceding chapter shows, according to the atomic coordinates, the salt bridge and the catalytic groups of chymotrypsin. The opening of the salt bridge drives the loss of the optimal conformation of the active centre; it results in important variations in the functional properties (see Part IV).

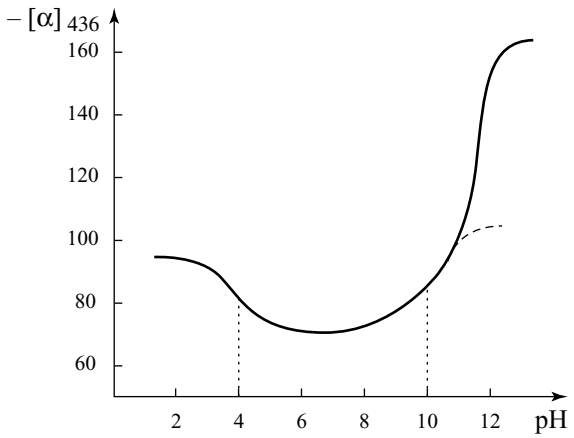


Fig. 9.4 The effect of pH on the conformational equilibrium of trypsin evaluated by variations in rotatory power at 436 nm
In the alkaline range, a more global denaturation prevails at the transition due to the deprotonation of the amino group

9.1.2. SUBSTRATE IONISATION STATES

The substrate itself can exist under different ionic forms in the enzyme activity range. Generally substrates are simpler molecules than enzymes and possess only a

limited number of ionisable groups. A direct titration permits the determination of the ionisation pK values independently of the presence of the enzyme.

◤ Let us consider the substrate:

$$SH_2^{++} \underset{}{\overset{K_{s1}}{\rightleftharpoons}} SH^{+} \overset{K_{s2}}{\rightleftharpoons} S$$

with its ionisation states. Several situations can arise. In one, the three forms are equally recognised by the enzyme without the charge interfering. In this case, the ionisation states of the substrate do not interfere with the activity curve as a function of pH. In another, a form is recognised preferentially by the enzyme, for example SH^+. The substrate fraction having the optimal affinity for the enzyme is in the form:

$$SH^+ = \frac{S_t}{1 + \frac{H^+}{K_{s1}} + \frac{K_{s2}}{H^+}}$$

The activity curve as a function of pH varies in exhibiting two values of pK, pK_{s1} and pK_{s2} which, if they have been identified by a direct titration, could be distinguished from pK values of enzyme group ionisation. When the affinity of the enzyme for the substrate depends on the ionisation state of the latter, this signifies that an electrostatic interaction interferes with the enzyme-substrate association and reflects itself in the variations in K_s as a function of pH. ◢

9.1.3. EFFECT OF pH ON ENZYME-SUBSTRATE ASSOCIATIONS

Certain enzyme-substrate associations involve some electrostatic interactions between one or several charges of the enzyme and the opposite charges of the substrate. Thus, the specific substrates of trypsin possess a positive charge on the ε-amino group of lysine or the guanidine group of arginine that interacts with the charge of aspartate 177 situated at the bottom of a hydrophobic pocket. The protonation of aspartate in acidic medium brings about a decrease in the enzyme affinity for its substrate.

◤ If one can know the parameter K_s (for example in the case of limiting acylation substrates for which $K_m = K_s$), the variations in K_s as a function of pH in the ionisation regions of this carboxyl group are represented by a sigmoidal curve as indicated in Fig. 9.5 opposite. In acidic medium the curve presents a plateau giving the value of $K_{s,max}$, that is in the absence of electrostatic interaction; in alkaline medium when the group is deprotonated one obtains $K_{s,min}$. The enzyme affinity for its substrate is maximal as it involves, in addition to some other interactions, the interaction between the charges of the enzyme and the substrate. Starting with these two values, it is easy to evaluate the electrostatic contribution to the energy of interaction. Indeed:

$$\Delta G_1 = RT \ln K_{s,max}$$

$$\Delta G_2 = RT \ln K_{s,min}$$

$$\Delta\Delta G = \Delta G_2 - \Delta G_1 = RT \ln K_{s,min}/K_{s,max}$$

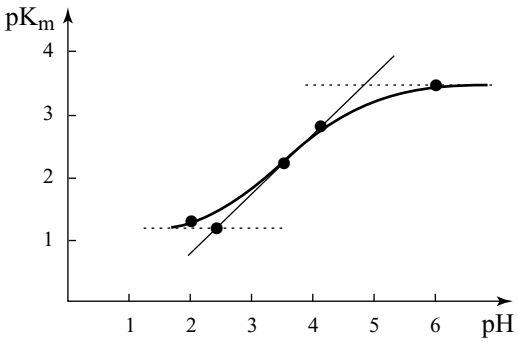


Fig. 9.5 Variations in K_m as a function of pH for a substrate of trypsin in which acylation is the limiting step in the reaction ($K_m = K_s$); DIXON plot

This value permits, in using COULOMB'S law, to calculate the approximate distance between the charges:

$$\Delta\Delta G = \frac{305 Z_1 Z_2}{d_e(6d_e - 7)} \times \frac{e^{-\kappa(r_0 - d_e)}}{1 + \kappa r_0}$$

In this expression Z_1 and Z_2 are the groups' charges that enter into the electrostatic interaction; d_e is the distance between these charges. The factor $(6d_e - 7)$ is a value approaching the dielectric constant. The exponential factor is a correction factor that takes into account the ionic atmosphere surrounding these groups, in which κ is the DEBYE and HÜCKEL constant equal to 0.7×10^7 for an ionic strength of 0.1; r_0 is the smallest ion approach distance estimated by a first approximation of 10.2 Å. This value includes the radius of the enzyme active group that is 1.7 Å; the radius of an ion included in the ionic atmosphere is 1.3 Å and the diameter of two molecules of water is 7.2 Å. ◢

It is generally preferable to resort to competitive inhibitor analogs of the substrate and to analyse the variations in pK_i as a function of pH (K_i, the inhibition constant being a true dissociation constant). In the case of trypsin inhibition by two specific competitive inhibitors, benzoyl arginine and tosyl arginine, the respective distances of 6.9 and 6.5 Å were determined in this manner (Fig. 9.6). For the inhibition of trypsin by benzamidine, the distance between the charges of the enzyme and of the substrate was evaluated to be 4.3 Å.

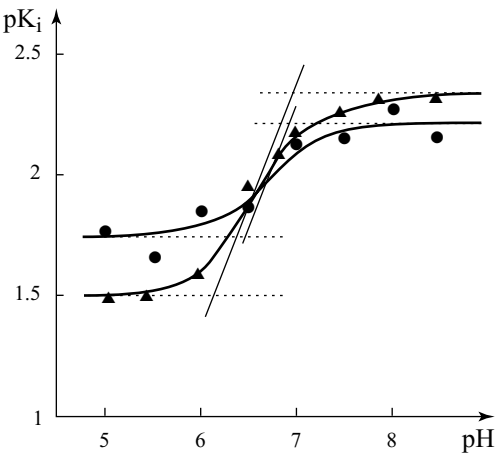


***Fig. 9.6 Variations in pK_i as a function of pH upon inhibition of trypsin hydrolysis of acetyl phenylalanine methyl ester by benzoyl argininamide* (triangles) *and tosyl argininamide* (circles)**

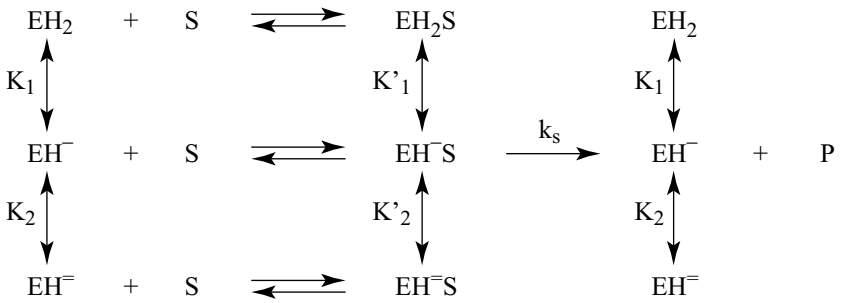
(Reprinted from *Biochim. Biophys. Acta*, **122**, BECHET J.J. *et al.*, Etude de l'inhibition de l'activité estérasique de la trypsine par des dérivés acides et amides des esters spécifiques, 101. © (1966) with permission from Elsevier)

9.1.4. EFFECTS OF pH ON THE IONISATION OF CATALYTIC GROUPS

Enzymes, like all proteins, contain a large quantity of ionisable groups. There exists a whole series of ionisation states of the enzyme and the distribution of the entire enzyme in its different ionic forms depends on the pH and the constants of ionisation of diverse groups. However, one had remarked for a while that the enzymatic activity is generally restrained over a fairly narrow range of pH, and one had deduced from this that only one of the ionic forms of the enzyme (or of a part of the enzyme) is catalytically active. Despite the size of the protein molecule, the number of ionisable groups interfering in the catalysis is generally restrained. In this paragraph, we will analyse the incidence of ionisation of these groups on the behaviour of kinetic parameters as a function of pH. Different authors including LAIDLER, ALBERTY, MASSEY, and DIXON have given the complete scheme of enzymatic reactions as a function of pH. To approach the methodology of such a study, we will consider first the simple case of an enzymatic reaction with a single substrate and involving a single enzyme-substrate intermediate of the MICHAELIS complex type; we will consider next the case of two intermediates, then that of several intermediates.

9.1.4.1. ENZYMATIC REACTIONS INVOLVING A SINGLE INTERMEDIATE

The general scheme of the reaction as a function of pH is written:



In such a scheme, only the ionic form EH^- of the enzyme and the form EH^-S of the complex are catalytically active. These forms are in equilibrium on the acid side with the forms EH_2 and EH_2S, respectively, and on the alkaline side with $E^=$ and $E^=S$.

The different ionic forms of the free enzyme are susceptible to associating with the substrate. K_1, K_2, K'_1, and K'_2 are the ionisation constants of the free enzyme and of the MICHAELIS complex. The reaction rate as a function of pH is in the general form:

$$v = \frac{k_s(pH)es}{K_m(pH)+s}$$

with:

$$k_s(pH) = k_{s,opt}\frac{1}{1+\frac{H^+}{K'_1}+\frac{K'_2}{H^+}}$$

and:
$$K_m(pH) = K_{m,opt} \frac{1 + \frac{H^+}{K_1} + \frac{K_2}{H^+}}{1 + \frac{H^+}{K'_1} + \frac{K'_2}{H^+}}$$

$k_{s,opt}$ and $K_{m,opt}$ are the kinetic parameters at optimal pH when the free enzyme is entirely under the form EH^- and the complex under the form EH^-S.

Study of variations of k_s as a function of pH

These expressions are useful under their logarithmic form. One has:

$$\log k_{s(pH)} = \log k_{s,opt} - \log\left[1 + \frac{H^+}{K'_1} + \frac{K'_2}{H^+}\right]$$

We generally obtain the profiles of pH of the form indicated in Fig. 9.7a.

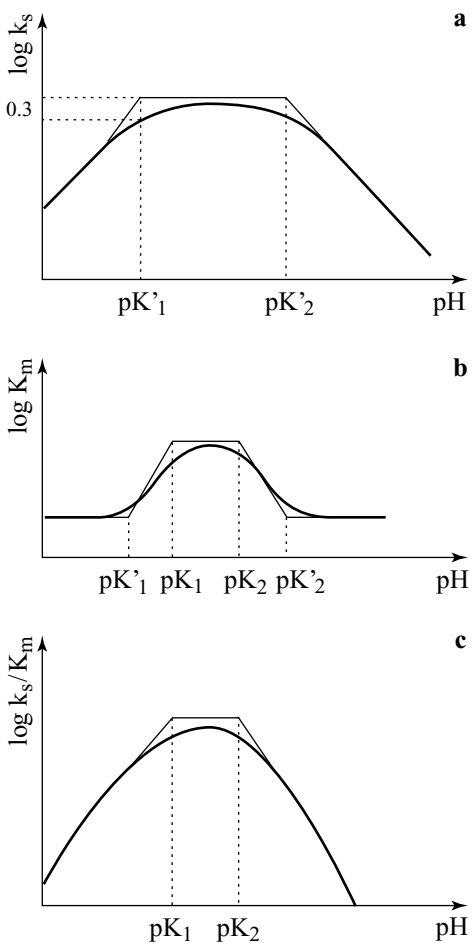


Fig. 9.7 DIXON plots
(a) *log k_s as a function of pH*
(b) *log K_m as a function of pH*
(c) *log k_s/K_m as a function of pH*

This curve permits obtaining the pK of the groups that are ionised in the MICHAELIS complex, meaning pK'_1 and pK'_2. It summons a certain number of observations known under the name of the DIXON rules.

The curve is formed by portions of straight lines joined by short curves, the curvature happens in the vicinity of the pK of ionisation. The slope of the lines is represented by integers, 1, 0, and –1 on the graph in Fig. 9.7a. The value of the slope corresponds with the number of protons implicated in the conversion of one ionic form to another. In the present case, a single proton is implicated in the passage of one ionisation state to another; the slope is equal to 1 or –1 following that the group is essential under deprotonated or protonated form. This results obviously from the preceding relationship. Indeed, at low pH $(H^+) > K'_1$ and *a fortiori* $(H^+) > K'_2$, so the equation simplifies into:

$$\log k_{s(pH)} = \log k_{s,opt} + pH - pK'_1$$

The line representing $\log k_{s(pH)}$ as a function of pH therefore has a slope of 1. At high pH $(H^+) < K'_2 < K'_1$ and one has:

$$\log k_{s(pH)} = \log k_{s,opt} - pH + pK'_2$$

The value of the pK of ionisation is given by the intersections of portions of lines. However, if the two values of pK'_1 and pK'_2 are not sufficiently distant, the plateau corresponding with the optimum is not determined with sufficient accuracy. Another imprecision occasionally comes from experimental difficulties in acidic or alkaline medium and it is not always easy to follow an insufficiency of experimental points and trace correctly the corresponding lines. In this case, the determination of the pK values meanwhile is possible. Indeed, in the vicinity of pK, the graph no longer follows a linear variation, but it curves with a downward concavity. This curve is situated at a vertical distance of 0.3 pH units below the intersection point of the lines if the ionisation of a single group interferes (see Fig. 9.7). Indeed, in the vicinity of pK'_1 for example, one has:

$$\log k_{s(pH)} = \log k_{s,opt} - \log\left(1 + \frac{H^+}{K'_1}\right)$$

At pK, $H^+ = K'_1$ and the equation becomes:

$$\begin{aligned} \log k_{s(pH)} &= \log k_{s,opt} - \log 2 \\ &= \log k_{s,opt} - 0.3 \end{aligned}$$

Variations in K_m as a function of pH

The expression of K_m is more complex as it yields the ionisation pK of catalytic groups of free enzyme and of the MICHAELIS complex. Its analysis follows nevertheless the same principles. The expression of $K_{m(pH)}$ in its logarithmic form is the following:

$$pK_{m(pH)} = pK_{m,opt} - \log\left(1 + \frac{H^+}{K_1} + \frac{K_2}{H^+}\right) + \log\left(1 + \frac{H^+}{K'_1} + \frac{K'_2}{H^+}\right)$$

It is interpreted by the curve given in Fig. 9.7b yielding four ionisation pKs. The rules of determining pK stated previously stay valid for the analysis of variations in pK_m as a function of pH. One supplementary rule occurs that derives from the equation above. In the vicinity of the pK of free enzyme, the graph presents a downward concavity; in the vicinity of pK of the complex, the concavity is upward. The analysis of variations in pK_m as a function of pH stays delicate. Indeed, the difference in the pK values between the free enzyme and the complex are often very weak and remain difficult to determine graphically. The pK′ values are given by the representation of log k_s as a function of pH.

There is a third type of graphical representation giving only the pK of the free enzyme. It results from studying variations in log (k_s/K_m) as a function of pH. Indeed, from preceding expressions of k_s and of K_m, it follows:

$$\left(\frac{k_s}{K_m}\right)_{pH} = \left(\frac{k_s}{K_m}\right)_{opt} \times \frac{1}{1 + \frac{H^+}{K_1} + \frac{K_2}{H^+}}$$

The corresponding curve is given in Fig. 9.7c.

Having determined the values of different ionisation pK that affect the enzymatic reaction, it is only possible to make some hypotheses on the nature of the groups implicated in the catalytic activity. One knows indeed the ionisation pK of principal groups titrated in the proteins. Table 9.1 recalls these data.

Table 9.1 Average values of pK and enthalpy of ionisation of protein side chain residues

Ionisable group	*pK*	*ΔH (kcal . mol^{-1})*
Carboxylate (Asp, Glu)	4–4.5	1.5
Imidazole (His)	6–6.5	6.9
α-amino	6–8	10–13
ε-amino (Lys)	9.2	10–12
Phenol (Tyr)	10	
Sulfhydryl (Cys)	8.5	
Guanidine (Arg)	12	12–13
Alcohol (Ser, Thr)	> 14	

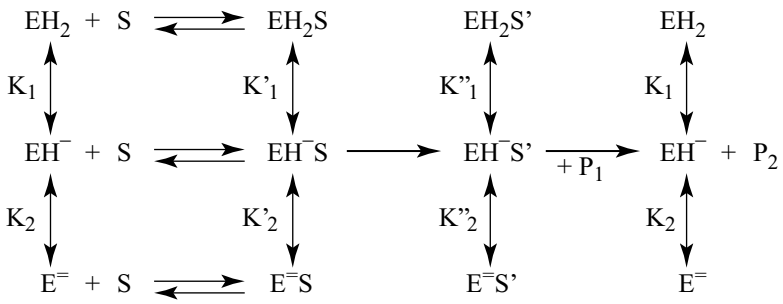
However, the kinetic study as a function of pH can only give an ambiguous response concerning the nature of the group; indeed, the pK of the ionisable groups in the proteins can be shifted depending on their environment. Hence for lactoglobulin, two carboxylate groups ionise with apparent pK neighboring neutrality; their

titration accompanies a conformational protein transition. For serine proteases, the apparent pK of ionisable groups participating in the establishment of the salt bridge that maintains the structure of the active site is displaced, as reported in Section 9.1.1.3 of this chapter. **The kinetic method gives mainly information on the role of the group.** It must always be accompanied by direct chemical determin ation of the implicated group or groups. It meanwhile permits making hypotheses that will direct the research of a reagent susceptible to labelling the essential residue. Moreover it requires, as we have already underlined, the use of physical methods to determine if the ionisable group is not implicated in a conformational event. The potentialities offered by site-directed mutagenesis techniques also permit such a determination.

9.1.4.2. ENZYMATIC REACTIONS IMPLICATING TWO INTERMEDIATE COMPLEXES

Like what was analysed in Part II, enzymatic reactions proceed via the formation of several intermediate complexes. The analysis of kinetic parameter variations as a function of pH then becomes more complex. In the classic case of two reaction intermediates and a single catalytic process ("bottle-neck mechanism"), like for example reactions catalysed by serine proteases that proceed via the formation of an acyl-enzyme intermediate, the analysis remains relatively simple.

◤ The reaction scheme as a function of pH can be written:



The constant K_1 and K_2 relate with ionisations of groups in the free enzyme, K'_1, K'_2 with those in the MICHAELIS complex and K''_1, K''_2 with those in the acyl-enzyme. In this scheme a single catalytic process is permitted, implicating the forms EH^-, EH^-S and EH^-S' that are the only catalytically active ionic forms; k_2 and k_3 are the constants of acylation and of deacylation, respectively.

Like what was presented in Part II (Chap. 5), the kinetic parameters of this scheme have complex values.

$$k_{cat} = \frac{k_2 k_3}{k_2 + k_3}$$

and:

$$K_{m,app} = K_m \frac{k_3}{k_2 + k_3}$$

The diverse constants depend on pH. Rather simple equations permit the expression of each value as a function of pH. Therefore k_2 and k_3 depend on the pK of the MICHAELIS constant and of the acyl-enzyme, respectively.

$$k_{2(pH)} = k_{2,opt} \frac{1}{1 + \frac{H^+}{K'_1} + \frac{K'_2}{H^+}}$$

$$k_{3(pH)} = k_{3,opt} \frac{1}{1 + \frac{H^+}{K''_1} + \frac{K''_2}{H^+}}$$

$$K_{m(pH)} = K_{m,opt} \frac{1 + \frac{H^+}{K_1} + \frac{K_2}{H^+}}{1 + \frac{H^+}{K'_1} + \frac{K'_2}{H^+}}$$

One knows that, whatever the pH is, there exists a simple relationship between the constants:

$$\frac{k_{cat}}{K_{m,app}} = \frac{k_2}{K_m}$$

It follows that the variations in this ratio as a function of pH permit the determination of the pK of groups that are ionised in the free enzyme:

$$\left(\frac{k_{cat}}{K_{m,app}}\right)_{pH} = \left(\frac{k_2}{K_m}\right)_{opt} \times \frac{1}{1 + \frac{H^+}{K_1} + \frac{K_2}{H^+}}$$

Whatever the rate-limiting step of the reaction is, it is therefore always possible to obtain the ionisation pK of groups in the free enzyme. In addition, with substrates for which k_2 is limiting, one can determine the ionisation pK for the MICHAELIS complex. For substrates in which k_3 is limiting, one can obtain the ionisation pK of catalytic groups in the acyl-enzyme. However, it is essential to ascertain that these stages remain limiting in all pH ranges studied.

9.1.4.3. ENZYMATIC REACTIONS INVOLVING SEVERAL INTERMEDIATE COMPLEXES

The preceding results can be generalised to n intermediates. When the kinetic scheme is such that there only exists a single catalytic process, the study of variations in the ratio k_{cat}/K_m as a function of pH permits the determination of ionisation pK of catalytic groups of the free enzyme. When certain steps are limiting, the variations in the rate constant of the limiting step as a function of pH permit the determination of ionisation pK of the complex upstream of the limiting step. If there is more than one catalytic process, i.e. if a catalytic group can be active under different ionisation states, the problem becomes much more complex. However, this situation is very unlikely.

9.2. CHEMICAL APPROACH TO STUDYING THE ACTIVE CENTRE OF ENZYMES

The kinetic and conformational approaches previously described drive to specify the role of ionisable groups, but concerning the nature of these groups the answer still remains ambiguous. In certain cases it can give some indications, yet it is necessary to provide a direct experimental proof.

9.2.1. PRINCIPLE OF CHEMICAL LABELLING

The side chains of a protein susceptible to interfering in catalysis and to be modified by a reagent are generally nucleophilic or electrophilic groups. It concerns essentially the imidazole group of histidine, the amine groups of lysine and arginine, the phenol group of tyrosine, the thiol group of cysteine, the sulfur of methionine, the carboxyl group of glutamate and of aspartate and, when it occurs at the active centre of enzymes, the functional alcohol group of serine and of threonine.

The choice of a reagent depends on the amino acid residue that one wants to modify. It can be directed by the results of kinetic studies performed at different pH values. It is important to know **that there do not exist specific reagents** that permit the modification of one type of residue at the exclusion of others, **but only selective reagents** bringing about the preferential modification of a category of groups in well-determined conditions. Besides, the protein milieu is heterogeneous with polar parts (essentially external) and apolar parts, creating locally a microenvironment that can considerably modify the reactivity of a group in the side chain. The surface of the protein, generally polar, interacting with the aqueous solvent, possesses very variable charge densities from one place to another. The protein interior is essentially constituted by non-polar residues forming a hydrophobic "core" with a weak dielectric constant. There is also inside a protein a large variety of local dielectric constants, electrostatic fields that make the same functional group able to present very different reactivities according to its localisation. All the effects of the microenvironment are of great importance for both the efficiency of enzymatic catalysis and the mechanism of action of the chemical reagents.

The particular microenvironment that exists at the active centre of an enzyme is responsible for the **hyperreactivity** of certain groups. For example, for serine proteases the catalytic serine is abnormally reactive, whereas the other serine residues do not present any reactivity. This hyperreactivity is created by the microenvironment particular to the active centre. Meanwhile, some other residues like those implicated in catalysis could present hyperreactivities with certain chemical reagents. Therefore, among the 19 tyrosines of papain, only tyrosine 123 reacts rapidly with diisopropyl phosphofluoridate. However, the catalytic SH group in the same enzyme and in other thiol proteases is not phosphorylated, but it is hyper-reactive

with cyanate. Glutamate 35 of lysozyme to which is attributed a catalytic role is the only carboxyl group of the enzyme that is resistant to chemical modifications.

The reactivity, the nucleophilic power, of a functional group is determined by the electron density of the implicated atoms, polarisability and steric factors that limit the access of the reagent. Hyperreactivity is therefore not linked to an exaltation of nucleophilic power itself, but to external factors of the functional group. All the external factors included in the effects of the microenvironment created by the protein matrix affect not only the reactivity of functional groups, but also those of the reagent. The principal effects are analysed successively.

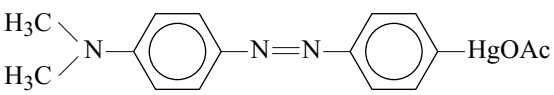
9.2.1.1. EFFECTS OF THE MICROENVIRONMENT ON THE PROTEIN FUNCTIONAL GROUPS

The reactivity of a functional group corresponds with its nucleophilic power that is generally but not always reliant on its basic properties, meaning its ionisation pK. There is a relationship between the nucleophilic power of a group and its pK of ionisation, the BRØNSTED relationship. Yet in a protein one observes important variations in the apparent pK of ionisable groups; these values are sometimes very different from those that are in the denatured protein or when these groups are carried by simple substituted amino acids. It is a question of "abnormal" pK. It results that the reactivity of a protein side chain toward a chemical reagent cannot be anticipated starting with data obtained from models.

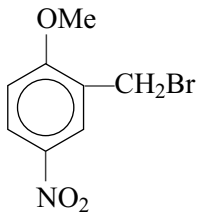
In some cases, the reactivity of a group was compared in the protein and in the model compound. Thus histidine at the catalytic site of ribonuclease is alkylated 500 times more quickly by bromoacetate at pH 5 than is N-acetylhistidine. The catalytic thiol group of papain is 3 000 times more reactive than cysteine is with cyanate and 15 000 times more with bromoacetate.

Polarity of the microenvironment

The polarity of the local environment affects reactivity in a way dependent on the reaction. The effect of an electric field in the vicinity of a charged group can modify the ionisation pK and reactivity of this group either in one direction or in the opposite direction according to the charges present. Diverse approaches have been proposed to evaluate the microenvironment polarity, in particular the use of reporter groups, using either ionisation pK titration data or spectral data of these molecules introduced as extrinsic probes of polarity. Hence the binding of azo-mercurials to the single SH of bovine serum albumin lowers the pK of the reporter group anilino nitrogen to 1.9 in the protein, whereas when the compound is coupled to cysteine it is 3.3.



It is possible to separate polarity effects from electrostatic effects by using certain reagents insensitive to pH effects but having chromatic properties that depend on the solvent polarity. In this way 2-methoxy-5-nitrobenzyl has been used by KOSHLAND in the case of chymotrypsin:



As a consequence of the presence of the aromatic nucleus, this compound binds to the binding site of chymotrypsin substrates and alkylates methionine 192. Its spectrum of absorption is very sensitive to the polarity of the medium; it presents a maximum at 288 nm in hexane and at 317 nm in dimethylsulfoxide. Molecules involving a spin label like a nitroxide radical have been also introduced as reporter groups. Studies in electron paramagnetic resonance then give information on the microenvironment and its variations over the course of an enzymatic reaction, or upon ligand binding for example.

Formation of hydrogen bonds

Some ionisable residues in proteins can form hydrogen bonds, which drive the displacements of ionisation pK and generally a decrease in reactivity. Different hydrogen bonds can be established between residues and side chains (Fig. 9.8).

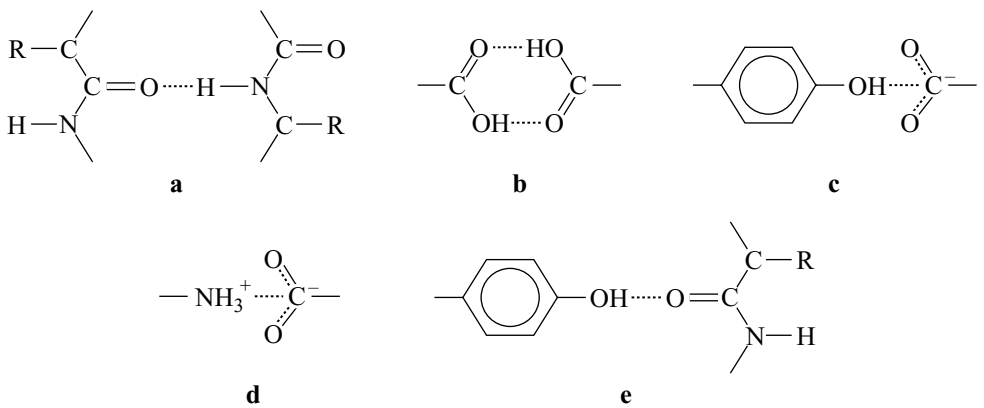


Fig. 9.8 Principal hydrogen bonds in proteins
*(**a**) interamide – (**b**) between two carboxylates – (**c**) between a charged residue and a protonated residue – (**d**) between two charged residues (ionic interaction) – (**e**) between the oxygen of the peptide bond carbonyl and a protonated residue*

Following that the implicated group acts as the donor or acceptor, its ionisation pK increases or decreases when it acquires a hydrogen bond. The equilibria are:

$$DH,A \overset{K_h}{\rightleftharpoons} DH\cdots A$$

$$DH \overset{K_1}{\rightleftharpoons} D^- + H^+$$

$$A + H^+ \overset{K_2}{\rightleftharpoons} AH^+$$

K_h, K_1, and K_2 being the constants of hydrogen bond formation and ionisation of groups of the donor DH and acceptor A, respectively. The apparent ionisation constant of the donor group is given by the equation:

$$K_{obs} = \frac{(D,A)(H^+)}{(DH,A)+(DH...A)}$$

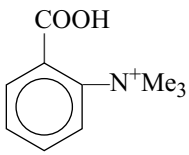
$$K_{obs} = \frac{K_1}{1+K_h}$$

(D,A) being the pair of donor and acceptor groups under their non-protonated form. The ionisation constant of the acceptor group is:

$$K_{obs} = \frac{K_2}{1+K_h}$$

Electrostatic effects

The presence of positive charges in the vicinity of an ionisable group disfavors its protonation, either by stabilisation of an ion pair (if the group is negatively charged), or by avoiding the formation of a new positive charge. Thus the pK of the carboxyl group in the following compound:



is 1.37 whereas it is 3.45 in the meta and para isomers.

The electrostatic effects can also create attractive or repulsive effects on the reagent.

Steric effects

Among the numerous effects due to the protein matrix, one notes sometimes a total steric protection rendering the group completely inaccessible whatever its size; only denaturation of the protein permits the group to react with a given reagent. Most often a partial protection is observed; the accessibility depends then on the size of the reagent. In some cases, the protein matrix limits freedom of rotation of side chains and therefore decreases their reactivity by preventing them from adopting an optimal orientation.

Effect of charge transfer

Spectroscopy techniques have revealed, in some cases, the participation of protein aromatic amino acids in charge transfer complexes with coenzymes, substrates or reagents. Charge transfers in systems with π electrons are susceptible to modifying reactivity in different directions.

Diverse effects

Diverse other factors susceptible to modifying the reactivity of protein amino acid side chains must still be mentioned. The formation of metallic complexes tends to weaken the reactivity of a nucleophilic group; indeed, the metal attractor has a tendency to diminish the electron density of the nucleophile. The establishment of reversible covalent bonds such as those driving the formation of a SCHIFF base or of a hemiacetal, tend to render the group provisionally non-functional or to diminish its reactivity.

9.2.1.2. EFFECT OF MICROENVIRONMENT ON THE REAGENT

The protein matrix has not only an effect on the intrinsic reactivity of amino acid residues, but it can also considerably modify the reactivity of the chemical compound used for the selective labelling in different manners and in a way that it is not always possible to foresee.

Selective adsorption of the reagent

Selective adsorption of the reagent on the enzyme before the reaction increases its reactivity by effect of proximity. This is equivalent to locally increasing its concentration. Formation of a complex between the protein and the reagent could induce a saturation effect on the reaction rate (to see later). This presents an advantage for the reaction that is particularly useful in chemical affinity labelling.

Steric effects

A partial steric protection of a residue by the protein matrix risks to notably modify the reactivity of the chemical compound; this protection depends on the size of the reagent. The presence of a substrate, an analog or a coenzyme can prevent the chemical compound from modifying a residue that usually is very reactive.

Electrostatic effects

Electrostatic interactions between the reagent and the charged protein groups or the electrostatic fields situated in the vicinity often create attractive or repulsive effects with the chemical compound following the nature of this last and the nature of the charges.

Proximity effects

The presence of other functional groups situated in the vicinity sometimes has the effect of increasing or on the contrary of diminishing the efficiency of the reagent.

Conformational effects

The presence of the reagent involves in some cases conformational restrictions limiting the rotational freedom of the residue and prohibiting it from taking the necessary conformation for the reaction to happen favorably.

The large number of factors capable of interfering makes the action of a reagent on a protein unforeseeable.

9.2.1.3. REQUIRED CONDITIONS FOR CONDUCTING A CHEMICAL MODIFICATION

Due to its complexity the protein structure makes the chemical approach delicate; also numerous controls are necessary for all the experimental steps. Taking into account all the implicated factors and the variety of proteins, there are not universal methods that permit deciding the choice of a reagent. In addition, when one wants to use a chemical reagent to label the active centre of an enzyme, it must not be forgotten that proteins have a relatively weak conformational stability and that they are susceptible to denaturation. However, chemical reactions on models are often carried out very efficiently in conditions that are hardly favorable to the protein structure. Between two reagents of a given group, one will choose that which is soluble in aqueous medium if it exists, or that which only necessitates a weak proportion of organic solvent; still it will be necessary to verify that this latter does not alter the conformation of the protein.

To be interpreted with a minimum of ambiguity, a chemical modification must only be carried out on a small number of groups, a single one if possible. The stoichiometric inactivation of an enzyme renders the interpretation much easier. *The labelling must be selective and limited.* It is important that the reaction conditions be the gentlest possible, both to limit the labelling and to preserve the native structure of the protein. The reaction can be oriented by different factors, in particular the pH. For example, numerous reagents of methionine that are non-selective in neutral medium become so in acidic conditions, because methionine does not ionise as the other residues that, in these conditions, are protonated and are no longer efficient nucleophiles. The variations in pH or in polarity of the medium can also reversibly modify the protein conformation and render accessible certain groups that are not so at optimal pH. Reversible stoichiometric labelling represents without any doubt, when it is possible, one of the most favorable situations.

The goal being to label functional residues of the active site and only those, the reagents used must be the most selective possible. In addition, it is important to find good conditions so that the labelling occurs selectively.

◤ The technical aspect of chemical modifications must be considered carefully. As a general rule, the study of the topology of an enzyme active centre can only be carried out with very pure proteins. It is preferable that at least the primary structure be known; this permits the identification not only of the nature of the modified residue, but also its position in the polypeptide sequence. This identification necessitates an important analytic work involving cleaving the protein into smaller fragments, the isolation of these peptides, the analysis of peptide maps, and the analysis of the peptide (or peptides) labelled by the reagent. For one uses radioactive reagents and one executes the analysis on the peptides thus labelled. The development of chromatography techniques, in particular the introduction of HPLC, has considerably increased resolution and the rapidity in separation of peptides. ◢

Knowledge of the 3-dimensional structure of a protein, allowing to "mimic" the chemical labelling by molecular modelling and use of energy minimisation methods, offers a remarkable tool for approaching a chemical modification as rationally as possible.

9.2.2. STRATEGY OF CHEMICAL MODIFICATIONS

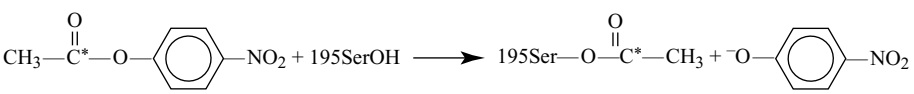
We will analyse successively different strategies used to attempt to selectively label the active centre of enzymes.

9.2.2.1. LABELLING BY A SUBSTRATE, A QUASI-SUBSTRATE OR A COENZYME

One of the reagents susceptible to binding at the active site and forming a covalent bond is the substrate itself or one of its less reactive analogs. Indeed, certain enzymatic reactions proceed via the formation of a covalent compound between one part of the substrate and a residue of the catalytic site (see Chap. 5). This intermediate is generally unstable in the conditions of optimal activity of the enzyme; its lifetime is very short. The labelling by a substrate is only possible in conditions, if they exist, that permit the accumulation of this intermediate.

◤ One of the most classical examples is the acylation of serine 195 of serine proteases by titrants of active sites, like paranitrophenyl acetate (pNPA) or paranitrophenyl guanido benzoate (NPGB) for trypsin. The use of these compounds for the titration of active sites of proteases has been presented previously (see Chap. 5). At pH lower than 5, the acetyl-enzyme formed during the hydrolysis of pNPA by trypsin or chymotrypsin is stable.

In using pNPA labelled with [^{14}C], one recovers the radioactivity on serine 195 after hydrolysis of the enzyme:



This labelling is therefore very selective since it only modifies serine 195. The enzyme thus labelled is totally inactive. Nevertheless its activity can be restored by alkalinisation to the optimal activity pH; in these conditions the acetyl-enzyme is hydrolysed and the active enzyme is regenerated.

Another example of the use of a quasi-substrate is the labelling of triose phosphate dehydrogenase by acetylphosphate. The physiological substrate of this enzyme is 1,3-diphosphoglycerate. Acetylphosphate is a quasi-substrate as paranitrophenyl acetate is for serine proteases.

In the case of enzymatic reactions that proceed *via* the formation of a phosphoryl-enzyme, the phosphorylated residue was identified. For phosphoglucomutase, the active centre serine implicated in the phosphorylation was labelled using glucose-6-phosphate ^{32}P as the substrate. The enzyme obtained after purification is already phosphorylated and catalyses the following reaction:

$$\text{G-1-P} \xleftarrow{\textit{phosphoglucomutase}} \text{G-6-P}$$

The addition of the substrate G-6-P favors the reaction in the reverse direction that can be thus split into:

$$\text{P-enzyme} + \text{G-6*P} \rightleftharpoons \text{enzyme} + \text{G-1,6*-diP}$$

$$\text{G-1,6*-diP} + \text{enzyme} \rightleftharpoons \text{*P-enzyme} + \text{G-1-P}$$

The implicated phosphorylserine was identified in this way (Fig. 9.9). The same type of direct labelling experiment by the phosphate of a specific substrate was achieved with phosphoglycerate mutase. There are some other examples of labelling by phosphorylation of a serine residue: the conversion of phosphorylase b in phosphorylase a by phosphorylase kinase in the presence of $\gamma[^{32}P]$ ATP. Alkaline phosphatase is phosphorylated during the catalytic cycle. These phosphoryl-enzymes are sufficiently stable so analysis of the group labelled by $[^{32}P]$ is possible.

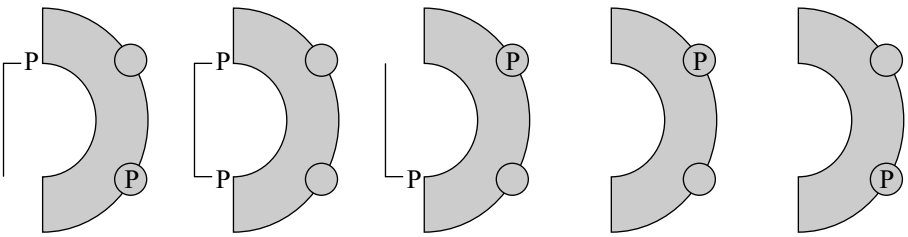
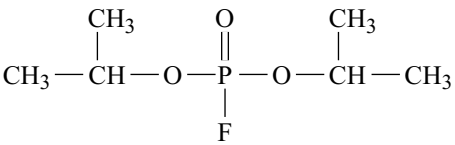


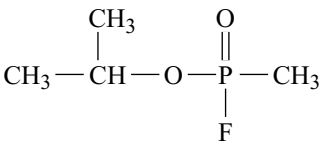
Fig. 9.9 Phosphorylation of the active site of phosphoglucomutase
(Adapted with permission from *Biochemistry*, **4**, HARSHMAN S. & NAJJAR V.A., 2526.

The labelling method by a quasi-substrate consists therefore of using a molecule that resembles a specific substrate. The preceding compounds are either true substrates or quasi-substrates that present the advantage of driving reactions that can be reversed. Some other quasi-substrate compounds of which the structure and reactivity differ from those of physiological substrates drive irreversible reactions. This is the case with certain phosphorylated derivatives that behave as powerful

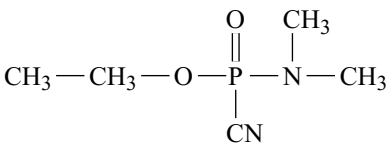
inhibitors of proteases and serine esterases; let us cite for example diisopropyl phosphofluoridate (DFP), sarin, tabun, diethyl p-nitrophenyl phosphate and tetraethyl pyrophosphate (TEPP) presented in Fig. 9.10. All these compounds form a phosphoryl bond with the single serine residue of the active centre of a great number of esterases. They are particularly powerful inhibitors of acetylcholinesterase, and in this way some of them were largely used as combat gas. There exist some cases where the covalent intermediate is unstable but may be stabilised by a supplementary reaction. Alkyl- and aryl-sulfonyl fluorides are powerful serine protease inhibitors. Phenylmethylsulfonyl fluoride (PMSF) is an inhibitor even more powerful than DFP. It is often used during enzyme purifications to avoid protein degradation by proteases present in the cell extracts.



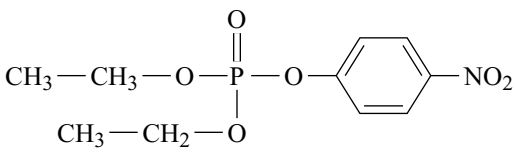
Diisopropylfluoro-phosphate (DFP)



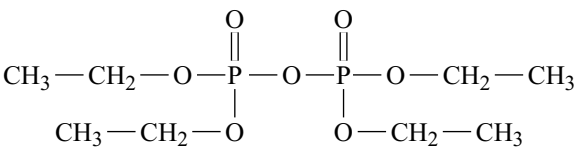
Sarin



Tabun



Diethyl p-nitrophenyl-phosphate



Tetraethyl pyrophosphate (TEPP)

Fig. 9.10 Several phospho derivatives, irreversible inhibitors of serine proteases and esterases

These inhibitors block the serine of the active site in an irreversible manner. Meanwhile to identify this, a supplementary chemical reaction must be used [GOLD, 1965]; phenyl methyl sulfonyl chymotrypsin is easily degradated with the loss of the phenyl methyl sulfonyl group. In the presence of a nucleophile such as 2-mercaptoethylamine in urea 8 M, the O-sulfonyl serine group is partially converted to S-amino ethyl cysteinyl which, being stable, was identified (Fig. 9.11 opposite). The stabilisation by an additional chemical reaction was used to identify the groups implicated in the catalytic site of enzymes which, over the course of the reaction, form a SCHIFF base with their substrate.

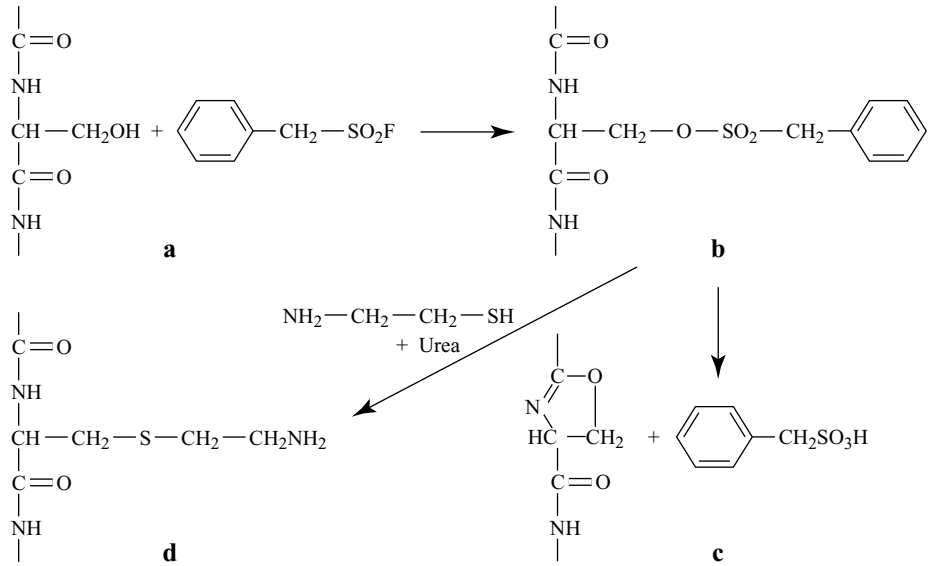
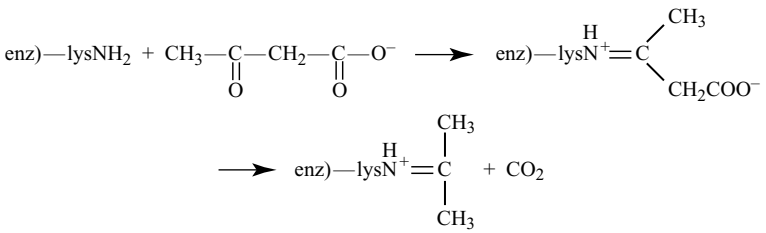


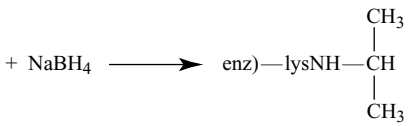
Fig. 9.11 Identification of serine labelled by PMSF

*(**a**) labelling reaction that leads to the compound (**b**), which can be easily degradated in (**c**) – in (**d**) use of 2-mercaptoethylamine in urea 8 M that transforms the O-sulfonyl serine group in S-aminoethyl cysteine*

◤ This is the case with acetoacetate decarboxylase that catalyses the decarboxylation of a keto acid, giving rise to a SCHIFF base following the reaction:



The unstable SCHIFF base can be stabilised by reaction with sodium borohydride:



The use of tritiated borohydride has permitted the isolation and identification of an isopropyl-ε-lysine showing the catalytic role of a lysine residue in this enzyme. The particular role played by a lysine at the catalytic site of aldolases and transaldolases was demonstrated in this manner.

Likewise the reduction by sodium borohydride has shown the role of a lysine in the binding of pyridoxal phosphate (PLP) in a certain number of enzymes, in particular in transaminases like aspartate amino transferase. The coenzyme PLP is bound to the protein by a lysine group as a SCHIFF base (Fig. 9.12 below). By an analogous ap-

proach, PLP allowed the identification of the active site of aspartate transcarbamylase even though it is not a cofactor of this enzyme. It behaves as a competitive inhibitor forming a SCHIFF base with a lysine residue localised at the active centre (KEMPE & STARK, 1975).

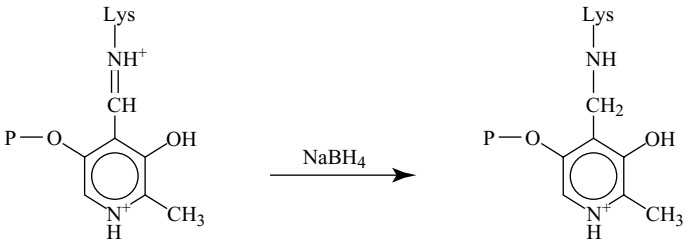


Fig. 9.12 Reduction by sodium borohydride of the SCHIFF base formed between a lysine of aspartate amino transferase and the pyridoxal phosphate ◢

However, there exist some cases where the use of a quasi-substrate does not result in labelling of a group implicated in the catalysis. Let us cite for example the labelling of chymotrypsin by bromoacetyl aminobutyrate. This reagent is supposed to acylate the single serine of the active site, which happens in a first step. Following this there exist two possible paths, either the simple deacylation like for an ordinary substrate, or the alkylation of another group, methionine 192, by the bromoethylketone with consecutive cleavage of the acyl-serine bond (Fig. 9.13 opposite). The labelled methionine is a non-catalytic residue; the serine is finally free and the enzyme thus modified presents still a weak catalytic activity.

9.2.2.2. AFFINITY LABELLING

Affinity labelling uses the structural characteristics of substrates to direct a reactive function (acylating or alkylating) to the active centre of an enzyme; these reagents are called *active-site directed reagents*. They require therefore a part possessing the structure of specific substrates on which is grafted a reactive function. The effective concentration of the reagent at the active centre is thus considerably increased.

In this case, like for quasi-substrates, the reaction proceeds *via* the formation of a complex between the reagent and the enzyme:

$$E + R \underset{}{\overset{K}{\rightleftharpoons}} ER \xrightarrow{k} E\text{-}R$$

R being the reagent and E-R the modified enzyme, K the dissociation constant of the complex ER and k the specific rate constant of the chemical reaction. The reaction rate is then given by a relationship implicating a Michaelian saturation:

$$v = \frac{ker}{K + r}$$

whereas usually one has a second order reaction:

$$v = ker$$

r being the concentration of the reagent and e that of the enzyme. If the reagent is in excess such that the variation in its concentration is negligible over the course of the chemical reaction, one has:

$$v = k'e \qquad \text{with } k' = kr = Ct$$

k′ is a constant of pseudo-first order (see Chap. 4).

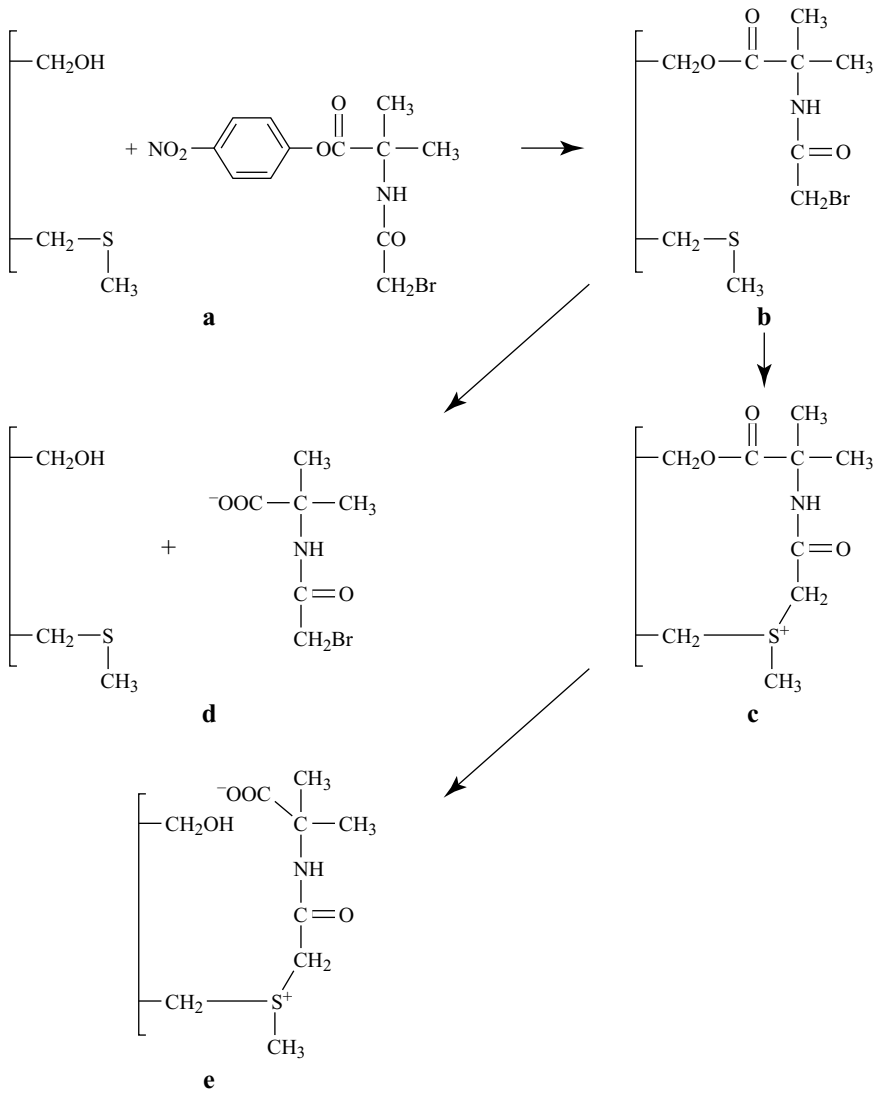


Fig. 9.13 Labelling of chymotrypsin by bromoacetyl amino isobutyrate

One of the first applications of this method permitted SCHOELLMANN and SHAW (1963) to demonstrate the catalytic role of His 57 in chymotrypsin. The reagent was tosyl phenylalanine chloromethyl ketone or TPCK. Figure 9.14 shows the analogy between a specific substrate, tosyl phenylalanine methyl ester (TPME) and TPCK. The latter provokes the progressive disappearance of the enzymatic activity with the correlated labelling of histidine 57. At the end of the reaction the inhibition is complete. The reaction is slowed by the presence of a competitive inhibitor like phenylpropionate (Fig. 9.14).

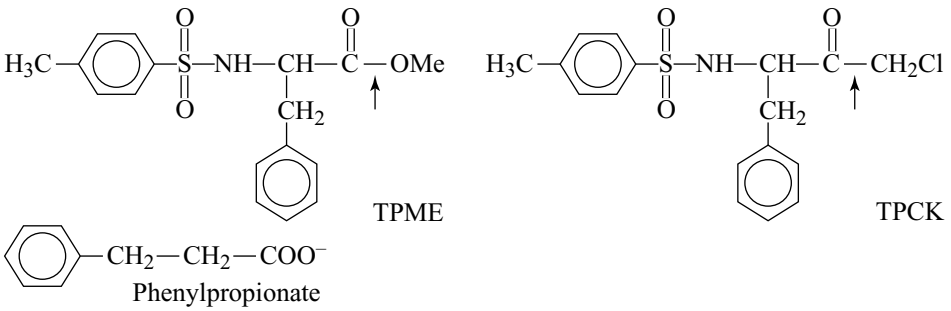


Fig. 9.14 TPCK, affinity reagent of chymotrypsin

TPCK does not bring about any inhibition of the activity of trypsin. On the other hand SHAW et al. (1965) showed that its homolog tosyl lysine chloro methyl ketone (TLCK) is specific for trypsin and does not inhibit chymotrypsin (Fig. 9.15). These compounds permit therefore a selective inhibition of one or the other enzyme. Thus some traces of trypsin present in a commercial preparation of chymotrypsin can be specifically inhibited by TLCK, and reciprocally TPCK is used to inhibit all possible traces of chymotrypsin in a preparation of trypsin.

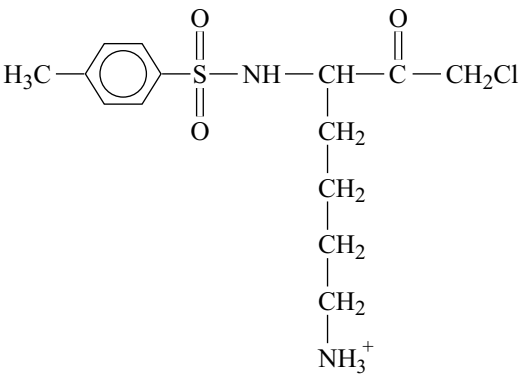


Fig. 9.15 Structure of TLCK, affinity reagent of trypsin

◤ Proteolytic enzymes were not only the first but also the most studied for directing the synthesis and use of affinity reagents. This strategy was progressively developed and such compounds were introduced to study other enzymes. Thus, at the same time, despite the difficulty to obtain pure acetylcholinesterase in sufficient quantity,

some affinity reagents of this enzyme were synthesised. Esters of methyl sulfonate possessing a quaternary ammonium such as the compound (a) (Fig. 9.16) inactivate the enzyme irreversibly. We have seen in the preceding paragraph that the quasi-substrates like organophosphates are powerful inhibitors of acetylcholinesterase. Certain carbamates like neostigmine (compound (c), Fig. 9.16) provoke their inhibition by carbamylation of the serine of the active centre.

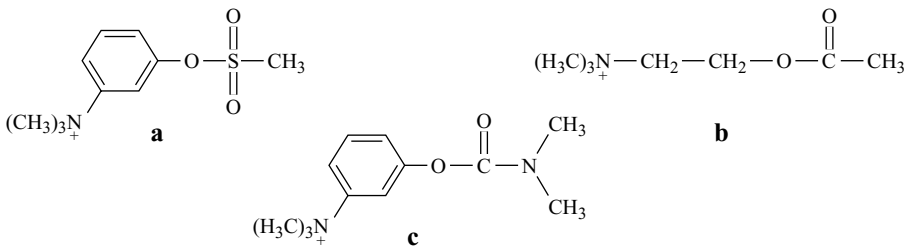


Fig. 9.16 Affinity reagents of acetylcholinesterase (a) and (c); the compound (b) is the substrate

Some active site-directed reagents of adenosine deaminase in which the bromoacetamide chain was branched in the *ortho, para* or *meta* position of 9-benzyl adenine (Fig. 9.17) were synthesised. The *ortho* and *para* isomers completely inhibit the enzyme whereas the substituted *meta* compound only provokes a partial loss of activity (80%); it is likely that, in this last case, the alkylation site is different from that of the *ortho* and *para* compounds. **Not only is it important that the reagent presents a specific part for directing it to the active centre, but also the orientation of the chemical function susceptible to label a catalytic group must be favorable.**

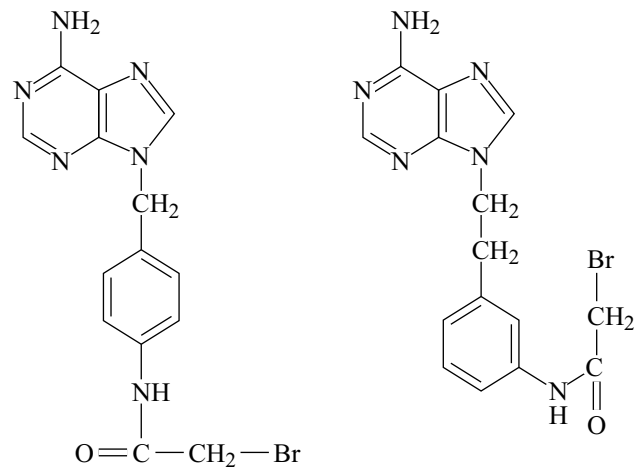


Fig. 9.17 Affinity reagents of adenosine deaminase

◤ Numerous active site-directed reagents were synthesised to inhibit glycosidases (see the review by LALLEGERIE et al., 1982). On some of them an epoxide function is grafted, which is the case for inhibitors of α- and β-glucosidases (compounds (a) and (b), respectively, in Fig. 9.18) and for an inhibitor of lysozyme on which has been grafted β-epoxy-propyl aglycone (compound c). In all these cases the labelled group was a carboxyl. Certain isothiocyanate derivatives, like β-D-gluco pyranosyl isothiocyanate (compound d) are powerful inhibitors of β-glucosidases.

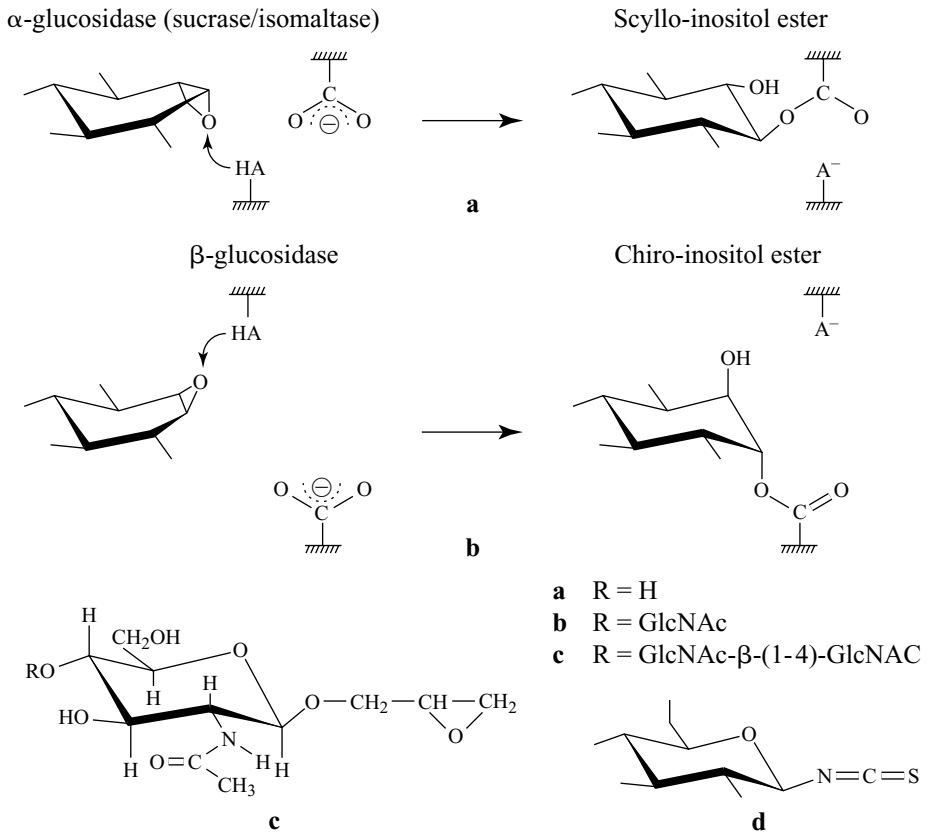


Fig. 9.18 Affinity reagents of several glycosidases ◢

Although they bind to the binding site of substrates, affinity reagents do not necessarily react with a catalytic group.

◤ Some reagents of chymotrypsin (Fig. 9.19 opposite) like benzyl bromide (a), and phenacyl bromide (b) modify methionine 192 which is near the active site but is not implicated in the catalysis, and the labelling only provokes 50% inactivation. The presence of specific substrates or inhibitors meanwhile brings about a protection of the enzyme with respect to the chemical labelling. **It is important therefore to interpret the results with caution.**

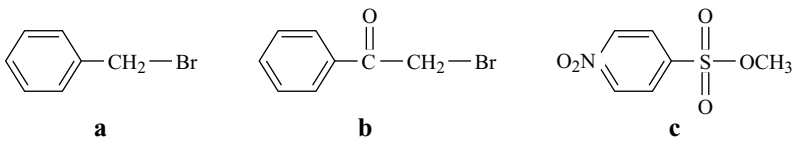


Fig. 9.19 Affinity reagents of chymotrypsin reacting with methionine 192

N-bromoacetyl-β-D-galactosylamine inhibits β-galactosidase in an irreversible way by alkylating methionine 500. Nevertheless this residue is not implicated in the catalysis. Bacteria grown in the absence of methionine, this being replaced by selenomethionine, synthesise a fully active β-galactosidase. ◢

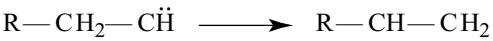
The use of affinity reagents is not limited only to the determination of catalytic groups. In some cases, it permits at the active site the introduction of a fluorescent reporter group or a spin label, comprising for example a nitroxide radical. Let us cite for example 2,2,5,5-tetramethyl-3-carboxypyrolidine-1-oxide phenyl ester that binds at the active site of chymotrypsin. This label allows following, by variation in fluorescence or by electron paramagnetic resonance (EPR), variations in the environment of the active centre under the influence of diverse effectors.

9.2.2.3. PHOTOAFFINITY LABELLING

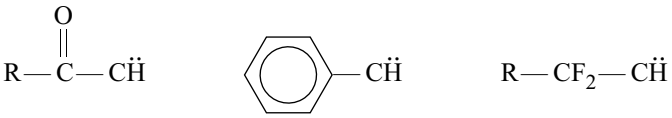
The principle of photoaffinity labelling resembles that of affinity labelling in the sense that the reagent possesses a specific part that permits it to interact with the active centre of the enzyme, thus serving as a vector to direct a chemical function near the active site. But the difference from affinity labels is that the chemical function is not initially reactive; it is susceptible to being activated by photolysis. The activation must be carried out in gentle conditions and at wavelengths such that the protein does not undergo photochemical damage. The derivative activated by photolysis must be highly reactive, not suffer molecular rearrangements producing a less reactive compound and have a very short lifetime. These reagents are capable of chemically modifying inert groups like methyls and methylenes in side chains of hydrophobic amino acids alanine, valine, leucine, isoleucine etc.

There exist different types of photoactive groups; these are precursors of carbenes and nitrenes. Irradiation of the precursors brings about the formation of excited electron states driving the appearance of reactive species.

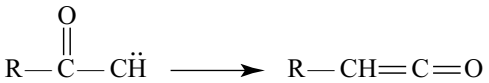
Carbenes react very rapidly with a great variety of chemical functions, either with a nucleophile group to give a carbanion, or by addition of multiple bonds including those of aromatic compounds, by insertion in single bonds including C—H bonds, or by abstraction of a hydrogen. If the adjacent carbon carries a hydrogen, this latter migrates easily bringing about the formation of a non-reactive compound:



Therefore, the adjacent atom must not carry a hydrogen. One generally chooses compounds of the type:

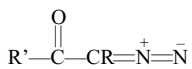
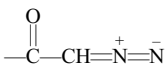
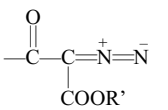
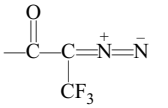
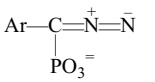
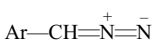
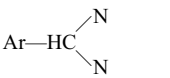


Acetocarbenes can undergo an intramolecular rearrangement or a WOLFF rearrangement driving a ketene:



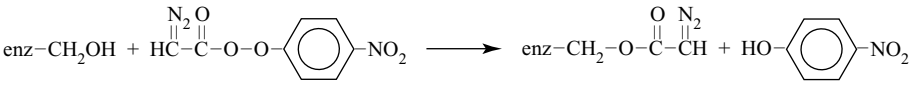
The ketene thus formed is reactive but can only interfere with its acylating function that reacts with a nucleophile group whereas carbenes present multiple reactivities. Precursors that produce carbenes by a photochemical reaction are indicated in Table 9.2.

Table 9.2 Principal precursors of carbenes (Reprinted from *Methods in Enzymol.*, **46**, BAYLEY H. & KNOWLES J.R., Photoaffinity labeling, 69. © (1977) with permission from Elsevier)

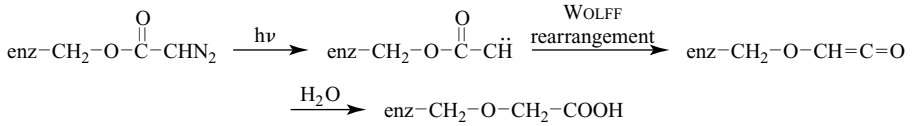
Type of reagent	*Formula*	*Stability in the dark in neutral solution*	*Sensitivity to intramolecular rearrangement after photolysis*
α-diazocetones		limited	high
α-diazocetyl-		limited	high
α-diazomalonyl		fair	reasonable
Trifluorobenzyl-phosphate		good	rather weak
α-diazobenzyl-phosphate		weak	—
Aryldiazomethane		weak	weak
Aryryldiazirine		good	weak

◤ Diazoacetyl derivatives give carbenes by photolysis at 370 nm. The first example of the application of this method by WESTHEIMER (1962) was carried out on the labelling of

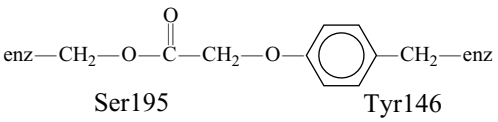
chymotrypsin. In fact, the authors used p-nitrophenyl-2-^{14}C-diazoacetate that reacts with serine of the active centre of chymotrypsin following the reaction:



At pH 6.2 diazoacetyl serine is stable for at least 48 hours. Its stoichiometric formation drives the inactivation of the enzyme. Photolysis of ^{14}C diazoacetyl chymotrypsin provokes incorporation of 20–25% of the radioactive reagent in the protein. Among amino acids carrying radioactivity, serine 195, tyrosine 146 and histidine 57 were identified. O-carboxymethylserine is formed after a WOLFF rearrangement following the reactions:

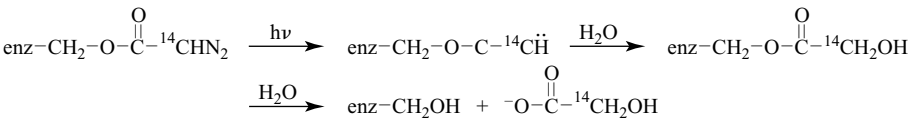


O-carboxymethyl tyrosine is formed by insertion of carbene on the hydroxyl group of tyrosine 146 of another molecule of chymotrypsin thus forming a dimer:



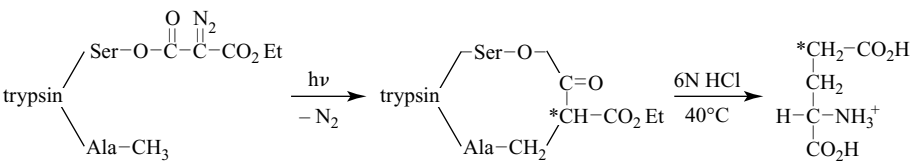
Ser195 Tyr146

One part of the enzyme is regenerated with the formation of glycolic acid according to the reactions:



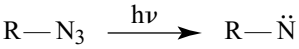
The diazoacetyl derivatives were also used as affinity reagents directed to the active centre of subtilisine and glyceraldehyde-3-phosphate dehydrogenase. Diazomalonyl derivatives were used with success in the case of chymotrypsin. They are more stable than the diazoacetyl derivatives and are less susceptible to undergoing WOLFF rearrangements.

Photolysis of diazomalonyl chymotrypsin as well as that of diazomalonyl trypsin leads to the same products as diazoacetyl derivatives, except for the acquisition of a glutamic acid with diazomalonyl trypsin following the reaction:



Arylazidomethanes are generally too reactive to be used with proteins. Aryldiazirines, more stable in the dark, are more interesting compounds than the preceding precursors.

Nitrenes participate also in a great variety of reactions, but their reactivity is weaker than that of carbenes. Nitrenes also result in the activation of a precursor by light:



There are several categories of nitrene precursors, acylazides, phosphorylazides, sulfonylazides, arylazides and nitroarylazides. The rearrangements that happen in the case of nitrenes are of the same type as those observed with carbenes. Therefore irradiation of 2-azido-biphenyl (Fig. 9.20) drives the intermediate of nitrene to the formation of carbazole.

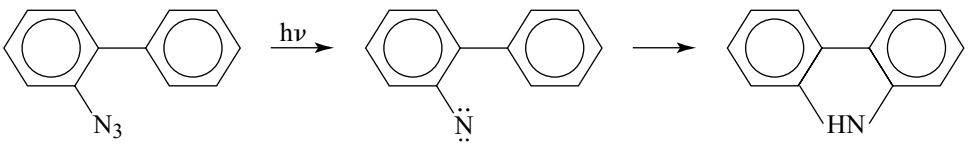


Fig. 9.20 Reaction of 2-azido biphenyl under the effect of irradiation

Among the nitrene precursors used to label enzymes, the most common are bound to nucleotides, such as 8-azido-cAMP, 8-azido-ATP, diazido-ATP (Fig. 9.21), and β-azido-NAD. They are used to label the active sites of kinases, ATPases and concerning the last reagent, NAD^+ dehydrogenases. Such compounds serve also in the study of receptors or membrane transporters, for example the receptors of nucleotides and cyclic AMP.

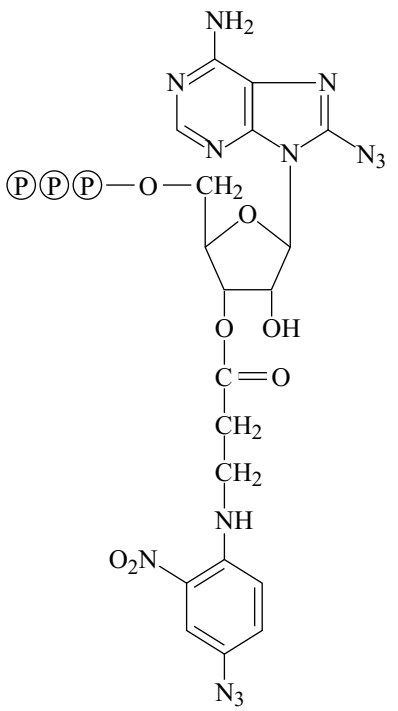


Fig. 9.21 Structure of diazido-ATP

In order to avoid non-specific labelling of the protein, one often resorts to scavengers added to the medium that trap the reagent molecules not bound to the active centre. For this purpose, p-amino benzoic acid, p-aminophenylalanine, dithiothreitol, ®-mercaptoethanol, tris or even amino acids are used.

9.2.2.4. SUICIDE REAGENTS

"Suicide" reagents are substrate analogs that are recognised by the enzyme, and the first transformation steps are the same as in the normal enzymatic reaction, followed by the generally covalent formation of an intermediate complex, EX. The reagent function is generated by the formation of this complex; it is only potential in the initial form S′ of the substrate. At this stage, the complex EX can evolve following two paths, either the formation of the product P′ with regeneration of the enzyme, or the irreversible formation of an inactive complex EI according to the scheme:

$$E + S' \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES' \xrightarrow{k_2} EX \begin{array}{l} \xrightarrow{k_3} E + P' \\ \xrightarrow{k_4} E\text{—}I \end{array}$$

The partition coefficient is given by the ratio of the rate constants k_3/k_4.

◤ Suicide substrates were used to study diverse enzymes. In the case of serine proteases, YON collaborators (1973) showed that 3,4-dihydro 3,4-dibromo methyl coumarine VII inhibits α-chymotrypsin by blocking both serine 195 and histidine 57 following the mechanism that is indicated in Fig. 9.22.

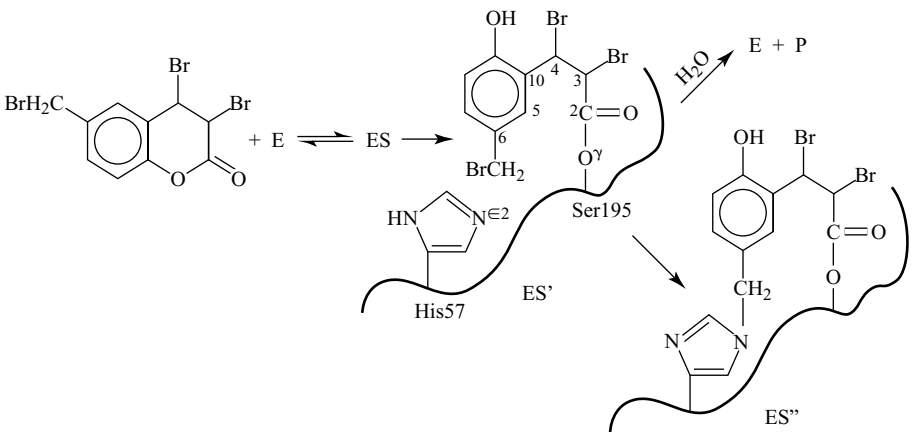


Fig. 9.22 Inhibition of chymotrypsin by a suicide substrate

Sulbactame was used as an inhibitor of β-lactamase. It behaves also like a suicide substrate. The mechanism of action was demonstrated by using the deuterated compound, 6,6-D_2-sulbactame. The different paths of transformation of this compound in the presence of the enzyme are represented in Fig. 9.23 below.

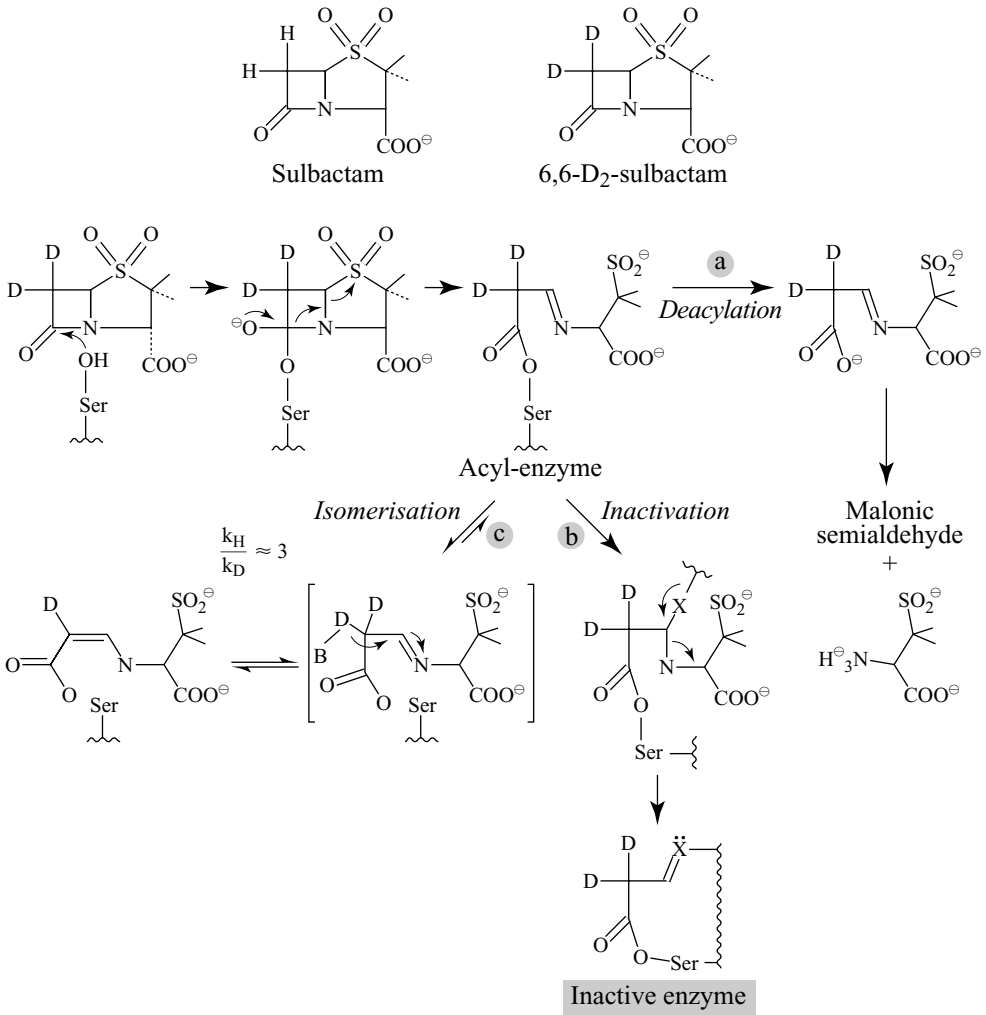


Fig. 9.23 Inhibition of β-lactamase by sulbactame ◢

The use of these reagents that form inactive compounds by blocking two of the catalytic groups permits not only the identification of these groups but also the evaluation of their distance according to the size of the reagents. **The use of "suicide" reagents is the object of important applications in molecular pharmacology.**

9.2.2.5. DIRECT LABELLING BY SELECTIVE REAGENTS

One does not always have reagents susceptible to being directed by affinity to the active centre of enzymes. Research and synthesis of these diverse types of compounds, affinity reagents, photoaffinity and suicide reagents is often laborious, and the examples we have cited represent favorable cases. The enzymologist often has no other recourse except the use of much less selective labelling. Direct labelling

accepts as selectivity criteria the preferential reaction of residues of the active site towards appropriate chemical reagents. Labelling procedures used consist of making reagents act, as selectively as possible, on the enzyme in solution. The reactivity of residues depends greatly on the pH and polarity of the solvent. Direct labelling must be accompanied by an analysis of modified residues, the determination of their position in the polypeptide chain and their importance in catalysis (to see later). In Sect. 9.2.2.7, some principal reagents of the amino acid side chains in proteins are presented.

9.2.2.6. DIFFERENTIAL LABELLING

In order to obtain a greater selectivity, one often resorts to differential labelling introduced by COHEN and WARRINGA (1953), then by KOSHLAND and collaborators (1959). Covalent labelling is carried out in two stages. In the first stage, often called the "pre-labelling stage", the protein is incubated in the presence of a large excess of substrate or a non-reactive analog of the substrate (a totally reversible competitive inhibitor); one chooses conditions such that the enzyme has a great affinity for the ligand. The goal of this stage is to protect the active centre against the reagent used for pre-labelling. Of course, the protective effect must be verified before undertaking the chemical modification. The enzyme thus protected is subjected to the action of the chemical reagent in great excess in order to modify all the residues of the same nature located away from the active site. The mixture is dialysed exhaustively or subjected to an exclusion chromatography in order to drive away the reagent in excess and to displace the inhibitor of the complex. In the second stage, the pre-labelled enzyme cleared of its inhibitor, with its active site becoming accessible, is subjected to the same chemical reaction as previously, but in the presence of the chemical reagent labelled by a radioactive isotope. All experimental conditions except the presence of the ligand must be identical to those in the first stage. It is necessary afterwards to identify the residue that carries the radioactivity. Each stage requires controls.

A quantitative description of the protection by the specific ligand in the differential labelling was presented by SINGER (1967). The rate of modification of a residue Y by a reagent R is given by the equation:

$$\frac{-dY}{dt} = k(Y)(R)$$

k is the second order rate constant, (Y) the concentration of unlabelled sites at time t, and (R) the concentration of the reagent in excess as to assure conditions of pseudo-first order. If x represents the fraction of residues Y that are modified, (1 – x) is the fraction of non-modified groups and one has:

$$\frac{dx}{dt} = k(1-x)(R)$$

The rate of modification in the presence of a protector ligand is the sum of the reactions of R with the free sites and the protected sites:

$$\frac{dY_p}{dt} = k(Y)(R) + fk(YP)(R)$$

The factor f is a measure of the degree of protection. The protection is total if f = 0; it is absent if f = 1. The binding equation leads to:

$$(YP) = K_P(P)(Y)$$

K_P being the association constant of P with Y. As previously mentioned, the labelling rate of the complex is the following:

$$\frac{dx_p}{dt} = \frac{k(1-x)(R)}{1+K_p(P)} \times [1+fK_p(P)]$$

The ratio of the chemical modification rates in the absence and presence of ligands is a function of the affinity of the enzyme for the ligand:

$$\frac{\frac{dx}{dt}}{\frac{dx_p}{dt}} = \frac{[1+K_p(P)]}{[1+fK_p(P)]}$$

9.2.2.7. PRINCIPAL REACTIONS OF AMINO ACID SIDE CHAINS

We present the most common reactions of different functional groups in the proteins. One cannot avoid here a slightly tedious enumeration; hopefully the reader pardons us. Recapitulative tables are given for the very practical goal of helping the experimenter in search of the adequate reagent to label a given amino acid.

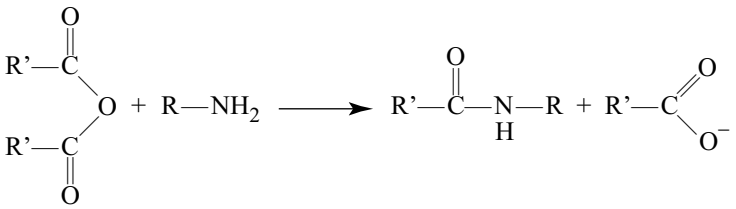
α- and ε-amino groups

The reagents of α- and ε-amino groups are the same. Sometimes however, it is possible to preferentially label the α-amino group by achieving the labelling in weak-ly basic conditions; the ε-amino groups react then more slowly than the α-amino groups as a consequence of their higher pK. Nevertheless there are exceptions when the α-amino residue is buried or participates in an interaction; this is the classical example of the NH_3^+ N-terminal group that, in serine proteases, participates in the formation of a salt bridge. Consequently its pK is shifted towards higher pH.

The principal reactions in which the α- and ε-amino groups participate are acylation, arylation, alkylation, the addition on a double bond and deamination.

◤ *Acylation*

Acylation is the attack of a trigonal sp^2 carbon by a nucleophile group, the nitrogen doublet of the amino group driving the formation of a bond of type R′–CO–NH–R. For this reaction, acid anhydrides are generally used:

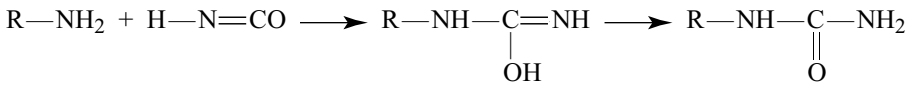


Depending on the nature of the radical R′, it is possible to conserve the positive charge (by using N-carboxyanhydride of α-amino acid), to introduce a negative charge (by using succinic, maleic, citraconic, or tetrafluoro succinic anhydrides), or even to neutralise the charge (by using acetic anhydride). These reagents are given in Table 9.3 below.

The amino group can be regenerated in different conditions according to the stability of the bond formed. The modification by maleic anhydride is easily reversed in acidic medium where the hydrolysis of the maleylamide bond is catalysed by the non-ionised carboxyl group. The half life of the maleyl-lysine at pH 3.5 and at 37ºC is around 10 h. Citraconylation is also reversible in acidic medium. Acylation by tetrafluorosuccinic anhydride gives a stable derivative in acidic medium, but it easily decomposes at pH 9.5.

These reagents are poorly selective; they are susceptible to reacting with other residues like cysteine, histidine and tyrosine. However these amino acids give unstable esters in alkaline medium.

Carbamylation by cyanate is a fairly selective reaction in alkaline medium:



Generally, carbamylation is carried out with a reagent of the form R—N═C═O. The reaction is performed around pH 8. Cysteine and tyrosine can also react reversibly at lower pH (pH 5-6).

Phenylisothiocyanate (PITC) drives an analogous reaction giving rise to the appearance of phenylthiocarbamyl with the ⟨-amino N-terminal group; this derivative cyclises in thiazolinone, then in phenyl thiohydantoin (PTH), bringing about a sequential degradation of proteins. This reaction is used in the EDMAN method that permits the sequence analysis of a polypeptide chain starting from the N-terminal extremity (see Table 9.3). Today one generally uses for microsequencing dimethylaminoazobenzene isothiocyanate (DABITC) that leads, by the same principle as previously, to the formation of a colored dimethylaminobenzene thiohydantoine (DABTH) derivative. The use of this reagent permits a very great sensitivity of detection and renders possible sequence determinations with less than a nanomole of polypeptide.

Guanidylation by action of O-methyl thiourea is performed principally on amino groups of lysine residues following the reaction:

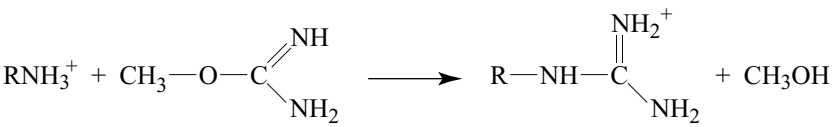
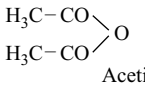
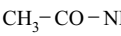
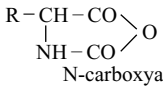
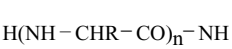
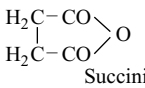
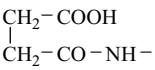
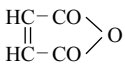
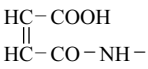
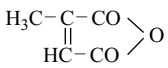
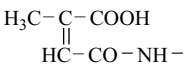
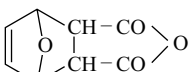
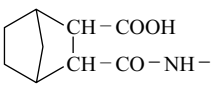
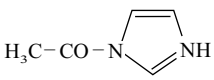
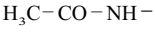
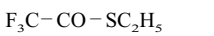
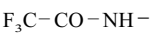
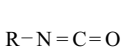
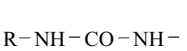
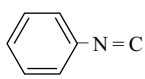
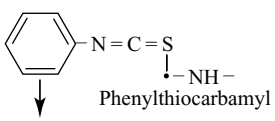
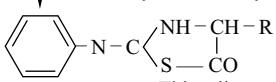
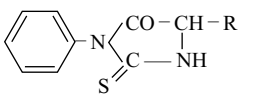
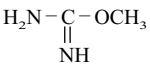
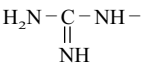
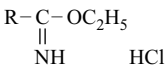
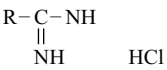
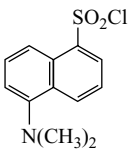
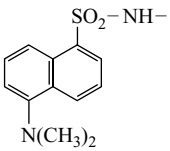
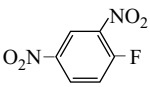
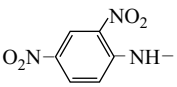
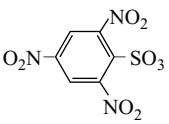
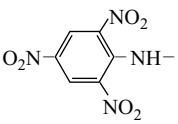
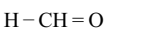
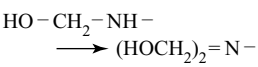
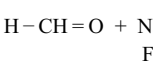
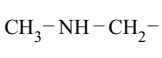
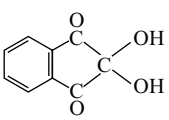
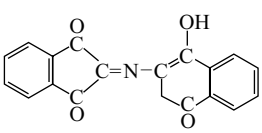
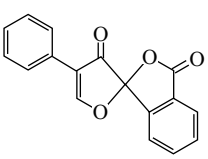
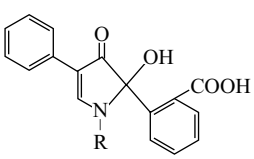
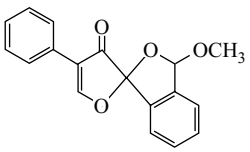
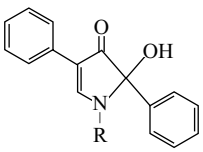
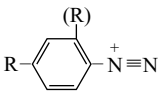
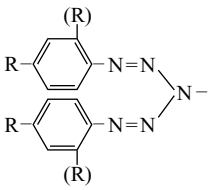


Table 9.3 Reactions of α- and ε-amino groups

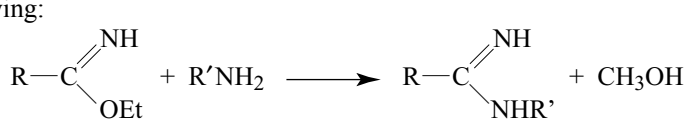
Reagent	*Product*	*Use*	*Comments*
 Acetic anhydride			
 N-carboxyamino acid anhydride		Introduction of polyR amino acids into the peptide	
 Succinic anhydride		Replacement of the positive charge of NH_2 by the negative charge of COOH	Stabilisation of insoluble polypeptides, limited trypsine cleavage next arginine
 Maleic anhydride		–	
 Citraconic anhydride		–	Regeneration of NH_2 in acetic media. Reagents differing in the stability and specificity of the bond
 Exo-cis-3,6-endoxo-Δ_4-tetrahydrophtalic anhydride		–	
 Acetylimidazole			
 Ethylthiotrifluoroacetate			
 Isocyanate		Reversible blocking of lysine ε-NH_2	Trypsin only cleaves at the side of arginine residues
 Phenylisothiocyantate (PITC)	 Phenylthiocarbamyl	Determination of the N-terminal	
	 Thiazolinone	Sequential degradation of proteins (EDMAN)	
	 Phenylthiohydantoin		

Reagent	*Product*	*Use*	*Comments*
		Replacement of NH_2 by the guanidyl group	
		Substitution of NH_2 without changing the charge	
 Dansyl chloride (DNSCl)	 (DNS−)	Determination of the N-terminal group and the ε-amines of lysine	Very stable fluorescent dansyl amino acids; high sensitivity of detection
 Dinitrofluorobenzene	 Dinitrophenyl− (DNP−)	Determination of the N-terminal group	Hydrolysis by HCl gives stable DNP amino acids (yellow)
 2,4,6-trinitrobenzene sulfonate (TNBS)	 Trinitrophenyl− (TNP−)	Colorimetry of α- and ε-amines	Hydrolysis by HCl gives TNP amino acids and regenerates amino acids
 Formaldehyde		Preparation of anatoxins	
 Formaldehyde + Sodium borohydride		Methylation of α- and ε-amines; introduction of ^3H or ^{14}C	
 Ninhidrine		Detection and colorimetry of NH_2	
 Fluorescamine		Detection and highly sensitive fluorimetry of NH_2	

Reagent	*Product*	*Use*	*Comments*
 2-methoxy-2,4-diphenyl-3(2H)-furanone (MDFP)			
$CH_2=CH-CN$ Acrylonitrile	$R-NH-CH_2-CH_2-CN$		
 Aryldiazo salts			Colorless product formed from free NH_2, contrary to colored products formed from histidine or tyrosine
$O=N-OH$ Nitrous acid	$-NH_2 \longrightarrow -OH + N_2$	Deamination	

It occurs around pH 10–11 in a very slow reaction. The positive charge of lysines is not lost by this modification. There is an increase in the length of the side chain that is transformed in homoarginine. Due to its weak reactivity and weaker nucleophile ability of its α-amino group, this reagent reacts selectively with ε-amino groups. More reactive S-methyl isourea shows little difference in reactivity towards α- and ε-amino groups.

Certain esters, thioesters or imidoesters are very reactive. Imidoesters are easily attacked by nucleophilic reagents and react selectively with amino groups of proteins between pH 8.5 and 9.5 to form amidines. **Amidination** leads to complete blockage of amino groups if relatively strong reagent concentrations are used. The reaction is the following:



Lysine residues thus blocked can be regenerated by incubation of the amidine derivative in the presence of ammoniac or hydrazine. Imidoesters in a longer chain such as dimethyl suberimidate or adipimidate are used as cross-linking reagents.

Sulfonylation is obtained by action of aliphatic or aromatic sulfonic acid halogenides on the amino groups of proteins. The most used reagent is S-dimethylamino-naphthalene-1-sulfonic acid chloride or dansyl chloride. Dansylation of amino groups gives fluorescent derivatives. For this reason, this reaction is used to introduce an extrinsic fluorescent label into a protein for conformational studies (see Table 9.3).

Alkylation and arylation

These reactions give rise to compounds of the type:

–NH–C or NH–Ar

They are brought about by action of aryl halogenides of which the most classic is dinitrofluorobenzene (DNFB) introduced by SANGER in 1945. The order of reactivity of the leaving group is:

$$F > Cl \sim Br > SO_3H$$

DNFB gives dinitrophenyl derivatives with α- and ε-amino groups. The reaction requires that the nucleophilic group not be protonated; at pH 8.5 and at lower pH, the ε-amino groups react more slowly than the α-amino groups. In ribonuclease therefore, dinitrophenyl lysine 41 is preferentially formed following the great reactivity of this particular group. Dinitrophenyl derivatives are stable in alkaline conditions and easily measurable by spectrophotometry. The compound formed with an amine absorbs in the visible spectrum with a maximum at 365 nm.

DNFB also reacts with other side chains in proteins, in particular the SH group of cysteines, the phenol group of tyrosines and the imidazole group of histidines. These amino acids substituted by the dinitrophenyl group can be regenerated by β-mercaptoethanol which displaces the reagent. This procedure therefore permits a selective labelling.

2,4,5-trinitrobenzene sulfonate (TNBS) also reacts with amino groups and is used as a reagent in colorimetric measurements. The absorption spectra of TNP lysine and TNP amino acids present a maximum around 345 nm with a shoulder at 420 nm. TNBS reacts little with the imidazole nitrogen or with the hydroxyls of tyrosine, serine and threonine; it does not react with arginine. It forms labile derivatives with SH groups.

There exist other reagents of amino groups used essentially for the spectroscopic detection of these residues; they give rise to colored derivatives like ninhydrine or fluorescent derivatives like fluorescamine or 2-methoxy-2,4-diphenyl-3H-furanone (MDPF) that allow a highly sensitive detection by fluorimetry.

Aliphatic aldehydes and ketones react rapidly and reversibly with protein amino groups. The products formed can be reduced to give stable alkylamino groups. In lightly alkaline conditions, residues other than the amino groups do not give stable derivatives. Formaldehyde reacts with lysine in the presence of sodium borohydride to give dimethyl lysine. If one uses tritiated borohydride, one introduces thus a radioactive labelling that permits the identification of the implicated lysine.

Addition reactions on a double bond

Nucleophilic addition of acrylonitrile was used to alkylate SH groups. The latter react in lightly alkaline conditions. At higher pH (9.5) the reaction with the amino groups becomes significant. At pH 9.2, all the lysine residues of ribonuclease are labelled by the reagent. The reaction is the following:

$$R–NH_2 + CH_2{=}CH{-}CN \longrightarrow R–NH–CH_2–CH_2–CN$$

Attack by an electrophilic reagent

Aryl azonium salts react with amino groups. The reaction proceeds rapidly to form a non-colored compound. In weakly alkaline conditions, the reaction of amino groups can become faster than those of tyrosine or histidine. These two last amino acids form colored compounds. Diazonium salts were extensively used for labelling antibodies.

Deamination by nitrous acid is one of the oldest methods of chemical modification of proteins. Although this reagent exhibits both nucleophilic and electrophilic properties; the latter is rarely observed with proteins.

Histidine residues

Histidine does not possess truly selective reagents. The best reagents of histidine in proteins are affinity reagents. Therefore, in conditions of well-determined pH, it is possible to operate a selective labelling with appropriate reagents, in particular diethylpyrocarbonate.

Acylation

Ethoxyformic anhydride or diethylpyrocarbonate reacts with imidazole; the derivative that results presents a characteristic absorption between 230 and 240 nm permitting to follow the reaction. At pH 4, ethoxyformic anhydride reacts principally with amino groups and the imidazole of histidine.

Arylation

Arylation by dinitrofluorobenzene (FDNB) gives an unstable derivative in acidic medium. This modification is very weakly selective.

Alkylation

Alkylation reactions by halo acids and amides at neutral pH represent the best approach to modifying histidines in native proteins. Alkylation by **iodoacetate or bromoacetate** brings about the formation of derivatives substituted in position N_1 or N_3 of imidazole or even bi-substituted compounds. With iodoacetate one obtains either 1-carboxymethyl histidine, 3-carboxymethyl histidine, or even 1,3-dicarboxymethyl histidine. This reaction is rarely selective except in the case of particularly reactive histidine residues. Thus, in ribonuclease, it was possible to label two histidines in the active centre (His 12 and His 119) and by working at distinct pH values to even selectively label either one or the other of these groups (CRESTFIELD et al., 1963). In carbonic anhydrase, the action of iodoacetate brings about a selective labelling of a single catalytic histidine. In these different cases, a single monocarboxymethyl derivative is formed.

Attack by an electrophilic reagent

Diazotetrazole reacts with histidine to give a colored bis-azo derivative that, as we have mentioned above, absorbs with a maximum at 480 nm, which permits proceeding to a spectrophotometric measurement. Similarly, diazosulfanilic acid or Pauly reagent gives a colored compound with histidine, permitting the detection.

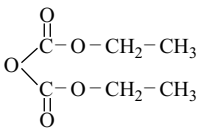
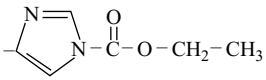
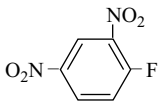
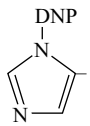
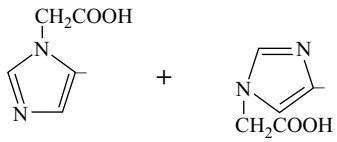
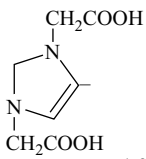
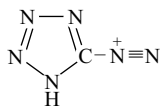
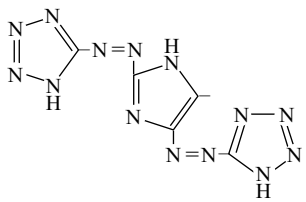
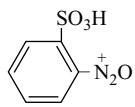
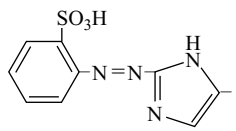
Iodation of histidine occurs generally with that of tyrosine. One can observe the formation of mono and diiodo histidine derivatives.

Photooxidation

Photooxidation of histidine results in irradiation of the protein in the presence of photosensitive compounds like methylene blue or Bengal rose. It is relatively selective in neutral medium and in rather gentle conditions.

Table 9.4 summarises these different reactions of histidine.

Table 9.4 Reactions of the imidazole of histidine

Reagent	*Product*	*Use*
 Ethoxyformic anhydride diethylpyrocarbonate		
 Dinitrofluorobenzene (DNFB)	 DNP-His	
ICH_2COOH Iodoacetic acid	 1-carboxy-methyl-His + 3-carboxy-methyl-His	
	 1,3-dicarboxymethyl-His	
 Diazotretrazole		Histidine colorimetry
 Diazo sulfanilic acid		Detection of histidine (PAULI reagent)
Light + methylene blue	Destruction of imidazole nucleus	Photo-oxydation

Tyrosine residues

The phenol group of tyrosine participates in the same types of reactions as amino groups, but in different conditions.

Acylation

Modification of tyrosines by **acetic anhydride** is reversible; the ester thus formed is unstable in alkaline medium, allowing, as we have previously mentioned, to selectively label amino groups at high pH. Acetyl-imidazole reacts with tyrosines with normal pK, giving an unstable ester in alkaline medium.

Arylation

Cyanuric fluoride (CNF), which presents a very high reactivity but possesses a weak selectivity, is hardly useful. It forms an ester with tyrosine that is stable in alkaline medium and of which the absorption spectrum is shifted towards blue as compared to that of tyrosine.

Fluorodinitrobenzene (FDNB) reacts with tyrosine to give the compound O-dinitrophenyl tyrosine.

Alkylation

Alkylation by **dimethylsulfate,** $SO_4(CH_3)_2$ brings about the formation of O-methyl tyrosine.

Action of electrophilic reagents

Nitration by **tetranitromethane** (TNM) is one of the most selective reactions with tyrosine. At pH 7-8 and with limited reaction times, nitration occurs in *ortho* of the OH group. One does not observe the formation of 3,5-dinitrotyrosine. The o-nitrophenolate ion is stable in acid hydrolysis. It has a maximum absorption at 428 nm, which permits following the reaction by spectrophotometry. The spectrum of absorption of 3-nitrotyrosine is very sensitive to the solvent polarity. The action of the reagent is dependent on the degree of tyrosine exposure. By selective reduction of ortho-nitrotyrosine by dithionite or sodium hydrosulfite, it is possible to convert tyrosine into 3-amino tyrosine. However there are parasitic reactions like the formation of intermolecular bridges, oxidation of cysteine and methionine, and the modification of tryptophan and histidine. Electrophoresis on a polyacrylamide gel in the presence of dodecyl sodium sulfate rapidly permits the detection of possible intermolecular bridges. Despite these inconveniences, the use of tetranitromethane knew much success for selectively labelling tyrosines, which occurs more rapidly than for other residues

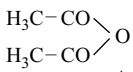
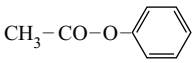
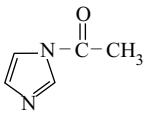
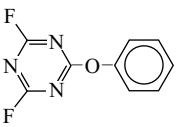
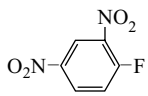
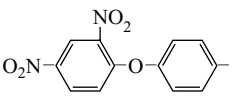
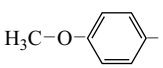
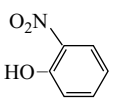
Iodation by iodine or iodic chloride is also a very useful reaction. The labelling occurs in ortho of the phenol group driving mono or diiodo (3-iodo or 3,5-diiodotyrosine) derivatives. Very reactive sulfhydryl groups are occasionally affected forming sulfenyl iodides. Outside of these substitution reactions, oxidative reactions implicating the methionine, cysteine, and tryptophan residues can occur. However, oxidations are predominant in acidic medium whereas substitution reactions are predominant in alkaline medium. Iodation can be brought about by an enzymatic reaction specific to tyrosines, the action of sodium iodide catalysed by peroxydase. Ioda-

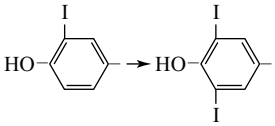
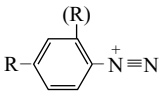
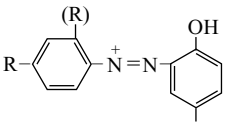
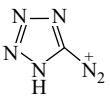
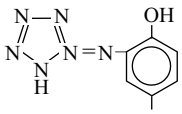
tion of tyrosines is very useful in labelling proteins with ^{125}I, in particular for radioimmunologic assays or for studies of receptors.

Aryl diazonium *salts* react with tyrosine to form derivatives of which the coloration is used in the detection of this amino acid. **Diazonium-1H-tetrazole** (DHT) reacts with tyrosine and also with histidine forming corresponding azoderivatives that absorb at 550 and 480 nm, respectively. The simultaneous spectrophotometric titration of two derivatives is therefore possible without ambiguity. The reaction rate increases with pH. Due to its instability, diazonium-1H-tetrazole is potentially explosive and must be manipulated with care. Concentrations greater than 0.2 M are considered dangerous. ◢

The different reactions of tyrosine are given in Table 9.5.

Table 9.5 Reactions of tyrosine

Reagent	*Product*	*Use*	*Comments*
 Acetic anhydride	 O-acetyl-tyrosine		
 N-acetylimidazole			
 Cyanuric fluoride			
 Dinitrofluorobenzene (DNFB)	 O-dinitrophényl tyrosine		Hardly selective
$SO_4(CH_3)_2$	 O-methyl tyrosine		
$C(NO_2)_4$	 3-NO_2 tyrosine	Specific nitration of tyrosine residues	Yellow 3-NO_2 tyrosine is resistant to HCI hydrolysis

Reagent	*Product*	*Use*	*Comments*
ICl_3 Iodine trichloride INa + lactoperoxydase		Introduction of ^{125}I into proteins	The enzymatic reaction is specific for tyrosines. It permits the selective labelling of cell surfaces proteins.
 Aryl diazonium salts		Detection of tyrosine	
 Diazonium-1H-tetrazole			

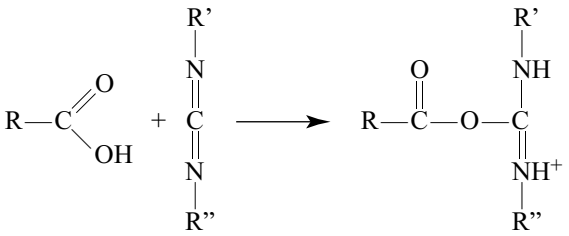
Carboxyl groups

The carboxylate ion is stabilised by resonance and is not very reactive. The reactive form is the protonated carboxyl COOH. The most common reactions used to modify carboxyl groups are, on the one hand esterification catalysed by an acid, and on the other hand coupling with a nucleophilic compound after activation by a carbodiimide soluble in water.

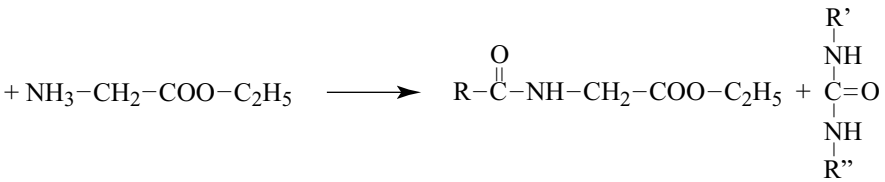
Carboxyl groups of aspartic and glutamic acids have pK values between 4.5 and 5. They can be esterified in relatively gentle conditions, either by methanol in diluted chlorhydric acid, or by diazo acid or diazo alcane derivatives. Diazomethane is a highly reactive compound and was used to esterify carboxylic acids. It is sometimes used to esterify carboxylic protein groups, but the reactions are often incomplete and other groups can also be modified. Diazo-acetate esters are more stable. The modification of SH groups constitutes a parasite reaction (WILCOX, 1967). Esterification can be reversed in alkaline medium.

◤ Carboxylic groups react with diazonium salts, isoxazolium salts, and triethyl oxonium fluoroborate. This last reagent was used to label carboxylic groups of the active centre of lysozyme by PARSON et al. (1969). The ester thus formed is labile, and the reaction is easily reversed around pH 7. Activation of carboxylic groups by N-alkyl-5-phenyl isoxazolium salts was described by the group of SHAW (1969). Therefore N-methyl-5-phenyl isoxazolium fluoroborate was used in the case of trypsin. It brings about complete inactivation of the enzyme; the activity can be recovered by treatment with hydroxylamine.

However, the most utilised reaction for labelling carboxylic residues is coupling with nucleophilic compounds, generally the ester or amide of an amino acid *via* the formation of a soluble carbodiimide intermediate in the presence of which a carboxylic group in aqueous milieu gives an O-acyl-isourea derivative:

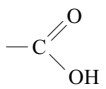
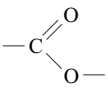
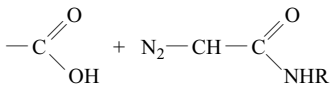
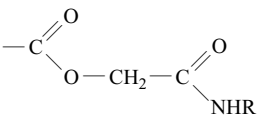
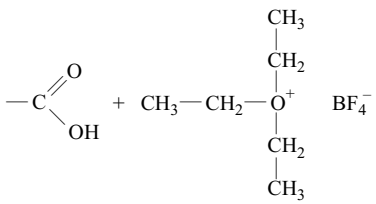
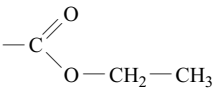
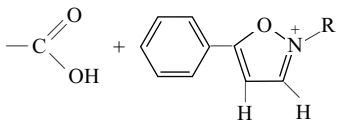
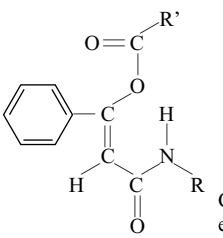
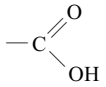


which with ethyl ester glycine gives an amine bond:



However, the carbodiimides are capable of reacting with other groups, in particular cysteines and tyrosines. Labelling of tyrosine can be reversed by treatment with hydroxylamine at pH 7 and 25ºC. These different reactions are regrouped in Table 9.6.

Table 9.6 Reagents of aspartate and glutamate carboxylate

Reagent	*Product*
 + CH_3OH Methanol	
 Diazoacetate	
 Triethyloxonium fluoroborate	
 N-methyl-5-phenyl-oxazoline fluoroborate	 Corresponding enolic ester
 + $R'-N=C=N-R''$ + $NH_2-CH_2-COOC_2H_5$ Carbodiimide + Glycine ethyl ester	$R-CO-NH-CH_2-COO-C_2H_5$

Hydroxyl groups of serine and threonine

Alcohol groups are usually hardly reactive due to their very high pK. Serine can become very reactive in a particular environment, such as in the active centre of serine proteases. For chemical modification one uses pseudo-substrates or affinity reagents, as previously mentioned. When it is reactive, serine can be acylated or phosphorylated. O-phosphoserine is naturally present in some enzymes. Occasionally, it is formed at the active site in an intermediate stage of the catalysed reaction; this is the case with phosphoglucomutase previously described. O-phosphothreonine is also found, although much less frequently.

Sulfhydril groups

Thiol groups of cysteines are very reactive under their form S^-. A great number of compounds are susceptible to react with cysteine; as we have underlined it, the labelling of cysteine often interacts as a parasite reaction and labels other residues at the same time. Since thiol groups are easily modifiable, certain reactions in which they participate are reversible, and their participation at the active site of some enzymes was revealed very early. In this chapter, the diverse reactions of thiols are rapidly reviewed to stress more particularly the reactions the most useful to enzymologists; their use depends of course on the goal that one has set.

◤ *Acylation and arylation*

Thiol groups react rapidly with acid anhydrides to give unstable derivatives in alkaline medium. The reaction of thiols with aromatic compounds such as chlorodinitrobenzene and bromo-nitro-imidazole is done more rapidly than the reactions of acylation, even in more acidic conditions.

Alkylation

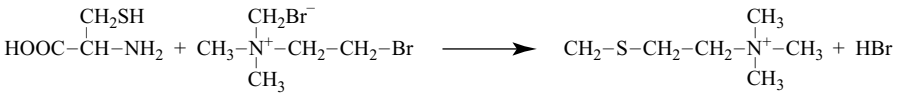
Carboxymethylation by acid halogenides is a very classical modification of thiols that are much more reactive than other residues that can react with these compounds. Reaction with iodoacetate gives rise to negatively charged carboxymethyl cysteine, following the reaction:

$$RS^- + I—CH_2—COO^- \longrightarrow R—S—CH_2—COO^- + I^-$$

Iodoacetamide is also highly used; its reaction with thiols brings about a non-charged carboxamido methyl cysteine derivative. The same reactions can be carried out with corresponding bromine derivatives.

N-ethylenimine reacts with proteins to converting cysteine residues in S(2-aminoethyl) cysteine. The reaction occurs in lightly alkaline conditions (RAFTERY & COLE, 1963). No other group reacts at this pH; in acidic medium, methionine reacts slowly to give the ion S(2-aminoethyl) methionine sulfonium (SCHROEDER *et al.*, 1967). This modification is interesting in the sense that S(2-aminoethyl) cysteine resembles lysine (with S in the place of CH_2), introducing into a protein a new trypsin cleavage site (WANG & CARPENTER, 1968).

Alkylation by **trimethyl ammonium bromide** converts cysteine to a derivative strongly basic, 4-thialamine, according to the reaction:



This trimethylaminoethyl compound is very stable and resists acid hydrolysis, permitting analyses driving the identification of cysteines that reacted.

Azobenzene-2-sulfenyl bromide, soluble in water, reacts selectively and rapidly with protein thiols in aqueous solution. Mixed disulphide thus formed is easily reduced by thiols (2-mercaptoethanol or dithiothreitol) or by sodium borohydride.

Addition reactions

Thiol groups participate in addition reactions with several reagents. Cyanoethylation by **acrylonitrile** is rapid and selective at pH 8; the reaction with amino groups occurs at higher pH. The reaction is carried out as follows:

$$-S^- + CH_2{=}CHCN \longrightarrow -S-CH_2-CH_2-CN$$

The most commonly used reagent is N-ethyl maleimide (NEM) that gives S-ethyl succinimido cysteine which is soluble in water and stable. NEM reacts rapidly with cysteine at neutral pH. The reaction can be followed by the decrease in absorbance at 305 nm. Indeed, NEM absorbs at this wavelength ($\varepsilon_M = 620$) whereas S-ethyl-succinimido-cysteine does not. In addition, S-ethylsuccinimido-cysteine gives by hydrolysis S-succinyl-cysteine that is stable in acidic medium therefore permitting the identification of the labelled cysteines. The reaction of cysteines with maleic anhydride also leads to S-succinyl-cysteine. NEM is a very selective thiol reagent that leads to an irreversible labelling.

The colored derivative **N-dimethylamino-3,5-dinitrophenyl maleimide** (DDPM) introduced by WITTER and TUPPY (1960) was used for the labelling of cysteines of many proteins.

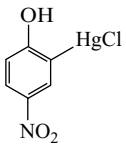
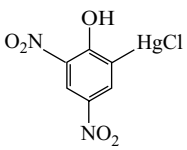
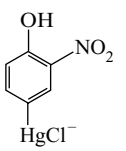
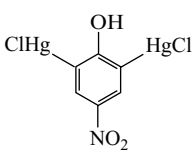
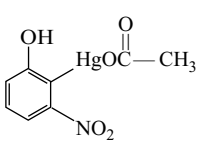
The addition reaction with 4-vinyl pyridine that brings about S-pyridyl ethyl cysteine is often used for identification tasks because the labelling is stable and resists acid.

Reaction with organomercurials

Reaction of thiols with organomercurials brings about the formation of mercaptides. One of the most common reagents is **p-chloro mercuribenzoate** (PCMB). The reaction can be easily followed by spectrophotometry. At pH 4.6, PCMB absorbs with a maximum at 234 nm and an extinction coefficient ($\varepsilon = 1.74 \times 10^4$ mol^{-1} . cm^{-1}). At 255 nm and at the same pH, the value of ε_M between the residues of PMB-cysteine and PCMB is on the order of 6 000–8 000 mol^{-1} . cm^{-1} according to the solvent used.

The mercuri-nitrophenol derivatives have the advantage of absorbing in the visible region of the spectrum (Table 9.7 below). They were frequently used as probes for sulfhydryl group environments in proteins as well as in the preparation of mercury derivatives to analyse proteins by X-ray diffraction.

Table 9.7 Spectral properties of some mercuri-nitrophenols
(from MCMURRAY & TRENTHAM, 1968)

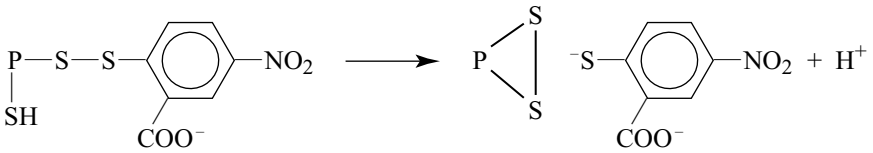
Compound	*Structure*	λ_{max} *(mμ)*	ε *(in 0.1 M KOH)* $(mol^{-1}.cm^{-1})$
2-chloromercuri-4-nitrophenol		405	1.74×10^4
2-chloromercuri-4,6-dinitrophenol		371	1.57×10^4
4-chloromercuri-2-nitrophenol		416	4.1×10^3
2,6-dichloromercuri-4-nitrophenol		410	1.74×10^4
2-acetoxymercuri-3-nitrophenol		410	1.7×10^3 *

* *In 0.1 M NaOH*

Exchange reactions with disulphide groups

One of the most classical reactions with thiol is their ease of exchange with disulphide groups to give mixed disulphides; these reactions are reversible. Introduced by ELLMAN (1959), one of the most familiar reagents is **5,5′-dithiobis (2-nitrobenzoate) or DTNB, or NbS2** in the official nomenclature. The reaction gives rise to a mixed disulphide with the appearance of a thionitrobenzoate ion that can be followed by variation in absorbance at 412 nm. One must take into account a light spontaneous hydrolysis of the reagent. Thionitrobenzoate ion (TNB^-) absorbs at 412 nm ($\varepsilon = 1.36 \times 10^4$ $mol^{-1}.cm^{-1}$ at pH 8), the mixed disulphide enzyme-thionitrobenzoate absorbs at 310 nm with an average extinction coefficient of 2.6×10^3 $mol^{-1}.cm^{-1}$, whereas DTNB presents a peak at 323 nm ($\varepsilon = 1.66 \times 10^4$ $mol^{-1}.cm^{-1}$). However,

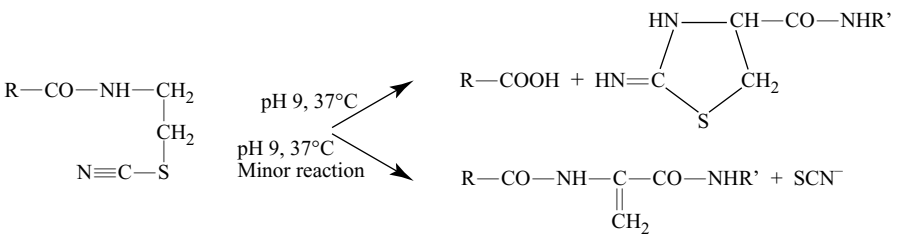
in proteins the presence of other thiols within the proximity of mixed disulphide can bring about the appearance of a second TNB^- ion with the formation of an internal disulphide bridge, bringing about an error in the titration:



- **4,4′-dithiopyridine** reacts with SH groups in a manner analogous to DTNB releasing one molecule of thiopyridine that can be measured by spectrophotometry. The reagent absorbs with a maximum at 247 nM ($\varepsilon_{247} = 1.63 \times 10^4$ mol^{-1} . cm^{-1}) and 4-thiopyridone at 324 nm ($\varepsilon_{324} = 1.98 \times 10^4$ mol^{-1} . cm^{-1}). There are other reagents of the same type; only the most common are presented here.

Cyanylation

Cyanylation of SH can be carried out by reaction with **2-nitro-5-thiocyanobenzoate** (NTCB) introduced by DEGANI et al. (1970), and DEGANI and PATCHORNIK (1974). Thiocyanoalanine is obtained. The substitution which is very specific leads to a derivative without steric hindrance; that is not the case for other thiol reagents. The reagent is obtained by treatment of DTNB by cyanogene bromide at pH 8. The reaction with thiols is followed by the variation in absorbance at 412 nm due to the thionitrobenzoate ion. JACOBSON et al. (1973) proposed a method using this reagent to obtain a selective cleavage of peptide bonds at the level of cysteines. The reaction is the following:



The cleavage is not carried out to 100%; a weak proportion is converted to a non-cleaved peptide that contains dihydroalanine.

Oxidation reactions

The action of an oxidising agent such as hydrogen peroxide or performic acid brings about the appearance of cysteic acid. Sodium sulfite transforms the cyteine residues into S-sulfonate derivatives (R–S–SO_3^-).

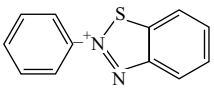
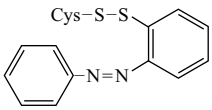
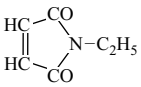
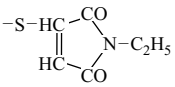
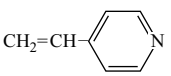
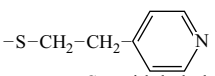
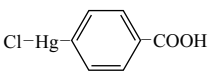
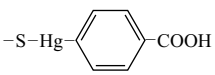
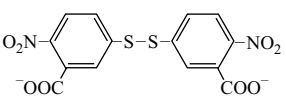
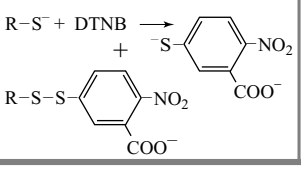
Tetrathionate gives a sulfonate S-sulfenyl derivative following the reversible reaction:

$$P\text{–}SH + SO_4O_6^{=} \longrightarrow P\text{–}S\text{–}S\text{–}SO_3^- + S_2O_3^{=} + H^+$$

often used to protect cysteines during chemical modification of other residues.

Table 9.8 below summarises the principal reactions of SH groups in proteins.

Table 9.8 Reactions of cysteine and cystine

Reagent	*Product*	*Use*	*Comments*
ICH_2COOH Iodoacetic acid	$-S-CH_2COOH$ Carboxymethyl cysteine (CM-Cys)	Substitution with introduction of a négative charge	CM-cys is resistant to acid hydrolysis
ICH_2CONH_2 Iodoacetamide	$-S-CH_2-CONH_2$	Introduction of a positive charge	Amino-ethyl cysteine is recognised by trypsine
H_2C(NH)H_2C Ethylene diamine	$-S-CH_2-CH_2-NH_2$		
H_2C(O)H_2C Ethylene oxide	$-S-CH_2-CH_2-OH$	Stabilisation without changing the charge	
$CH_3-N^+(CH_2-Br)(CH_3)-CH_2-CH_2-Br$ Trimethyl ammonium (2-bromoethyl) bromide	$-S-CH_2-CH_2-N^+(CH_3)_3$ 4-thiolaminine	Introduction of a positive charge	Very stable during HCl hydrolysis
 Azobenzene-2-sulfenylbromide		ε = 1,000 à 353 nm	
$CH_2=CH-C\equiv N$ Acrylonitrile	$Cys-S-CH_2=CH_2-CN$		
 N-ethylmaleimide (NEM)	 S-ethylsuccinimido cysteine	Spectrophotometric titration	
 Maleic anhydride	$Cys-S-CH(CH_2-CO)-CO-N-C_2H_5$ S-succinyl-Cys		
 4-vinyl pyridine	 S-pyridylethyl cysteine	Stabilisation and introduction of absorption in UV	The product is resistant to HCl hydrolysis
 p-chloromercuribenzoic acid (PCMB)		Formation of mercaptides	Cysteine can be regenerated by reduction
$R_1-S-S-R_1$ R_1 disulfide	$R-SH + R_1-S-S-R_1$ $\rightleftharpoons R-S-S-R_1$ (mixed disulfide) $\rightleftharpoons R-S-S-R$ (R disulfide)		
 5,5'-dithiobis (2-nitrobenzoate) (DTNB)		Thiols titration (ELLMAN reagent)	

Methionine

Methionine participates in numerous reactions with other amino acids. However, one can still obtain relatively selective labelling by placing it in acidic media (Table 9.9). Thus, alkylation by iodoacetate and iodoacetamide is carried out fairly selectively at low pH.

Table 9.9 Reactions of methionine

Reagent	*Product*	*Use*	*Comments*
$I-CH_2-COOH$ Iodoacetic acid	$-\overset{+}{S}(CH_3)-CH_2-COOH$	Substitution with introduction of a negative charge	
$I-CH_2-CONH_2$ Iodoacetamide	$-\overset{+}{S}(CH_3)-CH_2-CONH_2$	Substitution without changing charge	
Oxidation	$-S(=O)-CH_3$ Met-sulfoxide		Unstable in acid hydrolysis
Photo-oxidation	$-S(=O)_2-CH_3$ Met-sulfone (Met-SO_2)		Met-SO_2 is stable in acid hydrolysis
$BrC\equiv N$ Cyanogen bromide	$-AA_1-Met-AA_2-$ → $AA_1-HSL + H_2N-AA_2$ HSL = homoserine lactone → $-NH-CH(CH_2-CH_2-OH)-COOH$ Homoserine	Specific cleavage at the level of the methionine carboxyl	Homoserine lactone (HSL) gives homoserine by hydrolysis

◤ Oxidation by hydrogen peroxide occurs easily in acidic media; this yields a methionine sulfoxide. Photooxidation in acidic media or in the presence of a photosensitiser gives methionine sulfones. ◢

The reaction of methionine with cyanogene bromide permits the selective cleavage of the polypeptide chain at the level of the methionine carboxyl which is transformed into homoserine lactone, which by hydrolysis produces homoserine.

Arginine

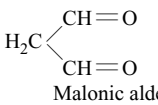
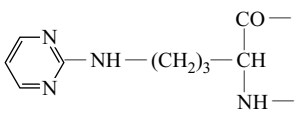
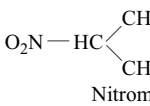
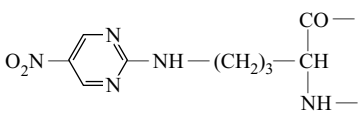
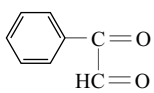
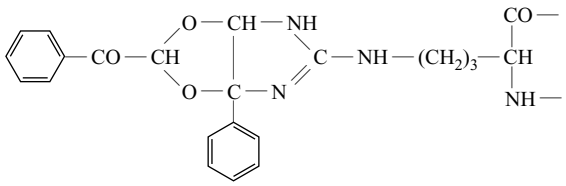
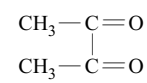
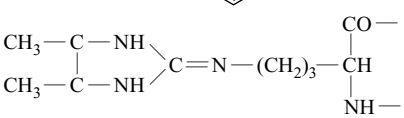
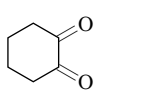
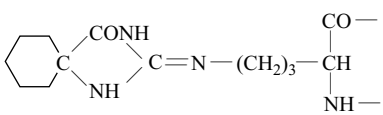
The arginyl residue possesses a high pK and is stabilised by resonance; it differs from other protein basic groups. The chemical modification of this residue is therefore very difficult. Acylation and nitration can only be carried out in extreme conditions, incompatible with protein stability. The only type of reaction practically used is condensation with dicarboxyl compounds that is still carried out at extreme pH values.

- Among the most useful reagents, one must cite **malonic or nitromalonic dialdehyde.** With nitromalonic aldehyde, a nitropyrimidine derivative is obtained; however the reaction is carried out at high pH (between pH 12 and 14), therefore only on denatured proteins.

  One of the very utilised reagents of arginine is **phenylglyoxal** introduced by TAKAHASHI (1968) that reacts rapidly at less extreme pH (pH 8). The derivative obtained is stable in acidic media and labile in alkaline media. Among the other arginine reagents, butanedione, its dimer or its trimer were used at pH values on the order of 8 to modify proteins. **1,2-cyclohexane dione** is used at higher pH.

  Table 9.10 indicates some of the reactions with arginine.

Table 9.10 Reactions of arginine

Reagent	*Product*	*Use*
 Malonic aldehyde		Arginine modified by this reactions is no longer accessible to trypsin.
 Nitromalonic dialdehyde		
 Phenylglyoxal		
 Butanedione		
 Cyclohexane dione		

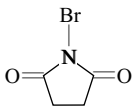
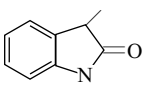
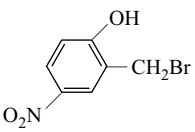
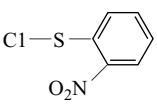
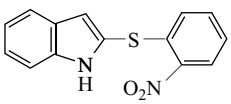
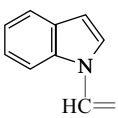
Tryptophan

Most of the reagents utilised to modify tryptophan react also with thiol groups which therefore must be protected beforehand. Oxidation of the indole nucleus is frequently used.

- Oxidation can be obtained by action of hydrogen peroxide, by ozonolysis in performic media in the presence of resorcinol; however this last reaction also modifies methionines. **N-bromosuccinimide** reacts rapidly with tryptophan to give an oxindole derivative that absorbs at 250 nm; although very utilised, this reaction is hardly selective. **2(2-nitrophenyl-sulfenyl)-3-methyl-3-bromoindolenine** (BNPS skatole) replaced N-bromosuccinimide; this compound also converts tryptophan into an oxindole derivative; it also reacts with methionine to give the corresponding sulfoxide derivative.

  Certain nitrobenzyl halogenides are very reactive towards tryptophan. **2-hydroxy-5-nitrobenzene bromide** is capable of alkylating tryptophan in gentle enough conditions. The compound formed has an absorption maximum at 410 nm that permits following its appearance. This reagent gives with methionine a derivative that spontaneously decomposes. Sulfenic halogenides such as **2-nitro or 2,4-dinitro-sulfenyl chloride** react in acidic media (pH < 3.5); they form compounds having a characteristic spectrum of absorption (Tab. 9.11). This reagent is also selective with tryptophans when the SH groups have been protected.

Table 9.11 Reactions of tryptophan

Reagent	*Product*	*Use*	*Comments*
Oxidation by H_2O_2 Light + methylene blue	Destruction of the indole nucleus	Photo-oxidation	
 N-bromosuccinimide	 Oxyindole		Generally brings about the cleavage of the peptide bond
 2-hydroxy-5-nitrobenzyl bromide	Complex reaction	High specific substitution for tryptophan	
 2-nitrophenylsulfenyl chloride (NPSCI)		Specific substitution of tryptophan in the absence of SH	
HCOOH + HCl Formic acid		Reversible substitution of tryptophan	

Another interesting modification is the formylation of tryptophan by formic acid in the presence of HCl; this modification is reversible in alkaline media. ◢

9.2.3. CRITERIA USED TO INTERPRET RESULTS

In order to determine if an amino acid residue is essential to the catalytic activity, it is necessary to define a certain number of criteria permitting the evaluation of the importance and role of a modified residue. **The diverse criteria used have their limits and one must not be satisfied with just one.** The principal criteria are stoichiometric inactivation, protection by the substrate, criteria following from kinetic analysis and of course, when it is possible to delete the chemical modification, the return of the enzymatic activity.

9.2.3.1. STOICHIOMETRIC INACTIVATION

When one modifies a residue essential to the active centre of an enzyme, one must observe a total inactivation for a reagent molecule bound to the active site; the inactivation must be stoichiometric. It will be convenient therefore to determine first of all the number of reagent molecules incorporated into the active site. In addition, it must be assured that there is a total loss of activity. Indeed, it is frequently observed that the labelling of a residue brings about a loss of activity down to a very weak value, sometimes less than 1%; in this case, the labelling of an essential residue cannot be concluded. To verify this, it is necessary to carry out a measure of activity with concentrations of the labelled enzyme two to three orders of magnitude greater than those used with the active enzyme.

If the loss of activity is not complete, the labelled group is not a catalytic group. It can be implicated in the ligand binding, the labelling bringing about an inactivation resulting from steric hindrance. The loss of activity can also result from a conformational variation, even weak, induced by the chemical modification. The binding of the reagent can occur on a non-catalytic residue located in the proximity of the catalytic residue in the three-dimensional structure, but which can be very distant in the sequence, rendering the interpretation difficult when the three-dimensional structure is not known.

9.2.3.2. SPECIFIC PROTECTION AGAINST INACTIVATION

Another criteria used correlatively with the preceding is protection by a competitive inhibitor or a mildly reactive substrate that blocks the entry to the active site. If the reagent labels a catalytic residue, one must observe a decrease in the labelling rate. However this criteria is very relative. The presence of the inhibitor will bring about a comparable effect if it concerns the labelling of groups neighboring the catalytic residue.

9.2.3.3. KINETIC ANALYSIS OF RESULTS

An indispensable and much more reliable criteria is that which results from the kinetic analysis of labelling and the loss of activity for each individual group.

Kinetic analysis for pseudo-first order reactions

Diagnostic rules permitting the determination of the role of a group labelled by a reagent were proposed by RAY and KOSHLAND (1961) in the case of pseudo-first order reactions, when the enzyme is treated by a large excess of reagent. The method developed consists of comparing the rate constants of the incorporation of the reagent and of the decrease in enzymatic activity. To illustrate the principle of analysis, the authors gave the example of a protein that possesses two histidine residues equally reactive towards a chemical compound. If the reaction is carried out to 50% of labelling, the distribution of molecules is that presented in Fig. 9.24; that is 25% of the molecules are not labelled by the reagent, 25% are labelled on residue 1, 25% on residue 2 and 25% on the two residues. Following the role of each of these two residues in the catalytic activity, the results observed will be different. If residue number 1 is essential, but not residue number 2, only the molecules E_1 and $E_{1,2}$ will be inactive. One will observe a 50% loss of activity; likewise if only residue 2 is essential. If the two residues are essential, the loss of activity will be 75%. If one of the two residues can substitute for the other, the loss of activity will only be 25%.

Fig. 9.24 Distribution of species having labelled histidine residues

In the case where there is a modification of a single essential group, the kinetics of labelling must correspond to the kinetics of activity loss. Indeed, if the protein is composed of diverse reactive groups, X_1, X_2, … X_n, the residues being modified with pseudo-first order constants k_1, k_2, … k_n, the fractions of residues not yet modified at time t will be: $x_1/x_{1,0}$, $x_2/x_{2,0}$, … $x_n/x_{n,0}$, respectively. The kinetics of labelling are given by the following equations:

$$x_1/x_{1,0} = e^{-k}1^t \quad ; \quad x_2/x_{2,0} = e^{-k}2^t \quad ; \quad \ldots\ x_n/x_{n,0} = e^{-k}n^t$$

The kinetic of disappearance of enzymatic activity is of the form:

$$A/A_0 = e^{-k_A t}$$

A being the activity remaining at time t, A_0 the initial activity, and k_A the rate constant. If a residue X_1 has been labelled with a constant k_1 and if $k_A = k_1$, the labelled residue is essential. If the two residues X_1 and X_2 have been labelled and if $k_A = k_1 + k_2$, the two residues are essential; the labelling of one or the other brings about a loss of activity. If the residue X_1 has reacted with the constant k_1 and if $k_A > k_1$, this residue is not a catalytic group (Fig. 9.25 below).

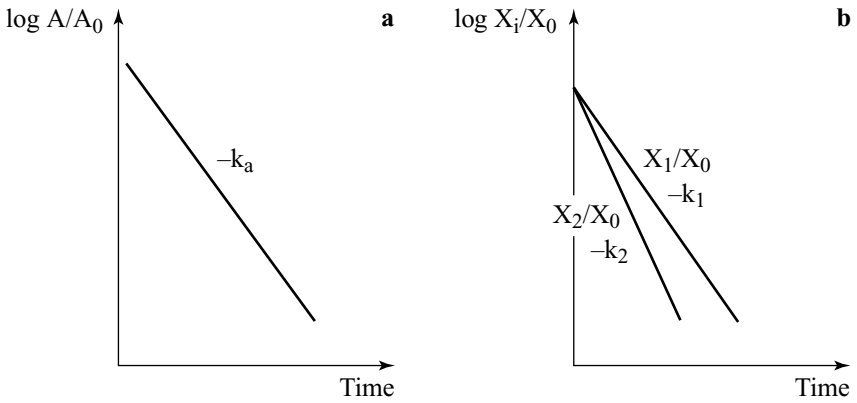
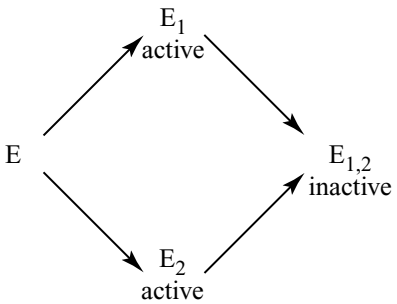


Fig. 9.25 Kinetics of residue labelling
*(**a**) kinetics of disappearance of enzymatic activity – (**b**) pseudo-first order kinetics for labelling of residues X_1 and X_2.*

RAY and KOSHLAND also considered two other cases, first that in which the enzyme would involve two residues X_1 and X_2 that can interfere in a complementary manner in the catalysis; this is the *either/or* (one or the other) case symbolised by the scheme:



The loss of activity as a function of time is then reliant on the rate of disappearance of groups X_1 and X_2 by the relationship:

$$A/A_0 = e^{-k_A t} = e^{-k_1 t} + e^{-k_2 t} + e^{-(k_1+k_2)\,t}$$

It seems very improbable that this case occurs at the active centre of enzymes.

The other case, more realistic, is that of the progressive denaturation of the enzyme bringing about the labelling of a non-essential residue:

$$\underset{\text{active}}{E} \xrightarrow{k_1} \underset{\text{active}}{E_1} \xrightarrow{k_2} \underset{\text{inactive}}{E_{1,2}}$$

In this case, the modification of the non-essential residue X_1 does not affect the activity of the enzyme but its conformation, so that the residue X_2 that is essential and non-reactive in the initial form becomes reactive in the form E_1. The kinetics of the reaction are:

$$A/A_0 = [k'_2/(k'_2 - k_1)]e^{-k_1 t} - [k_1/(k'_2 - k_1)]e^{-k_2 t}$$

Analysis of results in the most general case

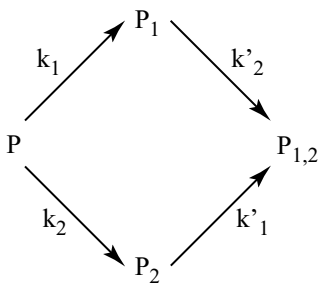
The analysis of RAY and KOSHLAND constitutes an important approach; however, it is limited to chemical modifications that are performed according to a pseudo-first order reaction, meaning in the presence of a large excess of reagent that constitutes a limitation. These conditions are present in a great number of cases, but for the others they are difficult to perform and control. TSOU (1962) proposed a more general treatment permitting the achievement of an analysis which applies whatever concentration of the reagent relative to titratable groups; this treatment applies to all cases and presents therefore a great interest when the treatment of RAY and KOSHLAND cannot apply.

TSOU considers different cases that are practically the same as those treated by RAY and KOSHLAND; he subdivides them following the relative reactivities of diverse categories of groups and the distribution of essential residues in these categories. He introduces a distinction based on the nature of the residues (His, Tyr, Lys etc.). The kinetic analysis however does not distinguish the nature of the groups, but only their reaction rate. TSOU does not make any hypothesis on the order of the reactions and considers, following the different cases treated, the variations in the residual activity fraction as a function of the fraction of non-modified essential groups. The case I which is the most classical relates to a single essential group; all the groups of the same type react with the same rate. If the number of groups essential to the activity is i, the activity fraction staying a, and the fraction of non-modified essential groups x_e, the following relationship must be verified:

$$a^{\frac{1}{i}} = x_e$$

The analysis consists of representing either log a as a function of log x, or a, $\sqrt{a}$, $\sqrt[3]{a}$, as a function of x until one obtains a linear graph.

◤ A generalisation of the treatment of TSOU was proposed by TENU and YON (unpublished results) of which the essential is described below. Different possibilities derive from a general model difficult to resolve in the case of n groups, but easy to represent and to resolve in the case of two residues. The scheme can be thus represented:



P_1 and P_2 represent the labelled species on the residues 1 and 2 respectively. In $P_{1,2}$, the two residues 1 and 2 are blocked by the reagent. One can be in the presence of two types of reactions, either a random mechanism with interactions ($k_1 \neq k'_1$; $k_2 \neq k'_2$), or without interactions ($k_1 = k'_1$; $k_2 = k'_2$), or an ordered mechanism ($k_2 = 0$ or $k_2 = k'_1 = 0$).

Random mechanisms

In the most general case of n residues, if x_i is the probability of the residue i being intact and y_i the probability that it is modified, these values are related by the expression:

$$x_i + y_i = 1$$

The number of species blocked on i residues is given by the relationship:

$$C_n^i = \frac{n!}{i!(n-i)!}$$

When all the rate constants are equal, the probabilities of formation of diverse species are:

$$x^n,\ C_n x^{n-1} y,\ \ldots\ C_n^i x^{n-1} y_i,\ \ldots$$

and:

$$(x+y)^n = x^n + \sum_{i=1}^{i=n} C_n^i x^{n-1} y_i$$

The chemical modification produces 2^n species. The conserved activity fraction corresponds with the sum of probabilities of diverse active species. Aside from this formulation one can find again different cases treated by RAY and KOSHLAND, and by TSOU.

Let us consider first the modification of essential residues, diverse residues reacting with whatever rates. This situation corresponds with the first case of RAY and KOSHLAND and with the third case of TSOU. One can show that:

$$a = \prod_i x_i^{b_i}$$

n_i being the number of residues modified with the constant k_i, x_i the fraction of intact residues n_i, b_i the number of essential residues among the n_i and r the number of residues blocked per enzyme molecule.

The residues are modified with the rate:

$$\frac{dx_i}{dt} = -k_i R^i x_i \quad ; \quad \frac{dx_j}{dt} = -k_j R^j x_j$$

R being the concentration of the reagent. The pseudo-first order conditions are fulfilled for i = j = 1. In the general case, one has:

$$\frac{\dfrac{dx_i}{x_i}}{\dfrac{dx_j}{x_j}} = \frac{k_i R^i}{k_j R^j}$$

There is a relationship between r and a:

$$r = \sum n_i - \sum n_i a^{k_i / \sum b_i k_i}$$

An interesting piece of information is given by the analysis of tangents to the origin; for a = 1, r = 0 of the curve a = f[r]. In the general case:

$$\left(\frac{dr}{da}\right)_{a=1} = -\frac{\sum n_i k_i}{\sum b_i k_i}$$

$$\left(\frac{da}{dr}\right)_{r=0} = -\frac{\sum b_i k_i}{\sum n_i k_i}$$

If the ratio $r/\Sigma n_i$ is not known with accuracy because there is an imprecision in a parameter (extinction coefficient of the protein or specific radioactivity of the reagent), it is nevertheless possible to determine the number of residues implicated Σb_i. Indeed, the representation $a^{\alpha} = f[r]$ is linear if $\alpha = 1/\Sigma b_i$ is independent of adopted parameters.

When all the residues react with the same rate $k_i = k_1$ and the relationship becomes:

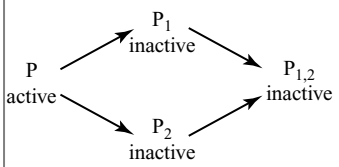
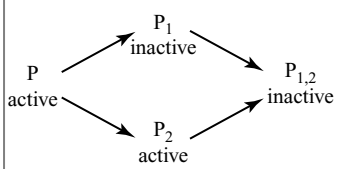
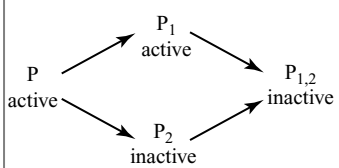
$$r = \sum n_i - \sum n_i a^{1/\sum b_i}$$

which can also be written: $$a^{1/\sum b_i} = 1 - r/\sum n_i$$

This is the first case treated by TSOU.

To illustrate this analysis, the simplest case of two residues is detailed below. One assumes that the two residues react with the rate constants k_1 and k_2 such that $k_1 < k_2$. It concerns a random mechanism without interactions between the reacting residues. In this type of mechanism, three possibilities must be considered, that the two residues are essential to the activity or only one, either residue 1 or residue 2 (Table 9.12).

Table 9.12 Case I: case I can lead to three possibilities

		Residue 1	*Residue 2*	*Activity*	
I.1	The two residues are implicated in the activity.	1 1 0 0	1 0 1 0	1 0 0 0	
I.2	Residue 1 is rapidly modified and is implicated in the activity.	1 1 0 0	1 0 1 0	1 1 0 0	
I.3	Residue 2 is slowly modified and is implicated in the activity.	1 1 0 0	1 0 1 0	1 0 1 0	

In the first case, one can show that:

$$r = 2 - a^{k_1/(k_1+k_2)} - a^{k_2/(k_1+k_2)}$$

if $k_1 = k_2$, the expression becomes: $a = (1 - r/2)^2$

r being the number of residues modified and a the residual activity. This expression corresponds with the first case of TSOU.

When residue 2 which reacts rapidly is the only group essential to the activity, the relationship between r and a becomes:

$$r = 2 - a - a^{k_1/k_2}$$

On the contrary, if the essential residue is that which reacts the most slowly, one has the following expression:

$$r = 2 - a - a^{k_2/k_1}$$

If $k_1 = k_2$, the two preceding expressions simplify into:

$$r = 2 - 2a$$

Figure 9.26 represents the variations in a as a function of r in these diverse cases and gives the value of the slope of initial and final tangents of the curve.

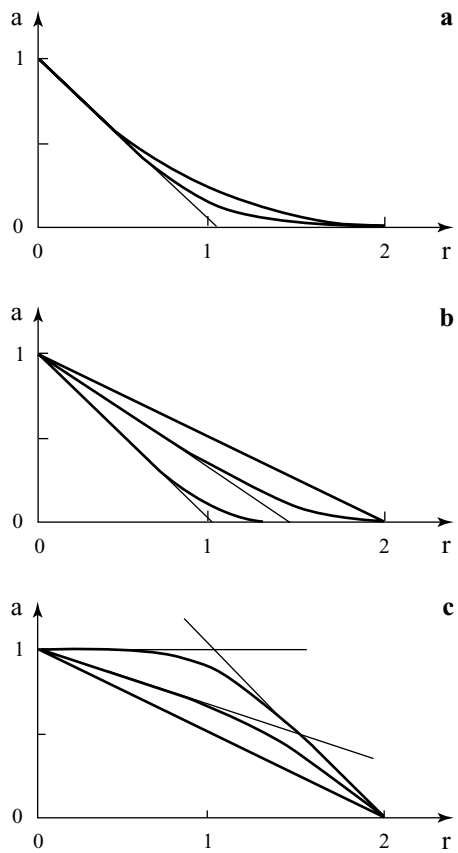


Fig. 9.26 Variations in a as a function of r in the random mechanism of labelling two residues
*Figure (**a**), (**b**) and (**c**) correspond respectively to the three cases of Table 9.12; the lines of Fig. (**b**) and (**c**) correspond to the case where $k_1 = k_2$*

In the case where the modification of all the essential groups is required for the loss of activity, which corresponds to the *either/or* case of RAY and KOSHLAND, the expression of a as a function of r becomes:

$$a = 1 - \frac{r^2}{4}$$

Ordered mechanisms

This type of mechanism can be symbolised by the following scheme:

$$\underset{\text{active}}{E} \xrightarrow{k_1} \underset{\text{active}}{E_1} \xrightarrow{k'_2} \underset{\text{inactive}}{E_{1,2}}$$

which is identical to the case of progressive denaturation treated by RAY and KOSHLAND in which the species $E_{1,2}$ is inactive. Yet the mathematical treatment is possible whatever the order of the reaction is. One defines two relations, one giving the expression of a, the other giving the expression of r. These expressions do not depend on the concentration of the reagent, but on the enzyme fraction (E/E_t) which reacted:

$$a = \frac{k_1}{k_1 - k'_2}\left[\left(\frac{E}{E_t}\right)^{k'_2/k_1} - \frac{k'_2}{k_1} \times \left(\frac{E}{E_t}\right)\right]$$

$$r = 2 + \left(\frac{2k'_2 - k_1}{k_1 - k'_2}\right) \times \frac{E}{E_t} - \frac{k_1}{k_1 - k'_2} \times \left(\frac{E}{E_t}\right)^{k'_2/k_1}$$

In these expressions, it is not possible to eliminate (E/E_t). Nevertheless the test of the slope of the tangents to the curve $a = f(r)$ brings about important conclusions. The tangent of the curve da/dr is given by the equation:

$$\frac{da}{dr} = \frac{k'_2\left[\left(\frac{E}{E_t}\right)^{(k'_2 - k_1)/k_1} - 1\right]}{2k'_2 - 1 - k'_2\left(\frac{E}{E_t}\right)^{(k'_2 - k_1)}}$$

This slope takes characteristic values for $r = 0$, when no residue has been blocked by the reagent, and for $r = 2$, when two residues are labelled.

If $k'_2 > k_1$, one has $\left(\frac{da}{dr}\right)_{r=0} = 0$ and $\left(\frac{da}{dr}\right)_{r=2} = -\frac{k'_2 - k_1}{2k'_2}$

If $k'_2 < k_1$ $\left(\frac{da}{dr}\right)_{r=0} = 0$ and $\left(\frac{da}{dr}\right)_{r=2} = -1$

Table 9.13 below summarises all these different cases.

Table 9.13 Compared analysis of the different cases for n = 2
(from TENU, Thèse Orsay, 1978)

Case	*Function a = f[r] or r = g[a]*	$\left(\frac{da}{dr}\right)_{r=0}$	$\left(\frac{da}{dr}\right)_{r=2}$
Random: I.1 Residues 1 and 2 are essential	$r = 2 - a^{k_1/(k_1+k_2)} - a^{k_2/(k_1+k_2)}$ $a = \left(1 - \frac{r}{2}\right)^2$ if $k_1 = k_2$	-1	0
Random: I.2 Residue 2 is blocked more rapidly than residue 1 and is essential	$r = 2 - a - a^{k_1/k_2}$ $a = 1 - \frac{r}{2}$ if $k_1 = k_2$	$-\frac{k_2}{k_1+k_2}$	0
Random: I.3 Residue 1 is blocked more slowly than residue 2 and is essential	$r = 2 - a - a^{k_2/k_1}$ $a = 1 - \frac{r}{2}$ if $k_1 = k_2$	$-\frac{k_1}{k_1+k_2}$	-1
Either/or *contribution of two residues:* The activity is intact when the two residues are not blocked	$a = 1 - \frac{r^2}{4}$ if $k_1 = k_2$	0	-1
Sequential scheme: $k'_2 > k_1$		0	$-\frac{k'_2 - k_1}{2k'_2}$
Sequential scheme: $k'_2 \leq k_1$		0	-1

For example, Fig. 9.27 is a representation of a as a function of r for the modification of histidine residues by diethyl pyrocarbonate in pyruvate kinase according to BORNMAN et al. (1974).

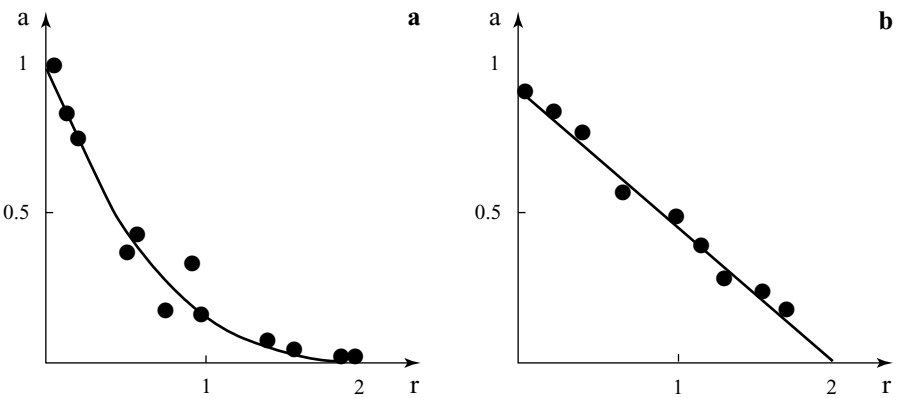


Fig. 9.27 Examples of analysis of results from chemical labelling of lysine residues of glucose-6-phosphate dehydrogenase (a) and of histidine residues of pyruvate kinase (b)

The results indicate that an essential histidine was modified by the reagent. In the case of glucose-6-phosphate dehydrogenase, two lysine residues were modified upon binding of pyridoxal phosphate (from MILHAUSEN & LEVY, 1975). The representation a = f[r] gives a curve, the tangent at the origin $(da/dr)_{r=0} = -1$ and $(da/dr)_{r=2}$ is null. This situation corresponds with the first case where the two residues are essential for the activity. The authors meanwhile had concluded a single essential residue since, when a residue is modified, more than 90% of the enzyme activity is lost. This example illustrates well the importance of rigorous analysis in the determination of residues essential to the activity of an enzyme. ◢

9.2.3.4. REVERSIBILITY OF THE CHEMICAL MODIFICATION AND THE LOSS OF ACTIVITY

In the case where the chemical modification can be reversed or more generally in the case where the reagent blocking an essential residue can be eliminated, one can recover the initial enzymatic activity. For example, in the case of serine acetylation by paranitrophenyl acetate in serine proteases such as α-chymotrypsin or trypsin, the chemical modification is stable in acidic media, but the increase in pH provokes the departure of the acetyl group; there is deacylation and the enzyme activity is restored. In the case of thiol proteases, the reversible labelling of the essential cysteine is possible and the enzyme recovers its activity after elimination of the chemical reagent.

However, the elimination of the chemical reagent employed to label a catalytic residue is not always possible. When it is, the criteria for reversibility constitutes an important element in the interpretation of the chemical modification.

9.3. USE OF MUTAGENESIS METHODS TO STUDY THE ACTIVE CENTRE OF ENZYMES

Introduced later, genetic methods, in particular site-directed mutagenesis techniques, offer another way of approaching the determination of groups essential to the enzyme activity. Site-directed mutagenesis techniques arose from the meeting of an American researcher, HUTCHINSON, and a Canadian researcher SMITH who were carrying out a year of sabbatical in the laboratory of SANGER at Cambridge at the moment where the sequencing of DNA was coming into focus (1977). These techniques permit the replacement of any amino acid of a protein by another and are perfectly adapted to the determination of amino acids implicated in enzymatic catalysis.

It is not our purpose to expand on the different methods of mutagenesis that are classical today and we redirect the reader to the specialised work: *Directed mutagenesis: a practical approach* (edited by MC PHERSON, Oxford University Press). We simply give the principle of one of these methods.

9.3.1. METHODOLOGY

The methodology is well developed today and represents a current technique concerning the modification of DNA. There exist different methods for replacing by genetic methods a determined amino acid in a protein. Figure 9.28 illustrates the principle of one of these methods, which utilises oligonucleotide synthesis.

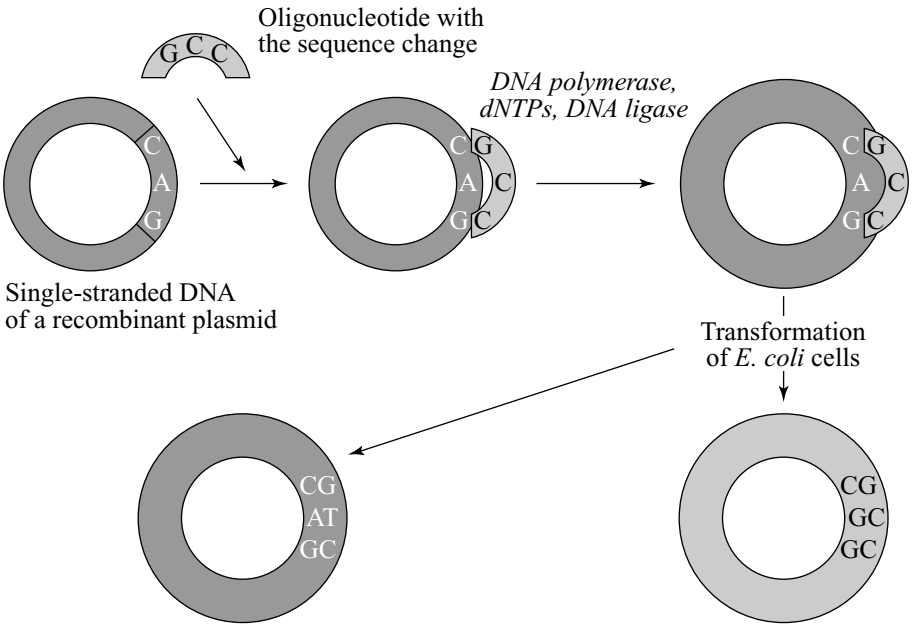


Fig. 9.28 Principle of the directed mutagenesis technique

- The DNA fragment corresponding to the protein gene that one wishes to mutate is cloned on a bacterial vector, generally a plasmid that permits obtaining a single-stranded form.
- The DNA fragment sequence is established and one defines the modifications that one wishes to introduce.
- The complementary oligonucleotide of the chosen sequence including the modification is synthesised by chemical methods.
- The oligonucleotide can be hybridised in vitro with a single-stranded vector. The hybrid is stable at a temperature less than the transition temperature T_m.
- DNA polymerase I recognises this hybrid as a substrate for replication in vitro, the oligonucleotide playing the role of primer and the single-stranded vector that of the matrix (template).
- This DNA is introduced into cells for transformation (*E. coli*, yeast…). It is replicated and gives rise to a mixed population, one part of the vectors carrying the

mutation, and not the other. Over the course of generations, there occurs a segregation.

- There remains selecting the colonies of mutant cells from colonies of wild type cells. One provides a simple method based on the following property: the hybrid formed between mutated DNA and the oligonucleotide has a greater thermostability than that formed by wild type DNA and the oligonucleotide.

There exist variations of this technique. An important condition for the application of the method is that the gene corresponding to the mutated protein expresses correctly and that the protein is not degradated by proteases. Although the modification produced in the DNA constitutes a well tested technique, the conditions of expression and degradation occasionally pose serious problems.

There currently exist variable genetic methods that are the object of numerous works. Aside from rational methods like site-directed mutagenesis, other methods proceed in a different manner like random mutagenesis and the techniques of directed evolution with selection of mutants as a function of their properties or the absence of them and further analysis of the mutation. These techniques however are essentially utilised for the remodelling of proteins in order to generate new properties. The selection is performed from bound ligands and uses methods like presentation on phage (phage display), or on ribosomes (ribosome display) or on RNA (RNA display). These methods are the subject of numerous reviews (TOBIN et al., 2000; O'NEIL & HOESS, 1995; HANES & PLÜKTHUN, 1997; WILSON & SZOSTAK, 1998). ◢

9.3.2. STRATEGY

The strategy used involves a preliminary, the choice of the protein modification. This is evidently dictated by the question which one wants to answer. If it concerns understanding the role of a group of a side chain in the catalysis or in the binding of the substrate, it will be convenient to consider first of all the three-dimensional structure of the enzyme studied and the position of the substrate or of an analog at the active centre. This strategy is equally valid for using chemical labelling methods when the structural data is available. All rational mutation implies knowledge of the structure. According to the position of the most likely implicated amino acids, the use of molecular modelling and energy minimisation methods will help to conceive the most sensible replacement. An important point to consider is that of internal dynamics of the protein. Indeed, the substitution of an amino acid can have important structural consequences in a region of the protein far from the mutation site. The use of genetic methods for determining catalytic groups of enzymes calls however some reservations. They do not permit like chemical labelling methods achieving a kinetic analysis of the results. In addition, the loss of activity engendered by the mutation can result in a secondary effect and does not necessarily indicate that the modified group is directly implicated in the catalysis.

9.3.3. SOME EXAMPLES

Since the introduction of the method, the work using directed mutagenesis to determine the groups essential to the activity of enzymes has rapidly increased. Also there is no question in this chapter of making an exhaustive review. Only some examples are given here in order to illustrate the potentials of the method. Other applications will be given relevant to particular enzymatic systems in Part IV.

β-lactamase possesses a serine in position 70 and a threonine in position 71 that are conserved in all species. The mutation Ser70 ⟶ Thr brings about a total loss of enzymatic activity showing the essential role of this residue for catalysis. The mutant Thr71 ⟶ Ser, on the other hand conserves still 15% of the activity of the wild type enzyme. The replacement of His170 by tyrosine in anthranilate synthase, the enzyme that catalyses the first reaction in the biosynthesis of tryptophan, brings about the loss of activity of the enzyme with the substrate glutamine, but not with ammoniac. This reaction proceeds normally *via* the formation of a covalent intermediate, glutaminyl-Cys84. The results obtained after directed mutagenesis indicate the essential role of histidine that seems to interfere like a general base catalyst in the glutaminylation of cysteine 84.

In the case of glycyl-tRNA synthetase, chemical modification experiments with N-ethylmaleimide had suggested that Cys395 is essential to the catalysis. A mutant Cys395 ⟶ Glu was constructed. It presented a weak but significant activity, representing about 10% of that of the wild type enzyme. The mutation was not affect-ing the rate of formation of aminoacyl adenylate, but merely the step depending on tRNA. These experiments showed that the cysteine residue is not a catalytic residue. Actually the mutation, like the labelling by NEM, limits the reaction by a steric effect.

The α subunit of tryptophan synthase catalyses the following reaction:

indole-3-glycerol phosphate ⇌ indole + glyceraldehyde-3-phosphate

The β subunit catalyses the formation of tryptophan starting with indole and serine. It was shown that the residue glutamate 49 of the α subunit is a group essential to the activity. Indeed its replacement by whatever one of the 19 other amino acids by directed mutagenesis brings about a totally inactive enzyme.

9.4. STRUCTURAL STUDIES BY RADIOCRYSTALLOGRAPHY AND NUCLEAR MAGNETIC RESONANCE OF THE ACTIVE CENTRE OF ENZYMES

The precise knowledge of protein structures resulting from development of crystallographic and nuclear magnetic resonance studies (NMR) constitutes an important contribution to the determination of active sites of enzymes and enzyme-substrate interactions. It permits specifying the position of enzyme groups in relation to diverse substrate atoms, therefore the knowledge of groups localised to the proximity

of the disrupted substrate bond and groups susceptible to interfering in the catalysis. Moreover it permits determination of structural changes that interfere when the substrate is bound to the enzyme. If crystallisation stays a limiting stage in the determination of the structure of enzymes, the crystallogenesis techniques have progressed considerably over the course of recent years going to the automation of methods, including weak volumes on the order of a nanoliter (KNIL et al., 2002). On the other hand, the use of genetic engineering permits the overexpression of genes and the production of large quantities of the protein studied. Nevertheless this overexpression frequently leads to the formation of aggregates called inclusion bodies. In favorable cases, the native protein can be obtained starting from these aggregates by dissolving in a denaturant and then refolding (YON, 2004).

However radiocrystallographic studies are difficult carried out with real substrates; the enzyme-substrate complex decomposes into reaction products during the acquisition of X-ray data. One must resort to substrate analogs or specific inhibitors. To obtain the complexes between an enzyme and its pseudo-substrates, there are several possibilities. The protein crystals contain around 50% of solvent and some substrates diffuse in the crystal without perturbing the enzyme structure. In other cases, it is possible to co-crystallise the enzyme in the presence of pseudo-substrates. In these conditions, if there are no profound modifications in the enzyme structure in the presence of its substrate analog, it is possible to solve the structure of the complex by the FOURIER difference technique. The method permits measuring differences between diffraction diagrams of crystals of the free enzyme and crystals of the enzyme having co-crystallised with its ligand. Therefore, the changes can be observed without necessarily solving the entire structure of the complex. The first attempts to determine the structure of enzyme-substrate complexes were based on the extrapolation of results obtained starting with enzyme-inhibitor complexes. In the case of lysozyme, a portion of the substrate playing the role of the inhibitor was bound to the enzyme, and the rest was determined starting from molecular models. Today molecular modelling methods constitute an effect-ive tool in the determination of enzyme-substrate interactions.

Starting from coordinates deposited in the Protein Data Bank (PDB), in using modelling software like for example SwissPDB Viewer accessible on the internet, it is possible to view the three-dimensional structure of a protein and to turn it in space in order to visualise the regions of interest.

Despite their power and their precision, radiocrystallography methods exclude neither the chemical methods nor the genetic methods for determination of active sites of enzymes; these approaches are complementary. The use of chemical labelling and the analysis which is derived from it stays very useful for determining the role of amino acids in contact with substrate atoms. In addition, the chemical approach was used to respond to an important question: is the structure of the active site of an enzyme the same in the crystal where constraints exist and in solution? Use of bifunctional reagents has permitted in certain cases to evaluate the distance between

functional groups of an enzyme in solution and to compare it to those derived from the structural analysis. In other cases the comparison was done starting with the nuclear magnetic resonance (NMR) analysis. In fact, in most systems, the structure of an enzyme is analogous in the crystal and in solution. However, when a protein oscillates between two or several conformations, crystallography can trap one of these conformations. This is observed for example in the case of allosteric enzymes, and in precise cases, the contribution of these methods has permitted the determination of steps of this transition (see Part V).

For certain enzymes, like lysozyme, the nature of catalytic groups was not known before structural studies. For others, structural studies confirmed and stated precisely the results of chemical studies; in some cases, they were refuted. In order to illustrate the complementarity between the two approaches, some examples are given in Chap. 12 where are treated in detail some enzymatic systems of which the three-dimensional structure is known. The chosen examples show the import-ance of structural studies in the determination of groups essential to the enzymatic activity and in the determination of catalytic mechanisms; they equally underline the limits. They also indicate the necessity of uniting the structural studies, the chemical and/or genetic approaches and the kinetic studies. It is only by the conjunction of these different methods that a catalytic mechanism can be elucidated.

High field NMR for the study of proteins was developed more recently. Beginning from the 1970s, NMR of the proton began to be used for the study of proteins. Then, more recently the study of nuclei ^{13}C and ^{15}N was considerably improved permitting the application of multidimensional heteronuclear spectroscopy methods. NMR has the advantage of applying to proteins in solution and no longer in the crystal; contrary to X-ray diffraction, it identifies the proton. However it requires important protein concentrations on the order of 0.5–1 mM. Taking into account this limitation, it is important to have specific probes in the study of protein-ligand interactions and the determination of active sites of enzymes. The most judicious is to use probes carried by the ligand after isotopic enrichment in ^{13}C or ^{15}N or even ^{19}F or ^{2}H after chemical modification. With phosphorylated compounds, the resonance spectroscopy of ^{31}P is very utilised. For example, the spectrum of ^{31}P of free NADPH and that of the coenzyme bound to a dehydrogenase give information on the conformation of the bound coenzyme and on its interactions with the enzyme. The results showed that the conformation of the bound coenzyme differs from that of the free coenzyme and revealed an electrostatic interaction with the positively charged histidine of the enzyme. Another study showed that the binding of 3'CMP to ribonuclease affects the position of protons C2 and C4 of histidine 12 and of histidine 119 that are catalytic residues.

For the methodological aspect of radiocrystallography and of NMR of proteins, we redirect the reader to two fundamental books: *Biophysical chemistry. Part II for the study of biological structure and function*, by C. CANTOR & P.R. SCHIMMEL (FREEMAN ed., 1980) and *Biologie structurale: principes et méthodes biophysiques*, by J. JANIN & M. DELEPIERRE (HERMANN ed., 1994).

In summary, the study of the topology of the active centre of enzymes requires an ensemble of complementary methods. The kinetic approach as a function of pH permits determining the role of groups implicated in the activity, but does not authorise the conclusion of the nature of these groups. The research on the nature of amino acid residues implicated in the catalysis necessitates the utilisation of chemical and genetic methods bringing about in parallel enzymatic studies.

Different strategies are employed for the chemical modification of residues of the active centre of an enzyme. The use of pseudo-substrates and of affinity or photoaffinity reagents when it is possible to prepare the adequate reagent brings about a specific labelling. However, when this strategy cannot be used, there exists an ensemble of more or less selective reagents for the group that one wants to label. In choosing sensibly the conditions of pH, ionic force, it is possible to selectively modify a precise residue. It is important to underline that chemical methods require numerous controls and a rigorous analysis in order to establish if the labelled residue is essential to the catalysis.

Directed mutagenesis methods that are highly developed constitute another way to approach very efficiently the determination of residues essential to the catalysis; they are more and more utilised. In this case equally, the interpretation of the results must be based on a rigorous analysis, and numerous controls of structural properties of the mutated protein must be carried out.

The combination of all these methods largely benefits from precise structural knowledge in order to rationally orient experimentation. The number of structures of proteins currently known is on the order of several tens of thousands and still increasing, offering therefore great possibilities to study enzymatic function. Despite these diverse approaches and the power of tools at our disposal, many mechanisms are not yet totally elucidated.

BIBLIOGRAPHY

BOOKS

CANTOR C.R. & SCHIMMEL P.R. –1980– *Biophysical chemistry. Part II: Techniques for the study of biological structure and function*, Freeman, San Francisco.

DIXON M. & WEBB E.C. –1964– *Enzymes*, Longmans, London.

GLAZER A.N., DELANGE R.J. & SIGMAN D.S. –1975– Chemical modifications of proteins. Selected methods and analytical procedures, in *Laboratory techniques in biochemistry*, T.S. WORK & E. WORK eds, North Holland American, Elsevier, New York.

JANIN J. & DELEPIERRE M. –1994– *Biologie structurale: principes et méthodes biophysiques*, Hermann, Paris.

LUNDBLAD R.L. & NOYES C.M. –1985– *Chemical reagents for protein modifications*, Vol. I and II, CRC Press, Boca Raton.

MC PHERSON M.J. –1990– *Directed mutagenesis: a practical approach*, Oxford University Press.

MEANS G.E. & FEENEY R.E. –1971– *Chemical modifications of proteins*, Holden Day Inc., San Francisco.

REVIEWS

BAYLEY H. & KNOWLES J.R. –1977– Photoaffinity labeling. *Methods Enzymol.* **46**, 69–114.

BLOW D.M. –1971– The structure of chymotrypsin, in *The Enzymes*, 3rd ed., Vol. III, Acad. Press, New York, 185–212.

BODLAENDER, P. FEINSTEIN G. & SHAW E. –1969– *Biochemistry* **8**, 4941.

BORNMANN L. & HESS, H. –1974– *Hoppe Seylers Z. Physiok. Chem.* **355**, 1073.

CHOWDHRY V. & WESTHEIMER L.H. –1979– Photoaffinity labelling of biological systems. *Annu. Rev. Biochem.* **48**, 293–325.

COHEN L.A. –1968– Group specific reagents in protein chemistry. *Annu. Rev. Biochem.* **37**, 695–726.

GLAZER A.N. –1970– Specific chemical modifications of proteins. *Annu. Rev. Biochem.* **39**, 101–130.

KOSHLAND D.E. –1960– The active site and enzyme action. *Adv. Enzymol.* **22**, 45–96.

KOSHLAND M.E., Englberger F. & Koshland D.E. –1959– *Proc. Natl. Acad. Sci. USA* **45**, 785.

LALLEGERIE P., LEGLER G. & YON J.M. –1982– The use of inhibitors in the study of glycosidases. *Biochimie* **64**, 977–1000.

MCMURRAY, CH. & TRENTHAM D.R. –1968– *Biochem. J.* **115**, 913.

O'NEIL K.T. & HOESS R.H. –1995– Phage display: protein engineering by directed evolution. *Curr. Opin. Struct. Biol.* **5**, 443–449.

PHILIPPS A.T. –1977– Differential labeling: a general technique for selective modification of binding sites. *Methods Enzymol.* **46**, 59–68.

SCHROEDER W.A., SHELTON J.R. & ROBERTSON B. –1967– *Biochim. Biophys. Acta* **147**, 590.

SINGER S.J. –1967– Covalent labeling of active sites. *Adv. Prot. Chem.* **22**, 1–54.

TENU J.P. –1978– Thèse Université de Paris-Sud Orsay.

TOBIN M.B., GUSTAFSSON G. & HUISMAN G.N. –2000– Directed evolution: the rational basis for irrational design. *Curr. Opin. Struct. Biol.* **10**, 421–427.

VALLEE B.L. & RIORDAN J.F. –1969– Chemical approaches to the properties of active sites of enzymes. *Annu. Rev. Biochem.* **38**, 733–794.

WALSH C.T. –1984– Suicide substrates, mechanism-based enzyme inactivators: recent developments. *Annu. Rev. Biochem.* **53**, 493–535.

WESTHEIMER F.H. –1962– *Adv. Enzym.* **24**, 441.

WILCOX P.E. –1967– Esterification. *Methods Enzymol.* **11**, 605–617.

WILSON O.S. & SZOSTAK J.W. –1998– In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.* **68**, 611–645.

WOLD F. –1967– Bifunctional reagents. *Methods Enzymol.* **11**, 617–640.

WOLFENDEN R. –1977– Transition state analogs as potential affinity labeling reagents. *Methods Enzymol.* **46**, 15–28.

YON J.M. –2004– Protein aggregation, in *Encyclopedia of molecular biology and molecular medicine*, Vol. V, R.A. MEYERS ed., J. Wiley VCH, New York.

SPECIALISED ARTICLES

ALBER T., PETSKO G.A. & TSERNOGLOU D. –1976– *Nature* **263**, 297, London.

ALBERTY R.A. & MASSEY V. –1954– *Biochim. Biophys. Acta* **13**, 347.

D'ALBIS A. & BECHET J.J. –1967– *Biochim. Biophys. Acta* **140**, 435.

BECHET J.J., DUPAIX A., YON J.M., WAKSELMAN M., ROBERT J.C. & VILKAS M. –1973– *Eur. J. Biochem.* **35**, 527.

BECHET J.J., GARDIENNET M.C. & YON J.M. –1966– *Biochim. Biophys. Acta* **122**, 101.

BLOW D.M., BIRKTOFT J.J. & HARTLEY B.S. –1969– *Nature* **221**, 337.

BODE W., CHEN Z., BARTELS K., KUNTZBACH C., SCHMIDT KASTEN G. & BARTURIK H. – 1983– *J. Mol. Biol.* **164**, 237.

BODLAENDER P., FEINSTEIN G. & SHAW E. –1969– *Biochemistry* **8**, 4941.

CHEVALLIER J. & YON J.M. –1966– *Biochim. Biophys. Acta* **112**, 116.

COHEN J.A. & WARRINGA M.G. –1953– *Biochim. Biophys. Acta* **11**, 52.

CRESTFIELD A.M., MOORE S. & STEIN W.H. –1963– *J. Biol. Chem.* **238**, 622.

CRESTFIELD., STEIN W.H. & MOORE S. –1963– *J. Biol. Chem.* **238**, 2413 and 2421.

DEGANI Y., NEUMANN H. & PATCHORNIK A. –1970– *J. Am. Chem. Soc.* **92**, 6969.

DEGANI Y. & PATCHORNIK A. –1974– *Biochemistry* **13**, 1.

ELLMAN G.L. –1959– *Arch. Biochim. Biophys.* **82**, 70.

ENGELBERGER F. & KOSHLAND D.E. –1959– *Proc. Natl Acad. Sci. USA* **45**, 1470.

FEINSTEIN G., BODLAENDER P. & SHAW E. –1969– *Biochemistry* **8**, 4949.

GOLD A.A. –1965– *Biochemistry* **4**, 897.

HANES J. & PLÜCKTHUN A. –1997– *Proc. Natl Acad. Sci. USA* **94**, 4957.

HARSHMAN S. & NAJJAR V.A. –1965– *Biochemistry* **4**, 2526.

HENDERSON R. –1970– *J. Mol. Biol.* **54**, 341.

HORTON H.R. & KOSHLAND D.E. –1965– *J. Am. Chem. Soc.* **87**, 1126.

HUGHES P.L., SIEKES L.C., BIETH J. & DIMICOLI J.L. –1982– *J. Mol. Biol.* **162**, 645.

JACOBSON G.R., SCHAFFER M.H., STARK G.R. & VANAMAN T.C. –1973– *J. Biol. Chem.* **248**, 6583.

KARTHA G., BELLO J. & HARKER D. –1967– *Nature* **213**, 862, London.

KEMPE D. & STARK G.R. –1975– *J. Biol. Chem.* **250**, 6861.

KNIL M.E., BODENSTAFF E.R., HOENMAEKER F.J. & ABHAHAMS J.P. –2002– *Enzyme Microb. Tech.* **30**, 262.

KOSSIAKOFF A.A. & SPENCER S.A. –1981– *Biochemistry* **20**, 6462.

LIPSCOMB W.N. –1973– *Proc. Natl Acad. Sci. USA* **70**, 37.

MASSEY V. & ALBERTY R.A. –1954– *Biochim. Biophys. Acta* **13**, 354.

MATTHEWS B.W., SIGLER P.B., HENDERSON R. & BLOW D.M. –1967– *Nature* **214**, 652.

MILHAUSEN M. & LEVY H.R. –1975– *Eur. J. Biochem.* **50**, 453.

OPPENHEIMER L., HESS G.P. & LABOUESSE B. –1966– *J. Biol. Chem.* **241**, 2727.

PARSONS S.M., JAO L., DALQUIST F.W., BORDES C.L., GROFF T., RACS J. & RAFTERY M.A. –1969– *Biochemistry* **8**, 700.

RAFTERY M.A. & COLE R.D. –1963– *Biochem. Biophys. Res. Commun.* **10**, 467.

RAY W.J. & KOSHLAND D.E. –1961– *J. Biol. Chem.* **236**, 1073.

SANGER F. –1945– *Biochem. J.* **39**, 507.

SCHOELLMANN G. & SHAW A.E. –1963– *Biochemistry* **2**, 252.

SHAW A.E., MARES-GUIA A. & COHEN W. –1965– *Biochemistry* **4**, 2219.

SHOTTON D.M. & WATSON H.C. –1970– *Nature* **225**, 811, London.

SINGH A., THORNTON E.R. & WESTHEIMER F.H. –1962– *J. Biol. Chem.* **237**, 3006.

STEITZ T.A., HENDERSON R. & BLOW D.M. –1969– *J. Mol. Biol.* **46**, 337.

TAKAHASHI K. –1968– *J. Biol. Chem.* **243**, 6171.

TSOU C.L. –1962– *Scientia Sinica* **11**, 1535.

WANG S.S. & CARPENTER F.H. –1968– *J. Biol. Chem.* **243**, 3702.

WITTER A. & TUPPY H. –1960– *Biochim. Biophys. Acta* **45**, 4.

Molecular and Cellular Enzymology

Volume II

Grenoble Sciences

Grenoble Sciences pursues a triple aim:
- to publish works responding to a clearly defined project, with no curriculum or vogue constraints,
- to guarantee the selected titles' scientific and pedagogical qualities,
- to propose books at an affordable price to the widest scope of readers.

Each project is selected with the help of anonymous referees, followed by an average one-year interaction between the authors and a Readership Committee whose members' names figure in the front pages of the book. Grenoble Sciences then signs a co-publishing agreement with the most adequate publisher.

Contact: Tél.: (33) (0)4 76 51 46 95 – Fax: (33) (0)4 76 51 45 79
E-mail: Grenoble.Sciences@ujf-grenoble.fr
Website: **http://grenoble-sciences.ujf-grenoble.fr/**

Scientific Director of Grenoble Sciences: Jean Bornarel, Professor at the Joseph Fourier University, France

Grenoble Sciences is supported by the French Ministry of Higher Education and Research and the "Région Rhône-Alpes".

"Molecular and Cellular Enzymology" is a translation of the original French book published by Grenoble Sciences in partnership with EDP Sciences (2005). The Reading Committee of the French version included the following members:

- Jean-Pierre Mazat, Professor at the Victor Ségalen University, Bordeaux 2
- Guy Lauquin, Professor at the Victor Ségalen University, Bordeaux 2
- Jacques Ricard, Professor at the Pierre et Marie Curie University, Paris 7
- Joël Lunardi, Professor at the Joseph Fourier University, Grenoble 1
- Jean Pelmont, Professor at the Joseph Fourier University, Grenoble 1
- Michel Van der Rest, Professor at the École Normale Supérieure de Lyon
- Bahram Houchmandzadeh, Senior Researcher at the CNRS, Laboratoire de Spectrométrie Physique, Grenoble

Translation performed by Philip Simister, Dr. in Biochemistry, University of Bristol, Kathleen Wood, Dr. in Physics for Life Sciences, Joseph Fourier University of Grenoble and Lisa Walter, Dr. in Pharmacology, University of Washington, Seattle.

Front cover illustration composed by Alice Giraud

Jeannine Yon-Kahn · Guy Hervé

Molecular and Cellular Enzymology

Volume II

Springer

Dr. Jeannine Yon-Kahn
Université Paris-Sud XI
UMR CNRS
Inst. Biochimie, Biophysique
91405 Orsay
Batiment 430
France
jeannine.yon-kahn@wanadoo.fr

Prof. Guy Hervé
Université Paris VI
CNRS UMR 7631
Labo. Biochimie des Signaux
96 bd. Raspail
75006 Paris
France
guy.herve@upmc.fr

Originally published in French: Jeannine Yon-Kahn et Guy Hervé, Enzymologie moléculaire et cellulaire, EDP Sciences, Collection Grenoble Sciences 2005, Tome 2, ISBN 978-2-86883-816-2

ISBN 978-3-642-01227-3 e-ISBN 978-3-642-01228-0
DOI 10.1007/978-3-642-01228-0
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2009934310

Cover design: Giraud, Alice

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

PREFACE

Enzymology lies at the boundary of several disciplines. It has benefitted from much progress: in physics, which has lead in particular to the determination of the three dimensional structures of biological macromolecules and to the study of their dynamics; in chemistry, with the development of analytical methods and processes of synthesis; and in biology, with the potential of genetic and cell biological techniques. The coincidence of these methods with those of molecular modelling today enables us to modify an enzyme in order to make it more stable or to alter its functional properties. These possibilities are of great importance not only for the precise knowledge of the mechanisms involved in catalytic events and in their regulation, but also to meet the need for applications in biotechnology. Furthermore, enzymology is of interest to diverse fields in biology and even in medicine where it contributes, amongst other things, to the development of diagnostics for, and soon the treatment of, genetic and metabolic diseases. Important developments are expected due to the entry of biology into the post-genomic era; it is now vital to process the information contained within genomes i.e. the structure and function of proteins coded for by genes. This is why we felt it necessary to bring together in a single work the ensemble of current knowledge relating to enzymology.

Molecular and Cellular Enzymology addresses not only experienced enzymologists but also all those, e.g. biologists, medical doctors, industrialists, who are confronted by enzymological problems during their fundamental or applied research. Equally, it addresses students who will find herein the basics, as well as the most recent developments in the subject. We hope that this work will be able to stimulate in these readers an interest in this field in which an urgent need for training exists in order to meet both the requirements of research and industrial endeavours. This book consists of several levels: practical aspects and elementary explanations are given for the benefit of non-specialists' understanding; specialists will also find topics more deeply expounded with the principal bibliographic references cited. The bibliography, however, is not exhaustive; the choice includes general books and review articles as well as some specialised articles. In order to facilitate the task of students, two typographies have been adopted. The main text corresponds to basic knowledge, whereas text in a smaller font, indented and indicated by ◤◢, provides more specialised information.

Part I describes the thermodynamics of enzymatic reactions. At the very beginning, in Chap. 1, there is a succinct reminder of the laws of thermodynamics at equilibrium, the experimental methods for determining the energetic parameters of reactions, as well as a study of coupled reactions. Chapter 2 treats protein-ligand association equilibria and their experimental study. It concerns as much enzymologists as biologists confronted with the problem of binding molecules to receptors. The third chapter is devoted to the study of living systems, as open systems far from equilibrium.

Part II concerns kinetic studies of enzymatic reactions in solution. Chapter 4 is a reminder of those elements of chemical kinetics indispensable for the understanding of enzyme kinetics. In Chap. 5, all aspects of the kinetics of enzyme reactions possessing Michaelian behaviour are worked through. The experimental methods for studying these reactions, along with procedures for data processing and interpretation, are presented in Chap. 6.

Part III covers the formation and structure of enzyme active sites. A discussion about the origin of enzymatic function is included (Chap. 7), as well as a presentation of the formation of the functional structure of enzymes (Chap. 8). The diverse approaches used to determine the topology of enzyme active sites are the subject of Chap. 9.

Part IV analyses catalytic function, detailing the mechanisms brought into play in the formation of enzyme-substrate complexes (Chap. 10), the principal catalytic mechanisms and their involvement in enzyme catalysis (Chap. 11). To illustrate these as an ensemble, some specific enzymatic systems are studied in Chap. 12, which represent diverse types of catalytic mechanism. They were chosen based on structural knowledge of the enzyme molecule and on the extent to which the relationship between the structure and functional properties has been determined.

Part V explores the regulation of enzymatic activity. It contains an analysis of the non-covalent regulation of allosteric enzymes, the kinetics of which are treated in Chap. 13 and their covalent regulation forms the focus in Chap. 14. Multi-functional enzymes and multi-enzymatic systems are discussed in Chap. 15.

Part VI is devoted to enzymology in a structured environment, whether it be artificially immobilised enzymes or enzymes linked to well-defined cellular structures. The theory of metabolic pathway control is also included.

This work thus brings together diverse aspects of enzymology, which only exist in separate books or articles. For practical reasons we have had no hesitation in recalling certain basic concepts in the different parts, so that the reader may –thanks to a detailed Index– consult the topic of interest without being obliged to study the preceding chapters. In a similar vain, the indexing will enable precise points of interest to be readily looked up.

ACKNOWLEDGEMENTS

The authors would like to thank deeply Drs Florence LEDERER, Charis GHELIS and Michel DESMADRIL for their careful reading of certain parts of this work and for their constructive suggestions, and in particular Dr Nicolas KELLERSHOHN who took great care in checking all the equations. These thanks also go to Profs Alain COZZONE, Bernard BADET, Christian CAMBILLAU and Jacques HAIECH for the documentation that they provided. The authors thank equally the members of the Reading Committee, Profs G. LAUQUIN, J.P. MAZAT, B. HOUCHMANDZADEH, M. VAN DER REST, J. LUNARDI, J. PELMONT and J. RICARD for having carefully read the manuscript and for communicating their constructive remarks and suggestions.

The authors wish also to thank Philip SIMISTER, Lisa WALTER and Katy WOOD for their contributions to the translation of the French version of this book and to Julie RIDARD and Sylvie BORDAGE from Grenoble Sciences for their skilful preparation and shaping of the two versions.

Jeannine Yon-Kahn

September 2009 Guy Hervé

TABLE OF CONTENTS

VOLUME I

VOLUME II

PART IV – THE CATALYTIC FUNCTION

PART V – REGULATION OF ENZYME ACTIVITY

PART VI – ENZYMOLOGY IN STRUCTURED ENVIRONMENT

PART IV

THE CATALYTIC FUNCTION

INTRODUCTION

The life of a cell, of an organism, depends on the multiplicity of the diverse chemical reactions which constitute metabolism. These reactions are carried out with appreciable rates in conditions where they could usually not occur, conditions of restricted temperature and near neutral pH. They occur in these conditions because they are catalysed by protein catalysts, the enzymes. These are either globular proteins soluble in aqueous media or membrane proteins. Their activity depends very tightly on their three-dimensional structure and their internal dynamics; every modification of this structure, even very slight, can bring about variations important to the catalytic activity, indeed even the loss of it. Knowledge of the three-dimensional structure of an increasing number of enzymes constitutes precious data specifying the mechanism of their catalytic action. In addition, metabolic reactions implicate extremely subtle rules. Cells, living beings are autoregulated.

Enzymes are characterised both by a **large catalytic efficiency** and by their **specificity**. Taking into account the diversity of reactions that take place in living organisms, it is not surprising to find a great number of enzymes in cells. Enzymes are catalysts; like all catalysts, they act in very weak quantities and are regenerated at the end of the reaction. The efficiency of the action of an enzyme is often defined by the **turn over number** or **molecular activity** (see Part II). It is the number of substrate molecules that an enzyme molecule can transform per second. According to different enzymes, this number can vary between 10^2 and 10^8 s^{-1}. Table I below gives an idea of the turn over number of several enzymes.

It appears that enzymes of electron transfer including catalase and raifort peroxydase are among the most efficient. Enzymes of group transfer have weaker molecular activities. Nevertheless, this table shows the very great catalytic efficiency of enzymes which, of all the catalysts, are the most efficient. Enzymatic reactions proceed at rates much greater than reactions catalysed by chemical catalysts. If one compares for the same chemical reaction, when possible, the rate of the reaction catalysed by an enzyme and by a chemical catalyst, one can find efficiency factors 10^8 to 10^{11} times greater for the enzymatic catalysis than for the chemical catalysis.

Table 1 Molecular activity of selected enzymes

Enzyme	*Molecular activity (s^{-1})*
Catalase	10^7
Cytochrome oxydase	1.3×10^6
Carbonic anhydrase	6×10^5
3-ketosteroid isomerase	2.8×10^5
Raifort peroxidase	1.5×10^5
Acetaldehyde reductase	2.9×10^4
Acetylcholinesterase	2.5×10^4
Oxaloacetate amino transferase	8.6×10^3
Lactate dehydrogenase	10^3
Chymotrypsin	10^2
Tryptophan synthase	2
Lysozyme	0.5

One of the most remarkable characteristics of enzymatic catalysis compared to chemical catalysis is its **specificity**. The specificity of an enzyme is of two orders: the specificity for a reaction and the specificity with regard to the substrate. Each enzyme can only catalyse a single reaction or one type of reaction; it has a well defined function. For example, a protease catalyses the hydrolysis of peptide bonds, a dehydrogenase a dehydrogenation reaction, a glycosidase the breaking of glycosidic bonds etc. Each enzymatic protein generally supports a single function. There exist however multifunctional enzymes, but these are generally constituted of the assembly of several subunits, each carrying a defined function, or are proteins comprised of several structural domains, each function relevant to a distinct domain (see Part V).

In addition, each enzyme "recognises" a well determined substrate or type of substrate. The required conditions for a compound to be a substrate of an enzyme can be more or less strict. Diverse enzymes have more or less large specificities. Several examples will be given. All pancreatic proteases catalyse the hydrolysis of peptide bonds, but trypsin attacks the bonds for which the carboxyl belonging to an L-lysine or an L-arginine, whereas chymotrypsin requires the presence of aromatic amino acids. Figure 1 opposite illustrates the difference in specificity of these two proteases. In addition, these are only active with respect to amino acids in the L configuration. Optic isomers of the D configuration are not substrates of these enzymes; however, they are recognised and can be bound to the active site with affinities comparable to those of their L isomers, but the peptide bonds are not (or are very slowly) hydrolysed. Enzymes are therefore stereospecific. Outside of this condition, the specificity of these enzymes is not absolutely strict; each of these proteases is capable of hydrolysing the substrates of the other, but with a much lower efficiency.

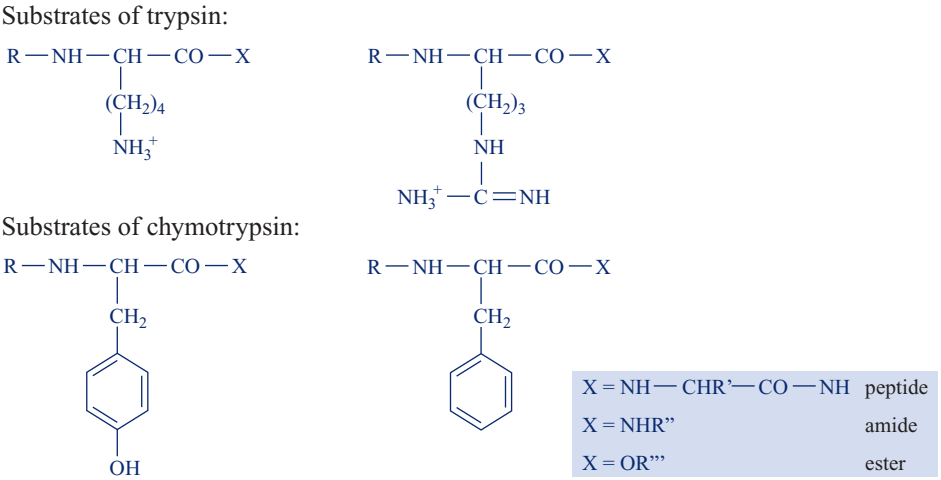


Fig. 1 Specificity of trypsin and chymotrypsin

Galactosidases are enzymes presenting a much stricter specificity. β-galactosidase hydrolyses β-galactosides (β-D-galactopyranosides), but not α-galactosides. Conversely, α-galactosidase hydrolyses α-galactosides but remains without action on β-galactosides. A very precise configuration is required so that the respective compounds are substrates of these enzymes. The inversion or the modification of one of the hydroxyls of the ring of β-galactosides suffices so that the compound is no longer a substrate of β-galactosidase; thus 2-methyl-β-D-galactosides, or β-D-glucosides resulting in the inversion of the position of the hydroxyl in 4, are no longer substrates of β-galactosidase. β-D-fucosides are poor substrates due to the replacement of the hydroxyl in 6 by a single proton. The replacement of oxygen in the glycosidic bond by a sulfur to form a β-D-thiogalactoside gives a compound that is not hydrolysed by β-galactosidase; it behaves as a competitive inhibitor (Fig. 2 below).

One of the interests of enzyme catalysis resides in the ability to distinguish between forms of homologous compounds when it is often impossible to separate them by chemical means. In particular, one must underline the importance of the specificity of enzymes in relation to enantiomorphic compounds. The enzyme generally only attacks one of the two optic isomers, whereas asymmetric syntheses are very difficult to carry out by chemical methods. PASTEUR succeeded in separating L and D tartric acids from a racemic mixture by using a mold which transformed only L-tartrate because it contained an enzyme, L-tartrate dehydratase. One knows today that there exists a L- and a D-tartrate dehydratase which hydrolyses respectively each of the two optic isomers with the exclusion of the other (Fig. 3 below).

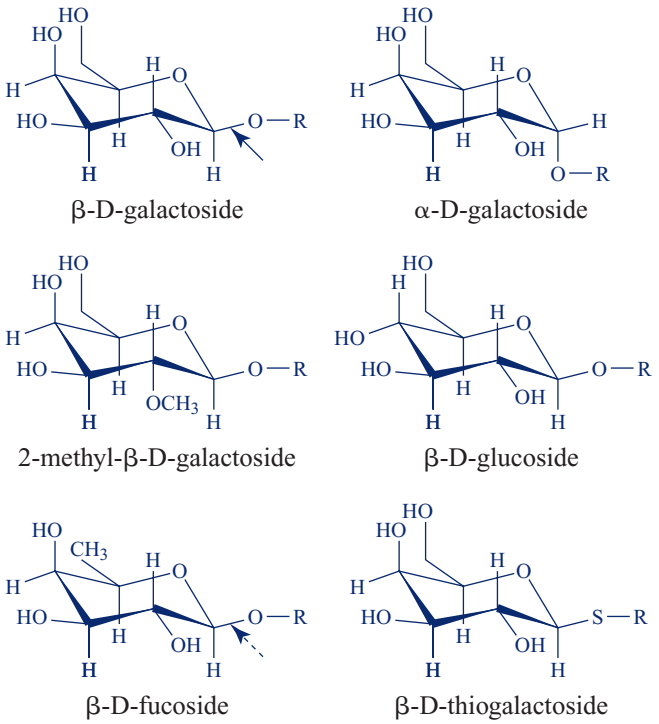


Fig. 2 Specificity of galactosidases

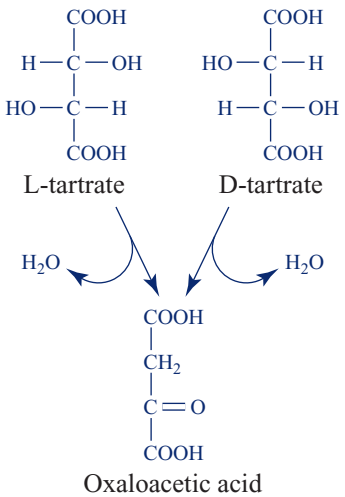


Fig. 3 Separation of L and D-tartrate by the corresponding enzymes

It is important to understand by which mechanisms enzymes, protein macromolecules, are the most efficient catalysts that exist and at the same time the most specific. The diverse factors that can account for this are analysed in this part in light of the most recent data.

10 – FORMATION OF ENZYME-SUBSTRATE COMPLEXES

To understand the mechanisms of enzyme catalysis, that is the mechanisms by which the enzyme lowers the energy barrier of the reaction that it catalyses, it is not sufficient to study only the catalytic act, but it is important to analyse the diverse events implied over the course of an enzymatic reaction which condition this catalytic act. The formation of intermediate complexes including non-covalent associations takes part in the catalytic function as does their decomposition. The formation of one or more intermediate complexes between the enzyme and the substrate is a necessary but not sufficient condition for the reaction to take place. It is therefore important to present the mechanisms of enzyme-substrate association; it is only for reasons of clarity that this first step of enzymatic reactions constitutes a chapter in itself.

The first authors emphasised the importance of the structural complementarity between the enzyme and its substrate to ensure their association, the enzyme constituting a sort of matrix (template) in which the substrate would come to insert itself. This rigid notion of the "lock and key" mechanism reflected well the static conception which was assumed during nearly a half century following studies of E. FISCHER on the stereospecificity of enzymatic reactions.

Flexibility and mobility of certain regions in proteins, in particular those that are localised at the active centre, currently bring about a much more dynamic vision of enzyme-substrate interactions. This is currently the notion of structural complementarity between the enzyme and the transition state of the substrate that is particularly underlined, because it is responsible for both specificity and catalytic efficiency. The affinity of enzymes is indeed much larger for the transition state than for the fundamental state of the substrate; this represents an important thermodynamic factor in enzyme catalysis.

This chapter includes first a presentation of the nature of forces implicated in enzyme-substrate complex formation, then an energy aspect, and finally a discussion on the mechanisms of formation of these complexes.

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_11, © Springer-Verlag Berlin Heidelberg 2010

10.1. NATURE OF FORCES INVOLVED IN ENZYME-SUBSTRATE ASSOCIATIONS

The forces responsible for the association between an enzyme and its substrate are the same as those present in associations between simple molecules. However, the complexity of the enzyme protein structure is such that multiple groups can be implicated and different types of forces can intervene simultaneously; in addition, the presence of a solvent such as water brings about by its nature an important contribution. The different forces susceptible to be involved in enzyme-substrate associations consist of:

- **Electrostatic interaction forces** that are produced between atoms or groups of atoms that possess a permanent electric charge; this concerns interactions between two ions, between an ion and a dipole or even between two permanent dipoles.
- **Induction forces** that result in the induction of a dipole in a non-polar group submitted to an electrostatic field with a permanent charge. One can distinguish the interactions between an ion and an induced dipole, or even between a permanent dipole and an induced dipole. Such interactions are sometimes classified under VAN DER WAALS interactions. Some authors include them in electrostatic interactions; at the limit the diverse forces considered are of an electrostatic nature.
- **Electrokinetic interactions or LONDON dispersion forces**. These are interactions between non-polar groups due to the mutual induction of fluctuating dipoles.
- **Short-range repulsions** which correspond to electrostatic repulsions between electrons following the recovery of electron clouds of two atoms.
- **Hydrogen bonds** that form when two electronegative atoms are joined by the intermediate of a hydrogen atom; they are also of an electrostatic nature.
- **The covalent bond** resulting from a shared pair of electrons between two atoms. The covalent bond is sometimes implicated in enzyme-substrate complexes; it appears then in a step consecutive to the formation of the MICHAELIS complex which by definition is non-covalent.

10.1.1. ELECTROSTATIC INTERACTION FORCES

10.1.1.1. INTERACTIONS BETWEEN TWO IONS OR COULOMB INTERACTIONS

A point charge in a homogeneous media possesses an electric field whose intensity diminishes as a function of the distance.

The force of the field is given by the equation:

$$F = \frac{q}{r^2 D} = \frac{Ze}{r^2 D}$$

In this expression q is the charge, r is the distance from the charge to the point where the field is measured, Z is the valence of the ion and e the electron charge. The dielectric constant D depends on the medium in which the charge is found and in particular on the polar nature of molecules that constitute the media. In vacuum, the dielectric constant is equal to one; it is 74.1 in water at 37.5ºC.

Ions in finite dimensions can be considered like point charges with the total charge being concentrated in their centre. If the charge is expressed in electrostatic units (esu) and the distance in cm, the force of the electric field is in dynes/esu. At a distance of 5 Å from an electron in vacuum, the field is equal to -4.8×10^{-10} esu/$(5 \times 10^8$ cm$)^2$, or -1.95×10^{-5} dynes/esu, the field being the same sign as the charge. Knowing that 1 dyne/esu equals 300 V . cm^{-1}, the field is -5.76×10^7 V . cm^{-1}. This example shows the intensity of the electric field neighboring an ion.

The force exerted on an ion by another ion situated at a distance r is equal to the force of the field exerted by one multiplied by the charge of the other:

$$F = q_2 F_1 = \frac{q_1 q_2}{r^2 D} = \frac{Z_1 Z_2}{r^2 D} e^2$$

The most important quantity for problems of association is the interaction energy between two separated charges. This energy can be evaluated by integration of the force between the charges from an infinite distance to the distance r:

$$E = \int_{\infty}^{r} F dr = \frac{q_1 q_2}{rD} = \frac{Z_1 Z_2}{rD} e^2$$

The interaction energy between the charges varies with the inverse of their distance. If the charges have the same sign, the energy is positive, and there is a repulsion. If they are of opposite signs, there is an attraction. As an example, to show the role of the dielectric constant, the interaction energy between two monovalent ions, an anion and a cation situated 5 Å apart in vacuum, is –66.3 kcal . mol^{-1}, whereas in water it is –0.85 kcal . mol^{-1}, that is very slightly higher than thermal motion.

10.1.1.2. INTERACTIONS BETWEEN AN ION AND A DIPOLE

An electric dipole results in the presence of two equal charges of opposite sign separated by a distance l. The size of a dipole is given by its moment:

$$\mu = q_r l$$

The electrostatic interaction between a dipole and an ion is the sum of the interactions of the ion with each dipole charge. The potential energy is given by the equation:

$$E = \frac{qq_r}{D}\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

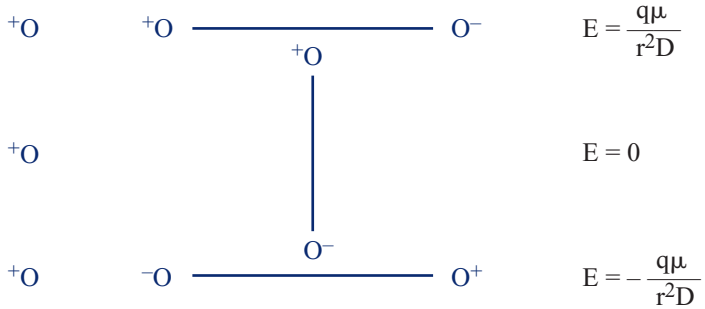
When the length l of the dipole is less than the distance r between the ion and the dipole centre, which is generally the case, the general equation can be simplified and becomes:

$$E = \frac{q\mu\cos\theta}{r^2 D}\left(1 - \frac{3l^2}{8r^2}\right)$$

The term in parentheses being around 1, the potential energy is:

$$E = \frac{q\mu\cos\theta}{r^2 D}$$

The potential energy depends on the orientation. The scheme below shows the interaction energy between an ion and a dipole with moment μ in the different positions, aligned and perpendicular:



If the dipole is free to rotate rapidly and uniformly, the average of the potential energy over time is zero. However, following the ion electric field, the dipole orients statistically towards the position of maximum attraction, meaning near the position of minimum potential energy. The degree of orientation depends on the relative sizes of the interaction energy and the thermal agitation, kT. A statistical dipole is produced with an average moment:

$$\bar{\mu} = \frac{\mu^2 q}{3kTr^2D}$$

The potential energy that results is:

$$E = -\frac{\mu^2 q^2}{3kTr^2D}$$

This relationship is particularly valid when the potential energy of electrostatic interaction is weaker than thermal motion. If it is much larger, and under the condition that the dipole is free to orient, the preceding expression reduces to:

$$E = -\frac{\mu q}{r^2 D}$$

This can be illustrated by some numerical examples. Let us consider a dipole in which the charges situated at a distance of 2 Å correspond to half of the electron charge. The dipole moment is:

$$\mu = 2.4\times10^{-10}\times2\times10^{-8} = 4.8\times10^{-18}\ \text{esu} = 4.8\ \text{debyes}$$
$$(10^{-18}\ \text{esu} = 1\ \text{debye})$$

The potential energy is expressed in ergs/molecule if q and μ are expressed in esu and r in cm. If a dipole of 1 debye is situated at a distance of 10 Å from a monovalent ion at 37.5ºC, the potential energy is:

$$E = -\frac{(4.8\times10^{-10})^2(10^{-18})^2}{3(1.38\times10^{-16})(310.5)(10^{-7})^4} = -1.87\times10^{-14}\ \text{erg / molecule}$$

or –0.269 kcal . mol^{-1} at 37.5ºC; at this temperature the thermal motion energy is 0.617 kcal . mol^{-1}, larger than the interaction energy. If the dipole is not free to rotate, but is found fixed in the position of optimal attraction, the interaction energy is then 0.691 kcal . mol^{-1}. The interaction energy between an ion and a dipole when the dipole is free to rotate only represents therefore 40% of the interaction energy when the dipole is fixed in its position of optimal attraction.

10.1.1.3. INTERACTIONS BETWEEN PERMANENT DIPOLES

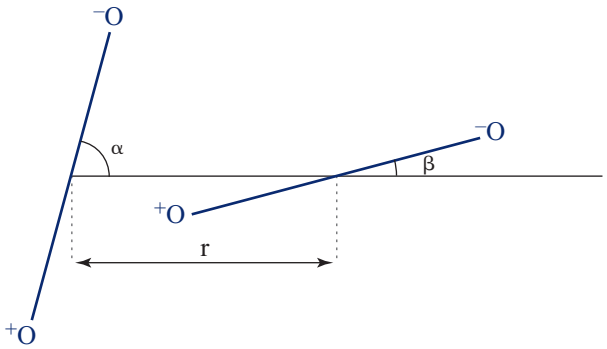
The potential energy of two dipoles depends not only on their distance, but also on their relative orientation. It is convenient therefore to consider first the case of two fixed dipoles, and then the case of two dipoles with rotational freedom.

Let us consider first the case of two fixed dipoles placed a distance r apart from each other, this distance being greater than the length l of the dipole. The potential energy is given by the MOELWYN-HUGHES equation:

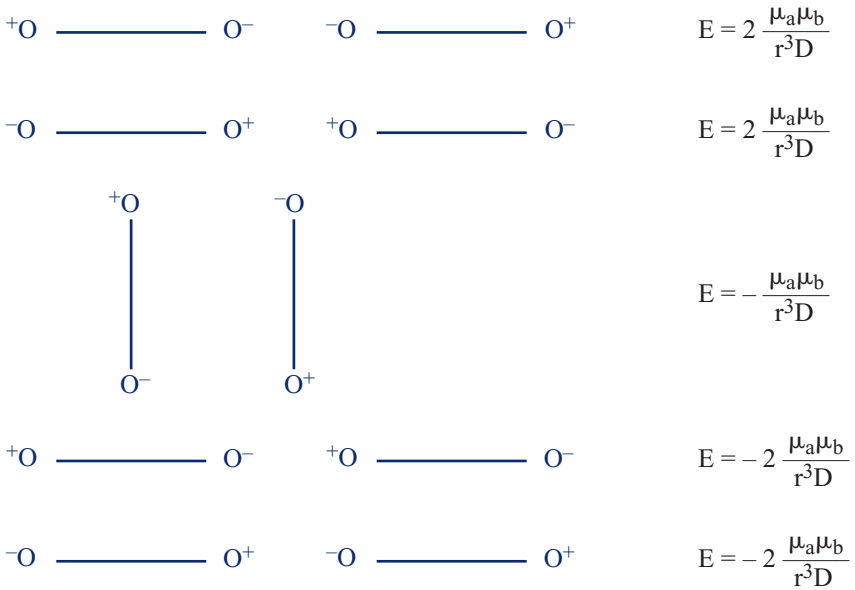
$$E = -\frac{\mu_a \mu_b}{r^3 D}(2\cos\alpha\cos\beta - \sin\alpha\sin\beta)$$

The scheme below indicates the different possible orientations:

- General case



- Aligned or parallel dipoles



In summary the potential energy of electrostatic interaction between two fixed dipoles not having any freedom of rotation is given by the general expression:

$$E = A\frac{\mu_a\mu_b}{r^3D}$$

The minimum potential energies are obtained when the dipoles are aligned with their opposite sign charges opposing.

When two dipoles are free to rotate, two situations can arise. In the first, dipole A is fixed and dipole B is free to rotate; this can correspond with the case of a dipole on the surface of an enzyme and a small molecule in solution. The potential energy is:

$$E = -\frac{\mu_a^2 \mu_b^2}{3kTr^6D}(1+\cos^2\theta)$$

θ being the angle that aligns the centres with the axis of the dipole. In the second situation, the two dipoles are free to rotate and their interaction is due to the polarisation of mutual orientation that they exert. The potential energy is of the form:

$$E = -\frac{2\mu_a^2 \mu_b^2}{3kTr^6D^2}$$

This expression is only valid when the potential interaction energy is less than the thermal motion; otherwise the interaction would be sufficiently strong to orient the dipoles in the position of minimal energy, meaning the maximal attraction. Therefore, when the potential energy is weak or on the order of thermal motion, it is expressed by the general expression:

$$E = -A\frac{\mu_a^2 \mu_b^2}{kTr^6D^2}\mu_b^2$$

These interactions are KEESOM interactions; their energy varies as $1/r^6$. They become therefore very weak as the distance increases.

10.1.2. INDUCTION INTERACTIONS

10.1.2.1. INTERACTION BETWEEN AN ION AND AN INDUCED DIPOLE

All the molecules subject to an electric field are polarised and there results a dipole whose moment depends on the force of the field and on the properties of the molecule. In the most general case, the total polarisation is given by the relation:

$$P = P_e + P_a + P_m$$

P_e is the electron polarisation due to the displacement of electrons in the field, P_a is the atom polarisation due to the displacement of atoms or groups of atoms in the molecule, and P_m is the orientation polarisation resulting from the preferential orientation in the electric field created by the ion of a molecule possessing a dipole. When the molecule does not possess a permanent dipole, the polarisation induced by the ion field reduces to $P_e + P_a$ and is called the induced polarisation P_i. The induced dipole moment is proportional to the force of the field:

$$\mu_i = \alpha_0 F$$

and it is oriented in the field, aligned with the ion. The proportionality constant α_0 is called the **polarisability** of the molecule.

The potential energy of a dipole in an electric field is given by the expression:

$$E = -\mu F \cos\theta$$

and like in the present case $\theta = 0$, so that the potential energy is:

$$E = -\alpha_0 F^2$$

However, the energy is necessary for inducing the dipole ($\alpha_0 F^2/2$), from where the expression of interaction energy between an ion and an induced dipole comes:

$$E = \frac{\alpha_0 F^2}{2} - \alpha_0 F^2 = -\frac{\alpha_0 F^2}{2}$$

The field of an ion is given by the equation:

$$F = \frac{Ze}{r^2 D}$$

and the resulting value of interaction energy between an ion and an induced dipole becomes:

$$E = -\frac{\alpha^0 Z^2 e^2}{r^4 D^2}$$

The values of polarisability α_0 are determined by measures of optical refraction, since the latter depends on electron and atom distortions induced by the alternating electric field from the radiation. The refraction varies with the wavelength of light used.

By determinations at two or more wavelengths, it is possible to calculate the refraction at zero frequency where one obtains the effect of a constant field. The molar refraction at zero frequency, R_0, relies on the polarisability by the equation:

$$R_0 = \frac{4\pi N \alpha_0}{3}$$

N being AVOGADRO's number. The values of R_0 are known for most of the atoms and the molar refraction of a substance is approximately equal to the sum of the atomic refractions of its constitutive atoms as shown by Landolt in 1862.

10.1.2.2. INTERACTIONS BETWEEN A DIPOLE AND AN INDUCED DIPOLE OR DEBYE INTERACTIONS

A permanent dipole can also polarise a neighboring molecule. The potential energy of a polarisable molecule in an electric field, as well as what was indicated above is driven by the equation:

$$E = -\frac{\alpha_0 F^2}{2}$$

The field force produced by a permanent dipole is:

$$F = \frac{\mu}{r^3 D}(1+3\cos\theta)$$

The potential energy is therefore:

$$E = -\frac{\alpha_0 \mu^2}{2r^6 D^2}(1+3\cos^2\theta)$$

When the polarised molecule is in the dipole axis, $\theta = 0$ and the potential energy becomes:

$$E = -\frac{2\alpha_0 \mu^2}{r^6 D^2}$$

If the polarised molecule is perpendicular to the dipole axis:

$$E = -\frac{\alpha_0 \mu^2}{2r^6 D^2}$$

from where the general DEBYE expression:

$$E = -A\frac{\alpha_0 \mu^2}{r^6 D^2}$$

A being a constant dependent on the orientation of the dipoles.

10.1.3. ELECTROKINETIC INTERACTIONS OR LONDON DISPERSION FORCES

Non-polar molecules or groups are susceptible to attracting each other although they do not apparently possess any localised charge. The nature of these forces remained unknown until LONDON in 1930 showed their relationship with optical dispersion and gave the expressions of these interactions derived from quantum mechanics.

A symmetrical molecule possesses at each instant a dipole moment following fluctuations in relative positions of the nucleus and electrons in the outer layer. This instantaneous dipole creates a field that polarises an adjacent molecule. The second molecule then possesses a fluctuating dipole that, in turn, polarises the first molecule. Thus a coupling establishes between the oscillations of the electrons of the two molecules, so that the statistical distribution of electrons is continually in favor of the attraction. There is not a simple dipole in a molecule, but several oscillating dipoles; since these dipoles have different moments, the interaction energy is obtained by integration of the ensemble for each molecule.

The potential energy of two neutral molecules is given by the LONDON equation:

$$E = -\frac{3\alpha_1\alpha_2(h\nu_i)_1(h\nu_i)_2}{2r^6[(h\nu_i)_1+(h\nu_i)_2]}$$

α being the polarisability and h PLANCK's constant (6.62×10^{-27} ergs . s^{-1}); ν_i is the frequency of electric oscillation responsible for the polarisability. The term $h\nu_i$ represents the ionisation potential I of the molecule. Thus, the potential energy is given by the expression:

$$E = -\frac{3\alpha_1\alpha_2 I_1 I_2}{2r^6(I_1+I_2)}$$

The existence of such forces implicates a tight contact between the two molecules because the energy varies as $1/r^6$.

10.1.4. SHORT-RANGE REPULSIVE INTERACTIONS

Besides attraction forces there are repulsive forces that maintain a separation between atoms and molecules. These non-specific repulsive forces become visible when atoms or molecules are found in very close contact. They result in repulsions between electron clouds of two molecules when they sufficiently approach each other. The repulsion energy is of the general form:

$$E = \frac{B}{r^b}$$

The total energy of a pair of interacting molecules is of the form:

$$E_{total} = E_{attraction} + E_{repulsion}$$

that is:

$$E = \frac{-A}{r^a}+\frac{B}{r^b}$$

a and b indicate the dependence as a function of the distance. Besides electrostatic interactions most of the previously described interactions vary as $1/r^6$, the coefficient a is therefore generally taken equal to 6. The value of the exponent b is large and difficult to evaluate with precision. There exist several approximations; the most utilised consists of attributing the value 12 to this exponent; this is **the LENNARD-JONES potential or 6-12 potential,** in which the potential energy is expressed by the equation:

$$E_{total} = \frac{-A}{r^6}+\frac{B}{r^{12}}$$

the coefficients A and B were evaluated for all atom pairs. The energy curve as a function of distance is given in Fig. 10.1 opposite. A recapitulation of different interactions is presented in Table 10.1 opposite.

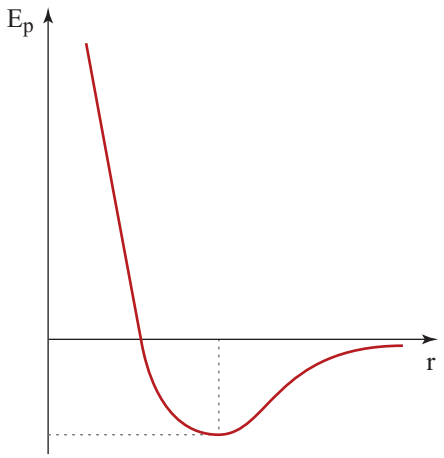


Fig. 10.1 Variation in potential energy as a function of the distance between reacting atoms

Table 10.1 Nature of interactions between molecules and corresponding energy values

1 – Electrostatic interactions	
a – between two ions	$E = \frac{Z_1 Z_2 e^2}{Dr}$
b – between an ion and a dipole	
– fixed position	$E = -\frac{Aq\mu}{r^2 D}$
– free to rotate	$E = -\frac{\mu^2 q^2}{3kTr^4 D^2}$
	$\max. E = -\frac{q\mu}{r^2 D}$
c – between two dipoles	
– fixed position	$E = -\frac{A\mu_a{}^2\mu_b{}^2}{r^3 D}$
– free to rotate	$E = -A\frac{\mu_a{}^2\mu_b{}^2}{kTr^6 D^2}$
	$\max. E = -2\frac{\mu_a\mu_b}{r^3 D}$
2 – Induction forces	
a – between an ion and an induced dipole	$E = -\frac{\alpha_0 Z^2 e^2}{2r^4 D^2}$
b – between a dipole and an induced dipole	$E = -A\frac{\alpha_0 \mu^2}{r^6 D^2}$
3 – Electrokinetic interactions	$E = -\frac{3\alpha_1\alpha_2 I_1 I_2}{2r^6(I_1 + I_2)}$
4 –Short-range interactions implicating attraction and repulsion	$E_{total} = \frac{-A}{r^6} + \frac{B}{r^{12}}$

10.1.5. THE HYDROGEN BOND

The hydrogen bond is established between two electronegative atoms bound by a hydrogen atom, which forms a covalent bond with one of the two electronegative atoms. It is written D—H···A, D being the donor and A the acceptor. In hydrogen bonds, the contact between the atoms is closer than in VAN DER WAALS interactions between the same atoms; this is due to strong electrostatic interactions between donor and acceptor. Such bonds are formed more easily with hydrogen than with other atoms due to the weak size of the hydrogen atom that permits it to have a close contact with other electronegative atoms. The optimal energy configuration is obtained when the atoms are collinear; the variation in linearity brings about a decrease in this energy. The hydrogen bond force decreases when the distance H···A increases; the energy profile as a function of the distance is analogous to that of VAN DER WAALS interactions represented in Fig. 10.1; it depends moreover on the angle formed by the atoms (Fig. 10.2).

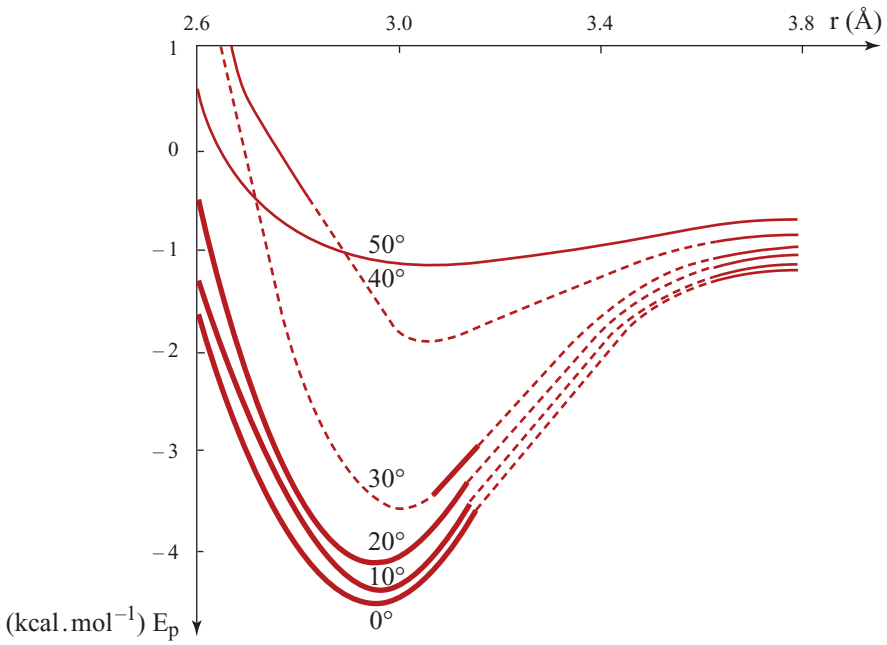


Fig. 10.2 Variation in energy of the interamide hydrogen bond as a function of the N···O distance for different values of the hydrogen bond angle

(From RAMACHANDRAN G.N., p. 3, Fig 2.

Different theories were proposed to explain the nature of the hydrogen bond. The most common assumes an electrostatic interaction between the two bound dipoles. The electronegative atoms D and A possess a negative fractional charge when they are bound to other atoms. Therefore the D—H bond constitutes a dipole and can be written D^--H^+, the charges being electron and proton charge fractions. Similarly the

A—B bond can be assimilated to a dipole A^--B^+. The interaction of these two bonds in their collinear position can be represented as follows:

$$D^-—H^+\cdots A^-—B^+$$

The dotted lines represent a dipole interaction. If this representation corresponds to a real mechanism of the hydrogen bond, the bond energy can be evaluated approximately by a dipole interaction calculation. In the case of the hydrogen bond of water O—H···O, the length of the bond is 2.76 Å and the distance between the dipole centres is 2.40 Å. The potential energy found for the dipole interaction is -6.61 kcal . mol^{-1}; it must be corrected for the repulsion that represents about 35% of the energy, resulting in a final energy of -4.5 kcal . mol^{-1}. The value determined experimentally is -4.5 kcal . mol^{-1}, therefore in agreement with the calculation. The hydrogen bond energy depends on the nature of the acceptor and the donor. It is generally between -2 and -10 kcal . mol^{-1} and is considerably weaker in water and polar solvents.

The hydrogen bond plays an important role in biological systems. The residues and side chains as well as the NH and C═O groups of the peptide backbone of an enzyme are often implicated in hydrogen bonds with the substrate. Hydrogen bonds sometimes stabilise the negative charges of the substrate developed in the transition state. For instance in serine proteases, the carbonyl from the oxygen of substrates interacts with two NH groups of the peptide backbone. Oxygen becomes negatively charged in the transition state and is found stabilised by dipole moments of two amide groups. All these aspects will be detailed in Chap. 12.

10.1.6. HYDROPHOBIC INTERACTIONS

Non-polar interactions are established between non-polar groups in water; they represent the tendency of non-polar compounds to pass from an aqueous phase to an organic phase. They result in the reorganisation of the structure of water in the presence of a non-polar compound. The molecules of water around a non-polarised compound are organised in order to contract the maximum number of hydrogen bonds between them. The entropy increase due to the disorganisation of the water molecule network structure constitutes the major contribution to the energy of hydrophobic interactions (NÉMÉTHY & SCHÉRAGA, 1962).

The hydrophobic character of a molecule is measured by the sharing between organic and aqueous phases, the most often between n-octanol and water. By such determinations, it was shown that several substituents of a molecule contribute in an additive manner to its hydrophobic character. There exists an empirical correlation between the accessible surface of an amino acid side chain and its energy of transfer from water to an organic phase (CHOTHIA, 1974). This corresponds to 20-25 kcal . mol^{-1} per $Å^2$ of surface area. The accessible surface of a protein is defined by the area engendered by the centre of a water molecule rolling along its surface.

10.1.7. THE COVALENT BOND

A simple covalent bond or σ bond exists between two atoms when two electrons of opposite spin are shared between the two atoms. The electron density is greater between the two nuclei and the electron resonance in this region contributes the greatest part (resonance energy) to the bond energy that varies between 50 and 150 $kcal \cdot mol^{-1}$ according to the implied atoms. However, covalent bonds do not have obligatorily a character of electronic symmetry. Indeed, when they join two different atoms, the bond usually possesses an ionic character, the electrons being more or less attracted by one of these atoms, conferring upon it an electronegativity. It is not essential here to enter into detail the chemical bond. It is necessary simply to mention that the enzyme-substrate interactions can use simple bonds or σ bonds, and also π bonds like with SCHIFF bases that form with transaminase and aldolase substrates.

10.1.8. DETERMINATION OF THE NATURE OF ENZYME-SUBSTRATE INTERACTIONS

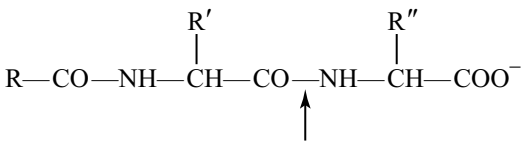
The different types of forces presented above are involved in the enzyme-substrate association. The first complex or MICHAELIS complex only implicates non-covalent interactions. It is however difficult to evaluate the energy contribution of hydrogen bonds and electrostatic and hydrophobic interactions. Indeed, in these associations, some interactions are disrupted, for example the interactions between water and the substrate, and water and the active site of the enzyme; others are established in such a way that the global energy represents a difference. These aspects are discussed in the following paragraph. In some enzymatic reactions, starting from the MICHAELIS complex, a covalent complex between a part of the substrate and a catalytic group of the enzyme is formed. This is the case with acyl-enzymes and phosphoryl-enzymes that were presented in Parts II and III.

The nature of interactions between an enzyme and its substrate can be analysed stepwise. The method consists of evaluating the particular contribution carried by a substrate analog, charged or not, or of a competitive inhibitor, then to compare the dissociation constants, and by consequence the differences in interaction energy in the formation of complexes between the enzyme and the substrates with and without substituents. One can use either direct binding studies when they are possible, or kinetic studies. The best conditions are achieved when it is possible to directly study the equilibrium of enzyme-substrate association by one of the procedures described in Part I and therefore evaluate the dissociation constant K_s. This is only achieved when there exist conditions under which the reaction does not interfere with the measurement, and this without the parameters of association being modified. Substrate analogs that behave like competitive inhibitors are frequently used. The test of values of K_i or K_s as a function of diverse factors, like size, shape, and

charge of the molecule of the substrate or inhibitor, gives information on the dimensions and the topology of the binding site of substrates, as well as on the electrostatic and hydrophobic interactions that are involved in the formation of the MICHAELIS complex.

When the direct study of interactions is not achievable, kinetic methods can be used. In general, as we have underlined in Part II, the MICHAELIS constant obtained experimentally is a complex parameter. In particular cases however, it represents the dissociation constant of the enzyme-substrate complex. In this case, it is necessary to verify it; the knowledge of the parameter $K_m = K_s$ can give information on the enzyme-substrate association. With enzymatic reactions that proceed *via* the formation of an acyl-enzyme, the value of K_s is obtained by a kinetic study with diverse substrates bringing about the same acyl-enzyme intermediate (see Chap. 5). When it is not possible to attain K_s, one can often resort to substrate analogs playing the role of competitive inhibitors, the inhibition constant K_i being a true dissociation constant. The study of variations in this parameter as a function of pH permitted in some cases the determination of the nature of ionisable groups implicated in the formation of the MICHAELIS complex as well as the electrostatic contribution to the interaction energy. Examples are given in Chap. 7.

This kind of approach has been used for a long time in enzymology. Among the first studies, one must cite those that were carried out with enzyme-substrate interactions in the case of carboxypeptidase. The research groups of NEURATH, SCHWERT, SMITH, LUMRY, and POLEGLASE studied the effect of the structure of various substrates and inhibitors on kinetic parameters of the enzymatic reaction. These studies showed first of all that the enzyme specificity brings about an attack on peptide bonds in the C-terminal position under the condition that the last amino acid R″ is a glycine, an alanine a phenylalanine, a tryptophan, a tyrosine, a leucine or an isoleucine and that it possesses a free carboxylate:



The residue R′ must belong to a glycine, an alanine, a glutamate, a methionine or a tryptophan. If R is replaced by H, the substrate is no longer hydrolysed.

◤ Table 10.2 below gathers the values of K_i obtained by these authors for different competitive inhibitors of this enzyme. The inhibitory activity of the first three compounds decreases as the distance between the indole ring and the carboxyl group increases. With another series, the benzoic, phenylacetic, β-phenylpropionic, and γ-phenyl butyric acids, this distance is still very large with a maximum effect for β-phenyl propionic acid. On the contrary, neither the *cis* form nor the *trans* form of cinnamic acid (C_6H_5—CH═CH—COOH) produce inhibition. Taking into account the preceding results, the absence of inhibition cannot be explained by the distance

between the phenyl group and the carboxylate, but rather by the fact that the cinnamic acids have a rigid planar structure such that the enzyme cannot associate at the same time with the phenyl group and the carboxylate; a non-planar configuration of the substrate must be necessary for its association with the enzyme. In the aliphatic series, the inhibitory ability increases with the length of the chain. The action of these different inhibitors is not sensitive to the ionic strength of the media and can only be interpreted by VAN DER WAALS interactions between the substrate and the active centre of the enzyme.

Table 10.2 Values of K_i for carboxypeptidase inhibitors

Inhibitors	*K_i (M)*
Indole propionic acid	0.0055
Indole butyric acid	0.033
Indole acetic acid	0.00078
Benzoic acid	0.14
Phenyl acetic acid	0.0045
β-phenyl propionic acid	0.0012
γ-phenyl butyric acid	0.02
Benzyl malonic acid	0.004
Cyclohexyl propionic acid	0.02
Naphtalene acetic acid	0.045
Propionic acid	0.1
Butyric acid	0.005
Valeric acid	0.0028
Caproic acid	0.00625
Isocaproic acid	0.0028

An analogous study with aspartate transcarbamylase and diverse inhibitors has shown that the two carboxyls of aspartate or its analogs must be in the *cis* configuration for the interaction with the active centre to be optimal; thus, contrary to maleate, fumarate is a good competitive inhibitor.

One particularly interesting example is that of acetylcholinesterase. This enzyme plays an important role in the transmission of the nerve influx since it catalyses the hydrolysis of the neurotransmitter acetylcholine. Well before the enzyme structure was known, the kinetic analysis of the enzymatic reaction had brought about an elaboration on a scheme of enzyme-substrate interactions implicating an anionic site where the quaternary ammonium of the substrate would interact, as well as the esterase site containing the serine that is acetylated in the first stage. The enzyme presents a great catalytic efficiency approaching that of a diffusion-controlled reaction. The three-dimensional structure of acetylcholinesterase of *Torpedo californica* was determined in 1991 with a resolution of 2.8 Å (SUSSMAN et al., 1991). The enzyme is a homodimer bound to the plasma membrane by a covalent bond that implicates the

C-terminal extremity of each protomer. After solubilisation by a phospholipase, the crystallised form stays homodimeric. The active site is localised at the bottom of a narrow deep groove 20 Å long. Like with serine proteases, it involves a catalytic triad with serine 200 and histidine 440, but the aspartate is replaced by glutamate 327. In addition, the orientation of the catalytic triad is the mirror image of that of chymotrypsin. The most striking phenomenon is that the quaternary ammonium is not bound to an "anionic site" possessing a negative charge, but interacts with some of fourteen aromatic residues that line the groove. The charge of the quaternary ammonium is interacting with the π electrons of the aromatic residues.

By a kinetic approach using the nucleophilic competition analysis on the hydrolysis of different substrates by trypsin with a series of primary aliphatic alcohols, $CH_3(CH_2)_{n-1}CH_2OH$, SEYDOUX et al. (1969) determined the hydrophobic interactions between the aliphatic chain of these alcohols and an apolar site of the enzyme; the interaction energy was evaluated as a function of the chain length (Fig. 10.3).

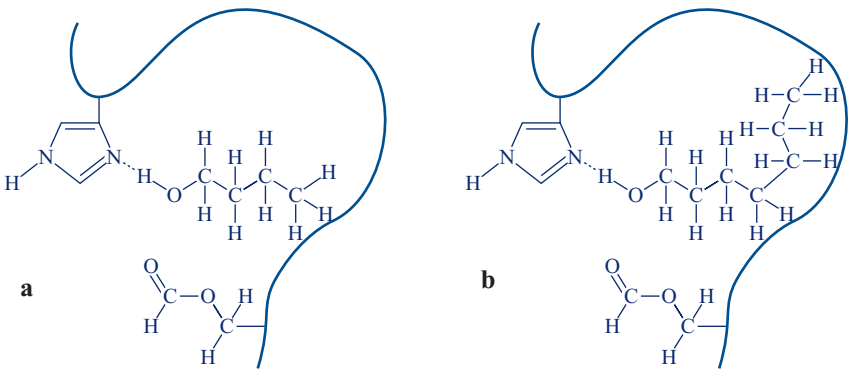


Fig. 10.3 Schematic representation of hydrophobic interactions between an apolar site of trypsin and the aliphatic alcohol chain

(a) *n-butanol; only the methyl group terminal interacts with the enzyme –* ***(b)*** *n-heptanol; the butyl group terminal interacts with the enzyme.* (Reprinted from *Biochim. Biophys. Acta* **171** SEYDOUX *et al.*, Hydrophobic interactions of some alcohols with acyl trypsins, 152. © (1969) with permission from Elsevier)

The crystallographic data, when they are known with sufficient resolution, or the NMR data give precise information on the nature of enzyme-substrate interactions. As an example Fig. 10.4 below shows the binding of NAD^+ to glyceraldehyde-3-phosphate dehydrogenase (from BIESECKER et al., 1977). The hydrogen bonds between the enzyme and the coenzyme as well as the residues of the enzyme in VAN DER WAALS contact with NAD^+ are indicated in this figure. The residues Leu 187 and Pro 188 form such a contact with the ribose part of adenosine. The binding sites of P_i and P_s anions correspond very probably to the binding sites of inorganic phosphate implicated in the stage of phosphorylation, and to the binding site of 3-phosphate of the substrate, respectively. Pi interacts by hydrogen bonds with the amino acids Ser148, Thr150 and 208, and P_s, with residues Thr179, Asn181 and Arg231 that form hydrogen bonds with the substrate.

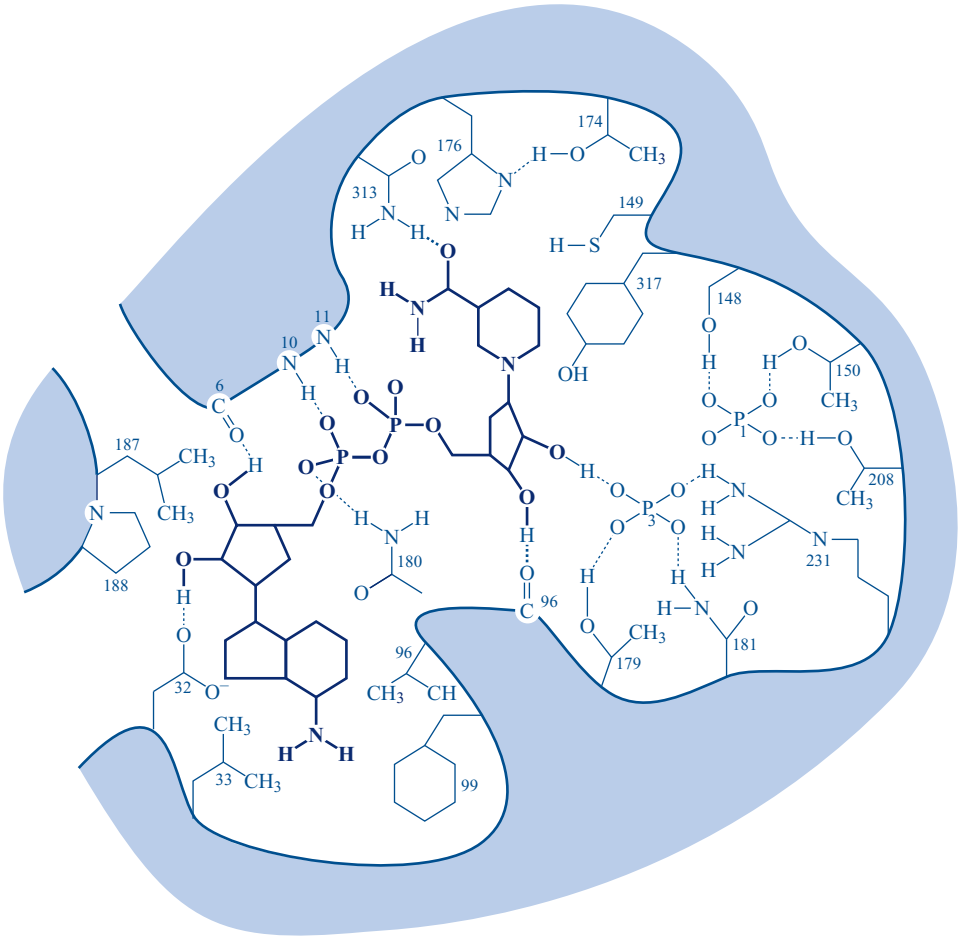


Fig. 10.4 Interactions between NAD$^+$ and glyceraldehyde-3-phosphate dehydrogenase, according to the crystallographic data

(Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **266**, BIESECKER G. et al., 332. © (1977))

The radiocrystallographic study of aspartate transcarbamylase in the laboratory of W. LIPSCOMB permitted also the determination of the amino acid side chains implicated in the interaction with an analog of two substrates, N-phosphonacetyl-L-aspartate (PALA). This complex interaction which involves two neighboring subunits is presented in Fig. 10.5 opposite. The oxygen atoms of the substrate phosphate group contract hydrogen bonds with OH groups of Ser52 and Thr55, the amide groups of the main chain of residues 53, 54 and 55 of the catalytic C1 subunit, and also with Ser80 and Lys84 of the C2 subunit. Likewise, the oxygens of the carboxylate group are interacting with groups of each of the subunits, Lys84 of the C2 subunit, Arg229 and Gln231 of the C1 subunit. The oxygen of the carbonyl interacts with His134 of the C1 subunit. Thus PALA contracts numerous interactions involving two catalytic subunits. Some other examples will be given in the following chapters pertinent to catalytic mechanisms.

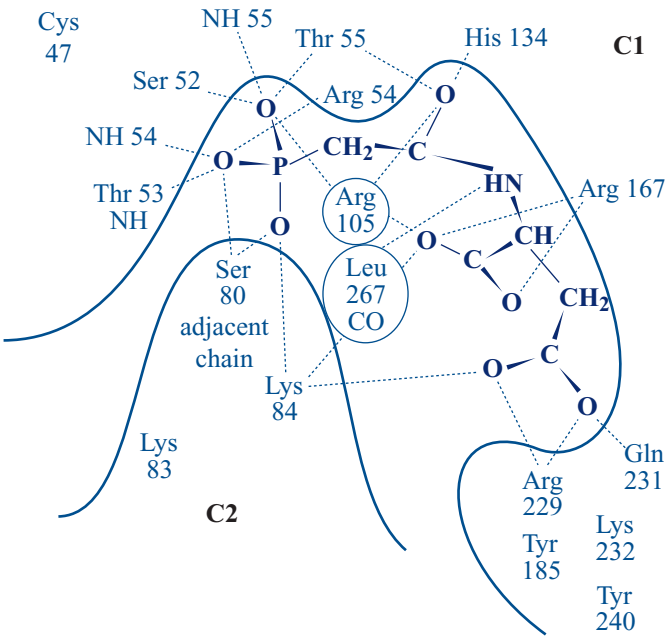


Fig. 10.5 Side chains implicated in the association of N-phosphonacetyl-L-aspartate (PALA) with aspartate transcarbamylase

The dotted lines represent the hydrogen bonds, salt bridges and other polar contacts at a distance of less than 3.5 Å. (Reprinted from *Biochem. Biophys. Res. Commun.*, **136**, VOLTZ K.W. et al., The binding of N-(phosphonacetyl)-L-aspartate to aspartate carbamoyltransferase of *Escherichia coli*, 822. © (1986) with permission from Elsevier)

10.2. ENERGETICS OF ENZYME-SUBSTRATE ASSOCIATIONS

When the dissociation constant of the enzyme-substrate complex K_s is known, it is possible to evaluate the thermodynamic constants that correspond to the formation of this complex. The free energy of complex formation is reliant on the constant K_s by the equation:

$$\Delta G = -RT \ln 1/K_s = RT \ln K_s$$

The values of ΔG obtained for the formation of the MICHAELIS complex are generally weak and always negative; **the formation of the first enzyme-substrate complex is accompanied by a decrease in free energy of the system, it is an exergonic process and therefore spontaneous.**

The study in variations in K_s as a function of temperature permits the evaluation of the variation in enthalpy, ΔH, corresponding to the formation of the complex, according to the equation:

$$\frac{-d \ln K_s}{d(1/T)} = \frac{\Delta H}{R}$$

Knowing ΔG at a given temperature and ΔH, it is possible to calculate ΔS, the variation in entropy corresponding to the formation of the intermediate complex.

The same parameters are obtained from the inhibition constants K_i, permitting thus the determination of the contribution of a substituent or a charge to the energy of enzyme-substrate interaction. Table 10.3 gives the energy parameters of enzyme-substrate or enzyme-inhibitor association for some enzymatic systems.

Table 10.3 Energy parameters corresponding to enzyme-substrate or enzyme-inhibitor associations

Enzyme	*Substrate or inhibitor*	*ΔG*	*ΔH*	*ΔS*
		(kcal . mol^{-1})		*(e.u.)*
Chymotrypsin	Methyl hydrocinnamate	–1.9	–5.3	–11.4
	Methyl-D-L-α-chloro-β-phenyl propionate	–2.6	–8.5	–19.8
	Methyl-D-β-phenyl lactate	–2.2	–12.0	–33.0
	Methyl-L-β-phenyl lactate	–2.8	–7.3	–15.1
	Benzoyl-L-tyrosine ethyl ester	–3.3	–8.4	–17.1
	Benzoyl-L-tyrosine	–0.9	–9.9	–30.0
Acetylcholinesterase	Acetylcholine	–5.5	0	+18.5
	Dimethyl amino ethyl acetate	–4.4	0	+14.6
	Methyl aminoethyl acetate	–2.9	0	+9.7
	Aminoethyl acetate	–2.5	0	+8.4
Carboxypeptidase	Carbobenzoxy-L-tryptophan	–3.4	+5	–28
	Carbobenzoxy glycyl-L-phenylalanine	–2.5	–0.4	–7
	Carbobenzoxy glycyl-L-tryptophan	–3.4	0	–11.5
Pepsin	Carbobenzoxy-L-glutamyl-L-tyrosine ethyl ester	–4.7	–1.4	+20.6
	Carbobenzoxy-L-glutamyl-L-tyrosine	–4.3	–3.0	+24.4
Urease	Urea	–3.2	–2.9	+0.9
ATPase	ATP	–7.5	+8	+52
β-galactosidase	β-D-galactosides			
	-phenyl	–5.4		
	-oNO2 phenyl	–4.7		
	-mNO2 phenyl	–4.67		
	-pNO2 phenyl	–6.18		
	-oNH2 phenyl	–4.54		
	-pNH2 phenyl	–4.73		
	-cinnamyl	–2.48		

Enzyme	*Substrate or inhibitor*	*ΔG ΔH* (*kcal . mol*$^{-1}$)		*ΔS* (*e.u.*)
	-methyl	–2.85		
	-naphtyl	–5.15		
	α-L-arabinosides			
	-oNO2 phenyl	–3.21		
	-β-D-fucoside			
	-oNO2 phenyl	–3.42		

In many enzymatic reactions, it is difficult to determine the dissociation constant of the complex; however it is possible to obtain information from the parameter k_{cat}/K_m. Its significance will be analysed in detail in the following chapter. The comparison of the k_{cat}/K_m values for two substrates that differ only by a substituent or charge permits deducing the energy contribution due to this substituent or charge.

This parameter includes both the binding energy of the substrate and the activation energy of the reaction:

$$\ln (k_{cat}/K_m) = \ln (kT/h) - (\Delta G_o^{\neq} + \Delta G_b)$$

$\Delta G_o^{\neq}$ is the activation energy characteristic of bond breaking.

For similar substrates, the differences in $\Delta G_o^{\neq}$ can be considered to be negligible; in other cases, they can be corrected by comparison with reactivities observed in non-enzymatic reactions. ΔG_b is the intrinsic binding energy of substrates (Fig. 10.6).

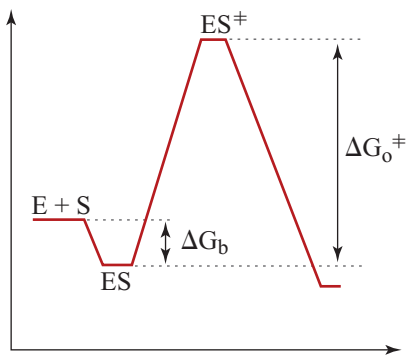


Fig. 10.6 Energy profile of an enzymatic reaction

Therefore, for two substrates s1 and s2 differing by a substituent or a charge, the difference in binding energy in the transition state is given by the following equation:

$$\delta\Delta G_b = -RT \ln \frac{(k_{cat}/K_m)_{s1}}{(k_{cat}/K_m)_{s2}}$$

Diverse evaluations of enzyme-substrate interaction energy were carried out on these bases. In the case of aminoacyl tRNA synthetases, the differences between the binding energy of specific and non-specific substrates and competitive inhibitors were obtained by comparing the values of k_{cat}/K_m for the phosphate exchange reaction. Thus, in comparing k_{cat}/K_m for the association of isoleucyl tRNA synthetase with isoleucine and valine, the association of valyl tRNA synthetase with valine and amino butyric acid, it was shown that a supplementary methylene group contributes 3 to 3.8 kcal . mol^{-1} to the association energy, which is much larger than the simple transfer of a methylene group of water to n-octanol (0.68 kcal . mol^{-1}), as Table 10.4 shows. The comparison of parameters for valine and alanine indicates a difference of 5.41 kcal . mol^{-1} between isopropyl and methyl groups whereas the corresponding value for the transfer from water to n-octanol is 4.6 kcal . mol^{-1}. It seems therefore that the hydrophobic interaction energies with the enzyme are greater than those observed in studies in solution. The same observation was made with chymotrypsin. FERSHT (1984) interprets the situation with chymotrypsin by the presence of at least sixteen water molecules at the binding site of the substrate and around the substrate. All happens as if the hydrophobic interaction for this enzyme corresponds with two normal interactions since there are two interfaces, one for the enzyme and the other one for the substrate, associated with water in an energetically unfavorable manner. In the case of synthetases, the explanation consists moreover of the formation of a cavity in the enzyme-substrate complex when valine occupies the binding site of isoleucine in isoleucine tRNA synthetase.

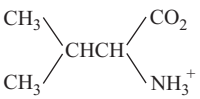
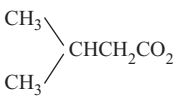
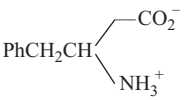
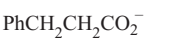
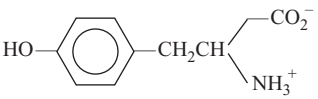
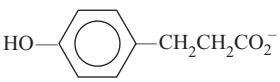
Table 10.4 Transfer energy values of selected groups
(from FERSHT, 1985)

Group	*$\delta\Delta G$ for the transfer of n-octanol to water (kcal . mol^{-1})*	*$\delta\Delta G$ for the transfer of aminoacyl tRNA synthetase to water (kcal . mol^{-1})*
—CH_3	0.68	3.2
—CH_2—CH_3	1.36	6.5
—$CH(CH_3)_2$	1.77	9.6
HO^-	–1.58	7

The role of hydrogen bonds in the association of tyrosyl tRNA synthetase with its specific substrate, tyrosine, was evaluated by comparison with the binding of phenylalanine. The affinity of the enzyme is at least 28 000 times stronger for tyrosine than for phenylalanine ($\delta\Delta G > 7$ kcal . mol^{-1}). The hydroxyl group forms two hydrogen bonds with the enzyme as well as VAN DER WAALS interactions. In the absence of the substrate, there very likely exists a water molecule bound to this site; it is displaced by tyrosine. For the association with phenylalanine, it is not known if the water molecule is displaced or if, from an energy point of view, it is more favorable for the enzyme to undergo some distortions to bind phenylalanine.

The energy contribution associated with the formation of salt bridges between the enzyme and the substrate was obtained by comparing the association energies of amino acids and the corresponding deaminated acids for aminoacyl tRNA synthetases. Table 10.5 shows that the presence of the NH_3^+ group contributes about 4 kcal . mol^{-1} to the enzyme-substrate association energy. This is interpreted by the formation of a salt bridge with a carboxylate group of the enzyme.

Table 10.5 Role of the NH_3^+ group in the interaction energy of aminoacyl tRNA synthetase *(from FERSHT, 1985)*

Compounds			$\delta\Delta G$ *(kcal . mol^{-1})*
	*versus*		4.5
	*versus*		4.4
	*versus*		4.3

Knowing the atomic coordinates of an enzyme-substrate complex or enzyme-substrate analog, it is conceivable to determine the interaction energies by a theoretical approach. To attempt this, it is necessary to know the protein structure at a very high resolution and to possess accurate data on the protein-ligand complex. The problem of the solvent however renders the approach very difficult. Diverse theoretical methods were proposed to analyse protein-ligand interactions. Molecular modelling (BASH et al., 1983) with the example of drug binding, methotrexate to dihydrofolate reductase, permitted specification of interactions between the two molecules, but without determining the interaction energy. The methods of conformational energy minimisation (BLANEY et al., 1982) were applied to determine differences in interaction energy between prealbumin and thyroxine and different analogs. MC CAMMON and collaborators (1984, 1986) proposed a theoretical approach, perturbation of the thermodynamic cycle, to calculate the difference in association energy of a receptor with two different ligands. The method takes into account problems of solvatation.

The examination of energy parameter values of Table 10.3 calls several remarks. The formation of enzyme-substrate and enzyme-inhibitor complexes is always exergonic; however the value of ΔG is generally weak. The value of enthalpy of associ-ation, ΔH, is generally not very high; it is sometimes zero. On the contrary, the enzyme-substrate association is always accompanied by great variations in

entropy, sometimes positive and sometimes negative. The global variation in entropy is the result of at least three contributions. The first corresponds to the formation of a single molecule, the complex, between two molecules, the enzyme and the substrate (or the inhibitor); consequently it must be accompanied by a decrease in entropy ($\delta\Delta S_1 < 0$). The second contribution results from the departure of the solvent around substrate molecules and from the active centre of the enzyme upon association, meaning a modification of the structure of water that goes from a structured state to a less structured state in liquid water ($\delta\Delta S_2 > 0$). Yet these two factors are insufficient to account for the observed entropy variations. Indeed, each time it was possible to evaluate the contribution due to the departure of water molecules upon association of the protein with a ligand, it did not suffice to account for the observed association entropy variation. This effect had been evaluated by PAULING in the study of antibody-hapten association that is a good model of enzyme-substrate interactions. It is necessary therefore to consider another contribution which results in conformational variations in the protein induced by its association with the substrate. The variation in corresponding entropy ($\delta\Delta S_3$) can thus be negative or positive.

The consideration of these energy parameters has brought about progressively, beginning from the 1950s, a revision in the representation of enzyme-substrate interactions.

10.3. MECHANISMS OF ENZYME-SUBSTRATE ASSOCIATION

A more dynamic image of enzyme-substrate association than the “lock and key” mechanism invoked by the first authors was progressively elaborated. However, many hypotheses were proposed, some resting on the model of a flexible enzyme and others conserving still a relatively rigid model.

10.3.1. INDUCED FIT THEORY

The induced fit theory formulated by KOSHLAND is a dynamic model that implicates a certain flexibility of the enzyme. It can be stated by the three following postulates:

- the association of the substrate with the enzyme induces in the geometry of the latter a reversible change in amplitude variable according to the proteins;
- for the enzyme activity to be effective, a very precise and very delicate orientation of the catalytic groups in relation to the corresponding substrate groups is necessary;
- the substrate induces its own orientation according to the change that it provokes in the enzyme geometry.

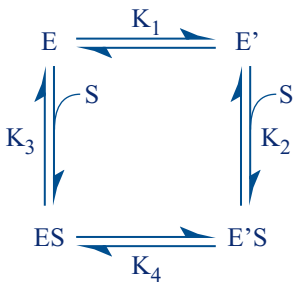
KOSHLAND supports his theory by diverse arguments. The values of thermodynamic parameters discussed above and in particular the values of enzyme-substrate association entropy assume that conformational protein variations accompany the formation of the complex.

The phenomenon of non-competitive inhibition that brings about the loss in activity can only be understood by a conformational enzyme modification. Indeed, a non-competitive inhibitor does not bind to the active centre; consequently its simple association with the enzyme should not bring about a loss in activity if the conformation of the latter is not modified.

The sequential mechanisms in the reactions of two substrates (see Part II) assumes a rearrangement in the local conformation of the enzyme upon binding of the first substrate, exposing the binding site of the second substrate not yet formed in the free enzyme.

The mechanisms of competitive inhibition of β-amylase by cycloamyloses and the formation of non-productive complexes were interpreted by THOMA and KOSHLAND (1960) by the induced fit theory. β-amylase hydrolyses the glycosidic bonds after the two last glucoses, meaning after maltose unity of linear chains of starch. Cyclohexa- and cyclopenta-amyloses behave like competitive inhibitors, but are not hydrolysed by the enzyme. THOMA and KOSHLAND interpreted this difference in behaviour as proof of the induced fit theory. The complex between the enzyme and the substrate is productive when the association occurs in such a way that the non-substituted hydroxyl in C4 forms a hydrogen bond with group X of the protein; this situation induces a change in conformation permitting catalytic groups A and B to be placed in the optimal position with respect to the glycosidic bond to be broken. This is not achieved in the non-productive complex and cannot occur with cycloamyloses (Fig. 10.7).

The induced fit theory can be described by the scheme below:



E is the inactive form of the enzyme, E′ the active form predominant in the presence of the substrate, but not significantly present in its absence; $K_1 << 1$ and $K_4 >> 1$, therefore $K_2 > K_3$, meaning $-\Delta G_2 > -\Delta G_3$. Since $K_1K_2 = K_3K_4$, one has $K_4 > K_1$; in the presence of the substrate the form E′S is favored. This scheme will be discussed in detail in Chap. 12.

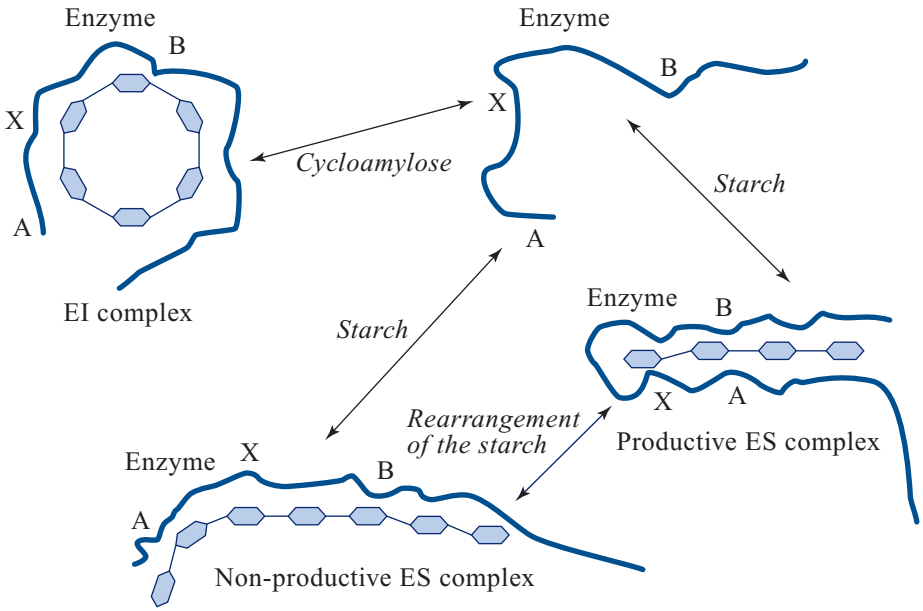


Fig. 10.7 Competitive inhibition of β-amylase by cyclodextrane
(Reprinted with permission from *J. Am. Chem. Soc.*, **82**, KOSHLAND, 3332.
© (1960) American Chemical Society)

There are today many experimental facts indicating the existence of conformational variations of an enzyme consecutive to the binding of the substrate. The first arguments were often obtained by spectrophotometric measurements. For example, the association of an inhibitor like (NAG)3 to lysozyme is accompanied by a modification in the environment of tryptophan residues, the association of a substrate analog with carboxypeptidase brings about a difference spectrum in the region of tyrosines.

The structural data confirmed and specified many of these conclusions. For instance, upon binding of (NAG)3 to lysozyme, a small displacement of residues Trp62 and 63 is observed. In the case of carboxypeptidase, the binding of a substrate analog induces a rotation of 12 Å of Tyr248. For many enzymes, crystallographic studies revealed an open structure in the absence of substrates and a closed structure when the substrates are associated with the enzyme. These enzymes generally have a polypeptide chain folded into two structural domains and the closed structure is generated by a relative movement of two domains (hinge bending motion) that represents a particular mode of induced fit. This situation was observed for several NAD dehydrogenases and kinases, as we will see in particular examples.

In reality the term induced fit is open for discussion. Such as the square scheme is represented, it would be more suitable to consider the stabilisation by the substrate of a conformation that was very minor in its absence. Due to the internal motions in proteins, there exist many slightly different conformations presenting weak energy differences or protein sub-states in fluctuating equilibrium. Some of these less

populated states in the free enzyme would be stabilised by ligand binding. **This would have a selective effect rather than an instructive effect.**

10.3.2. "RACK" OR "STRAIN" THEORY (DISTORTIONS OR CONSTRAINTS IN THE SUBSTRATE)

This theory assuming a relative enzyme rigidity had been proposed by HALDANE (1930) and then by PAULING (1946); it was reintroduced in 1954 by LUMRY and EYRING. It consists of assuming that the active site has a structure complementary to the transition state rather than a fundamental state of the substrate or the product. The substrate, in binding to the enzyme, is found in a constrained state and the bond to break undergoes a distortion that labilises it. Thus HALDANE, recovering the image of the key in the lock, assumes that the key is not correctly adapted to the lock that exerts on it a certain constraint; this constraint in the substrate geometry would have the effect of destabilising the bond to be cleaved. This concept is stated today under a slightly modified form of **stabilisation by the enzyme of the transition state of the substrate,** which in fact is not a new concept.

It is accepted today that the structure of an enzyme active site is complementary to the transition state of the substrate rather than of its fundamental state. There exists a great number of experimental and theoretical arguments to support this concept. The transition of the chair form of the saccharide ring in the fourth position in the substrates of lysozyme deduced from crystallography studies has often been presented as an example of the constraint in the geometry of the substrate upon formation of the enzyme-substrate complex. More recent studies on the association of (NAG)6 with lysozyme by a theoretical approach use refined atomic coordinates and no longer use molecular models. WARSHEL and LEVITT (1976) suggest that the substrate binds to the enzyme in a chair configuration and only takes the half-chair configuration in the transition state. For serine proteases, the structural data indicates that there is no conformational variation of the enzyme upon substrate binding, but a complementarity of the active site for the transition state. One currently possesses numerous data showing that enzymes have a greater affinity for the transition state analogs than for the substrate.

These examples as well as the energy importance of the complementarity of the active centre of an enzyme to the transition state rather than to the fundamental state of the substrate are detailed in the following chapter (Sect. 11.4.3) in the discussion on the particularities of enzyme catalysis.

10.3.3. "DYNAMIC RACK" THEORY

The "dynamic rack" represents a compromise between the concept of the flexible enzyme and that of the rigid enzyme, meaning between the induced fit theory and the "constraints" theory in substrate geometry. It assumes that an enzyme sufficiently

flexible would undergo conformational changes upon substrate binding that would, in turn, bring about constraints in the geometry of the substrates.

In summary, enzyme-substrate associations involve diverse interactions of weak energy, electrostatic interactions, hydrophobic interactions, formation of hydrogen bonds, or salt bridges. These are exergonic processes of which the principal contribution is entropic. They are frequently accompanied by the departure of water molecules and by reversible conformational variations of the enzyme.

The energy consequences of these different mechanisms for the efficiency of the enzyme catalysis are analysed in the following chapter.

BIBLIOGRAPHY

BOOKS

FERSHT A.R. –1977– *Enzyme, structure and function*, 1st ed., W.H. Freeman and C°, New York.

FERSHT A.R. –1985– *Enzyme, structure and function*, 2nd ed. W.H. Freeman and C°, New York.

HALDANE J.B.S. –1930– *Enzymes*, Longmans, Green and C°, London.

REVIEWS

JENCKS W.P. –1975– Binding energy specificity and enzyme catalysis: the Circe effect, *Adv. Enzymol.* **43**, 220-410.

KOSHLAND D.E. –1960– The active site and enzyme action, *Adv. Enzymol.* **22**, 45-97.

SPECIALISED ARTICLES

BASH P.A., PATTABIRAMAN N., HUANG C., FERRIN T.E. & LANGRIDGE R. –1983– *Science* **222**, 1325.

BIESECKER G., HARRIS J.I., THIERRY J.C., WALKER J.E. & WANACOTT A. –1977– *Nature* **266**, 328, London.

BLANEY J.M., WEINER P.K., DEARONG A., KOLLMANN P.A., JORGENSEN E.C., OATLEY S.J., BURRIDGE J.M. & BLAKE C.C.F. –1982– *J. Am. Chem. Soc.* **104**, 6424.

CHOTHIA C. –1974– *Nature* **248**, 338, London.

FERSHT A.R. –1985– *Nature* **314**, 235, London.

LEATHERBARROW R.J., FERSHT A.R. & WINTER A. –1985– *Proc. Natl Acad. Sci. USA* **82**, 7840.

LUMRY R. & EYRING H. –1954– *J. Phys. Chem.* **58**, 110.

LYBRAND T.P., MCCAMMON J.C. & WIPFF G. –1986– *Proc. Natl Acad. Sci. USA* **83**, 281.

NEMETHY G. & SCHERAGA H.A. –1982– *J. Phys. Chem.* **66**, 1773.

PAULING L. –1946– *Chem. Engng. News* **24**, 1375.

RAMACHANDRAN G.N. –1973– Accuracy of Potential Functions in the Analysis of Biopolymer Conformation, in E.D. BERGMANN & B. PULLMAN eds, *Conformation of biological molecules and polymers* (Jerusalem Symposia on Quantum Chemistry and Biochemistry, **5**, 1), The Israel Academy of Sciences and Humanities, Jerusalem.

SEYDOUX F., YON J.M. & NEMETHY G. –1969– *Biochim. Biophys. Acta* **171**, 152.

SUSSMAN J.L., HAREL M., FROLOW F., OEFNER C., GOLDMAN A., TOKER L. & SILMAN I. –1991– *Science*, **253**, 872.

TEMPE B.L. & MCCAMMON J.A. –1984– *Comput. Chem.* **8**, 281.

THOMA J.A. & KOSHLAND D.E. –1960– *J. Am. Chem. Soc.* **82**, 3329.

VOLTZ K.W., KRAUSE K.L. & LIPSCOMB W.N. –1986– *Biochem. Biophys. Res. Commun.* **136**, 822.

WARSHEL A. & LEVITT M. –1976– *J. Mol. Biol.* **103**, 227.

WELLS T.N.C. & FERSHT A.R. –1985– *Nature* **316**, 656, London.

WELLS T.N.C. & FERSHT A.R. –1986– *Nature* **325**, 1881, London.

11 – CATALYTIC MECHANISMS

The mechanisms of enzyme catalysis do not differ from those implicated in chemical catalysis. Enzymes however, due to the complexity of their three-dimensional structure, generally involve a greater number of efficiency factors than simple chemical catalysts. The principle of enzyme catalysis, like all catalysis, is to lower the energy barrier of a reaction, thereby increasing the rate of appearance of the products. There are two ways of lowering the energy barrier: either destabilising the initial state (1), or stabilising the transition state (2) in relation to the non-catalysed reaction (0), as is illustrated in Fig. 11.1. Catalysts can weaken strong substrate bonds or stabilise fractional bonds formed in activated complexes (transition states) in different manners: either by electrostatic interactions (polar effects), by covalent interaction with unpaired electrons, by creating distortions in the bond geometry, or by increasing resonance effects or internal inductive effects. In enzyme catalysis these diverse factors can intervene in the same reaction. The functional groups implicated in enzyme catalysis are of the same nature as those that are involved in chemical catalysis. They concern nucleophilic or electrophilic groups, acids or bases. It is important therefore to consider first the principles of chemical catalysis.

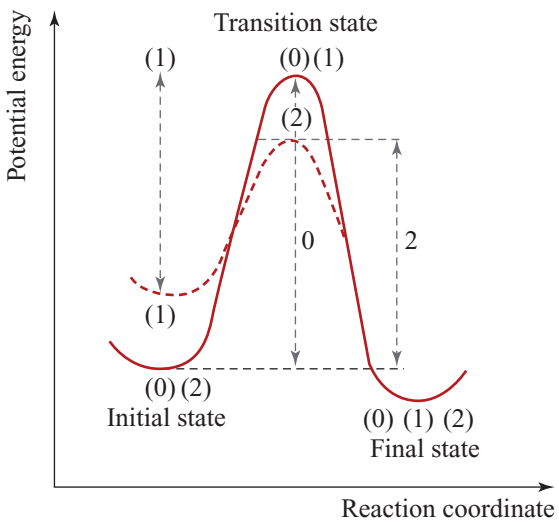


Fig. 11.1 Energy diagram of a reaction in the absence and presence of a catalyst

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_12, © Springer-Verlag Berlin Heidelberg 2010

To present the mechanisms implicated in enzyme catalysis, it is useful to introduce a classification of enzymes that differs from usual classifications and relies on functional considerations. The first class includes enzymes that function without a coenzyme; for these the chemical reaction takes place directly by the intermediate of protein catalytic groups. It can involve side chain residues having a nucleophilic power, such as the imidazole of histidine, the carboxylate of aspartate or glutamate, amine groups, and the functional alcohol of serine or threonine. Protein dipoles can also intervene to polarise the bonds that will be broken. Some of these enzymes require moreover the presence of a metallic cation in coordination with the protein by intermediate atoms like oxygen, nitrogen or sulfur. This is the case with metalloproteases in which the metal plays a catalytic role; carboxypeptidase is an example that will be analysed in detail in Chap. 12.

In the second category are groups of enzymes functioning with a dissociable coenzyme like NAD^+ which is present in a great number of dehydrogenases. From a kinetic point of view, the coenzyme behaves as a second substrate. The chemical reaction is mediated by this coenzyme; nevertheless, certain protein groups can be implicated.

The third category includes enzymes that involve a prosthetic group, that is a non-protein part that is very strongly associated with the protein and non-dissociable; this most often involves a metalloporphyrin. In this case, the chemical reaction uses the metal whose valence or spin state changes over the course of the reaction, resulting in catalysis by a transition metal. Oxydases and peroxydases belong to this category.

Taking into account the diversity of enzymes and the reactions that they catalyse, most of the different cases of chemical catalysis are found in enzyme catalysis.

11.1. CHEMICAL CATALYSIS

11.1.1. DEFINITIONS AND GENERAL PRINCIPLES

11.1.1.1. MECHANISMS OF BREAKING A COVALENT BOND

The breaking of a covalent bond can take place following different mechanisms if one considers what becomes of the electrons from the bond that will be broken. Two types of reactions must be considered, **the heterolytic or polar breaking and the homolytic or radical breaking.** The heterolytic or polar breaking is made of asymmetric bond breaking. It can be schematised in the following manner:

whether the doublet pertains to B or A in the reaction products after breaking. Therefore, when a C—H bond undergoes a heterolytic cleavage, there exist two possible mechanisms; either the electrons stay on the carbon giving rise to a carbanion intermediate and a proton:

or the electrons leave with the hydrogen forming the hydride ion and yielding a carbocation:

The homolytic or radical cleavage is carried out symmetrically with sharing of the doublet and the appearance of two free radicals:

The radical species are generally unstable.

Many biochemical reactions are produced by heterolytic cleavage. However, some enzymatic reactions, in particular those catalyzed by oxidases and peroxidases, produce radical species.

11.1.1.2. NUCLEOPHILIC AND ELECTROPHILIC REACTIONS

Reagents that can catalyze chemical reactions are polar reagents designated as nucleophilic or electrophilic agents according to their role in the reactions. When a reagent gives a doublet to a carbon from the broken bond in an organic molecule, it is called a **nucleophile,** or anionoid. Nucleophilic agents include negatively charged ions, molecules containing atoms with free doublets and molecules carrying polarised or polarisable bonds, for example :

M being a metal.

Over the course of an organic reaction, if a reagent accepts a carbon doublet, it is called an **electrophile** or cationoid. Electrophilic reagents can be positive ions, molecules containing atoms without their octet (LEWIS acids), or molecules containing polarised or polarisable bonds like the following compounds:

The terms base and nucleophilic reagents, the same as for acid and electrophilic agents, have often been confused and can still be for certain reactions. Nevertheless

the expressions nucleophilic agents and electrophilic agents refer to carbon rather than other elements. The term basic alludes to the affinity of a reagent for protons or LEWIS acids, whereas the term acidic applies to the behaviour of protons and LEWIS acids towards reagents containing a free doublet.

11.1.1.3. TRANSITION STATE THEORY

Transition state was already defined in Chap. 4. It corresponds with the peak of the potential energy curve as a function of the reaction coordinates. This is an unstable state in which the bond is forming or breaking. There also exist intermediate states in which the bonds are completely formed and which correspond to troughs in the potential (Fig. 11.2). The stabilisation of the transition state is implicated in catalytic mechanisms; it permits the lowering of the reaction energy barrier.

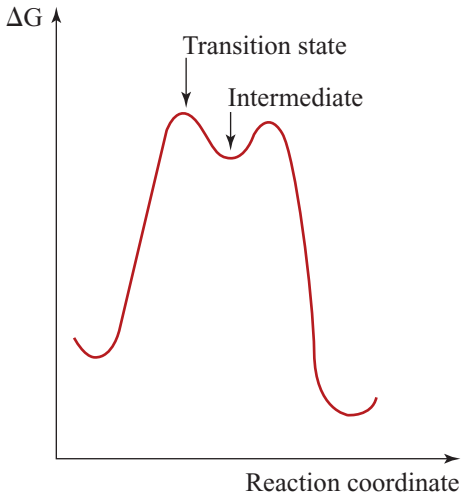


Fig. 11.2 Energy diagram of a chemical reaction proceeding via the formation of an intermediate near the transition state

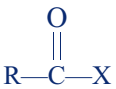
In the analysis of structure activity relationships, the HAMMOND postulate states that if there is an unstable intermediate in the path of the reaction, the transition state must have a structure close to this intermediate. The detection of an intermediate permits therefore to predict the structure of the transition state and to anticipate the types of stabilisation that it requires. The HAMMOND postulate is not applicable to biomolecular reactions.

11.1.2. NUCLEOPHILIC CATALYSIS

In proteins, several amino acid side chains carry nucleophilic groups susceptible to participate in catalysis. This is the case for the serine hydroxyl in serine proteases and in acid and alkaline phosphatases, for the thiol group of cysteine in thiol proteases and in glyceraldehyde-3-phosphate dehydrogenase, and for the OH^- bound to Zn^{++} in carbonic anhydrase. The imidazole of histidine, although it acts in many

cases as a basic catalyst (see below), interacts as a nucleophile in phosphoglycerate mutase. The OH group of tyrosine can also participate in a nucleophilic catalysis.

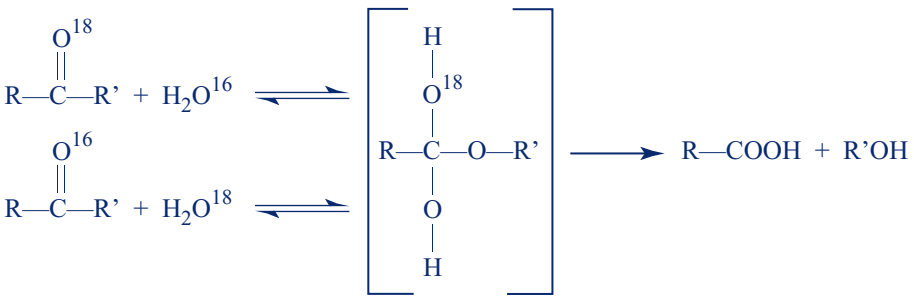
To present nucleophilic catalysis, it is interesting to take as an example of substrates carboxylic acid derivatives of the general form:



because the hydrolytic enzyme substrates correspond to derivatives of this type, the bond to break being a peptide, amide or ester bond. The nucleophilic attack occurs on the carbon of the C—X bond according to a mechanism that uses an addition intermediate; the reactivity depends both on the structure of the substrate and on that of the nucleophilic compound.

11.1.2.1. FORMATION OF THE ADDITION INTERMEDIATE: THE TETRAHEDRAL COMPLEX

The formation of addition intermediates in biomolecular substitution reactions of carboxylic acid derivatives rests essentially on the discovery of a competition between hydrolysis and isotope exchange. This was shown for the alcaline hydrolysis of ethylbenzoate that is a second order reaction showing an acyl oxygen cleavage; this mechanism implicates the formation of an unstable addition intermediate compound according to the reaction:



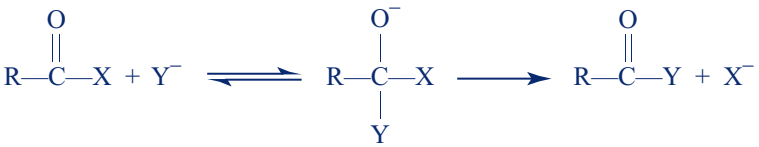
The competition between these two reactions is a proof of the existence of the intermediate compound. The nucleophilic attack occurs approximately in the direction perpendicular to the carboxylic group in the transition state.

The existence of an intermediate in reactions of this type implies a kinetic pathway as follows:

One can treat the kinetics in the steady state, k_2 and k_3 being of the same order of magnitude, a situation general enough in biomolecular substitutions of carboxylic acids. One defines $\alpha = k_2/k_3$.

11.1.2.2. EFFECT OF STRUCTURE ON REACTIVITY

The effect of structure on reactivity must be considered both on the nucleophile that attacks and on the derivative that is being attacked by this nucleophile, which can affect both k_1 and the ratio $\alpha = k_2/k_3$. Let us consider indeed the reaction:

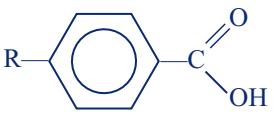


An increase in the electron attractivity of R increases the reactivity of the bond to be cleaved and by consequence increases k_1. It results in a relationship between the reaction rate and the substituent properties given by the HAMMETT law in the case of aromatic compounds and by the TAFT law in the case of aliphatic compounds.

When it concerns an equilibrium like for example the dissociation of an acid, the HAMMETT relationship is written:

$$\log K_R/K_0 = \sigma\rho$$

σ is the parameter that relates to the substituent and ρ the parameter characteristic of the reaction; K_R is the dissociation constant of the compound possessing the substituent R, and K_0, that of the standard compound. As an example, let us consider the compound



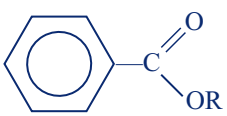
whose standard is benzoic acid (R = H); its dissociation $\rho = 1$ represents the standard reaction, and one compares with other substituted derivatives to determine:

$$\sigma = \log K_R/K_0$$

For the reaction rate, one has an analogous relationship carrying on the rate constants:

$$\log k_R/k_0 = \sigma\rho$$

When the substituents are known as well as the corresponding parameters, one can determine ρ. For example for compounds of the type:



When the substituent is attracting more electrons than H, σ is positive; if the substituent is attracting fewer electrons than H, σ is negative. ρ is positive if the reaction is facilitated by the substituent attracting electrons and negative if it is more difficult.

HAMMETT'S law is applicable to substituents in *para* and *meta*, but not in *ortho*. The polar influence of the substituent is a combination of inducing effects and mesomeric

effects. It cannot apply to saturated aliphatic compounds where one cannot neglect the steric effects.

TAFT'S law applies to acid and alcaline hydrolysis of aliphatic esters; it is expressed by an analogous relationship:

$$\log k_R/k_0 = \sigma^*\rho^*$$

To calculate the values of σ^* corresponding to the substituent R, one compares the rates of acid and alcaline saponification of esters (R—COO—CH_3), the reference being acetate. One can then calculate σ^* corresponding to R:

$$\sigma^* = \frac{1}{2.48}\left[\log\left(\frac{k}{k_0}\right)_B - \log\left(\frac{k}{k_0}\right)_A\right]$$

A and B refer to acid and alcaline hydrolysis, respectively. One takes the value 2.48 to make the TAFT parameters coincide with those of HAMMETT.

These relationships are in general well verified for a great number of compounds. The resonance effects implicating the radical R tend to stabilise the fundamental state with respect to the transition state that must be electronically similar to the addition intermediate tetrahedral compound. The influence of substituents and their effect on the reaction rate can be classified. Table 11.1 which summarises the situation shows that the effect on k_1 is predominant in all the cases.

Table 11.1 Influence of substituents on the reaction rate

Increase in ability to attract electrons on the substituent	*Effect on k_1*	*Effect on k_2/k_3*	*Effect on k_{obs}*
R	↗	↗	↗
X	↗	↘	↗
Y	↘	↗	↘

When a reaction proceeds *via* a nucleophilic attack, the limiting step is the transfer of an electron pair. The rate depends on the ability of the nucleophile to transfer its electron pair; consequently, it is connected to the basic property of the compound by the BRØNSTED relation:

$$\log k = \beta\, pK_a + Ct$$

pK_a represents the ionisation constant of the conjugated acid:

$$B: + H_3O^+ \rightleftharpoons BH^+ + H_2O$$

Figure 11.3 below illustrates this relationship concerning the reactivity of several nucleophilic compounds on paranitrophenyl acetate. The sign and the magnitude of the parameter β are indications of the charge of the transition state.

The nucleophilic power of a compound depends therefore on the basic force and the facility to expel the leaving group, which is determined by both its pK and its stability in solution.

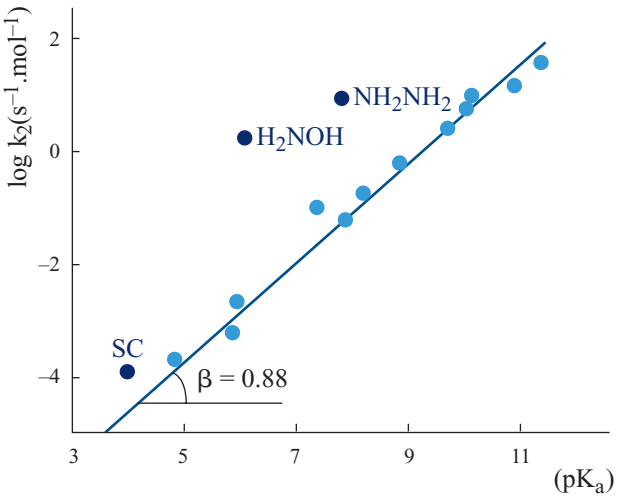


Fig. 11.3 BRØNSTED representation for the nucleophilic attack of p-nitrophenyl acetate by primary and secondary amines

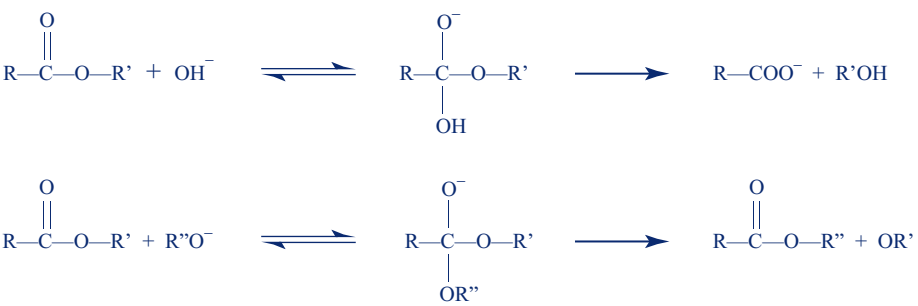
The semi-carbazide (SC), the hydroxylamine and the hydrazine are more reactive than the value of their pK_a foresees. (Reprinted with permission from *J. Am. Chem. Soc.*, **90**, JENCKS W.P. & GILCHRIST M., 2622. © (1968) American Chemical Society)

11.1.2.3. POSSIBLE MECHANISMS OF NUCLEOPHILIC CATALYSIS

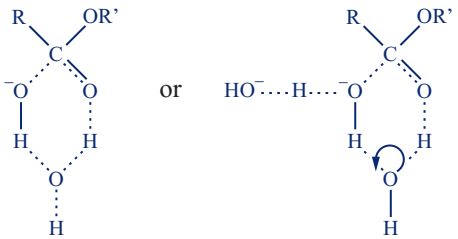
- A BRØNSTED base defined as a proton acceptor can have two functions in organic reactions:
  - either it acts as a nucleophile with respect to a carbon atom,
  - or it accepts protons and reacts according to its classical function as a base that is analysed in Sect. 11.1.3.

These two different functions of a BRØNSTED base bring about two possibilities for its catalytic action, and these two possibilities are encountered in reactions of carboxylic acid derivatives.

Nucleophilic attack occurs upon bond cleavage by the OH^- ions or in alcoholysis, according to the schemes:



In reality the mechanism is a little more complicated because there is an intervention of one or two water molecules; different possibilities have been proposed:



The description of the catalysis implies the existence of at least two consecutive reactions. Two mechanisms are possible:

- either the unstable intermediate is present in an infinitesimal quantity with an energy diagram like the one indicated in Fig. 11.4a below; the kinetics of such a system can be treated by approximation of the steady state;
- or the intermediate attains a concentration having a finite value compared with the concentration of reagents and products (diagram in Fig. 11.4b). In this case, the standard free energy of the intermediate cannot be much higher than that of the reagents:

$$\delta\Delta G < 5 \text{ kcal} \,.\, \text{mol}^{-1}$$

For example reactions catalysed by imidazole can occur either by an intermolecular catalysis or by an intramolecular catalysis (to see later). In intermolecular catalysis, there is a slow formation of a tetrahedral addition intermediate that decomposes rapidly (Fig. 11.4b, curve 1) whereas in intramolecular catalysis, the formation of the intermediate is rapid but its decomposition is slow (Fig. 11.4b, curve 2).

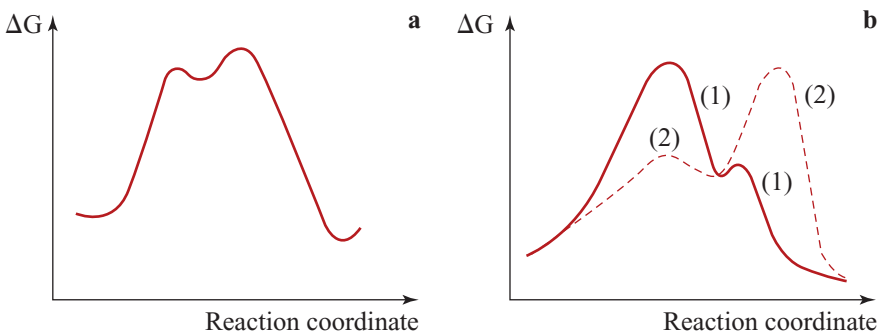
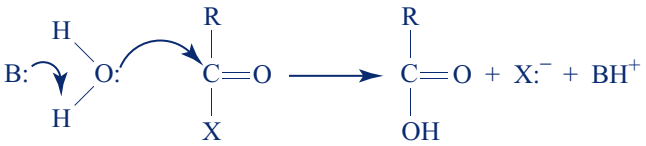


Fig. 11.4 Nucleophilic catalysis
*(**a**) the intermediate is unstable*
*(**b**) the intermediate reaches a significant concentration*

11.1.3. GENERAL BASE CATALYSIS

In general base catalysis, the catalyst reacts by an attack on hydrogen and not on the carbon like in nucleophilic catalysis. Nevertheless, in the two cases, the catalyst is implicated in the slow step of the reaction. The intervention of a general base catalysis in some reactions is shown by the dependence of the rate as a function of

pH; indeed, the catalytic form is the free base B:. All conjugated bases can act as general base catalysts. Let us consider for example the reaction:



For such a reaction mechanism in aqueous solution, the rate equation can be written:

$$v = k(B:)(S)(H_2O) = k'(B:)(S)$$

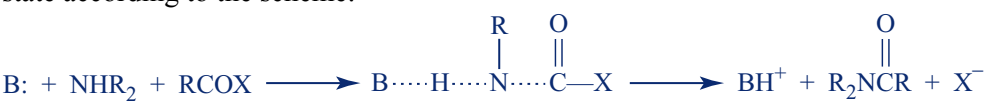
with $k' = k(H_2O)$ and (S) being the substrate concentration. At each pH, the concentration of the base (B:) is given by the relationship:

$$(B:) = (B_0)\frac{K_a}{K_a + H^+}$$

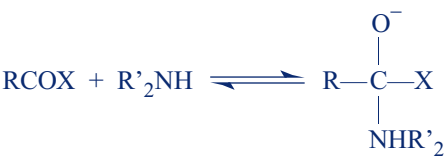
where B_0 is the total concentration of the free base (B:) and of the conjugated acid (BH^+), K_a being the ionisation constant; consequently:

$$v = k'(S)(B_0)\frac{K_a}{K_a + H^+}$$

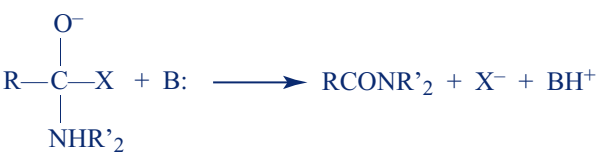
In the case where a base B: and a nucleophile other than water intervene, for example NHR2, three possibilities of general base catalysis can be involved. In the first case, the base carries out the attack of the proton whereas deprotonated nitrogen attacks the carbon by a nucleophilic process, driving a trimolecular transition state according to the scheme:



In the second case, the base catalyst acts in the cleavage of the tetrahedral addition compound:

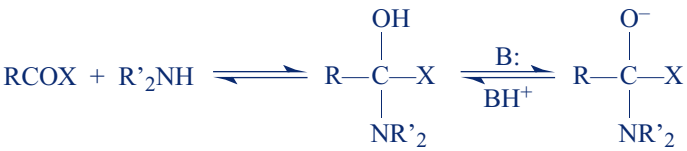


This first step is followed by a slow step:

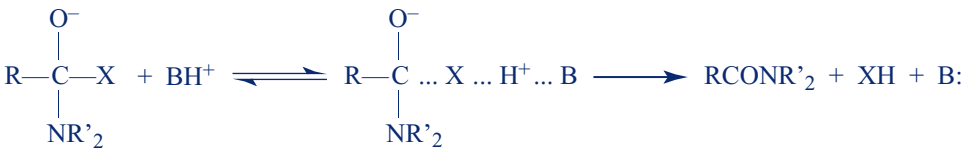


The action of the base on the tetrahedral intermediate compound can be an elimination of the type E_2, the separation of the proton from the nucleophile being concerted with the departure of the X group as an electronegative ion.

In the third case, the general base catalysis proceeds *via* a combination of a prototropic preequilibrium and a general acid catalysis (to see later). There is an equilibrium reaction between B: and BH^+. The mechanism can be represented by one or the other of the schemes below. Following one of the schemes:



This first rapid step is followed by a slow step:



Following the other scheme:

Just as before, this first rapid step is followed by a second slow step:

It is difficult to distinguish between these different mechanisms that proceed *via* the same transition states except for the distribution of atoms.

11.1.4. ELECTROPHILIC CATALYSIS

In electrophilic catalysis, the role of the catalyst and substrate are exactly the inverse of what happens in nucleophilic catalysis; the electrophilic catalyst extracts a pair of electrons from the substrate. However, a step implicating electrophilic catalysis is often preceded by a step in which the catalyst acts as a nucleophile to bind to the substrate. Besides, an electrophilic catalysis in one direction can be nucleophilic in the reverse direction of the reaction.

Some enzymatic reactions involve an electrophilic catalysis over the course of the reaction. The electrophilic groups are either acids conjugated to amino acids susceptible to interacting as nucleophiles, or metallic ions or coenzymes. An interesting example of catalysis by metals is that of the transfer of phosphate of ATP to water or to other acceptors. Metalloenzymes also offer examples of electrophilic catalysis by metals. Coenzymes like thiamine pyrophosphate (Fig. 11.5 below) or pyridoxal phosphate react through an electrophilic catalysis.

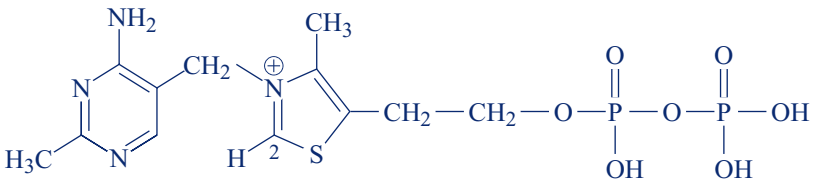


Fig. 11.5 Thiamine pyrophosphate

Thiamine pyrophosphate catalyses the decarboxylation of ketonic acids:

$$RCOCOO^- \longrightarrow CO_2 + RCHO$$

It is used as a coenzyme in transcetolases. It acts in an electrophilic catalysis by the cationic nitrogen atom of the thiazolium ring which is bound to carbon in position 2 by a double bond. Due to the proximity of this atom, nitrogen brings about the stabilisation of a carbanion which results in the rapid exchange with the solvent (Fig. 11.6). It seems that it is the carbanion that intervenes to attack the ketonic carbon and cleave the C—C bond.

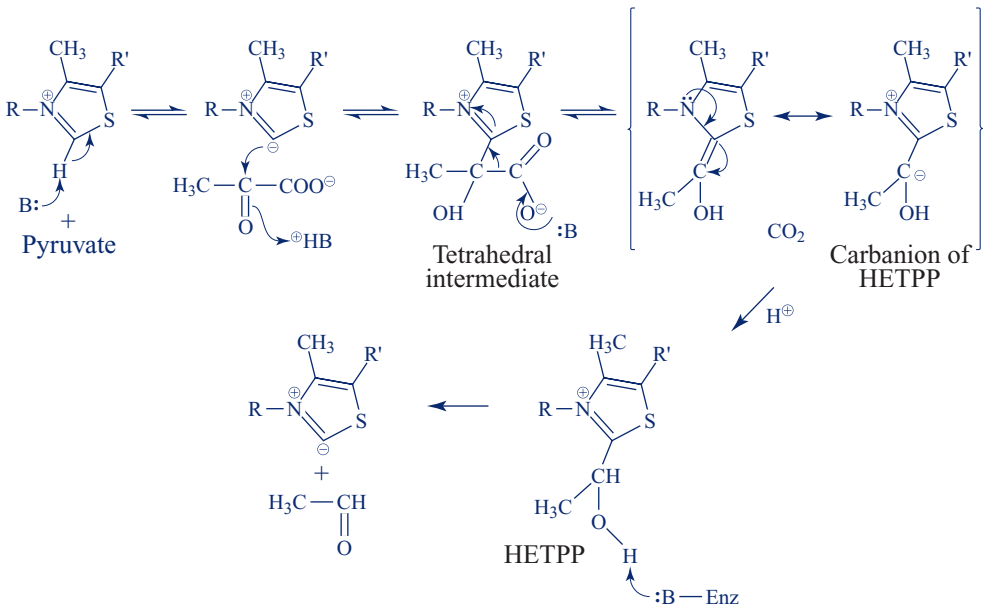


Fig. 11.6 Decarboxylation of pyruvate by thiamine pyrophosphate HETPP: hydroxyethyl thiamine pyrophosphate

Pyridoxal phosphate is the coenzyme of transaminases and some decarboxylases. In transamination reactions, it produces first a condensation of pyridoxal phosphate with an amino acid, which brings about the formation of a SCHIFF base. The pyridinic ring of the coenzyme attracting electrons interferes like an electrophilic agent according to a mechanism that will be detailed in Chap. 12.

11.1.5. GENERAL ACID CATALYSIS

Catalysis by an acid depends also on the strength of the acid. In the case of the BRØNSTED relationship, it is written:

$$\log k = A - \alpha\, pK_a$$

pK_a being the dissociation constant of the acid and α a proportionality constant analogous to the coefficient β defined previously for nucleophilic catalysis.

It was shown that acid catalysis is in fact a catalysis by hydronium ions H_3O^+. In dilute acid solution, the rate of hydrolysis of amides is proportional to the concentration of hydronium ions. The hypothesis of a protonated intermediate was proposed, according to a mechanism of the following type:

$$RCONH_2 + H^+ \rightleftharpoons RC(OH){=}NH_2^+ \xrightarrow{H_2O} \text{products}$$

The apparent rate constant goes through a maximum due to the protonated form of the catalyst and the activity of water. Finally the rate is:

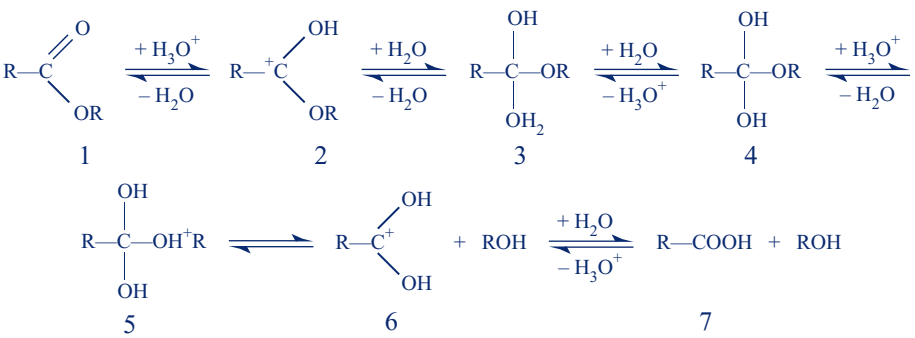
$$v = k(S)(H_3O^+)$$

This expression is valid for acid hydrolysis of amides and esters. The effect of the activity of water in these reactions was proved by their study in non-aqueous solvents, for example water acetone mixtures in which water was the limiting factor.

The reaction rate as a function of pH is given by:

$$v = \frac{k(S)(B_0)(H^+)}{K_a + (H^+)}$$

◤ The study of reactions in D_2O carries also a confirmation of the interference of the activity of water in this mechanism. The ion deuterium in D_2O is an acid three times weaker than the hydrogen ion in water. The concentration of the substrate is therefore larger; the hydrolysis rate is also higher. The ratio k_D/k_H is around 1.5, but H_2O is a better nucleophile than D_2O; this explains why the ratio is 1.5 and not 3. On the basis of these experiments with D_2O, it is possible to conceive the scheme of reactions that proceed during acid catalysis:



The path of the reaction corresponds with the energy diagram in Fig. 11.7.

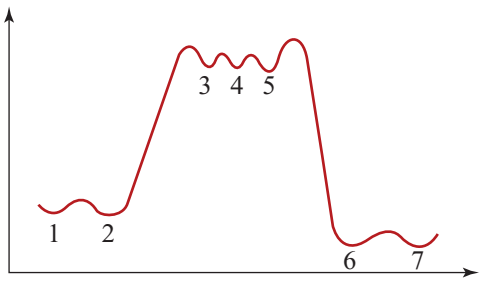
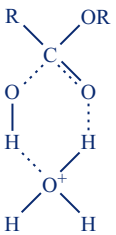


Fig. 11.7 General acid catalysis

The transition state very likely has a cyclic structure of the following type:



There are different cases of general acid catalysis involving different mechanisms. In the first case considered, the transfer of the proton from the acid, in the BRØNSTED definition, to the substrate constitutes a slow process that gives either the final product or an intermediate transforming into the final product:

$$S + AH^+ \longrightarrow SH^+ + A \quad \text{(slow reaction)}$$
$$SH^+ \longrightarrow \text{products} \quad \text{(rapid reaction)}$$

In a second type of mechanism, there is a rapidly formed hydrogen bond between the substrate and the acid; this first rapid step is followed by a slow reaction which generates the products:

$$S + AH^+ \rightleftharpoons S\cdots AH^+ \quad \text{(rapid reaction)}$$
$$S\cdots AH^+ \rightleftharpoons \text{products} \quad \text{(slow reaction)}$$

Upon the formation of a complex by a hydrogen bond between an acid catalyst and a carboxylic derivative, the catalyst changes the electron distribution of the substrate molecule. The formation of hydrogen bonds is evidenced by infrared spectroscopy.

In a third type of mechanism, a protonation equilibrium between the acid and the substrate is produced followed by a step where the base conjugated to the acid intervenes; this last step determines the reaction rate:

$$S + BH^+ \longleftrightarrow SH^+ + B:$$
$$SH^+ + B: \longleftrightarrow \text{products}$$

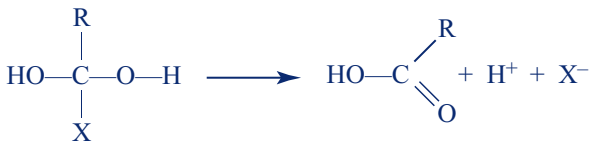
Acid catalysis can be considered as a mechanism in which the introduction of a positive charge in the substrate molecule modifies the electron distribution and makes the reaction possible.

11.1.6. GENERAL ACID-BASE CATALYSIS

As previously mentioned, protein amino acid side chains having a conjugated acid function can serve as acid catalysts and conjugated bases as base catalysts. In general acid-base catalysis both acid and base catalysis occur. It is very frequent in enzymatic reactions.

Let us consider for example the attack of a carboxylic acid derivative in a general acid-base catalysis:

This first slow step is followed by a rapid decomposition of the intermediate:



The reaction rate is given by the relationship:

$$v = k(S)(OH^-)(BH^+)$$

$$v = \frac{kK_W(S)(B_0)}{(H^+)+K_a}$$

with $K_w = (H^+)(OH^-)$.

When a reaction occurs by the intermediate of a base catalyst other than OH^- ions, for example B_1:, and of an acid catalyst B_2H^+, the reaction rate is given by:

$$v = \frac{k(S)(B_1)_0(B_2)_0K_{a2}(H^+)}{[K_{a1}+(H^+)][K_{a2}+(H^+)]}$$

K_{a1} and K_{a2} being the respective dissociation constants.

Tautomerisation of a SCHIFF base formed by pyridoxal and an amino acid offers an example of concerted general acid-base catalysis (Fig. 11.8 below); this mechanism will be detailed in Chap. 12.

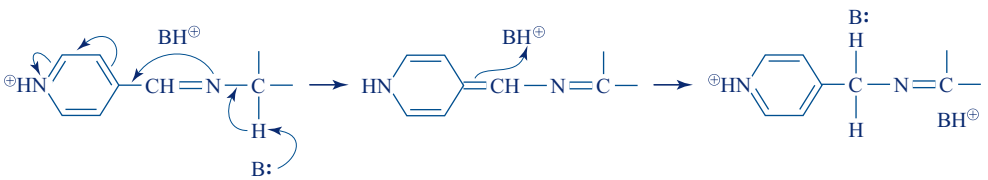


Fig. 11.8 Tautomerisation of a SCHIFF base between pyridoxal phosphate and an amino acid

Figure 11.9 shows the pH profiles corresponding to general acid catalysis, general base catalysis and general acid-base catalysis. In this last case, the reaction rate as a function of pH is represented by a bell-shaped curve that recalls the pH profiles of enzymatic reactions.

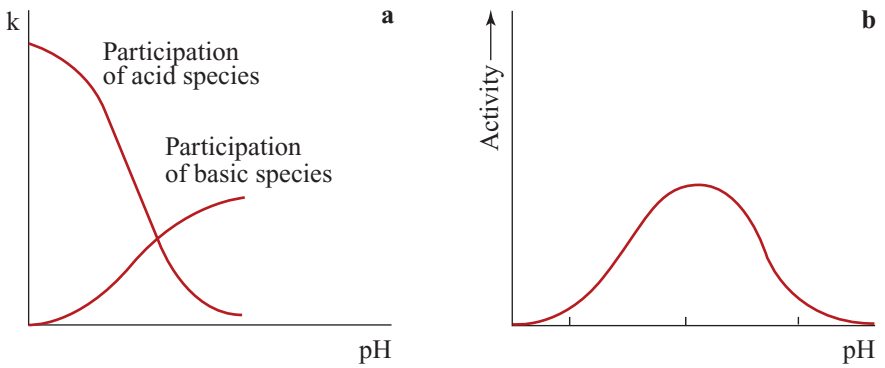


Fig. 11.9 pH profiles
(a) *of acid catalysis and base catalysis –* ***(b)*** *of acid-base catalysis*

Rapid techniques, in particular relaxation methods developed by EIGEN (to see in Chap. 6) permitted measuring the rates of proton transfer between an acid and a base conjugated in aqueous solution for a great quantity of compounds:

$$HA + B: \underset{k_r}{\overset{k_a}{\rightleftharpoons}} BH^+ + A^-$$

k_a and k_r being the rate constants in the directions indicated by the arrows. The transfer rates of a proton of the imidazolium ion to water and from water to imidazole are on the order of 2×10^3 s^{-1}. This value being close enough to the molecular activity of numerous enzymatic reactions implicating a proton transfer, it was proposed that the proton transfer is really the limiting step of the reaction. However, the exchanges with the ionic species of buffers in solution occur with much larger rates permitting enzyme catalytic groups to equilibrate with the solvent at rates on the order of 10^7 to 10^8 s^{-1} and even faster (FERSHT, 1985).

The different catalytic mechanisms just described are rather well known in some simple organic reactions; they are of great importance in enzyme catalysis.

11.2. ISOTOPE EFFECTS

The methods used to determine catalytic mechanisms in enzyme reactions are the same as those used in organic chemistry. They consist of researching the nature of reaction intermediates and determining the limiting steps. These methods include the study of isotope exchange. An example was given previously, in which the existence of a tetrahedral addition complex in the nucleophilic attack of carboxylic acid derivatives was established. They include also the study of isotope effects.

It is frequently observed that the reaction rate in which deuterium is transferred is much slower than the corresponding reaction in which hydrogen is transferred. This isotope effect is used to determine if, in a reaction, the transfer of hydrogen is the determining step. However, it is convenient to distinguish different types of isotope effects: primary isotope effects, the effect due to the solvent and secondary isotope effects.

11.2.1. PRIMARY ISOTOPE EFFECTS

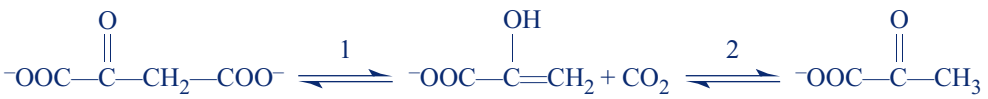
11.2.1.1. DEFINITION

Primary isotope effects occur when there is a cleavage of a bond between the substituted atom and another atom; thus in the following bonds:

C—D	compared to	C—H
C—^{15}N	compared to	C—^{14}N
C—^{18}O	compared to	C—^{16}O
C—^{13}C	compared to	C—^{12}C

The magnitude of the effect observed gives information on the reaction mechanism and the structure of the transition state.

Decarboxylation of oxaloacetate catalysed by Mn^{++} ions can be represented by the following reactions:



In this reaction CO_2 is enriched in ^{12}C with regard to the corresponding isotope ^{13}C and the ratio $^{12}C/^{13}C$ is 1.06. Conversely, in the corresponding enzymatic reaction there is no isotope effect. This indicates that in the non-enzymatic reaction the C—C cleavage constitutes the limiting step, whereas in the reaction catalysed by oxaloacetate decarboxylase, the limiting step is ketonisation (2), that is the transformation of enol in pyruvate.

For reactions implicating hydrogen, for example the C—H cleavage, the rate ratio k_H/k_D between hydrogen and deuterium is generally on the order of 6 to 10. Such

ratios represent normal isotope effects. No theory provides an explanation for larger isotope effects observed in some cases. The separation of a hydrogen from methanol or from acetate to form gaseous H_2 corresponds to a ratio k_H/k_D on the order of 20 to 22. On the contrary, in some reactions the isotope effect corresponds to weaker values of the ratio k_H/k_D that varies between 0.3 and 1.9. For instance, aminolysis of phenyl acetate by glycine, which is an example of general base catalysis, occurs with an isotope ratio k_H/k_D of 1.1; in the hydrolysis of saccharose this ratio is 0.49.

11.2.1.2. ENERGY OF PRIMARY ISOTOPE EFFECTS

Theoretical interpretation is complex. In fact the most important reason for the existence of the isotope effect between hydrogen and deuterium is the difference in the minimum energy point between the C—H bond and the C—D bond as it is schematised in the energy diagram (Fig. 11.10a). According to quantum mechanics the energy levels of H bound to C are quantified. The lowest level energy state of C—H is not the lowest of the curve, but is situated at ½ hν above the potential energy well. The vibration frequency depends on the mass according to HOOKE'S law:

$$\nu \propto 1/\sqrt{\text{mass}}$$

The stretching frequencies of the C—H bond and the C—D bond are respectively equal to 2 900 and 2 100 cm^{-1}, corresponding to 4.15 and 3 $kcal \cdot mol^{-1}$; it is therefore the difference in mass between hydrogen and deuterium that is the predominant factor for determining the vibration frequency of these atoms. If the hydrogen atom is transferred to an acceptor B, this stretching is lost; C—H and C—D reach the same level of maximum energy (Fig. 11.10b). Consequently, the activation energy differs between the two isotopes by 1.15 $kcal \cdot mol^{-1}$. This energy difference corresponds to a rate ratio k_H/k_D equal to 7 at ordinary temperature. In the case of O—H and O—D this ratio would be approximately 10.

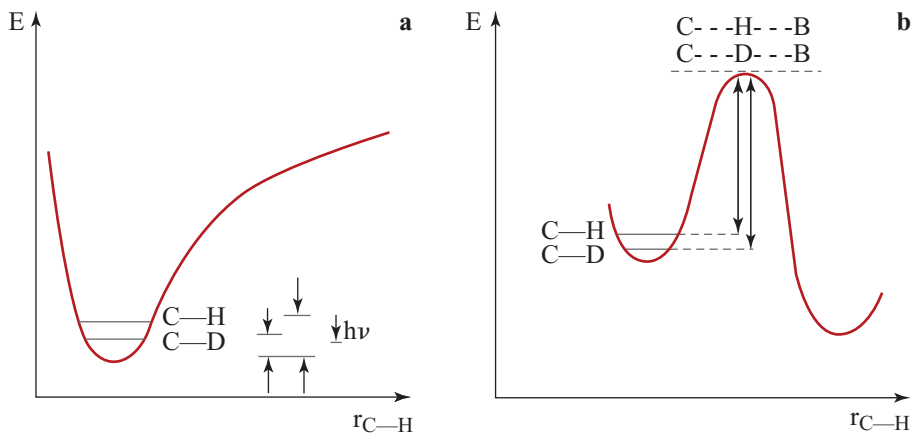


Fig. 11.10 Principle of isotope effects

This explanation takes only into account the effect of the stretching vibration, which is in fact a rather gross approximation. There can also be changes in the vibrations of bond deformation (bending), and in these conditions a new stretching vibration state appears in the transition state. In addition, the reasoning presented above includes another simplification in the sense that the simple C—H vibration is considered isolated in its environment within the core of a polyatomic molecule. However, a given vibration is probably more or less coupled to other vibrations, so that changes in frequency when hydrogen is replaced by deuterium are not always predictable. Whatever the case may be, **the major effect comes from the difference in minimum energy levels.** A quantitative generalisation which can be made also for secondary isotope effects (to see Sect. 11.2.3) results from the fact that species containing deuterium are more stable and the transition states are in narrower potential energy wells.

11.2.1.3. ISOTOPE EFFECTS WITH TRITIUM

Isotope effects are larger with tritium than with deuterium since tritium has a larger mass than deuterium. The mechanism of the isotope effect is similar. **The difference in the lowering of the minimum energy point is directly connected to the mass,** so that the theory permits predicting the following relationship between the rate constants:

$$\log \frac{k_H}{k_{^3H}} = 1.44 \log \frac{k_H}{k_D}$$

This relationship is frequently verified experimentally.

Another possible effect in the transfer of hydrogen is the tunnel effect. The uncertainty principle makes the indetermination in the specification of the proton position sufficiently large such that the probability of H^+ passing on the other side of the energy barrier without crossing over the summit is not negligible. The tunnel effect is less likely for D and ^3H which are larger. This difference can explain the ratios $k_H/k_D > 7–10$. On the other hand, the tunnel effect does not modify the relationship between rate constants for tritium and deuterium.

11.2.2. SOLVENT EFFECTS: EQUILIBRIUMS IN H_2O AND D_2O

The rates of reactions such as the hydrolysis of saccharose which occurs by an acid catalysis are two to three times greater in D_2O than in H_2O for a given acidity. As previously mentioned, one of the mechanisms of acid catalysis involves the following steps:

- a first equilibrium of substrate protonation according to a rapid reaction:

$$S + H^+ \underset{}{\overset{K_{SH^+}}{\rightleftharpoons}} SH^+$$

with the dissociation constant $K_{SH^+} = (S)(H^+)/(SH^+)$;

- the slow transformation of the protonated substrate into products:

$$SH^+ \overset{k_2}{\rightleftharpoons} \text{products}$$

If there is not a transfer of protons in the second step, k_2 will have the same value in D_2O and in H_2O. The difference in the global rate will result then from the difference in concentration at equilibrium of the acid form of the substrate that is two-fold greater in D_2O than in H_2O; indeed, at an given acidity SD^+ in D_2O is an acid weaker than SH^+ in H_2O.

The minimum energy point of an acid A—H such that:

$$A—H \rightleftharpoons A^- + H^+$$

and that of A—D are such that the dissociation of A—D is more difficult than that of A—H. Yet, the difference is smaller for strong acids than for weak acids. The isotope effect on acidity of a series of alcohols and phenols is illustrated in Fig. 11.11 which shows a linear relationship as a function of their pK. The change in concentration equilibrium between acids and bases is given by the relationship (JENKS, 1969):

$$pH = pD + 0.4$$

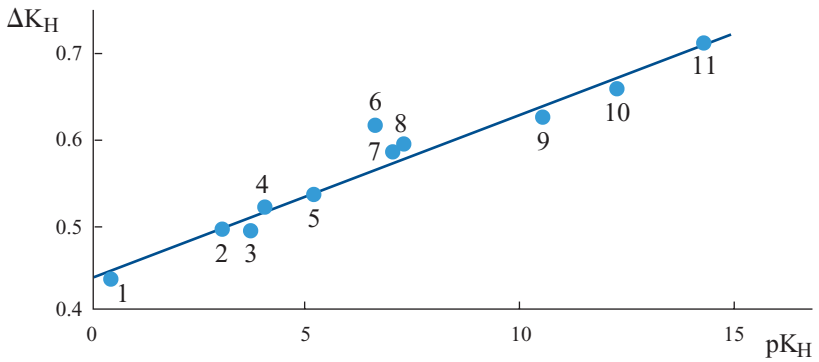


Fig. 11.11 Isotope effect on acidities of phenols and alcohols as a function of their pK_H
(1) picric acid – (2) 4-chloro-2,6-dinitrophenol – (3) 2,6-dinitrophenol – (4) 2,4-dinitrophenol – (5) 2,5-dinitrophenol – (6) 3,5-dinitrophenol – (7) p-nitrophenol – (8) o-nitrophenol – (9) hydroxyquinone – (10) 2,2',2''-trifluoroethanol – (11) 2-chloroethanol
(Reprinted from *Catalysis in Chemistry and Enzymology*, JENCKS W.P., 251, 1969, with author's permission)

In fact the dissociation of an acid in water is better described by the following equilibrium:

$$A—H + H_2O \rightleftharpoons A^- + H_3O^+ \quad \text{or} \quad H_9O_4^+$$

The solvated proton has a significant minimum energy point, and the replacement of H by D can considerably change the dissociation for acids stronger than H_3O^+.

The ionisation of a substrate in D_2O can be perturbed in unpredictable directions. D_2O is 23% more viscous than H_2O, and that introduces a change in the structure

of the solvent and its polarisability. Likewise the O—D bond is 0.04 Å shorter than the O—H bond, which can also bring about important effects.

There is another approach proposed by BUNTON and SHINER (1961) to estimate the solvent effects. It is based on the comparison of stretching frequencies and energy minima of hydrogens in the acids and the molecules of the solvent associated with the acid and the conjugated base (Fig. 11.12):

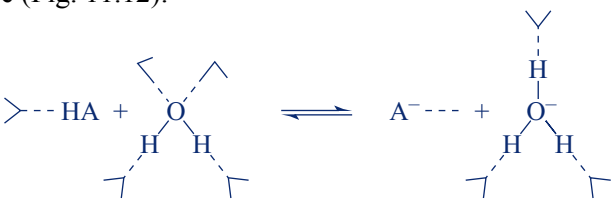


Fig. 11.12 Association of solvent molecules with the acid and the conjugated base

The resulting relationship is the following:

$$\frac{K_{H_2O}}{K_{D_2O}} = \text{antilog}\,\frac{\nu_H - \nu'_H}{12.53T}$$

Table 11.2 gives the calculated values compared with the experimental values. In fact this treatment is often difficult to apply, because one does not always know the nature of the hydrogen bonds; in addition, for some reactions, it is not possible to neglect bond bending.

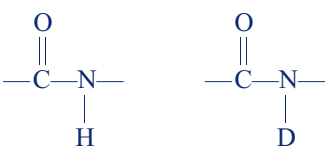
Table 11.2 Isotope effect of the solvent D_2O on the dissociation of several weak acids

(Reprinted with permission from *J. Am. Chem. Soc.*, **83**, BUNTON C.A. & SHINER V.J., 42. © (1961) American Chemical Society)

Acid	*pK_a*	*$K_{H,O}/K_{D,O}$ observed*	*$K_{H,O}/K_{D,O}$ calculated*
Water	15.74	6.5	4.6
β-trifluoroethanol	12.4	4.4	4.5
HCO_3^-	10.25	4.4	3.8
p-nitrophenol	7.21	3.67	3.3
Acetic acid	4.74	3.33	3.1
Chloroacetic acid	2.76	2.76	2.8

11.2.3. SECONDARY ISOTOPE EFFECTS

◤ Secondary isotope effects result from isotope substitutions of atoms that are not among the reacting atoms, but are adjacent. For example, a secondary isotope effect can be observed upon substitution:

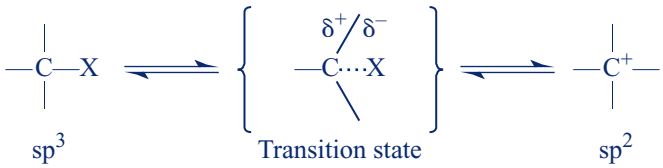


However these effects are the object of controversy because it is not always easy to know if the substituted atom does not react in the transition state.

Secondary isotope effects are often divided into induced effects, steric effects and resonance effects. This denomination does not correspond to the classical sense of these terms, because the observed effects are caused by vibrational frequency variations and not by changes in the surfaces of potential.

11.2.3.1. CHANGE IN FREQUENCY OF BONDS BETWEEN NON-REACTING ATOMS

One such effect is produced for example in reactions with the carbonium ion; the reacting carbon undergoes a change in hybridation from sp^3 to sp^2 to which is associated a change in the vibration frequencies of neighboring atoms:



If the hydrogen or the deuterium is the neighboring atom, a secondary isotope effect resulting in a change in the minimum energy points will be observed.

11.2.3.2. INDUCED EFFECTS

The substitution of a hydrogen by a deuterium in a non-reactive position is susceptible to change the chemical evolution of a molecule in a direction that can be predicted, the deuterium being a better electron donor than hydrogen. Thus, deuteroformic acid is a weaker acid than formic acid.

11.2.3.3. HYPERCONJUGATION

The NMR of ^{19}F in fluorotoluene is a direct method for observing differences in the electron environment of the fluoride atom (Fig. 11.13).

Fig. 11.13
p-fluorotoluene

Thus, measurements with p-fluorotoluene and deuterated p-fluorotoluene (CD_3) indicate that the latter is a weaker electron donor than the CH_3 group, which is contrary to an induced effect and is interpreted as a hyperconjugation weaker for hydrogen than for deuterium, resulting from the energy loss at the minimum energy point in hyperconjugation.

11.2.3.4. STERIC EFFECTS

In general steric effects are due to the difference in lengths of C—H and C—D bonds, the first being 0.005 Å greater than the second. These effects are generally weak.

11.2.3.5. SOLVENT EFFECTS

Besides isotope effects reliant on the establishment of hydrogen bonds between the solute and the solvent, there are differences in solvatation in water and in deuterated water which cannot be attributed to hydrogen bonds. These non-specific effects are weak and have a tendency to decrease the reaction rate in D_2O as compared to H_2O. Such a situation likely results from the fact that D_2O is a more structured liquid than H_2O and, as was previously mentioned, its viscosity is 23% greater such that solvent-solute interactions that depend on the structure of the solvent are different. The solvatation of non-polar solutes is more difficult to explain than that of polar solutes. Hydrocarbides have a greater solubility in D_2O than in H_2O, around 10%. Other effects than the solvent structure, like polarisability for example, can interfere. ◢

11.2.4. MAGNITUDE OF ISOTOPE EFFECTS

There are many reactions implicating a proton transfer that do not show the normal ratio $k_H/k_D = 6–10$ of a primary isotope effect. Besides the explanations already mentioned, in particular the tunnel effect (see Sect. 11.2.1), other explanations were proposed. The first relates to asymmetric transition states. If the transition state is asymmetric, the hydrogen will be able to move in the transition state and consequently there will be a difference between hydrogen and deuterium. A second cause of these variations in the ratio k_H/k_D is the existence of bond bending frequency.

A third cause of abnormal isotope relationships results from non-linear transition states. Usually in the transition state the atoms are colinear:

In certain cyclic reactions however, there can be curvature:

Changes in vibration frequency and minimum energy points in non-linear transition states are different than those that are produced in linear transition states in which the loss of bond stretching vibration is generally rather large. When there is only a loss in bending frequency in the transition state, the observed isotope effect is weak, and the ratios k_H/k_D are on the order of 2–3.

A fourth case in which weak, even inversed, isotope ratios are observed is that of the proton not being at an energy maximum in the transition state. If the proton remains bound to the reagent or to the product and therefore does not reach the energy of the transition point, the transition state of the reaction is not that of the proton transfer. The clearest situation is that in which the reaction is controlled by diffusion; the limiting step is then diffusion and not proton transfer. The isotope effect that results is that of diffusion and not of the reaction, increased by effects on the equilibrium that precede the reaction. This generally corresponds to rate maxima on the order of 10^{10} $mol^{-1} . s^{-1}$.

Generally, normal isotope effects can be easily interpreted and help in the understanding of reaction mechanisms. In general, the existence of a "normal" primary isotope effect indicates a hydrogen transfer in the transition state when one can prove that it does not result from an effect on the equilibrium constant in a state that precedes the limiting step. However, the absence of an isotope effect does not signify that the hydrogen transfer is not an essential step to the reaction mechanism; in some reactions of acid catalysis, inverse solvent isotope effects can be observed. The situation is more complicated and every generalisation on the transfer of protons becomes difficult when this situation deviates from "normal" effects. The interpretation of isotope effects remains sometimes difficult, and it is necessary to keep every conclusion circumspect.

11.2.5. ISOTOPE EFFECTS ON ENZYME REACTIONS

The study of isotope effects on enzyme reactions requires to consider several possibilities. First, it is convenient to distinguish between the effects on K_m and the effects on V_m. The effects observed on K_m do not have significance themselves, the MICHAELIS constant being generally a complex constant. In the case where K_m represents the dissociation constant of the enzyme-substrate complex, no isotope effect should be observed. An isotope effect on K_m indicates an isotope effect on one of the rate constants. But the absence of the isotope effect on K_m does not signify that this parameter is an equilibrium constant. The variations of V_m also result from changes in the specific rate constants. If there is a limiting step, the isotope effect on V_m will reflect the effect on this step and the interpretation is the same as for a simple chemical reaction. In most cases however, it will be convenient to obtain the individual rate constants to identify an isotope effect.

However, in the case of an enzyme reaction involving a chemical intermediate arising from the MICHAELIS complex, such as an acyl-enzyme, with the following reaction pathway:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES_1 \xrightarrow{k_2} \underset{+\,P_1}{ES_2} \xrightarrow{k_3} E + P_2$$

As previously mentioned, the ratio k_{cat}/K_m is:

$$\frac{k_{cat}}{K_m} = \frac{k_2}{\dfrac{k_{-1} + k_2}{k_1}}$$

Using the analysis developed by NORTHROP (1975), one can show that the isotope effect of deuterium on k_{cat}/K_m is:

$$\frac{(k_{cat} / K_m)_H}{(k_{cat} / K_m)_D} - 1 = \frac{(k_{2H} / k_{2D}) - 1}{(k_2 / k_{-1})_H + 1}$$

The expression of isotope effects with tritium is given by an analogous relationship. It follows that the comparison of isotope effects of deuterium and tritium are expressed by the relationship:

$$\frac{[(k_{cat} / K_m)_H / (k_{cat} / K_m)_D] - 1}{[(k_{cat} / K_m)_H / (k_{cat} / K_m)_T] - 1} = \frac{(k_{2H} / k_{2D}) - 1}{(k_{2H} / k_{2T}) - 1}$$

Thus the isotope effects of deuterium and tritium that can be determined experimentally will concern only the constant k_2. If no isotope effect is observed, this will signify that k_2 is not the limiting step of the reaction.

It is convenient to distinguish the isotope effects implicating non-exchangeable hydrogen atoms from those rapidly exchangeable with the solvent. For experiments in D_2O, it is important also to know the pH dependence of the isotope effect on the kinetic parameters of the reaction. In enzymes, conformational effects are sometimes observed, in particular modifications of hydrogen bonds in the presence of D_2O which are difficult to encompass.

In Chap. 12, examples will be given concerning particular enzyme systems.

11.3. PRINCIPAL TYPES OF REACTIONS CATALYSED BY ENZYMES

The chemical reactions catalysed by enzymes can be classified into four main categories:

- group transfer reactions,
- oxydoreduction reactions,
- isomerisations, eliminations and rearrangements,
- and finally, reactions implicating the formation or cleavage of carbon-carbon bonds.

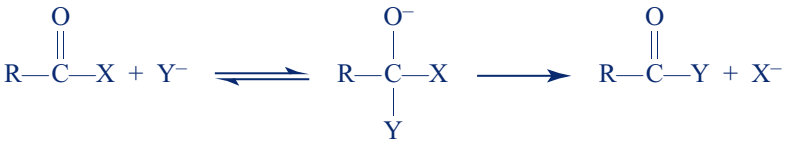
It is important to describe first these diverse reaction types before analysing the characteristics of enzyme catalysis.

11.3.1. GROUP TRANSFER REACTIONS

In this reaction type, an electrophilic part of the substrate, for example an acyl, phosphoryl or glycosyl group, is transferred to an acceptor nucleophile. This latter can be oxygen of the water molecule; there is then degradative hydrolysis. One such mechanism is implicated in enzyme hydrolysis processes of macromolecules (proteins, DNA and RNA, polysaccharides, lipids), and also in the hydrolysis of smaller molecules like acetylcholine. Conversely, in the cell, other molecules carrying a nucleophilic group such as oxygen, nitrogen or sulfur serve as the acceptor in synthesis processes.

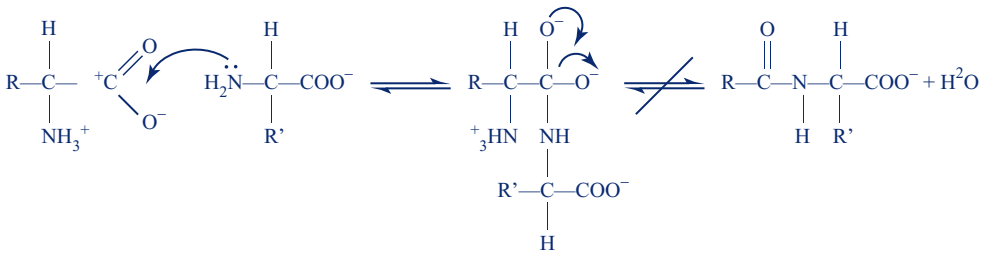
11.3.1.1. ACYL TRANSFER REACTIONS

Reactions of acyl transfer were taken like the previous example to describe mechanisms implicated in nucleophilic catalysis. These are reactions of the following type:



These reactions proceed through the formation of a tetrahedral intermediate as was shown in Sect. 11.1.2.1. When the nucleophile is the oxygen of water, such a mechanism leads to hydrolysis. When the acceptor is an amino acid or a peptide, and the nucleophile is amino nitrogen, there is a transpeptidation.

Protein biosynthesis which implicates the formation of peptide bonds cannot be done by a direct condensation as the following reaction would be:

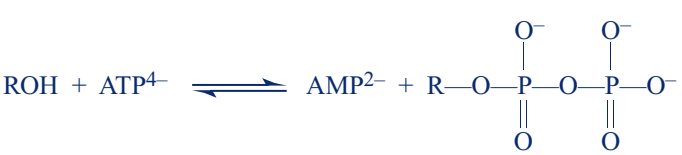


This last reaction can take place only if the carboxylate is activated. For protein biosynthesis in the cellular milieu, the activation of the carboxylate occurs *via* its conversion in phosphoric ester which uses ATP; it is the formation of the aminoacyl adenylate. The phosphate group can be easily eliminated *via* a tetrahedral addition complex. In reality it occurs following transfer reactions of phosphoryl groups.

11.3.1.2. PHOSPHORYL GROUP TRANSFER

The formation of aminoacyl adenylate consists of a nucleophilic attack on the α phosphate of ATP, releasing the pyrophosphate (Fig. 11.14 opposite). The following step is the charge of aminoacyl adenylate on specific tRNA which implies a nucleophilic attack on the carbon of the phosphoric ester bond by the hydroxyl oxygen in position 2′, with liberation of AMP and formation of aminoacyl tRNA. Then a very rapid migration in position 3′ occurs. Finally, these aminoacyl esters form the peptide bond by nucleophilic attack on the carbon of the ester bond by the amino group nitrogen with liberation of the deacylated tRNA.

The attack on the β phosphate of ATP drives a pyrophosphate transfer:



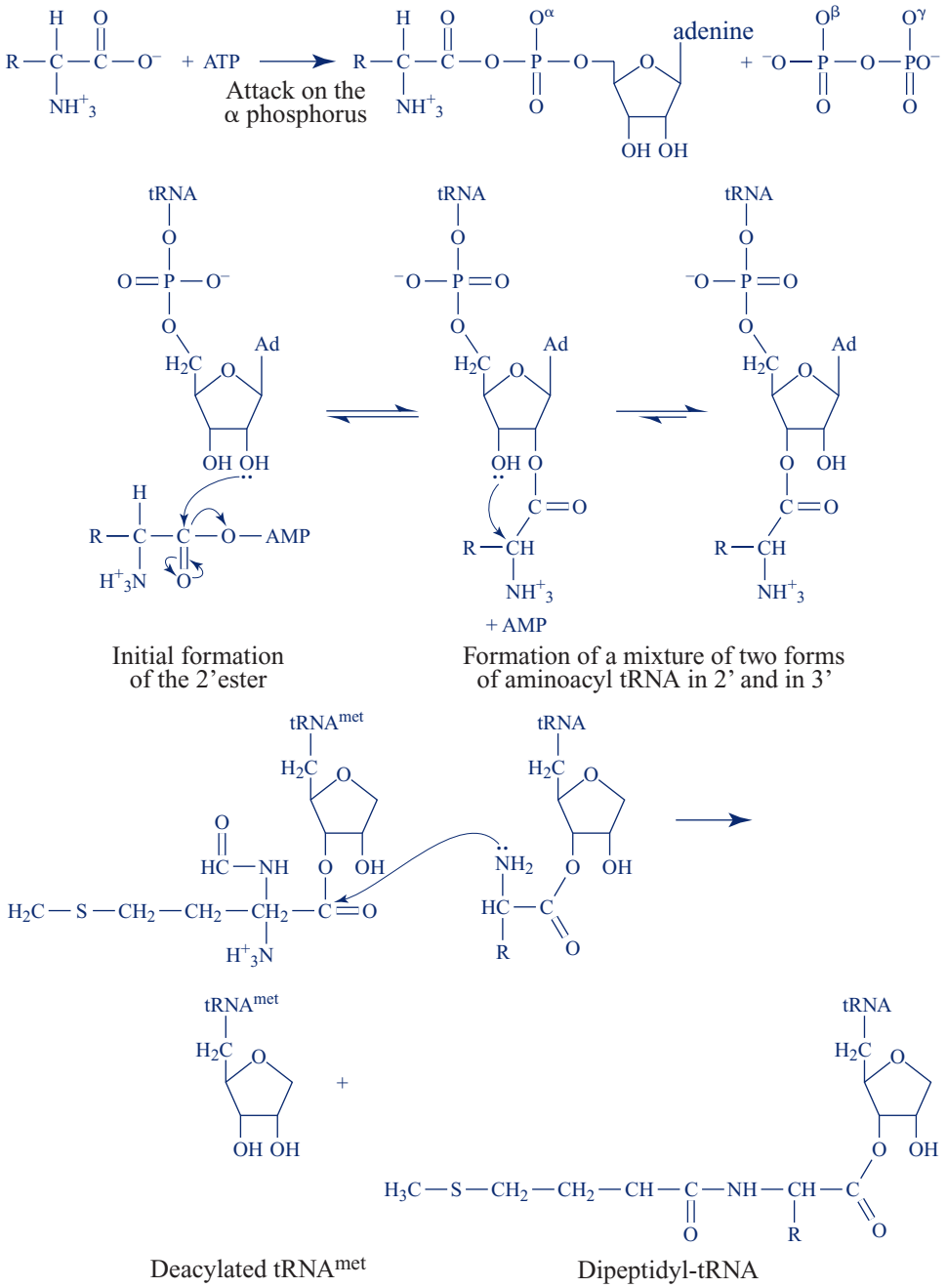


Fig. 11.14 Formation of a peptide bond in the reactions catalysed by tRNA synthetases *The series of reactions includes the formation of aminoacyl AMP, aminoacyl tRNA with specific tRNA, and the condensation of methionyl tRNA^{met} with aminoacyl tRNA to form the peptide bond with release of the tRNA^{met}*

Such a reaction is rare in enzymology; however it takes place in reactions catalysed by 5-phosphoribosyl pyrophosphate amidotransferase (PRPP amidotransferase), thiamine pyrophosphate synthetase (TPP synthetase), hydroxymethyl pteridine pyrophosphate synthetase (HPPP synthetase), 3′-pyrophosphoryl 5′-guanosine diphosphate synthetase (PPGPP synthetase) and 3′-pyrophosphoryl 5′-guanosine triphosphate synthetase. Figure 11.15 illustrates the transformation of phosphoribose (α anomer) in 5-phosphoribose-1-phosphate.

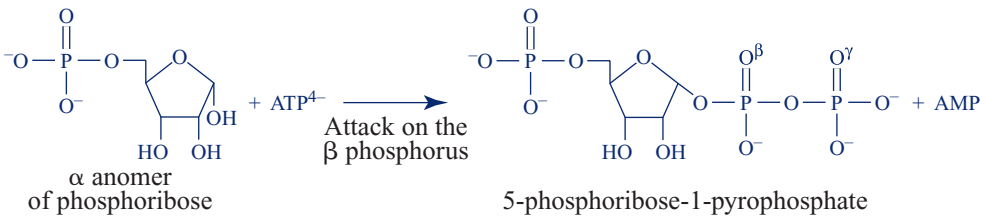


Fig. 11.15 Transformation of 5-phosphoribose in 5-phosphoribose-1-phosphate (attack on the β phosphate)

The attack on the γ phosphate is more frequently observed. The transformation of glucose in glucose-6-phosphate, which constitutes the first reaction of the glycolytic pathway catalysed by hexokinase offers an example (Fig. 11.16).

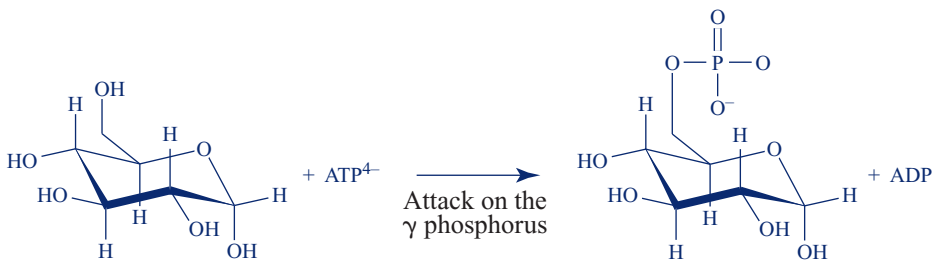


Fig. 11.16 Attack on the γ phosphate of ATP during the transformation of glucose in glucose-6-phosphate catalysed by hexokinase

11.3.1.3. GLYCOSYL GROUP TRANSFER

The transfer of glycosyl groups is carried out in organisms by glycosyl transferases. In degradation reactions, the acceptor is water; there is a hydrolysis of glycosidic bonds. In synthesis reactions, first an active glycosyl derivative intermediate is formed that uses ATP or another triphosphate nucleotide, thus making possible the attack of carbon C_1 by a nucleophile (Fig. 11.17 opposite).

11.3.2. OXYDOREDUCTION REACTIONS

The processes of oxydoreduction and their importance in biological systems where oxidation is often coupled to energy production were presented in detail in Part I.

The notion of oxydoreduction was analysed there. In the present chapter, only the chemical mechanisms involved in these reactions are considered.

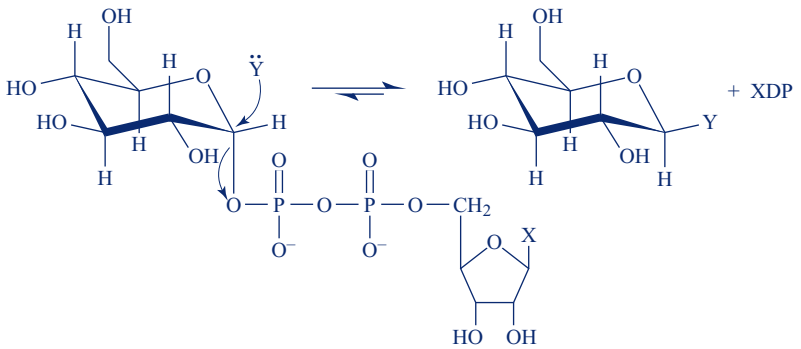
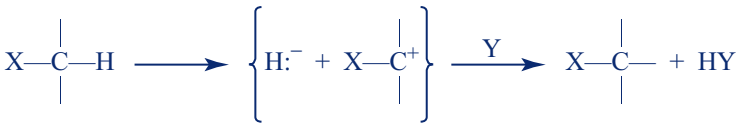


Fig. 11.17 Glycosyl group transfer reaction

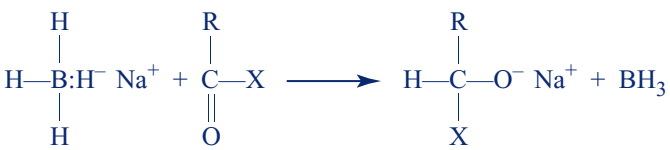
In most enzyme reactions, the oxydoreduction proceeds either by a twice $1e^-$ transfer step mechanism or by a once $2e^-$ transfer step mechanism. In reactions proceeding by a twice $1e^-$ transfer step mechanism, a free radical intermediate that is a very reactive species is formed. One can detect these intermediates only if they are stabilised as free radicals. This occurs in some enzyme reactions having for coenzyme a flavine or a compound of quinoid structure like vitamin C, vitamin E or vitamin K, as well as coenzyme Q (Fig. 11.18 below).

This type of mechanism occurs also in metalloenzymes that possess a transition metal participating in the reaction. Some examples are given in Chap. 12.

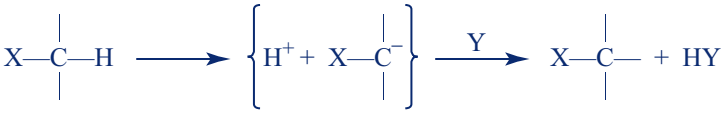
The most classical reactions are those that proceed by a once $2e^-$ transfer step mechanism. Two mechanisms are then possible. In the first case, there is a transfer of hydride with the appearance of a carbocation intermediate:



For example, the reduction of carbonyl compounds by sodium borohydride proceeds by a transfer of hydride:



In the second case, there is an abstraction of the proton with the appearance of a carbanion:



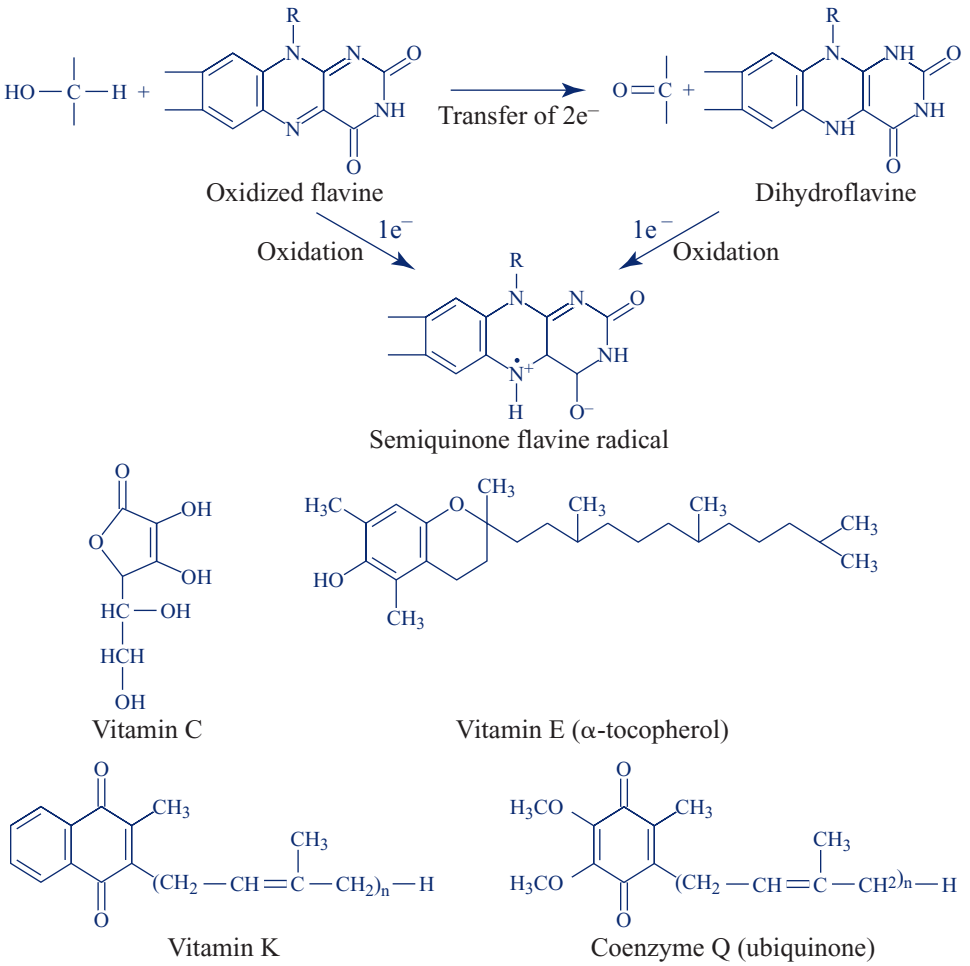


Fig. 11.18 Coenzymes involved in oxydoreduction reactions

Enzymes that have as coenzymes nicotinamide nucleotides such as NAD^+ or $NADP^+$ catalyse the reactions that proceed by transfer of hydride (Fig. 11.19), the hydride ion is transferred in C_4 from the pyridine ring driving the formation of hydropyrimidine nicotinamide NADH. This type of mechanism explains the mode of action of NAD^+ dehydrogenases that will be detailed in Chap. 12.

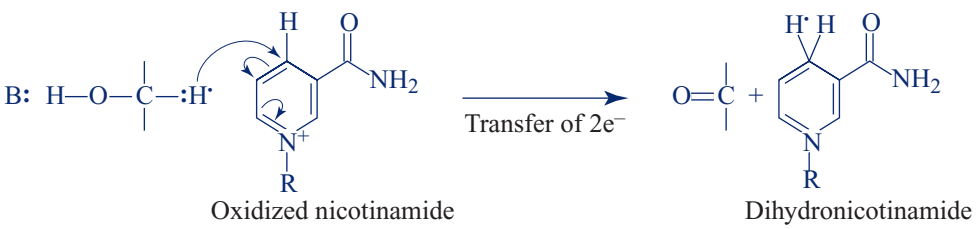


Fig. 11.19 Reactions of hydride transfer on NAD^+

Mechanisms proceeding by proton transfer with the appearance of a carbanion were described for different systems. Some oxydoreduction reactions catalysed by enzymes possessing a flavinic coenzyme can occur through the formation of a carbanion intermediate. Flavine can act either in a once $2e^-$ transfer step process or in a twice $1e^-$ transfer step process with the formation of a semi-quinone intermediate (see Fig. 11.18). The hydrogen transfer reaction between a halogen-lactate and a halogen-pyruvate catalysed by yeast flavocytochrome b_2, which is a flavine dehydrogenase, proceeds by a mechanism of proton extraction with the formation of a carbanion; this was evidenced by the group of LEDERER (1983, 1984, 1985). After an extensive kinetic study implicating the analysis of isotope effects, the mechanism indicated in Fig. 11.20 was proposed.

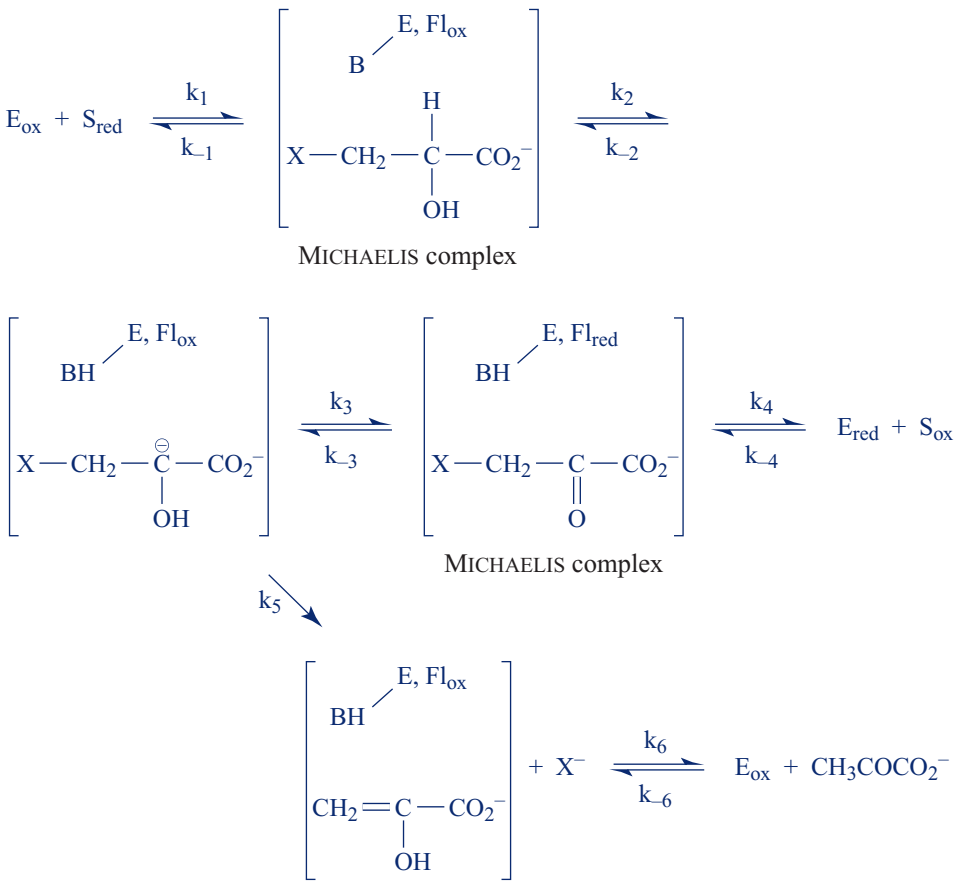
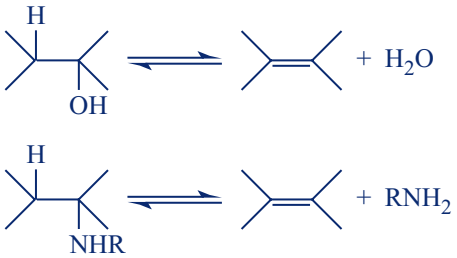


Fig. 11.20 Mechanism of transfer between halogen-lactate and halogen pyruvate catalysed by flavocytochrome b_2 showing that the enzyme is capable of catalysing the reaction in the two directions

The elimination of halogen X brings about a departure of an intermediate that uses a carbanion (from URBAN & LEDERER)

11.3.3. ELIMINATION, ISOMERISATION AND REARRANGEMENT REACTIONS

Enzyme reactions that catalyse elimination transform saturated carbons into unsaturated carbons. For example, the loss of water, ammonium or their derivatives corresponds to this type of reaction that proceeds by the loss of hydrogen by one of the carbons and the loss of oxygen or nitrogen by the adjacent carbon:



Enzymes that catalyse the elimination of water are dehydratases like aconitase, fumarase, enolase, crotonase and all the dehydratases that have more complex molecules as their substrate. For example, enolase reversibly catalyses the transformation of 2-phosphoglycerate (isomer D) into phosphoenolpyruvate in the presence of Mg^{++} ions, according to the reaction indicated in Fig. 11.21, which proceeds by a carbanion intermediate. Carbon-nitrogen lyases catalyse the cleavage of carbon-nitrogen bonds, either in an elimination reaction of ammonium like amino acid ammonium lyases, or in elimination reactions of a NHR group like for example L-arginosuccinase which converts L-arginosuccinate into L-arginine and fumarate. In metabolism, carbon-nitrogen lyases intervene in degradation reactions whereas enzymes like arginosuccinase are implicated in biosynthesis.

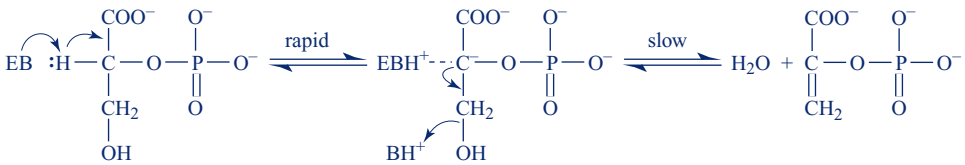
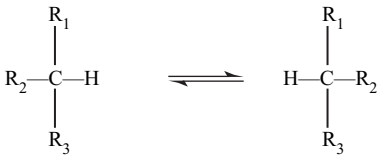
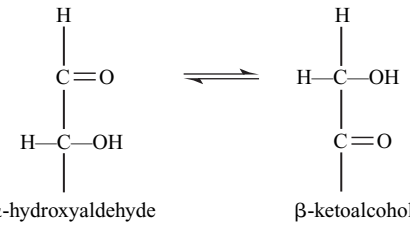
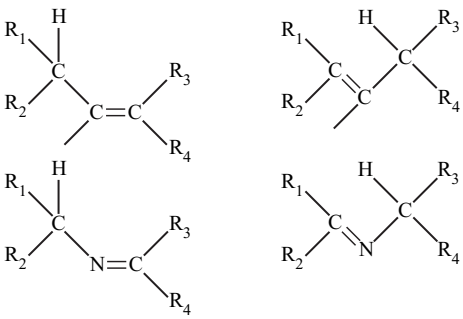


Fig. 11.21 Transformation of 2-phosphoglycerate into phosphoenolpyruvate catalysed by enolase, with the formation of a carbanion intermediate

Isomerisation reactions are catalysed by specific isomerases. They implicate hydrogen displacements in 1,1 or in 1,2 or even in 1,3. As indicated in Table 11.3 opposite, these different types of reactions correspond to specific enzymes. Displacements 1,1 consist simply of the inversion of an asymmetric carbon and are catalysed by racemases or epimerases. Proline racemase converts L-proline in a racemic mixture of D- and L-proline with an equilibrium constant of 1. The mechanism used by this enzymatic reaction involves a carbanion intermediate according to the scheme in Fig. 11.22 opposite.

Table 11.3

Type of displacement	*Examples*	*Corresponding enzymes*
1,1		Epimerases, racemases
1,2	 α-hydroxyaldehyde β-ketoalcohol	Aldose-ketose isomerases
1,3		Allylic isomerisations
		Azaallylic isomerisations

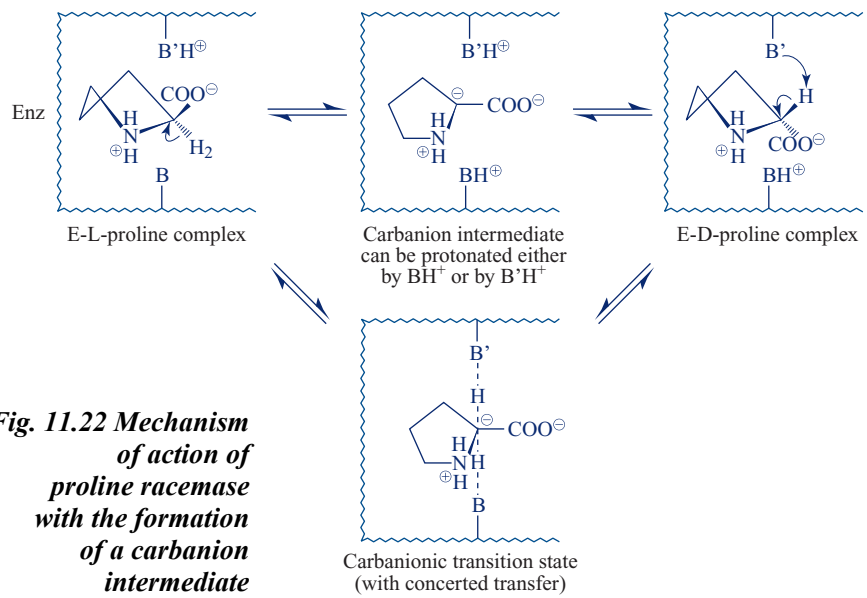


Fig. 11.22 Mechanism of action of proline racemase with the formation of a carbanion intermediate

1,2 displacements take place between adjacent carbons, one being oxydated, the other reduced. Isomerisation reactions of hydroxyaldehydes in keto-alcohols catalysed by aldose-ketose isomerases enter into this category. These transformations proceed through the formation of a cis-enediol intermediate according to the mechanism indicated in Fig. 11.23. Triose phosphate isomerase and phosphoglycerate mutase belong to this class of enzymes.

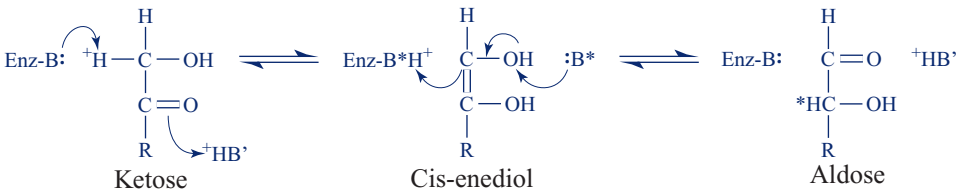


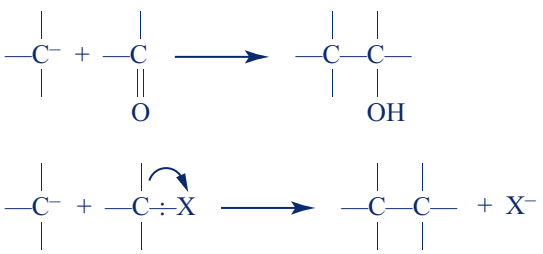
Fig. 11.23 Isomerisation of hydroxyaldehyde in keto-alcohol catalysed by aldose-ketose isomerases

The third category of reactions consists of the transfer of hydrogen in 1,3 according to a mechanism of the type:

with the formation of a carbanion intermediate. Some enzymes like $D^{4,3}$ ketosteroid isomerase, aconitate isomerase, and vinyl acetyl CoA isomerase catalyse allylic isomerisations, others like transaminases that function with pyridoxal phosphate catalyse azaallylic isomerisations.

11.3.4. FORMATION OR BREAKING OF CARBON-CARBON BONDS

Enzymatic reactions leading to the formation of a C—C bond generally involve a carbanion intermediate that acts in the nucleophilic attack of the other electrophilic carbon. The carbon atoms attacked by the carbanion are either of the C═O groups (aldehydes, ketones, esters or CO), or tetrahedral carbon atoms whose substituant X is a leaving group:



Carboxylation of pyruvate into oxaloacetate catalysed by transcarboxylase, which is a biotin enzyme, is an example of this type of reaction. The first step is a carboxylation of biotin; it is dependent on ATP according to the global reaction:

$$ATP + HCO_3^- + \text{enz-biotin} \longleftrightarrow ADP + P_i + \text{enz-biotin-}CO_2^-$$

The second reaction is the following:

$$\text{enz-biotin-}CO_2^- + \text{pyruvate} \longleftrightarrow \text{enz-biotin} + \text{oxaloacetate}$$

The mechanism proposed by O'KEEFE and KNOWLES (1986) for the carboxylation of pyruvate in oxaloacetate is a stepwise mechanism utilising a basic group of the enzyme that captures a substrate proton with the appearance of a carbanion; this latter, in a second step, attacks the carboxyl group of N-carboxybiotin (Fig. 11.24). The formation of the carbanion was demonstrated by the analysis of the isotope effect with tritium. However, if the experiments permitted rejecting mechanism a, they did not allow discrimination between mechanisms b and c.

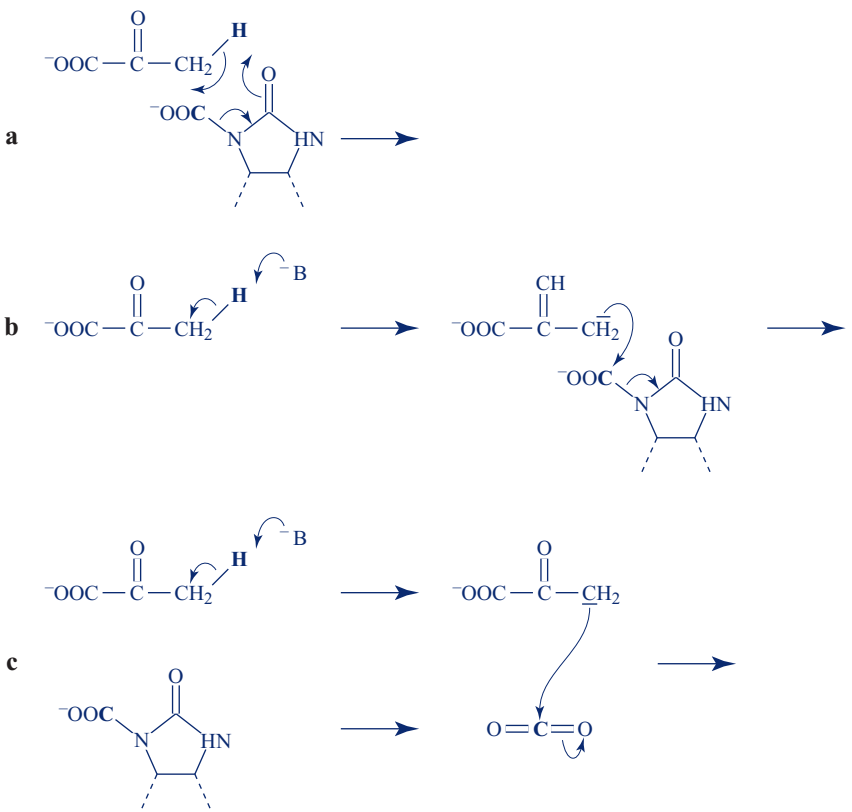


Fig. 11.24 Three mechanisms of enzyme biotin-dependent decarboxylation
Carbon and hydrogen implicated in the isotope effects are in **grey** *characters* (Reprinted with permission from *Biochemistry*, **25**, O'KEEFE S.J. & KNOWLES J.R., 6077. © (1986) American Chemical Society)

Many reactions which use the formation of C—C bonds represent key steps in the biosynthesis pathways of biological molecules.

Other enzymes catalyse the cleavage of the C—C bond; this is the case, for example, of decarboxylases which also have an important role in metabolism. There are three main types of decarboxylases. Some have as substrates α-keto acids like oxaloacetate decarboxylase. Others, such as isocitrate dehydrogenase have as substrates α-hydroxy acids. In the third group, one finds enzymes having for substrates α-keto acids like pyruvate decarboxylase α-keto acid dehydrogenase.

11.4. PARTICULARITIES OF ENZYME CATALYSIS

Enzyme catalysis uses mechanisms of chemical catalysis that are described above. However, it is much more efficient, as shown in Table 11.4; moreover, it is characterised by specificity. These two notions, efficiency and specificity, cannot be separated.

Table 11.4 Comparison of chemical and enzyme reaction rates

Enzyme	$k_n\ (s^{-1})$ *(chemical reaction)*	$k_{cat}/K_m\ (M^{-1}\ .\ s^{-1})$
Alcaline phosphatase	1×10^{-15}	5.7×10^{3}
Acetylcholinesterase	1.1×10^{-8}	1.6×10^{8}
Adenosine deaminase	1.8×10^{-10}	1.4×10^{7}
Cytidine deaminase	3.2×10^{-10}	7.5×10^{5}
Urease	3.0×10^{-10}	2.3×10^{5}
Chorismate mutase	2.6×10^{-15}	1.1×10^{6}
Carbonic anhydrase	3.7×10^{-2}	1.2×10^{8}

For the most efficient enzymes, the ratio k_{cat}/K_m, which represents the apparent second order rate constant for the reaction of the free enzyme with its substrate to give the reaction product and regenerate the free enzyme, is on the order of magnitude of the molecule diffusion rate, which is 10^8 to 3×10^8 $M^{-1}\ .\ s^{-1}$ (Table 11.5 opposite).

In order to understand the characteristics of enzyme catalysis a question must be addressed: how does the three-dimensional structure of the protein macromolecule, by its complexity, generate an ensemble of factors that contribute to assuring the functional efficiency of the catalyst as well as the selectivity, and modulate it following metabolic requirements of the cell? These diverse factors are presented and discussed successively, but it is from the ensemble of their contributions that arise the catalytic performances of an enzyme. The particularities of enzyme catalysis result from several main properties:

- enzyme catalysis is an intramolecular catalysis;
- enzyme catalysis is a polyfunctional catalysis;

- the structure of the active centre of the enzyme is complementary to the transition state of the substrate(s);
- the protein matrix creates a particular microenvironment favorable for the reaction;
- enzyme catalysis involves many reaction intermediates.

Table 11.5 Enzymes for which the ratio k_{cat}/K_m is close to the diffusion rate
(Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **314**, FERSHT A.R. *et al.*, 235. © (1985))

Enzyme	*Substrate*	k_{cat} (s^{-1})	K_m (*M*)	k_{cat}/K_m ($M^{-1}.s^{-1}$)
Crotonase	Crotonyl CoA	5.7×10^{3}	2×10^{-5}	2.8×10^{8}
Triose phosphate isomerase	Glyceraldehyde-3-phosphate	4×10^{4}	4.7×10^{-4}	2.4×10^{8}
Acetylcholinesterase	Acetylcholine	1.4×10^{4}	9×10^{-5}	1.6×10^{8}
Fumarase	Fumarate	800	5×10^{-6}	1.6×10^{8}
	Malate	900	2.5×10^{-5}	3.6×10^{7}
β-lactamase	Benzyl-penicillin	2×10^{3}	2×10^{-5}	1×10^{8}
Carbonic anhydrase	CO_2	1×10^{6}	0.012	8.3×10^{7}
	HCO_3	4×10^{5}	0.026	1.5×10^{7}
Catalase	H_2O_2	4×10^{7}	1.1	4×10^{7}

11.4.1. ENZYME CATALYSIS IS AN INTRAMOLECULAR CATALYSIS

The formation of an enzyme-substrate complex, the MICHAELIS complex, the first one in the reaction pathway, makes the enzyme catalysis an intramolecular catalysis. As a consequence, the reaction becomes of the first order and not of the second order, and the local concentration of catalytic groups increases. In addition, the orientation effects which can result, following the nature of the substrates and the enzyme specificity, optimise the catalysis. **The ensemble of these factors drives an entropy gain as compared to an intermolecular catalysis.** Finally, the different mechanisms implicated in enzyme-substrate associations, induced fit and constraints in the geometry of the substrate, have energy consequences which will be discussed.

11.4.1.1. ENZYME REACTIONS ARE FIRST ORDER REACTIONS

Chemical reaction rates discussed previously, which concern general acid catalysis or general base catalysis, are second order reactions. Following the formation of the MICHAELIS complex, enzyme reactions are first order reactions, the catalytic acid or base groups being carried by the same molecule as the substrate, the complex ES.

Thus instead of having: $v = k_2(E)(S)$

one has: $v = k_1(ES)$

That creates a local concentration effect which can result in a considerable amplification.

11.4.1.2. CONCENTRATION EFFECT

The effective concentration of catalytic groups in the enzyme-substrate complex is one of the most important factors in catalytic efficiency. The local concentration of the catalyst becomes indeed very high. To illustrate this effect, several model chemical reactions were proposed. For example, the hydrolysis of aspirin and of some of its derivatives shows the effective local concentration (Fig. 11.25) according to FERSHT and KIRBY (1967, 1968). The hydrolysis of the ester bond of aspirin is a general base intramolecular catalysis which proceeds as indicated in Fig. 11.25a. The comparison of the rate of this reaction with that of the reaction catalysed by an external base indicates that the effective concentration of the carboxylate ion, which plays the role of the base catalyst, is 13 M.

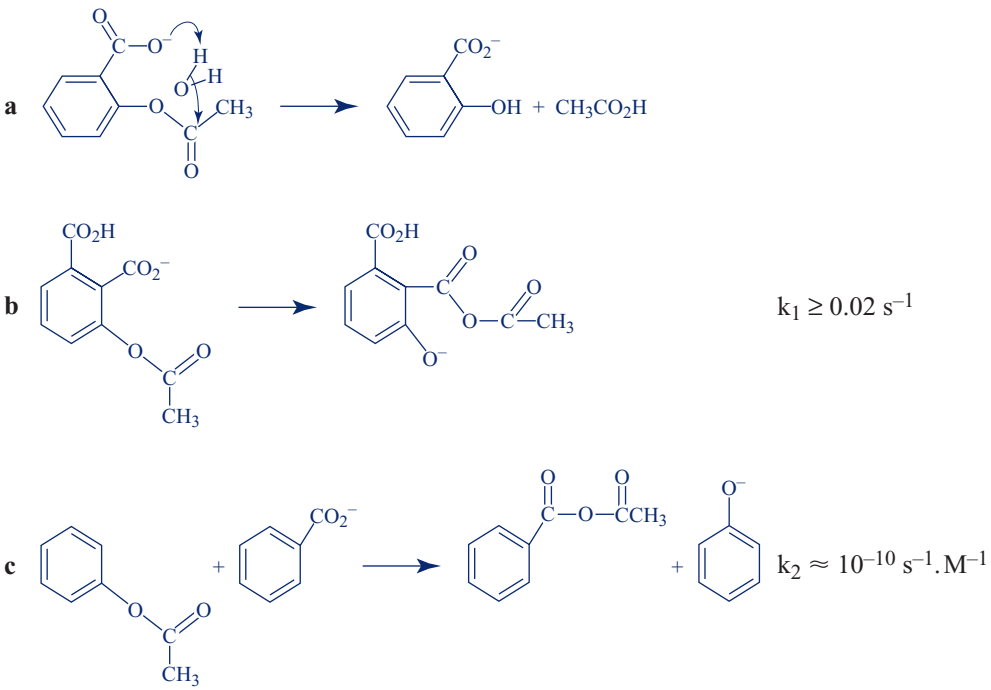


Fig. 11.25 Hydrolysis of aspirin by water (a), and of its derivatives by a nucleophilic group according to an intramolecular (b) and an intermolecular (c) catalysis

(Reprinted from *Enzyme, Structure and Mechanism*, FERSHT A.N., 43. © (1985) with permission from W.H. Freeman and Company/Worth Publishers)

The effective concentrations of the nucleophile are sometimes much more important if the compounds have a constrained geometry. If one compares the transfer rate of acyl in aspirin derivatives by intramolecular and intermolecular catalysis,

one obtains an effective concentration of the base catalyst COO^- on the order of 2×10^8 M. The intramolecular reaction (Fig. 11.25b) proceeds with a rate constant $k_1 > 0.02\ s^{-1}$. In the case of an intermolecular catalysis, the second order rate constant is $k_2 = 10^{-10}\ M^{-1} . s^{-1}$ (Fig. 11.25c), from where an effective concentration of COO^- equals $k_1/k_2 > 2 \times 10^8$ M in intramolecular catalysis.

Another example studied by BRUICE and BRADBURY (1965) is that of the transfer rate of acyl in succinates. The intramolecular reaction (Fig. 11.26a) proceeds with a first order rate constant $k_1 = 0.8\ s^{-1}$, the intermolecular reaction with a second order rate constant, $k_2 = 4 \times 10^{-6}\ M^{-1} . s^{-1}$. The effective concentration of COO^- in the intramolecular reaction is therefore 2×10^5 M.

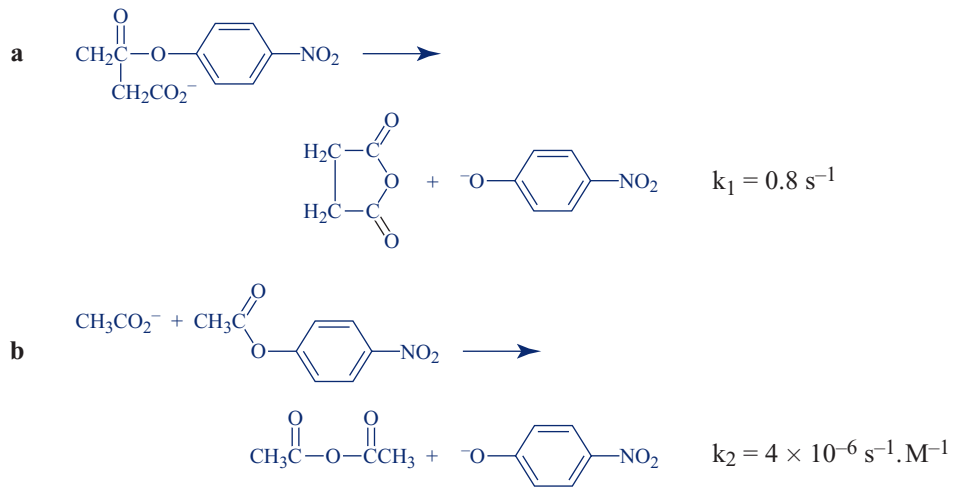


Fig. 11.26 Transfer of acyl in succinates

(Reprinted from *Enzyme, Structure and Mechanism*, FERSHT A.N., 43. © (1985) with permission from W.H. Freeman and Company/Worth Publishers)

BENDER evaluated at 10 M the local concentration of imidazole of histidine in the active centre of chymotrypsin. Diverse models were developed in an attempt to explain the efficiency factor carried by intramolecular catalysis in the case of this enzyme and serine proteases of the same family. One of the models is that of hydrolysis of paranitrophenyl acetate by imidazole and some of its derivatives which proceeds by a nucleophilic attack. The substrate is hydrolysed by chymotrypsin and by imidazole. In intermolecular catalysis by imidazole, the reaction proceeds through the formation of an unstable intermediate compound, N-acetyl imidazole, which was observed by spectrophotometry at 245 nm. The step implicating the appearance of paranitrophenol can be compared with enzyme acylation and the decomposition of N-acetyl immidazol, with the deacylation (Fig. 11.27 below). The first step has a rate 10^3 times weaker than the enzyme reaction; the second step 50 times weaker.

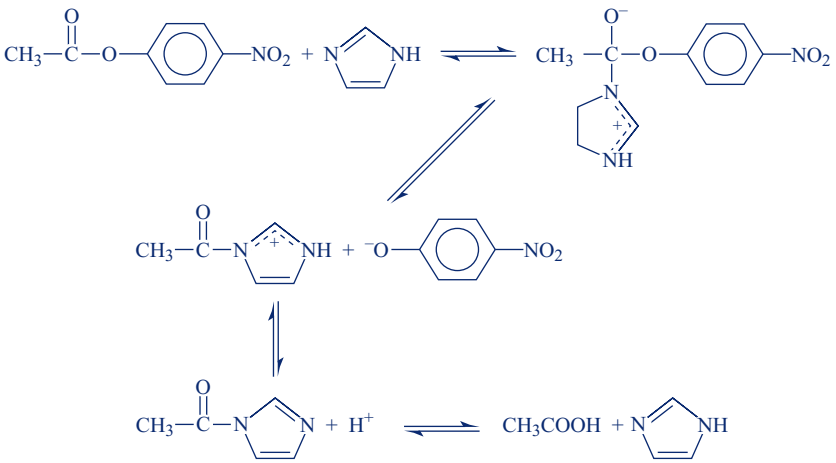


Fig. 11.27 Hydrolysis of paranitrophenyl acetate by imidazole

A model compound of intramolecular catalysis was synthesised; it is p-nitrophenyl-γ-4-imidazolyl butyrate in which imidazole and its paranitrophenyl ester bond are carried by the same molecule. The reaction is schematised in Fig. 11.28. As previously, steps 1 and 2 can be compared to enzyme acylation and deacylation. In step 1 the efficiency is 100 times greater than in the intermolecular reaction, but there are few differences in the second step.

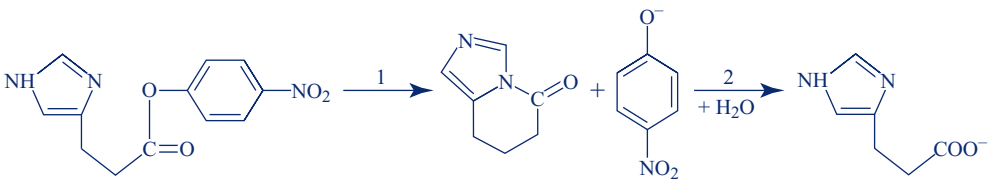


Fig. 11.28 Intramolecular hydrolysis of p-nitrophenyl-γ-4-imidazolyl butyrate

11.4.1.3. ORIENTATION EFFECTS

Several authors, in particular JENCKS, BENDER and principally KOSHLAND, have underlined the fact that the formation of the enzyme-substrate complex is not only accompanied by an effect on local concentration, but drives optimal orientation effects important for the catalysis. It favors the optimal orientation of catalytic groups of the enzyme relative to the atoms of the bond to break when it concerns a "good" substrate.

In 1962, KOSHLAND achieved a calculation to estimate the contribution due to the effective concentration of substrates at the active site of an enzyme as compared to an intermolecular reaction. For this the author assumes that the concentration of reacting species at the active site of the enzyme is as large as if the solution was only composed of reacting molecules. The comparison of rates estimated in such a

way with the observed rates shows that this calculation does not entirely explain the acceleration of the enzyme catalysis. The rate ratios are on the order of 10^9 to 10^{11} (Table 11.6), in the case of reactions catalysed by hexokinase and glycogen phosphorylase for example.

Table 11.6 Comparison of enzyme reaction rates with theoretical rates calculated from the theory of collisions

(Reprinted from *PNAS*, **66**, STORM D.R. & KOSHLAND D.E., Source for the Special Catalytic Power of Enzymes: Orbital Steering, 445, 1970, with authors' permission)

Enzyme	*Substrate*	*V_0 observed* $(mol.L^{-1}.min^{-1})$	*V_E observed* $(mol.L^{-1}.min^{-1})$	*V_E/V_0 calculated*	*$\frac{(V_E/V_0)obs.}{(V_E/V_0)cal.}$*
Hexokinase	Glucose 0.003 M ATP 0.002 M	10^{-13}	1.3×10^{-3}	1.8	$>7\times10^{9}$
Phosphorylase	Glucose-1-P 0.016 M Glycogen 10^{-5} M	5×10^{-15}	1.6×10^{-3}	2.5	$<1.4\times10^{11}$

An analogous deviation is observed in comparing the enzyme reaction rates and the reaction rates of model analogs (Table 11.7) (KOSHLAND & STORM, 1970).

Table 11.7 Comparison of enzyme reaction rates with those of their non-enzyme analogs

(Reprinted from *PNAS*, **66**, STORM D.R. & KOSHLAND D.E., Source for the Special Catalytic Power of Enzymes: Orbital Steering, 445, 1970, with authors' permission)

Enzyme	*Non-enzyme reaction*	*Enzyme reaction rate $V_e(s^{-1})$*	*Non-enzyme reaction rate $V_0(s^{-1})$*	*V_e/V_0*
Lysozyme	Hydrolysis of acetal catalysed by a base	6×10^{-1}	3×10^{-9}	2×10^{8}
Chymotrypsin	Hydrolysis of amide catalysed by a base	4×10^{-2}	1×10^{-5}	4×10^{3}
β-amylase	Hydrolysis of acetal catalysed by a base	1×10^{2}	3×10^{-9}	3×10^{11}
Fumarase	Hydration of alkene catalysed by an acid and a base	6×10^{2}	3×10^{-9}	2×10^{11}

The proximity factors are evaluated at 55 M by the authors. To account for the efficiency of the enzyme catalysis, these latter suggest that in the enzyme-substrate complex, the reacting atoms of the substrate and of the catalytic groups are in an optimal orientation.

The theory of *orbital steering* stated by STORM and KOSHLAND (1970, 1972) is an attempt at a quantitative explanation of these orientation effects. The authors propose the requirement of a precise alignment of electron orbitals of the reacting atoms assuring the overlap of these orbitals in the transition state for the enzyme to form a productive complex with a specific substrate. These experiments conducted by the authors on model chemical reactions led to the estimation that the acceleration of rates of the combination of two substrates driven by two catalytic groups can attain a factor of 10^9 to 10^{15} when the orbitals are correctly aligned. Among the models studied was the formation of cyclic lactones. Figure 11.29 illustrates this effect in the case of those models for which, according to the nature of the substituents, there is a variation in bond angles betweens the atoms.

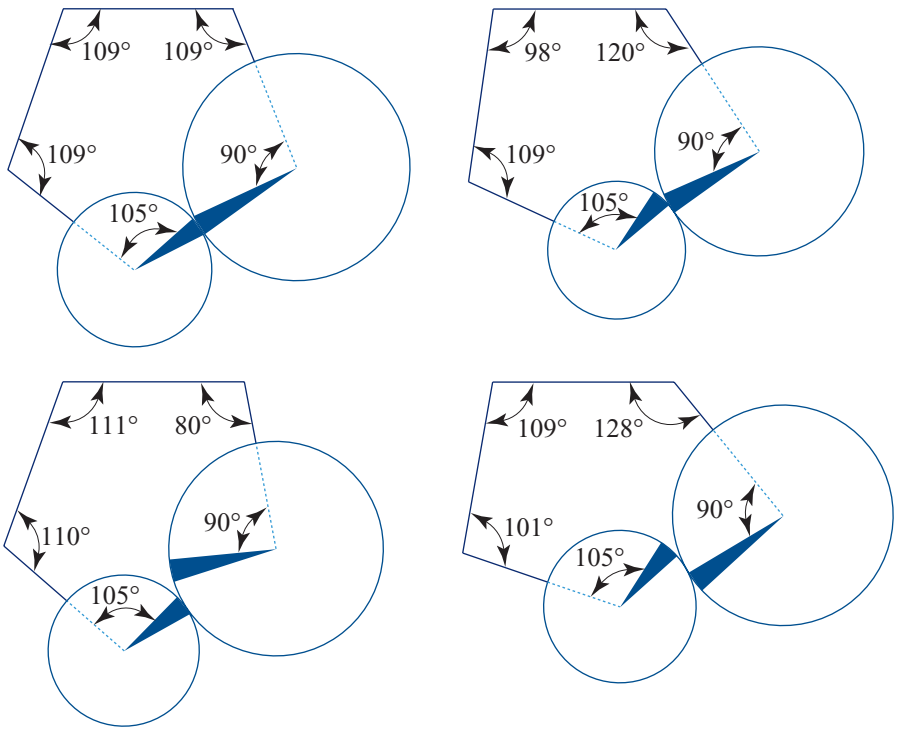


Fig. 11.29 Illustration of orientation effects (orbital steering)
Representation of the variation of angles on the orientation of reacting atoms
The dark part represents the solid angle in which the atoms can react; these parts coincide totally or not at all. The rate will be optimal in the configuration represented in the upper left*, weaker in that on the* right*. In the configurations represented in the lower part, the reaction will not occur*

When the overlap is total, the reaction rate is optimal; it is less if the overlap is partial. It does not occur at all if the solid angles do not coincide. KOSHLAND estimates that a factor of 10^4 can be obtained by an overlap of only 10° in the angles of reactivity. The enzyme function would therefore favor an optimal overlap in the complex formation with the two substrates (Fig. 11.30).

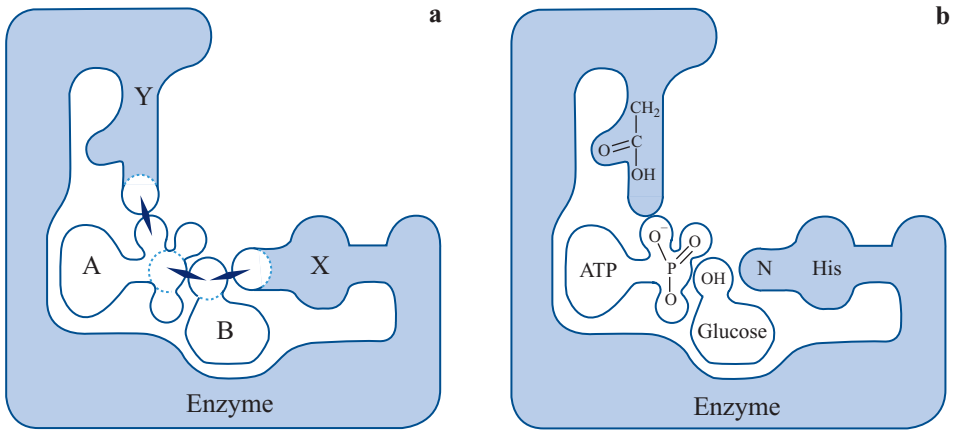


Fig. 11.30 Orientation effects (a) in an enzyme reaction with two substrates, (b) case of hexokinase

(Adapted from *PNAS*, **66**, STORM D.R. & KOSHLAND D.E., Source for the Special Catalytic Power of Enzymes: Orbital Steering, 445, 1970, with authors' permission)

More recently, FERSHT (1984) considers that the theory of *orbital steering*, which has the merit of underlining the role of orientation effects, overestimates their importance. PAGE (1973) and FERSHT (1984) estimate that a distortion of 10° in a carbon-carbon bond, even formed, brings about only a constraint of 2 kcal . mol^{-1}; a distortion of 5° only costs 0.68 kcal . mol^{-1}. Conversely, FERSHT insists on the effect of effective concentration and considers that the value of 55 M adopted by KOSHLAND is underestimated. Nevertheless, the theory of *orbital steering* has the advantage of underlining the importance of orientation effects in the case of an enzyme catalysing the reaction of a "good substrate", effects which correspond to a loss of entropy (see Sect. 11.1.4).

The role of orientation effects in the catalytic specificity and efficiency of an enzyme was illustrated by BENDER and collaborators (1966) in the case of hydrolysis by chymotrypsin of some substrates of variable specificity. The author compares the values of thermodynamic parameters of activation corresponding to hydrolysis by chymotrypsin of four substrates, the first two specific and the other two non-specific, in the deacylation step. He establishes that the values of $D H^{\neq}$ are practically the same for all substrates whereas significant differences are observed in the values of $D S^{\neq}$ (Table. 11.8 below). The absolute value of activation entropy increases when the substrate specificity decreases. The difference $\delta D S^{\neq}$ is 23 e.u. between the most specific and the least specific substrates.

Table 11.8 Thermodynamic parameters corresponding to the deacylation step during the hydrolysis of several substrates of chymotrypsin (Reprinted with permission from *Biochemistry*, **25**, O'KEEFE S.J. & KNOWLES J.R., 6077. © (1986) American Chemical Society)

Substrates	$\Delta H^{\neq}$ *(kcal . mol^{-1})*	$\Delta S^{\neq}$ *u.e.*
N-acetyl-L-tyrosine ethyl ester	10.3	–13.4
N-acetyl-L-tryptophan ethyl ester	12.0	–19.8
Transcinnamoyl imidazole	11.0	–29.6
p-nitrophenyl acetate	9.7	–35.9

BENDER considers it reasonable to attribute these differences to variations in rotational entropies corresponding to torsion oscillations around substrate single bonds. The entropy values corresponding to torsion oscillations around single bonds were calculated by SCHERAGA; they vary between 2.5 and 7.5 e.u.; the experimental values are situated between 4 and 6 e.u. Taking the value of 6 e.u., BENDER estimates that four bonds must be correctly oriented and rigidified during the activation in the case of the least specific substrate, whereas in the case of more specific substrates the bonds are in the adequate position already in the fundamental state of the enzyme-substrate complex (Fig. 11.31). BENDER translates therefore the enzyme specificity into terms of rotational entropy, meaning orientation.

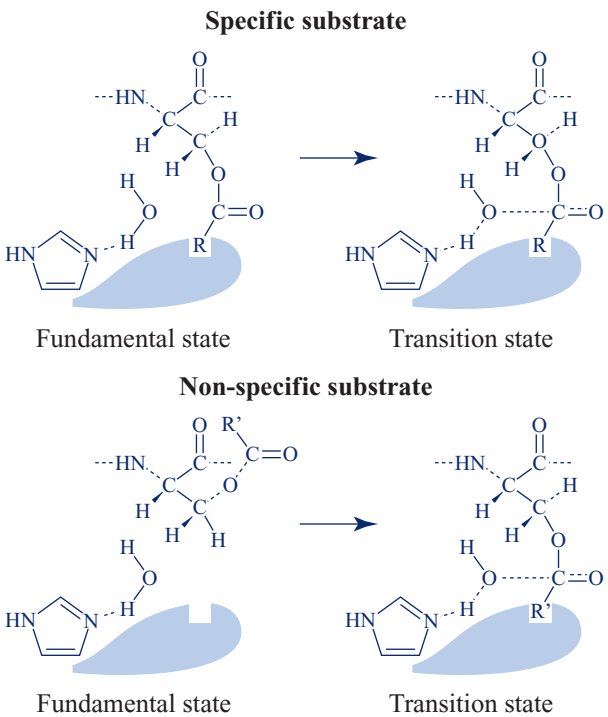


Fig. 11.31 Activation of the acyl-enzyme in the case of a specific substrate and in the case of a non-specific substrate

(Reprinted with permission from *J. Am. Chem. Soc.*, **86**, BENDER M.L. *et al.*, 3714. © (1964) American Chemical Society)

11.4.1.4. ENTROPY EFFECTS

The energy advantage of an intramolecular catalysis as compared to an intermolecular catalysis lies mainly in the entropy contribution. When two molecules associate to form a single molecule, the order of the system increases, which corresponds to a decrease in entropy. In addition, in the enzyme-substrate complex as compared to free molecules, the atoms are found more constrained; there is a loss of a certain number of degrees of freedom of vibration, translation and rotation. Thus, a free molecule has three degrees of translational freedom and three degrees of rotational freedom. When two molecules are condensed to form a single molecule, there is a loss of three degrees of freedom of translation and of rotation. Figure 11.32 illustrates this situation. PAGE and JENCKS (1971) evaluated the contributions carried largely by entropy of translation and of rotation. Translational entropies are on the order of 30 $cal \cdot mol^{-1} \cdot K^{-1}$ at 298 K for small molecules; they increase proportionally to the volume occupied by the molecules. The rotational entropy is on the order of 3 to 5 $cal \cdot mol^{-1} \cdot K^{-1}$ for small molecules, but can reach 30 $cal \cdot mol^{-1} \cdot K^{-1}$ for organic molecules of a larger size. Vibrations contribute to entropy only very weakly.

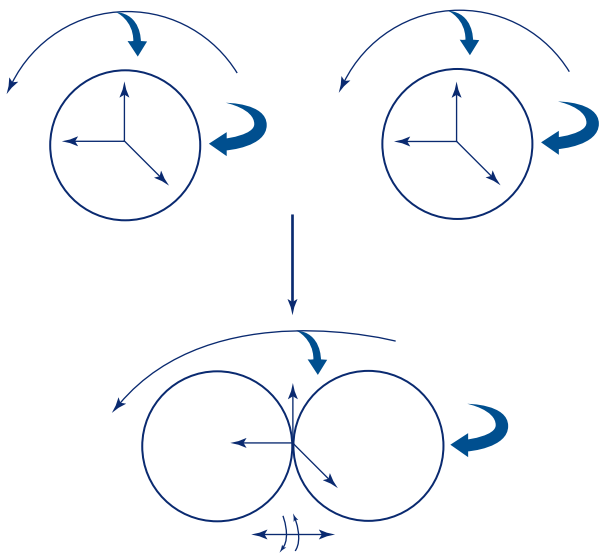


Fig. 11.32 Decrease in number of degrees of freedom of rotation and of translation upon association of two molecules

Considering the dependence of reaction rate on the entropy factor:

$$k_v = \frac{kT}{h} e^{-\Delta H^{\neq}/RT} e^{\Delta S^{\neq}/R}$$

it appears that a loss of degrees of freedom of rotation and of translation representing 45 $cal \cdot mol^{-1} \cdot K^{-1}$ brings about an increase in the rate constant by a factor of

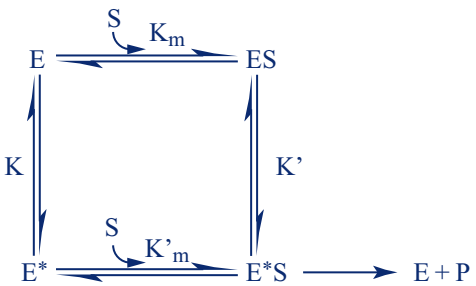
6×10^9, which means that the effective concentration of a neighboring group is 6×10^9. Before the work of PAGE and JENCKS, most authors considered the effective concentration of a neighboring group in an intramolecular catalysis to be 55 M, which is the concentration of water in an aqueous solution, corresponding to an entropy variation of 8 cal . mol^{-1} . K^{-1}. In fact this factor is much larger.

One of the important contributions in enzyme catalysis is therefore entropy. The formation of the MICHAELIS complex in adjusting the catalytic groups of the enzyme to those of the substrate represents a considerable entropy gain. The fact that the catalytic groups themselves are carried by the same molecule also contributes to entropy gain. In this regard, one can consider that the information contained in DNA, that codes for the enzyme and that can be expressed in terms of "neguentropy", is used for catalysis.

11.4.1.5. ROLE OF INDUCED FIT AND CONSTRAINTS

In this chapter, induced fit is uniquely considered in the case of monomeric enzymes. The theory of KOSHLAND, or more exactly the model of KOSHLAND, NÉMÉTHY and FILMER (1966) applied to oligomeric enzymes with allosteric behaviour will be treated in Part V with the aspects of regulation.

The induced fit theory assumes that the enzyme in the free state is not in its optimal conformation. Due to its flexibility, it reaches this optimal conformation by its adjustment to the substrate upon its binding (see Chap. 10). This has the effect of inducing a favorable orientation of the catalytic groups towards the substrate atoms implicated in the reaction. Thus as FERSHT underlined it (1984), this means to induce in the enzyme the conformation which is complementary to that of the substrate transition state (see Sect. 11.4.3). However, this author precised that the induced fit represents an energy disadvantage. Indeed, according to the scheme:



E and ES being the inactive forms of the enzyme and the complex, E* and E*S the active catalytic forms, the observed value of the ratio k_{cat}/K_m is given by the relationship:

$$\left(\frac{k_{cat}}{K_m}\right)_{obs} = \frac{k_{cat}}{K_m}\,\frac{K}{1+K}$$

K << 1 since in the absence of the substrate the form E predominates; therefore:

$$\left(\frac{k_{cat}}{K_m}\right)_{obs} = \frac{k_{cat}}{K_m} K$$

The observed value of the ratio k_{cat}/K_m is smaller than if all the free enzyme would be in the good conformation. Besides, if $K' >> 1$, meaning that practically all the enzyme bound to the substrate is in the good conformation, one has

$$k_{cat)obs} = k_{cat} \quad \text{and} \quad K_{m)obs} = K_m/K$$

The value of k_{cat} remains identical whereas K_m increases, decreasing the ratio k_{cat}/K_m.

In fact, as was previously shown, the existence of induced fit was proven experimentally for several enzymes. It is interesting to ask what its role is. It is not excluded that such a mechanism intervenes to select between several substrates and that the enzyme conformation only adjusts correctly to that which is the most specific. Thus, when diverse substrates are found in competition in the cell, the enzyme would only catalyse the transformation of the "good substrate" and would remain inactive or much less active with respect to the others; concurrent reactions could not occur or would occur with a very weak efficiency, the active site being "poorly adjusted". In this regard, even for monomeric enzymes with michaelien behaviour, the induced fit would constitute a selection mechanism rather than a catalytic advantage. It is however permitted to think of the energy disadvantage associated with induced fit as being compensated for by resulting orientation effects.

Strain, i.e. constraint or distortion in the substrate geometry following its association with the enzyme, represents another type of mechanism that implies a certain rigidity of the enzyme. These constraints have the effect of carrying the substrate into a state close to the transition state. One such mechanism, although it brings about an increase in the catalytic constant, does not change the ratio k_{cat}/K_m as remarked by FERSHT (1984). The increase in the catalytic constant is represented by the energy necessary to distort the substrate; this is reflected in the value of K_m in the simple case where $K_m = K_s$.

11.4.2. ENZYME CATALYSIS IS A POLYFUNCTIONAL CATALYSIS

The polyfunctional character of enzyme catalysis, i.e. the concerted attack on the bond to break by many catalytic groups, also constitutes an important efficiency factor. The discovery of several groups essential to the catalytic act in enzymes, in particular the identification of a serine and a histidine in serine proteases, has prompted the study of models of bifunctional catalysis, the same catalyst carrying two reactive functions. The model of SWAIN and BROWN (1952) already represented a particularly significant example of this type of approach. The authors studied the mutarotation of tetramethyl glucose catalysed by hydroxypyridine. The reaction proceeds at a rate 7 000 times greater than when it is carried out in the presence of an equivalent concentration of pyridine and phenol. The presence on the same

molecule of the acid catalyst and of the base catalyst allowing the concerted attack on the bond to be broken (Fig. 11.33) represents an important efficiency factor.

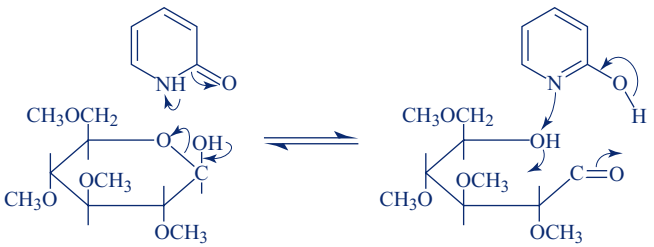


Fig. 11.33 Model of SWAIN and BROWN
((1952) J. Am. Chem. Soc. 74, 2434-2438)

At that time, numerous models of polyfunctional catalysis were studied and compared to corresponding models of monofunctional catalysis. Among them, one can mention the hydrolysis of several aspirin derivatives. The hydrolysis of compound 1 or 1′ (Fig. 11.34) is particularly efficient.

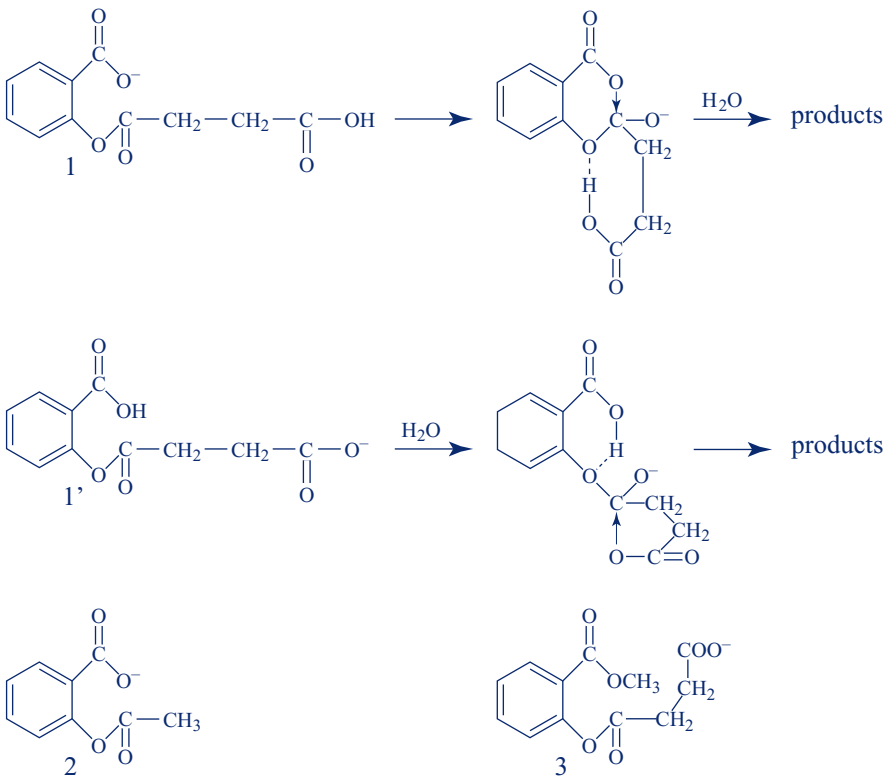


Fig. 11.34 Hydrolysis of aspirin derivatives

The profile of the reaction rate as a function of pH is in agreement with one of the two mechanisms indicated in Fig. 11.34 which implicates a nucleophilic attack on the carbon and an electrophilic attack on the oxygen of the C—O bond in an intramolecular catalysis. The hydrolysis of compound 1 is much faster than that of derivatives 2 and 3 which possess only a nucleophilic centre. The hydrolysis of derivative 1 according to the first mechanism is 24 000 times faster than that of derivative 2, and according to the second mechanism which implicates the ionisation state 1′, it is more efficient than that of derivative 3.

Thus, in both intermolecular catalysis like in the model of SWAIN and BROWN and intramolecular catalysis like in the case cited above, bifunctional catalysis represents a gain in efficiency as compared to monofunctional catalysis. All enzyme reactions involve the participation of at least two catalytic groups: Ser195 and His57 in serine proteases, Asp52 and Glu35 in lysozyme, Cys25 and His159 in papain, His231 and Zn^{++} in thermolysine, Asp32 and Asp215 in pepsin, His12 and His119 in ribonuclease etc. **Enzyme catalysis is therefore intramolecular and polyfunctional.**

11.4.3. COMPLEMENTARITY OF THE ENZYME FOR THE TRANSITION STATE OF THE SUBSTRATE

The "lock and key" mechanism proposed by the first authors insisted on the notion of enzyme complementarity for its substrate. This idea stayed commonly assumed for a long time, the substrates having a good affinity for the enzyme being considered "good substrates"; even today, it is not rare to find statements of this type in the literature. However, in 1930 HALDANE and then in 1940 PAULING suggested that the enzyme-substrate energy association was used to distort the substrate and carry it into a state close to the transition state. The idea of complementarity of the enzyme for the transition state of the substrate was thus developed and has ultimately found experimental support thanks to the conjunction of crystallographic studies, molecular modelling methods, site-directed mutagenesis experiments, and enzymatic and associated chemical studies.

11.4.3.1. ENERGY ASPECTS

The fact that the complementarity of the enzyme for the transition state rather than for the fundamental state of the substrate is a catalytic advantage is thermodynamic evidence. Indeed, the higher the enzyme affinity for its substrate, the more stable the complex ES, and therefore the higher the energy barrier is to attain the active state. This can easily be shown in the case of a simple scheme of an enzyme reaction for a single substrate, a single intermediate complex with the quasi-equilibrium approximation ($K_m = K_s$):

$$E + S \underset{}{\overset{K_m}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} E + P$$

with: $$DG_s = RT \ln K_m$$

and: $$k_{cat} = \frac{kT}{h} e^{-\Delta G^{\neq}/RT}$$

The reach of the transition state is written:

$$E + S \underset{}{\overset{k_{cat}/K_m}{\rightleftharpoons}} ES^*$$

It is represented by the ratio k_{cat}/K_m with the activation energy $DG_T^{\neq}$ such that:

$$DG_T^{\neq} = DG_s + DG^{\neq}$$

Thus: $$RT \ln (k_{cat}/K_m) = RT \ln (kT/h) - DG^{\neq} - DG_s$$

The corresponding diagram is represented in Fig. 11.35.

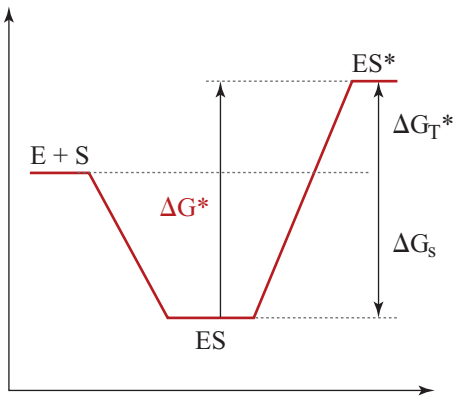


Fig. 11.35 Energy diagram of a simple enzyme reaction
(Reprinted from *Enzyme, Structure and Mechanism*, FERSHT A.N., 323. © (1985) with permission from W.H. Freeman and Company/Worth Publishers)

An illustration of the importance of enzyme complementarity to the transition state of the substrate was presented by FERSHT (1984). The author assumes that the addition of a supplementary group to the substrate increases its energy of association to the enzyme by a value DG_R. He examines the diverse possible situations and their consequences. In conditions where the substrate concentration is saturating ($v = k_{cat}\, e_0$) if the presence of this group stabilises both the MICHAELIS complex and the complex in the transition state, the activation energy remains unchanged (Fig. 11.36-1 opposite). This does not represent any gain for the catalysis. The stabilisation of the MICHAELIS complex alone (Fig. 11.36-2) increases the activation energy; the situation is unfavorable for catalysis. If on the contrary, only the transition complex is stabilised (Fig. 11.36-3), the activation energy is decreased, and the reaction rate increases; such a situation is favorable to catalysis. An analogous demonstration can be made in the case where $s <<< K_m$. The complementarity of the enzyme for the transition state of the substrate implies that the value of k_{cat}/K_m is maximal. If the value of K_m is very weak, the association of the enzyme with the initial state of the substrate being maximal, the complex ES will be stable, and the activation energy will be increased and k_{cat} decreased. If however, the association energy is optimised in the transition state, K_m will increase and the gain in association energy will increase the value of k_{cat}.

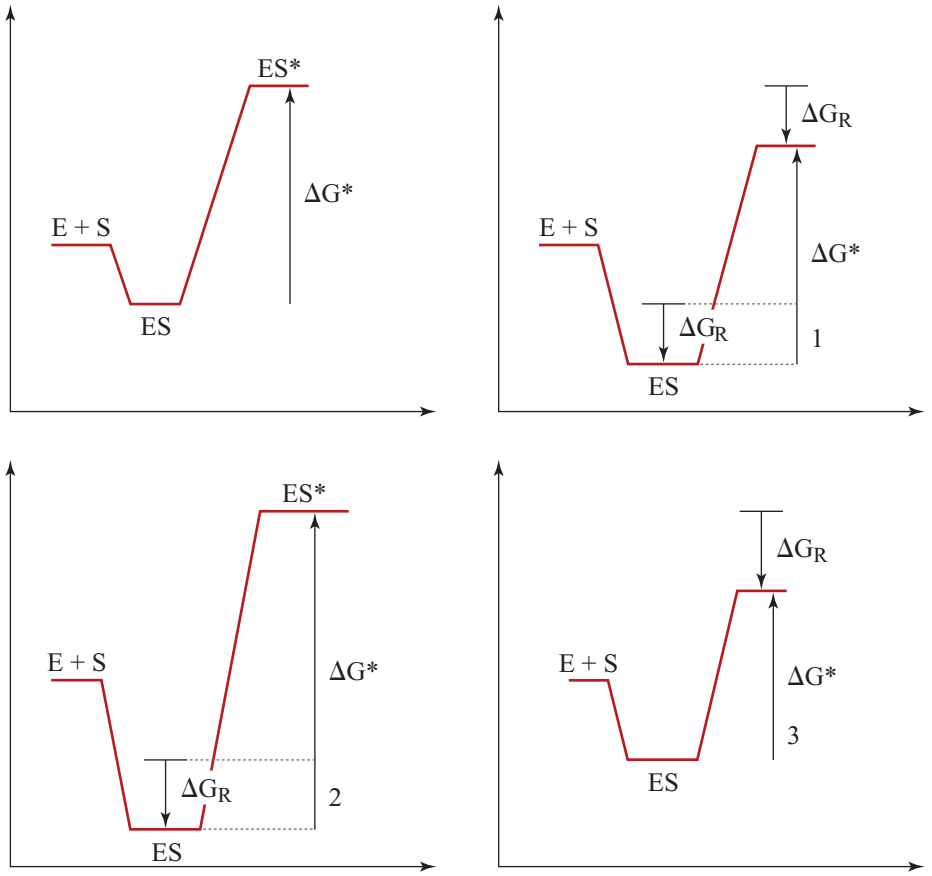


Fig. 11.36 Consequence of the addition of a supplementary group to the substrate on the activation energy

(Reprinted from *Enzyme, Structure and Mechanism*, FERSHT A.N., 323. © (1985) with permission from W.H. Freeman and Company/Worth Publishers)

It is therefore clear in terms of **thermodynamics that one of the major factors of catalytic efficiency is the complementarity of the enzyme for the transition state of the substrate and not for its fundamental state.** This property is also responsible for the specificity. In support of these thermodynamic considerations, there is a great number of experimental arguments which demonstrate that enzymes have a structure complementary to the transition state of specific substrates. **The complementarity of the enzyme for the transition state of its substrate accounts for both specificity and efficiency of enzyme catalysis.**

11.4.3.2. KINETIC PARAMETERS CORRESPONDING TO SEVERAL ENZYME REACTIONS

Many kinetic arguments which are interesting to examine show how enzyme-substrate association energy is used for enzyme catalysis in several systems such as serine proteases, pepsin and β-galactosidase of *E. coli*. In reactions catalysed by

chymotrypsin and elastase, it appears in Table 11.9 that in most cases elongation of the substrate structure which increases the ratio k_{cat}/K_m produces an increase in k_{cat} without noticeably decreasing the value of K_m. In these different cases $K_m = K_s$ since it concerns hydrolysis of amide bonds with limiting acylation. One similar situation is observed for the hydrolysis by pepsin of a series of substrates (Table 11.9).

Hydrolysis by chymotrypsin and trypsin (Table 11.10 opposite) as well as by β-galactosidase (Table 11.11 opposite) of substrates of different specificities shows that variations in the ratio of k_{cat}/K_m result in larger variations in the value of k_{cat} than in the value of K_s.

Table 11.9 Interconversion of association and activation energies

(Reprinted from *Enzyme, Structure and Mechanism*, FERSHT A.N., 300 and 304. © (1985) with permission from W.H. Freeman and Company/Worth Publishers)

Enzyme and substrate	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($s^{-1}.M^{-1}$)
Chymotrypsin (pH 7.9 and 25°C)			
Ac-Tyr-NH_2	0.17	32	5
Ac-Tyr-Gly-NH_2	0.64	23	28
Ac-Tyr-Ala-NH_2	7.5	17	440
Ac-Phe-Gly-NH_2	0.06	31	2
Ac-Phe-Ala-NH_2	0.14	15	10
Elastase (pH 9 and 37°C)			
Cleavage at –NH_2			
Ac-Ala-Pro-Ala-NH_2	0.09	4.2	21
Ac-Pro-Ala-Pro-Ala-NH_2	8.5	3.9	2.2×10^3
Ac-Gly-Pro-Ala-NH_2	0.02	33	0.5
Ac-Pro-Gly-Pro-Ala-NH_2	2.8	43	64
Pepsin (pH 3.5 and 37°C)*			
Phe-Gly	0.5	0.3	1.7×10^3
Z-Phe-Gly	25	0.11	2.2×10^5
Z-Ala-Gly	145	0.25	5.8×10^5
Z-Ala-Ala	282	0.04	7×10^6
Z-Gly-Ala	409	0.11	3.7×10^6
Z-Gly-Ile	13	0.07	1.8×10^5
Z-Gly-Leu	134	0.03	4.2×10^6
Phe-Gly-Gly	6	0.6	1×10^4
Z-Phe-Gly-Gly	127	0.13	9.8×10^5

* *cleavage of the Phe-Phe bond in the peptide A-Phe-Phe-OP4P, A being one of the substrates indicated in the list and OP4P the 3-(4-pyridyl)propyl-1-oxy.*

Table 11.10 Kinetic parameters corresponding to hydrolysis of specific and non-specific substrates by chymotrypsin and trypsin

Enzyme and substrate	k_2 (s^{-1})	k_3 (s^{-1})	K_s (*mM*)	k_{cat}/K_m ($s^{-1}.M^{-1}$)
Chymotrypsin (pH 7.5, 25°C)				
N-Ac-Gly-OMe	0.49	0.14	3.38×10^{3}	0.13
N-Ac-But-OMe	8.8	1.7	417	21
N-Ac-Norval-OMe	35.6	5.93	100	360
N-Ac-Norleu-OMe	103	19	34	3×10^{3}
N-Ac-Phe-OMe	796	111	7.6	1×10^{5}
N-Ac-Tyr-OEt	5×10^{3}	200	17	3×10^{5}
Trypsin				
a) Bz-L-ArgNH_2	0.54	24	2.7×10^{-3}	2×10^{2}
a) To-L-ArgNH_2	0.75	95	7.5×10^{-3}	1×10^{2}
b) Bz-L-ArgOEt	2.10^{4}	24	2.2×10^{-3}	9.2×10^{6}
b) To-L-ArgOMe	4.8×10^{4}	95	7.5×10^{-3}	6.4×10^{6}
c) L-LysOMe	3.4×10^{2}	6.7	1.4×10^{-3}	2.6×10^{4}
c) L-LysOEt	1.9×10^{2}	6.7	1.4×10^{-3}	1.3×10^{4}
b) Ac-L-TyrOEt	36	193	47×10^{-3}	7.6×10^{2}
b) Ac-L-PheOMe	55	173	1.1×10^{-1}	5×10^{2}
b) To-L-OrnOMe	5.4	> 5.4	1.6×10^{-2}	3.4×10^{2}
b) Bz-D-ArgOEt	0.28	0.14	2.2×10^{-3}	1.2×10^{2}
c) L-TyrOMe	0.68	> 0.68	1.8×10^{-1}	3.8
b) Ac-GlyOEt	3.2×10^{-2}	0.6	8.8×10^{-1}	3.6×10^{-2}

a) from J. CHEVALLIER & J. YON; b) from F. SEYDOUX & J. YON; c) from F. SEYDOUX.

Table 11.11 Kinetic parameters corresponding to the hydrolysis of different substrates of β-galactosidase of E. coli
(pH 7.0; 25°C, 10^{-3} M SO_4Mg; 0.146 M ClNa) (from YON, 1976)

Substrate	k_2 (s^{-1})	k_3 (s^{-1})	K_s (*M*)	k_{cat}/K_m ($s^{-1}.M^{-1}$)
b-D-galactosides				
Phenyl	45		10^{-4}	4.5×10^{5}
oNO_2-phenyl	2 100	1 200	3.1×10^{-4}	6.8×10^{6}
mNO_2-phenyl	1 900	1 400	3.6×10^{-4}	5.3×10^{6}
pNO_2-phenyl	90		0.3×10^{-4}	3×10^{6}
oNH_2-phenyl	56		4.5×10^{-4}	1.2×10^{5}
pNH_2-phenyl	90		3.3×10^{-4}	2.7×10^{5}
2-4-diNO_2-phenyl	> 5 000	1 200		7.5×10^{6}

Substrate	k_2 (s^{-1})	k_3 (s^{-1})	K_s (*M*)	k_{cat}/K_m ($s^{-1}.M^{-1}$)
3-5-diNO_2-phenyl	> 5 000	1 135		1.8×10^6
Cinnamyl	285		1.5×10^{-2}	1.9×10^4
Methyl	61	1 200	8×10^{-3}	7.6×10^3
Naphtyl	280	1 135	1.6×10^{-4}	1.75×10^4
a-**L-arabinoside**				
oNO_2-phenyl	44		4.3×10^{-3}	1×10^4
b-**D-fucoside**				
o-NO_2-phenyl	4		3×10^{-3}	1.3×10^3
b-**D-glucoside**				
o-NO_2-phenyl	0.14		6.2×10^{-3}	2.2×10^1

There are three forms of trypsin presenting an enzymatic activity, α-, β- and ψ-trypsin. β-trypsin is formed from a single polypeptide chain; α-trypsin results from the cleavage of this chain between Lys131 and Ser132. ψ-trypsin consists of an additional cleavage between Lys176 and Asp177; its catalytic activity is strongly decreased as is shown in the value of the ratio k_{cat}/K_m (Table 11.12). Here also, variations in the value of K_s remain weak whereas the catalytic constant decreases. This situ-ation appears therefore rather frequently, that the substrate varies or that the enzyme is modified.

Table 11.12 Kinetic parameters corresponding to the hydrolysis of different substrates for the three forms α-, β- and ψ-trypsin

(Reproduced from *Eur. J. Biochem.,* **47**, FOUCAULT G. et al., Comparative Kinetic Properties of alpha, beta and Psi Forms of Trypsin, 295. © (1974) with permission of Blackwell Publishing Ltd)

Enzyme and substrate	k_2 (s^{-1})	k_3 (s^{-1})	K_s (*M*)	k_{cat}/K_m ($s^{-1}.M^{-1}$)
Bz-L-ArgOEt				
α-trypsin	2×10^4	26.6	2.3×10^{-3}	10^7
β-trypsin	2×10^4	25.8	2.3×10^{-3}	10^7
ψ-trypsin	0.45	0.45	1.8×10^{-3}	224
To-L-ArgOMe				
α-trypsin	4×10^{-4}	76	10^{-3}	4.55×10^6
β-trypsin	5×10^{-4}	96	10^{-3}	6×10^6
ψ-trypsin	1.6		3.9×10^{-3}	0.41×10^2
Ac-L-TyrOEt				
α-trypsin	52		6×10^{-2}	770
β-trypsin	56		4.7×10^{-2}	760
ψ-trypsin	0.15		6×10^{-2}	2.3

These examples indicate how the enzyme-substrate association energy is utilised to accelerate the reaction. The situation that corresponds to trypsin hydrolysis of benzoyl-L-arginine ethyl ester and its corresponding D isomer is interesting in this regard. The substrate carrying L-arginine is rapidly hydrolysed ($k_{cat}/K_m = 9.2 \times 10^6$ $s^{-1} . M^{-1}$) whereas the D isomer is hydrolysed very slowly ($k_{cat}/K_m = 1.2 \times 10^2$ $s^{-1} . M^{-1}$). The efficiency ratio between the two substrates approaches 10^5. However the enzyme possesses the same affinity for the two isomers. But as is shown in the energy diagram in Fig. 11.37, one notes a great difference in the transition state of the complex ES and also of the acyl-enzyme for the two substrates. The complementarity of the enzyme for the fundamental state is identical whereas it differs considerably in the transition state (SEYDOUX & YON, 1971).

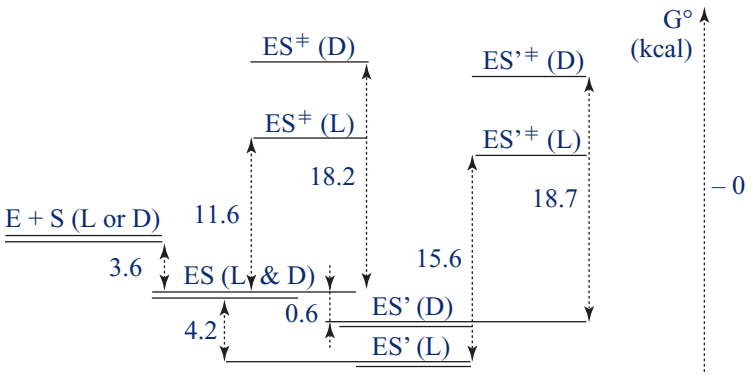


Fig. 11.37 Energy diagram for the trypsin hydrolysis of benzoyl-L-arginine ethyl ester and its corresponding D derivative ◢

11.4.3.3. AFFINITY OF ENZYMES FOR TRANSITION STATE ANALOGS

Another argument is supplied by the great enzyme affinity for compounds analogous to the substrate transition state. This approach had already been suggested by PAULING in 1940. As was shown in Chap. 9, this property was used for chemical labelling of the active centre of enzymes.

Today it is used in pharmacological applications since these compounds are powerful enzyme inhibitors. Some examples are given below.

Inhibitors of glycosidases

Compounds analogous to the transition state were synthesised for several glycosidases; their mechanisms of action were studied. Lysozyme catalyses the hydrolysis of glycosidic bonds in 1 ⟶ 4 of synthetic polymer substrates of N-acetyl glucosamine (NAG) or of N-acetyl muramic acid (NAM). Until recently, one of the intermediates of the hydrolysis reaction was supposed to implicate the formation of a carbocation stabilised by the charge of Asp52; in this intermediate, glycopyranoside shifts from a chair configuration to a half-chair configuration. The mechanism of action of lysozyme will be detailed in Chap. 12. Compound 1 obtained by

synthesis and which possesses a lactone group in position 4 resembles transition state 2 of the substrate (Fig. 11.38). It is associated with the enzyme with a dissociation constant $K_d = 8.3 \times 10^{-8}$ M, whereas for the substrate $K_s = 10^{-5}$ M. The affinity of the enzyme for the analog of the transition state is therefore around 100 times higher than for the substrate.

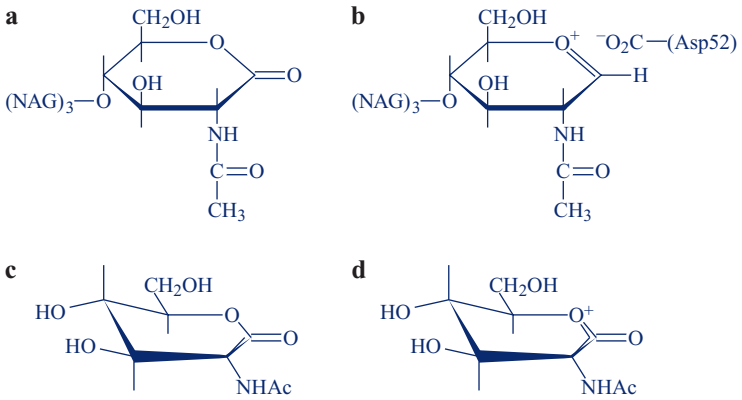


Fig. 11.38 Substrate transition state analogs, inhibitors of several glucosidases
*(**a**) lactone derivative inhibitor of lysozyme − (**b**) transition state of the substrate of lysozyme − (**c**) lactone derivative inhibitor of β-N-acetyl glucosaminidase − (**d**) transition state of the substrate of β-N-acetyl glucosaminidase*

For another glycosidase, β-N-acetyl D-glucosaminidase, the affinity is around 4 000 times greater for the lactone (Fig. 11.38c) analog of the transition state (Fig. 11.38d) than for the substrate. β-galactosidase of *E. coli* hydrolyses the β-D-galactoside bonds in 1 ⟶ 4. Its affinity for the substrates is reflected by the values of K_s given in Table 11.11 of the preceding paragraph. An analogue of the transition state, D-galactal, has been synthesised; it possesses a very high affinity for the enzyme ($K_d = 5 \times 10^{-6}$ M) whereas the values of K_s for the galactoside substrates are situated around 10^{-4} M.

Inhibitors of proline racemase

Racemisation of proline is carried out according to a mechanism that was previously described. Over the course of the reaction symmetric carbon becomes trigonal. On this basis, transition state analog compounds were synthesised (compounds b and c of Fig. 11.39 opposite). They present an affinity for the enzyme 100 times greater than proline.

Inhibitors of cytidine deaminase

The affinity of the enzyme for tetrahydrouridine (Fig. 11.39d) is 10 000 times greater than for uridine (Fig. 11.39e) and ammonia which are the products of the reaction. The transition state of the substrate is thought to be a tetrahedral intermediate (Fig. 11.39f).

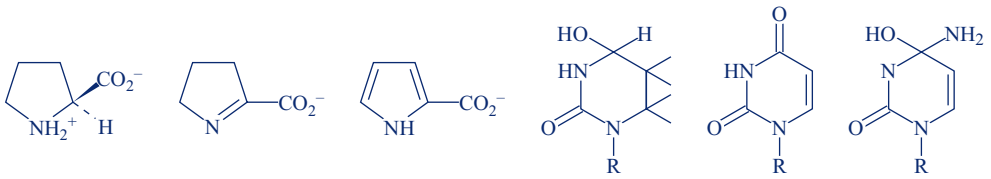


Fig. 11.39 Inhibitors of proline racemase and of cytidine deaminase
*(**a**) proline – (**b** and **c**) analogs of the transition state of proline, inhibitors of proline racemase – (**d**) tetrahydrouridine – (**e**) uridine – (**f**) transition intermediate*

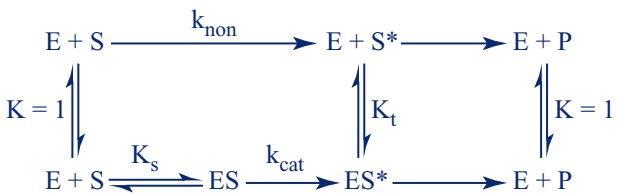
Inhibitor of aspartate transcarbamylase

N-(phosphonacetyl)-L-aspartate (PALA), an analog of the two substrates of aspartate transcarbamylase was synthesised by COLLINS and STARK (1969). This compound is a powerful inhibitor of the enzyme in all organisms examined so far. The value of the inhibition constant K_i is very weak in relation to K_m of these enzymes with respect to their substrates. PALA was the object of detailed investigations for inhibiting the proliferation of tumor cells and was used in the clinic.

In general the affinity of enzymes for analog compounds of the transition state is on the order of 100 to 1 000 times higher than for their specific substrates. These compounds resemble the presumed transition state; it is likely that the affinity of enzymes for the true transition state of substrates is even higher.

11.4.3.4. ESTIMATION OF MINIMAL AFFINITIES OF ENZYMES FOR TRANSITION STATES OF SUBSTRATES

An estimation of an enzyme affinity for the transition state of its substrate was proposed from the comparison of a non-catalytic reaction to the corresponding enzymatic reaction (CANNON et al., 1998; CANNON & BERKOVIC, 1998). This was justified by the fact that the equilibrium constant between the substrate and the product does not include any direct contribution from the enzyme concentration. The scheme below compares the transformation of a substrate and product *via* its transition state S* for the non-catalytic reaction and the catalytic reaction; k_{non} represents the unimolecular rate constant of the non-enzymatic reaction:



Assuming that the catalytic efficiency of an enzyme results only from its complementarity for the transition state of the substrate, it is possible to deduce a hypothetical equilibrium constant, K_T, between the enzyme and the transition state of the substrate in the simple case where $K_m = K_s$. One has:

$$K_T = \frac{k_{cat} / K_m}{k_{non}}$$

Table 11.13 gives an idea of the hypothetical dissociation constant of the complex ES*, K_{Tdiss}, for several enzymatic systems.

Table 11.13 Minimal enzyme affinities for the transition state of the substrate

Enzyme	k_{non} (s^{-1})	k_{cat}/K_m ($M^{-1}.s^{-1}$)	K_{Tdiss} (M)
Alcaline phosphatase	1.0×10^{-15}	5.7×10^{3}	2×10^{-19}
Acetylcholinesterase	1.1×10^{-8}	1.6×10^{8}	7×10^{-17}
Adenosine deaminase	1.8×10^{-10}	1.4×10^{7}	8×10^{-16}
Cytidine deaminase	3.2×10^{-10}	7.5×10^{5}	4×10^{-16}
Urease	3.0×10^{-10}	2.3×10^{5}	1×10^{-15}
Triose phosphate isomerase	4.3×10^{-10}	9.1×10^{6}	5×10^{-13}
Chorismate mutase	2.6×10^{-15}	1.1×10^{6}	2×10^{-11}
Carbonic anhydrase	3.7×10^{-2}	1.2×10^{8}	3×10^{-11}

These values indicate much higher affinities than those observed for transition state analog inhibitors.

11.4.3.5. STRUCTURAL ARGUMENTS

The theory of enzyme complementarity for substrate transition states receives today convincing confirmation due to structural arguments and the possibility of visualising this property. The convergence of crystallographic studies, molecular modelling, site-directed mutagenesis experiments that permit substituting one amino acid for another in pre-determined positions and the analysis of kinetic and thermodynamic parameters of the enzymatic reaction in these diverse conditions, brings the proof of this important characteristic of enzymatic catalysis.

Several enzymatic systems were and still are the object of such an approach. One of the first most careful analyses was carried out by the group of FERSHT at the Imperial College of London on tyrosyl-tRNA synthetase. As was described in this chapter, this enzyme catalyses the aminoacylation of tRNA, a reaction proceeding in two steps, the first being the formation of tyrosyl adenylate:

$$E + Tyr + ATP \longleftrightarrow E\text{-}Tyr\text{-}AMP + PP_i$$

The three-dimensional structure of the enzyme of *B. stearothermophilus* was solved by X-ray diffraction, as well as that of the complex formed by the enzyme and tyrosyl adenylate. Hydrogen bonds which can be established between the enzyme and the substrate were determined on these bases (Fig. 11.40 opposite). Tyrosine, by its OH group, contracts two hydrogen bonds, one with the side chain of Asp176, the other with that of Tyr34 ; by its amino group, it participates in three other hydro-

gen bonds with the groups Asp38, Asp178 and Tyr34. The C═O group forms a hydrogen bond with the amide group of Gln195. AMP interacts with the main chain of Arg381 by the phosphate group; the ribose contracts hydrogen bonds with Cys35, the main chain (C═O) of Gly38 and that (NH) of Gln192 as well as with the side chain of Thr51; adenine interacts with His48.

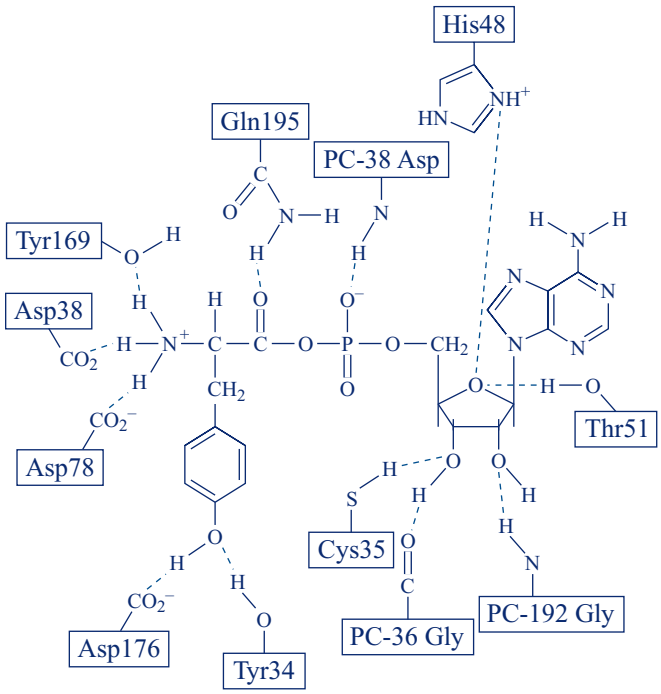


Fig. 11.40 Presumed hydrogen bonds between tyrosyl-tRNA synthetase and tyrosyl-adenylate according to crystallographic studies
(PC: polypeptide chain)
(Reprinted from *Enzyme, Structure and Mechanism*, FERSHT A.N., 323. © (1985) with permission from W.H. Freeman and Company/Worth Publishers)

From these data, different mutants were prepared. Kinetic parameters corresponding to the reaction catalysed by these mutants were determined and analysed. The results (Table 11.14 below) led to the evaluation of the contribution to the enzyme-substrate interaction energy of the different hydrogen bonds. In a rather general manner, those that use a charged group of the enzyme or substrate contribute more to the binding energy of the substrate (3.5 to 4.5 kcal . mol^{-1}) than those that form between non-charged groups.

Moreover, kinetic studies showed that mutants His ⟶ Gly48 and Cys ⟶ Gly35 have weaker k_{cat} values and slightly higher K_m values than the wild type enzyme. The residue His48 interacts with ATP, but this interaction is stronger when Thr51 is replaced by a proline. The substitution of His45 by Asn45 practically does not change K_m; however it considerably decreases k_{cat}. This residue probably does not

interact with tyrosyl adenylate but is very likely in a position suitable for interacting with the phosphate group of ATP.

Table 11.14 Activation of tyrosine by tyrosyl-tRNA synthetase and its mutants
(Reprinted from *Enzyme, Structure and Mechanism*, FERSHT A.N., 300 and 304. © (1985) with permission from W.H. Freeman and Company/Worth Publishers)

Enzyme	k_{cat} (s^{-1})	K_{mATP} (mM)	K_{mTyr} (μM)	$[k_{cat}/K_m]_{ATP}$ ($s^{-1}.M^{-1}$)	$[k_{cat}/K_m]_{Tyr}$ ($s^{-1}.M^{-1}$)
Wild type	8.35	1.08	2.23	7 730	3.74×10^{6}
Tyr ⟶ Phe34	6.86	1.2	4.4	5 720	1.56×10^{6}
Cys ⟶ Gly35	2.95	2.6	2.7	1 130	1.09×10^{6}
Cys ⟶ Ser35	2.52	2.4	2.6	1 050	9.65×10^{5}
His ⟶ Asn48	7.90	1.4	3.8	5 640	2.08×10^{6}
His ⟶ Gly48	2.00	1.3	3.2	1 540	6.25×10^{5}
Thr ⟶ Ala51	8.75	0.54	2.0	16 200	4.36×10^{6}
Gln ⟶ Gly195	0.19	2.5	100	76	1.90×10^{3}
His ⟶ Asn45	0.0031	1		3	
Thr ⟶ Pro51	12.4	0.058	1.7	213 790	7.29×10^{6}

The mechanism driving the formation of tyrosyl adenylate goes through an inversion of α phosphate. Since no covalent intermediate is formed over the course of the reaction, there occurs an in-line attack of the phosphorous by the carboxylate of tyrosine that plays the role of the nucleophile, and the pyrophosphate is released (Fig. 11.41).

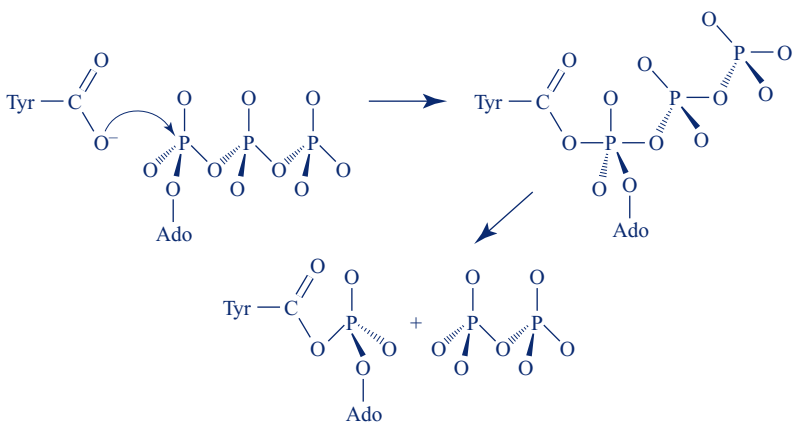


Fig. 11.41 Mechanism of formation of tyrosyl adenylate

This implies that in the transition state, the phosphorous is pentacoordinated. From data concerning on the one hand the presumed structure of the transition state, and on the other hand the structure of tyrosyl adenylate, FERSHT and colleagues used molecular modelling to construct a model permitting the localisation of the γ phosphate of ATP (Fig. 11.42 opposite). They deduced that this group contracts a hydro-

gen bond with the side chains of His45, in agreement with the preceding kinetic results, and of Thr40. On these bases, the authors constructed several mutants by site-directed mutagenesis; in addition to the mutant His ⟶ Gly45, they obtained the mutant Thr ⟶ Ala40 and the double mutant His ⟶ Gly45, Thr ⟶ Ala40. Kinetic studies in pre-steady state conditions clearly show that the modification of these groups contributes little to the binding energy of ATP, but is involved in the stabilisation of the transition state (Table 11.15). It follows that for the double mutant, the rate of formation of tyrosyl adenylate is decreased by a factor of 3×10^5, whereas the value of K_s for ATP is only affected by a factor of 5. Figure 11.43 below illustrates the catalytic mechanism at the active centre of the enzyme.

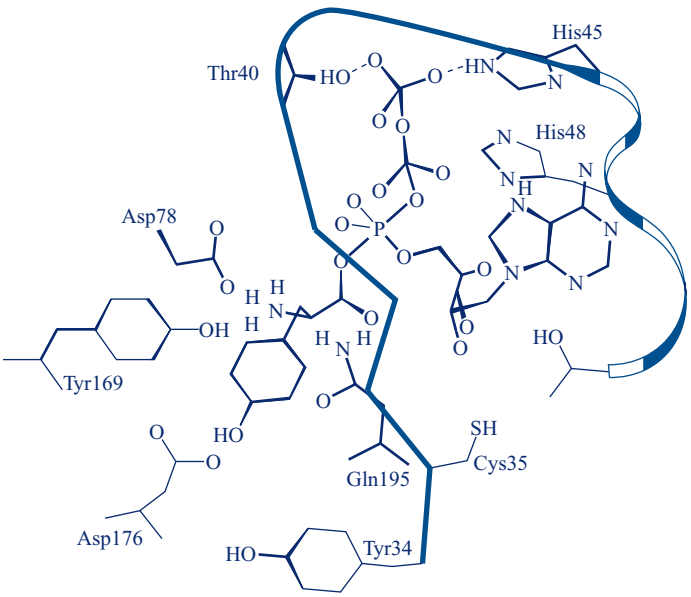


Fig. 11.42 Modelling of the transition state for the formation of tyrosyl adenylate
(Reprinted from *PNAS*, **82**, LEATHERBARROW R.J. *et al.*, Transition-state stabilization in the mechanism of tyrosyl-tRNA synthetase revealed by protein engineering, 7840, 1985, with authors' permission)

Table 11.15 Kinetic parameters for the formation of tyrosyl adenylate determined at the pre-steady state (Reprinted from *PNAS*, **82**, LEATHERBARROW R.J. et al., Transition-state stabilization in the mechanism of tyrosyl-tRNA synthetase revealed by protein engineering, 7840, 1985, with authors' permission)

Enzyme	k_3 (s^{-1})	K_{sTyr} (μM)	K_{sATP} (mM)
Wild type	38	12	4.7
His ⟶ Gly45	0.016	10	1.2
Thr ⟶ Ala40	0.0055	8.0	3.8
Double mutant			
His ⟶ Gly45 Thr ⟶ Ala40	0.00012	4.5	1.1

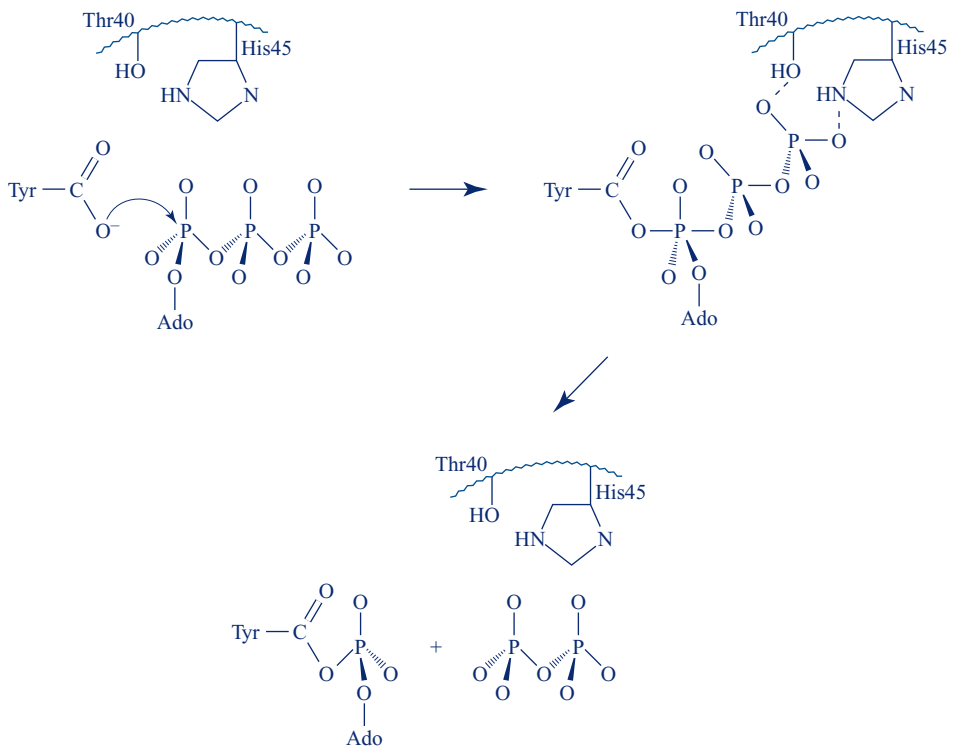


Fig. 11.43 Mechanism of formation of tyrosyl adenylate catalysed by tyrosyl-tRNA synthetase. It underlines the interactions of γ phosphorus with Thr40 and His45

Studies of this type were developed on other enzymatic systems. It is interesting to cite, for example, a slightly different approach used on alcohol dehydrogenase. The transition state corresponding to the transfer of hydride in the reaction catalysed by this enzyme was predicted theoretically by means of quantum mechanics. It was then positioned by modelling the active centre of the enzyme where it was perfectly adjusted. The stabilisation energy by the protein was calculated. The transition state is found stabilised whereas the substrate and the products of the reaction are destabilised by the geometrical constraints of the enzyme. This study will be detailed in Chap. 12 where the mechanism of action of alcohol dehydrogenase will be presented.

The ensemble of these results shows that **enzymes have evolved to bind to the transition state of their specific substrate much more strongly than to their fundamental state in order to make the catalysis more efficient.** Most values of the dissociation constants, K_s, are indeed on the order of 10^{-4} to 10^{-3} M, sometimes even greater. This particularity of enzyme catalysis that had been observed fifty years ago is now generally accepted thanks to structural studies that permit visualisation and specification of the roles of different interactions.

11.4.3.6. ONE APPLICATION OF THIS PARTICULARITY: ABZYMES

At this stage it is interesting to mention an application that utilises the great affinity of an enzyme for the transition state of its specific substrate(s). This property drove the conception of catalysts that are specific antibodies obtained by using as haptens transition state analogs of the substrate in known enzymatic systems. These "abzymes", whose name results from condensing the terms "antibody" (Ab) and "enzyme", present significant catalytic activities and great specificity. At the beginning, two groups of researchers, one at the Scripps Clinic (TRAMONTANO et al., 1986) and the other at the University of Berkeley (POLLACK et al., 1986), succeeded independently in selecting antibodies endowed with a catalytic activity.

The models chosen were carried out on the hydrolysis of carboxylic esters that are substrates of hydrolytic enzymes like esterases, serine proteases or metallopeptidases. The catalytic mechanism involves the formation of a tetrahedral addition complex in the transition state; the transition state analog compounds bind to the enzyme with a very high affinity and behave as inhibitors. Metallopeptidases, for example, are inhibited by substrate analogs having a phosphoryl or tetrahedral phosphonyl group (Fig. 11.44b) in place of the amide bond which is broken in the enzymatic reaction (Fig. 11.44a). Diverse phosphonate esters were synthesised by the group of TRAMONTANO (1986) and used as haptens for the fabrication of antibodies. Among the different compounds tested, one gave antibodies endowed with catalytic activity to hydrolyse corresponding carboxylic esters. The results showed a very strict specificity of the abzyme thus prepared and an activity significant in experimental conditions although weaker than that of corresponding enzymes.

Fig. 11.44 Abzymes capable of hydrolysing carboxylic esters
*(**a**) transition state of a metallopeptidase substrate*
*(**b**) transition state analog*

POLLACK et al. (1986) developed a similar model corresponding to the hydrolysis of p-nitrophenyl aminoethyl carbonate (Fig. 11.45a below) which is an analog of substrates of acetylcholinesterase. In order to show the principle of abzyme preparation, some details on this study are given. The hydrolysis proceeds according to the mechanism indicated in Fig. 11.45.

The analog of the transition state is p-nitrophenyl phosphoryl choline; it was used to produce monoclonal antibodies. The antibody MOPC167 hydrolyses this compound according to michaelien kinetics with a catalytic constant of 0.4 min^{-1} and a K_m value equal to 208 mM. Compound 2 is not hydrolysed, and compound 3 is only weakly hydrolysed. Compound 4, which has an ethyl group substituted for a methyl in the quaternary ammonium, is hydrolysed with a better catalytic constant whereas the value of K_m is increased (Table 11.16 below). The value of the ratio k_{cat}/K_m is slightly higher than that corresponding to the hydrolysis of compound 1.

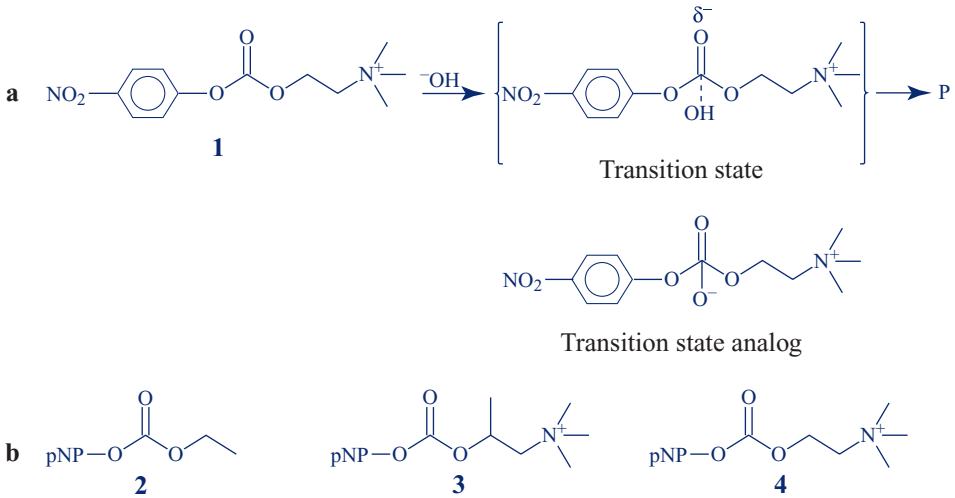


Fig. 11.45 Abzymes capable of hydrolysing p-nitrophenyl N-trimethyl aminoethyl carbonate

*(**a**) mechanism of hydrolysis of the substrate until the appearance of the reaction products P – (**b**) p-nitrophenyl carbonates used to determine abzyme specificity*

(From *Science*, **234**, POLLACK S.J. et al., Selective chemical catalysis by an antibody, 1570. © (1986) reprinted with permission from American Association for the Advancement of Sciences)

Table 11.16 Value of catalytic parameters of the abzyme MOPC16

(From *Science*, **234**, POLLACK S.J. et al., Selective chemical catalysis by an antibody, 1570. © (1986) reprinted with permission from American Association for the Advancement of Sciences)

Compound	k_{cat} *(min*$^{-1}$*)*	K_m	k_{cat}/K_m *(s*$^{-1}$*. M*$^{-1}$*)*
1	0.4 ± 0.04	208 mM	32
2	0		
3	0.2 ± 0.04	2 mM	1.6
4	1.7 ± 0.15	650 mM	43.6

These results clearly show the high specificity of abzymes. In the cases studied however, the catalytic activity remains weak compared to that of enzymes although the carboxylic ester bonds are relatively labile. The obtention of these first abzymes has initiated numerous studies; this domain of research was widely developed. Abzymes susceptible to selectively hydrolysing peptide bonds or catalysing reactions for which no enzyme was known were made. This approach presents a double interest. On the one hand in the fundamental framework, it underlines the importance of the complementarity of enzymes for the transition state of their substrate because that provides a specific catalysis. On the other hand, it offers interesting perspectives for the production of enzymes more specific than natural ones and for the fabrication of new types of enzymes. ◢

11.4.4. EFFECTS OF MICROENVIRONMENT

11.4.4.1. ELECTROSTATIC EFFECTS

Proteins possess numerous charges as well as dipoles that create electrostatic fields of which the intensity is highly variable according to the different protein sites. In addition, these charges fluctuate as a function of diverse motions that animate the molecule. The influence of charges at the active centre of an enzyme and their role in catalysis have been recognised for a long time. The first attempts to explain electrostatic effects in enzyme catalysis were made in 1938 with the approach of STEARN, which today can seem very naïve but at the time represented an original approach. STEARN calculated by means of quantum mechanics the activation energy required to hydrolyse a peptide bond. He took as a model an ideal peptide bond, considering only the atoms C, O, N, and H in vacuum and completely isolated in this context. He attempted moreover to evaluate the effect resulting from the approach of a particular enzyme dipole on this energy. He only considered the reacting groups of atoms and a fixed dipole of the protein. He thus evaluated at 17 kcal . mol^{-1} the decrease in the activation barrier produced by a C—O dipole placed in an adequate position at 5 Å from the centre of the group of reacting atoms. This approach represents the first attempt to explain the role of electrostatics in enzyme catalysis. It was very simple and very static but original at that time.

Since then, many groups of theoreticians have proposed methods to evaluate the electrostatic effects in enzyme catalysis. Diverse approaches of enzymatic reactivity using quantum mechanics were tried (LOEW & THOMAS, 1972; UEYEMA et al., 1973; SHEINER et al., 1975). However, all these treatments were too simplified and considered the reaction in vacuum. The heterogeneity of the protein milieu and the difficulty in evaluating the local dielectric constant made it difficult to evaluate these electrostatic contributions. Nevertheless, more and more precise knowledge of the three-dimensional structure of enzymes on the one hand, and the performance of calculation methods on the other hand with the increase in the power of computers, gave a renewed interest in this type of study. Currently more and more attempts try to predict the enzymatic reactivity and to specify the reaction mechanisms of enzymes by taking into account electrostatic effects.

In 1976, WARSHEL and LEVITT proposed an approach that took into account the solvent molecules surrounding the enzyme-substrate complex. The chosen system was lysozyme. In the method used, the authors considered the entire enzyme-substrate complex. They calculated by quantum mechanics the energy and the charge distribution of all the atoms that are directly implicated in the reaction. The surface of the potential energy of the rest of the system including the water molecules was evaluated by classical methods. The dielectric model was based on a direct calculation of the electrostatic field due to dipoles induced by atoms of the protein and to dipoles created by the water molecules oriented towards the vicinity of the protein (Fig. 11.46 below). Particular attention was given to the stabilisation of the carbocation. The results of the calculation showed that the carbonium ion that was

thought to form in the transition state has an energy decreased by 9 kcal.mol^{-1} by the presence of the charge on Asp52 as compared to the fundamental state. The decrease in the energy barrier corresponds to an acceleration of the reaction rate by a factor of 3×10^6. However, if the calculated differences are reasonable, the method overestimates the energy values. Yet, as we will see in the following chapter, recent experimental data shows that lysozyme proceeds by another mechanism using a covalent intermediate. On the basis of theoretical data, LEVITT and WARSHEL proposed to describe the mechanism of action of lysozyme by evoking "electrostatic constraint" rather than steric constraint, which remains valid. Under another formulation and without having carried out a theoretical study, an analogous conclusion had been formulated by JENCKS (1975).

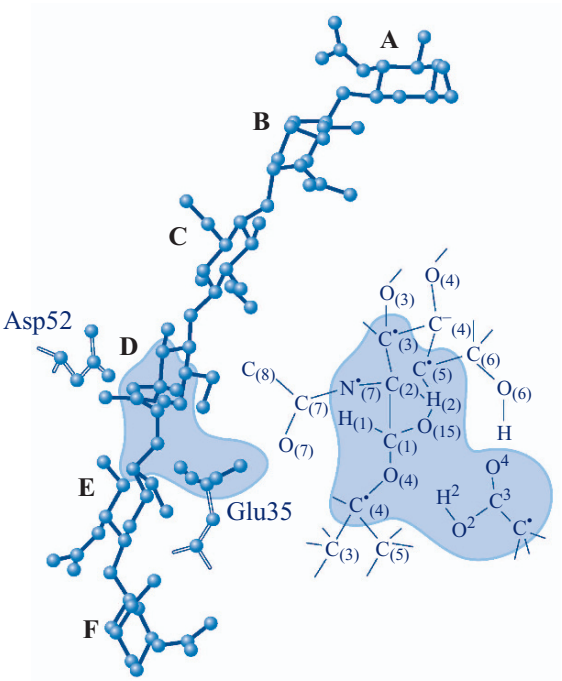


Fig. 11.46 Theoretical approach of catalysis by lysozyme
The shaded grey *regions correspond to the parts treated by quantum mechanics* (Reprinted from *J. Mol. Biol.*, **103**, WARSHEL A. & LEVITT M., Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme, 227. © (1976) with permission from Elsevier)

◤ The difficulty of the problem to treat reactions in solution resides in a correct estimation of solvatation energies that cannot be obtained from macroscopic models. New microscopic approaches are indispensable to include solvent effects. Another method was proposed by WARSHEL and WEISS (1980), WARSHEL (1980, 1981a) using the empirical method of empirical valence bond (EVB) which is a modification of the valence bond (VB) method and which had been forsaken after the development of molecular orbital methods. By this approach, the authors compare the surface potentials of the reaction in solution and at the active centre of an enzyme. The process consists of evaluating the ion resonance forms of the covalent bond by adding the energy of solvatation to determine the surface potential of the reaction. This latter is calibrated using experimental data from the reaction in aqueous solution. The calibrated surface potential serves to evaluate the activation energy of the reaction in solution and to compare it to that of the enzymatic reaction which is evaluated by

replacing the solvatation of ion resonance forms by their interactions with the active site of the enzyme. Thus, the energy of the reaction catalysed by the protein is expressed by the relationship:

$$E_2^p = E_2^s + (G_{sol}^p - G_{sol}^s)$$

E_2^s being the energy of the reaction in solution, G_{sol}^p and G_{sol}^s, the energy of solvatation in the protein and in solution, respectively.

Several enzymatic processes were treated by this approach, the ionisation of acid groups of active sites of enzymes, the stabilisation of ion pairs in enzymes and in solution, proton transfer reactions and reactions implicating a general acid catalysis. Taking into account only permanent dipoles can lead to significant errors. It is important therefore to take together into account charges, permanent dipoles and induced dipoles. The results of these calculations drive some authors to conclude that enzymes can be considered as "supersolvents" that stabilise ion transition states more efficiently than they do in aqueous solution.

TAPIA proposed a theoretical approach based on the same principle as that of LEVITT and WARSHEL. He used quantum mechanics for the group of reacting atoms and the classical model to evaluate the electrostatic field. This approach differs somewhat from the preceding by the calculation methods; in particular TAPIA used the quantum theory ab initio. By this approach, he presented a molecular and electron description of factors intervening in enzyme catalysis and applied it to chymotrypsin and alcohol dehydrogenase (TAPIA et al., 1982; TAPIA, 1987, TAPIA & EKLUND, 1986). These studies based on precise structural data involve a simulation of the structure of water by the Monte Carlo method, a simulation of fluctuations of the active site by molecular dynamics methods, and quantum mechanics being used for the reaction itself. Similar studies were carried out by KARPLUS and colleagues (1991, 2001) to elucidate the catalytic mechanism of triose phosphate isomerase (see the following chapter). These methods which associate quantum mechanics and molecular mechanics, referred to as QM/MM for quantum mechanics/molecular mechanics, are currently highly utilised.

If the theoretical methods involve even more approximations, the ensemble of these studies is worth underlining for the important role of electrostatic effects in enzyme catalysis. The protein matrix, by the presence of charges, dipoles, induced dipoles, mutual polarisation of groups intervening in the reaction and groups of the protein create an electrostatic microenvironment at the active site permitting the stabilisation of charges in the transition state of the substrate. The structure of the protein, for example the presence of α helices or β sheets appropriately oriented in the vicinity of the active site is at the origin of important electrostatic fields (HOL, 1985). The calculation procedures applied to the study of enzyme catalysis were developed and benefited from the increase in the power of computers. Molecular dynamics methods that take into account fluctuations are applied to this type of study.

WARSHEL (1998) attributes a prominent role to electrostatic effects in enzymatic catalysis. From the ensemble of his work, he concludes that the activation energy, $\Delta G^{\neq}$, results from the electrostatic stabilisation of the transition intermediate which is due to a polar environment pre-organised in the structure of the enzyme. He considers the energy of enzyme refolding to be used to pre-orient dipoles at the active

site. His reasoning relies on the fact that mutations that increase $\Delta G^{\neq}$ also increase the energy of refolding. ◢

Thus, electrostatic fields created by the protein at the active centre are complementary to the charge distribution of the substrate in the transition state.

11.4.4.2. ROLE OF WATER MOLECULES

In solution, the effects of the solvent and the microscopic effects due to water molecules at the active centre of enzymes have an influence on the reaction rate. Displacement of water molecules around charges of the substrate or the enzyme upon their association increases the local electrostatic effects. Some model reactions provide examples of acceleration of the reaction rate by desolvatation. Thus decarboxylation of a pyruvate addition compound and of a thiamine pyrophosphate analog (see Fig. 11.6) is carried out with a rate 10^4 to 10^5 times greater in ethanol than in water. The increase in rate is brought on by a smaller charge localisation in the transition state in the presence of ethanol, whereas it is stabilised in water.

Structural studies revealed for some enzymes the presence of water molecules at the active centre. In carboxypeptidase A, the replacement of a water molecule by the substrate decreases the dielectric constant in the vicinity of Zn^{++} and increases its potential to polarise the acyl group of the substrate for a nucleophilic attack (LIPSCOMB et al., 1969). In alcohol dehydrogenase, a water molecule is bound by coordination with the catalytic Zn^{++}. The substrate (alcohol) replaces the water molecule when it associates with the enzyme. In lysozyme, the carboxylate of Asp52 is bound to two water molecules that must be displaced by association with the saccharide ring at site D; it has been suggested that the presence of these water molecules is responsible for the weak binding of the substrate to this site. In chymotrypsin and homologous serine proteases, the water molecule that reacts in the deacylation step was localised by crystallography. These diverse aspects will be illustrated and detailed in Chap. 12.

Molecular dynamics simulations based on structural data were carried out on the hydrated active site of ribonuclease A and on enzyme complexes with a dinucleotide substrate such as CpA as well as an analog of the transition state uridine vanadate (BRÜNGER et al., 1985). A striking result from this study concerns the role of water molecules. In the free enzyme a network of water molecules stabilises the positive charges of residues of the active site that include Lys7, 41, and 66, Arg39 and doubly protonated His119. The bridge of water molecules between these charges stabilises an unfavorable energy conformation. In addition, water molecules form hydrogen bonds with protein groups that interact with the substrate and are displaced upon formation of the MICHAELIS complex. In the free enzyme the catalytic residues His12 and His119 are separated by three water molecules. The authors evaluate up to 23 the number of water molecules thus evacuated by the binding of CpA. Elsewhere, between His119 which interacts with the oxygen $O^{5'}$ of the sugar

of uridine vanadate and the vanadate ion, there are water molecules implicated in general base catalysis during the hydrolysis step.

Water molecules interfere therefore in different ways in enzyme catalysis. Stabilising charges at the active centre of the free enzyme, they make the formation of the complex ES more difficult; these water molecules must be evacuated during the formation of the transition intermediate. The desolvatation upon formation of the enzyme-substrate complex would have the effect of an electrostatic destabilisation (JENCKS, 1975). DEWAR and STORCH (1985) suggested that enzymes operate on mechanisms of desolvatation. However, WARSHEL et al. (1989) think that the effect is described more correctly as substitution of the solvent by the protein. For hydrolytic enzymes, precisely localised water molecules intervene directly in the hydrolysis steps that implicate a general base catalysis.

11.4.4.3. ROLE OF THE HYDROPHOBIC ENVIRONMENT

The microenvironment of the active centre of an enzyme can involve a high charge density, but also regions of hydrophobic character. Following the nature of the reaction, it can be facilitated in a hydrophobic medium as was previously underlined (see Sect. 11.4.4.2 and Chap. 9). However, the current data insist more on the induced electrostatic effects than on effects of the local dielectric constant that is very difficult to evaluate.

11.4.4.4. LOW BARRIER HYDROGEN BONDS

In 1993 and 1994, it was proposed for the first time that low barrier hydrogen bonds (LBHB) would play a role in enzyme catalysis. This hypothesis is strongly supported by CLELAND et al. (1998). The force of a hydrogen bond depends on its length, its linearity, the nature of the microenvironment and the degree of proximity of pK values between the donor and the acceptor groups. In water, hydrogen bonds have a length of 2.8 Å, and the enthalpy of formation DH is –5 kcal . mol^{-1}; they are weak. In gaseous phase, hydrogen bonds between heteroatoms having close pK values can be short and very strong with values of DH on the order of –25 to –30 kcal . mol^{-1}. The distance O—O in H—O···H—O—H is 2.28 Å. In organic solvents, the value of DH can reach –20 kcal . mol^{-1}. LBHB have other properties than a weak distance between heteroatoms; the chemical displacement in NMR is situated near the lower fields. In enzymatic catalysis, LBHB would operate by the conversion of weak hydrogen bonds between the enzyme and the substrate in the fundamental state to LBHB in the transition state. The enzyme would convert a weak hydrogen bond in LBHB by changing the pK value of the substrate that would become close to that of the enzyme residue with which it contracts the hydrogen bond. This would accelerate the reaction by at least five orders of magnitude. The authors illustrate their concept by several examples among which are lactate dehydrogenase, serine proteases, ketosteroid isomerase, triose phosphate isomerase, citrate synthase and mandelate racemase.

This hypothesis was refuted by WARSHEL (1998). This author considers LBHB to be effective only in a non-polar environment, and he considers experiments achieving a distinction between hydrogen bonds and LBHB unable to be interpreted without ambiguity and signalling several contradictions. Some NMR studies, in particular those based on previous authors' data, were interpreted differently. It seems therefore, without any irrefutable proof, that this interpretation is far from meeting a consensus.

11.4.5. REACTION INTERMEDIATES IN ENZYME CATALYSIS

Most enzymatic reactions proceed *via* the formation of several intermediate complexes, implicating either enzyme modifications, substrate modifications, or both (see Chap. 5). The reaction schemes are:

$$E + S \rightleftharpoons ES_1 \rightleftharpoons ES_2 \rightleftharpoons \cdots \rightleftharpoons ES^* \rightleftharpoons E + P$$

with sequential formation of different intermediates which, according to the reactions and the associated rate constants, are found more or less occupied and more or less stable. In a series of sequential reactions, the binding energy is used at each step, but the efficiency of the passage:

$$E + S \longrightarrow ES^*$$

remains unchanged since: $k_{cat}/K_m = k_2/K_s$

Some of these intermediates are unstable and therefore do not accumulate, in particular those that are near the transition state. For several enzymes, the reaction proceeds *via* a stable chemical intermediate (Table 11.17). These intermediates generally do not accumulate during the reaction with the physiological substrate in vivo. They were evidenced in vitro with synthetic substrates or in particular experimental conditions. For instance, the reaction catalysed by glyceraldehyde-3-phosphate dehydrogenase involves the formation of a covalent intermediate acyl-enzyme (thioester). However, in vivo the concentration of glyceraldehyde-3-phosphate has a weak K_m ($K_m/s = 15$ in brain, 23 in muscle) such that the intermediate does not accumulate in the normal functioning conditions of the organism. Conversely, proteases implicated in diges-tion are often found in the presence of strong substrate concentrations; in certain cases, however, activation phenomena by the substrate were described even for the hydrolysis of proteins, which can limit the accumulation of intermediates (MONNOT et al., 1966).

Table 11.17 Enzymes and intermediates

Enzyme	*Substrate*	*Intermediate*	*Accumulation*
Chymotrypsin (serine proteases)	Peptides	Acyl-enzyme	–
Carboxypeptidase	Peptides	?	
Papain (thiol proteases)	Amides	Acyl-enzyme	–

Enzyme	*Substrate*	*Intermediate*	*Accumulation*
Liver esterase	Aliphatic esters	Acyl-enzyme	–
Acetylcholinesterase	Acetylcholine	Acyl-enzyme	+
Alcaline phosphatase	Monophosphorylated esters	Phosphoryl-enzyme	+
Lysozyme (glycosidases)	Polysaccharides	Glycosylenzyme	–

The accumulation of a stable acyl-enzyme is not an energy advantage, as is illustrated in the energy diagram corresponding to a reaction catalysed by chymotrypsin (Fig. 11.47). It has the consequence of decreasing K_m bringing about a saturation in the presence of weak substrate concentrations, but increasing the energy barrier. The formation of a covalent bond between the enzyme and a part of the substrate decreases the entropy of the bound substrate. A specific substrate has an appropriate position at the active centre, and due to this the acylation and deacylation rates are optimal. For non-specific substrates, the acylation rate decreases more rapidly than the deacylation rate; this is responsible for a change in the limiting step during ester hydrolysis. Deacylation is limiting for the most specific esters, and acylation becomes limiting for the least specific esters.

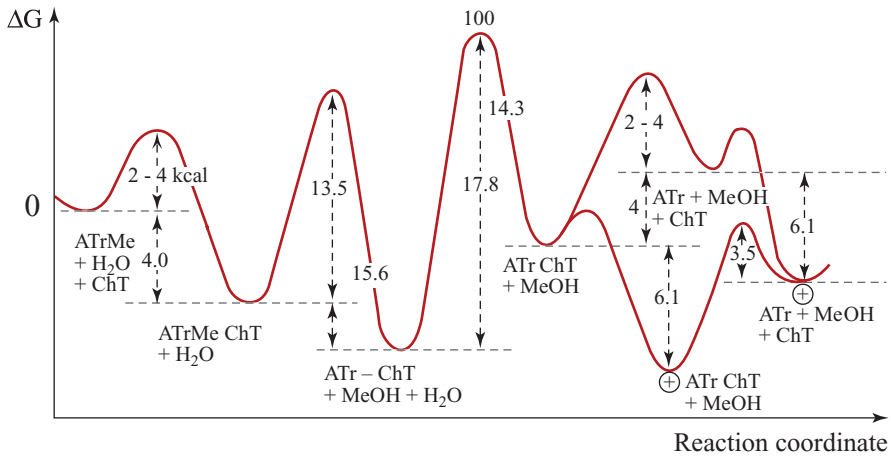


Fig. 11.47 Energy diagram of a reaction catalysed by chymotrypsin
(ATrMe: Acetyl tyrosine methyl ester)

The formation of an acyl-enzyme, if it is not a thermodynamic advantage, presents however a selective advantage permitting the enzyme, in the presence of many substrates, to react more rapidly with the most specific ones. The formation of an acyl-enzyme permits also, in certain conditions in the absence of water, to favor reactions of transpeptidation. In addition, as JENCKS (1975) underlined, the selection in the course of evolution of enzymes with a unique reacting site by a nucleophilic

attack on the acyl-enzyme bond in deacylation, was probably a process simpler than the selection of enzymes with two distinct regions for the leaving group and the nucleophile which would be necessary in the absence of the covalent intermediate.

11.5. CONCLUSIONS

The diverse factors analysed above contribute in a synergistic manner to the efficiency of enzyme catalysis. The formation of the MICHAELIS complex between the enzyme and the substrate, which implicates numerous non-covalent interactions, drives an intramolecular catalysis which increases the local concentration of the catalytic groups. It favors the reaction by entropic effects. The energy of formation of the enzyme-substrate complex is used in catalysis. Enzymatic catalysis possesses another catalytic advantage, its polyfunctional character. **The determining contribution in the specificity and efficiency of enzyme catalysis resides in the structural and electrostatic complementarity between the enzyme and the transition state of the substrate.** This particularity results from the protein structure and its conformational dynamics. In the following chapter, several examples are described in order to show how the ensemble of structural data joined to enzymatic studies permits the analysis and understanding of the mechanisms of enzyme catalysis.

BIBLIOGRAPHY

BOOKS

BERNHARD S. –1968– *Enzyme, structure and function*, BENJAMIN Inc., New York.

BRUICE T.C. & BENKOVIC S. –1966– *Bioorganic mechanisms*, Benjamin Inc., New York.

FERSHT A.N. –1984– *Enzyme, structure and mechanism 1rst ed*. Freeman, New York.

FERSHT A.R. –1985– *Enzyme, structure and mechanism*, 2nd ed., Freeman and Co, New York.

JENCKS W.P. –1969– *Catalysis in chemistry and enzymology*, Mc Graw Hill Book and Co, New York.

WALSH C. –1979– *Enzymatic reaction mechanisms*, Freeman and Co, San Francisco.

GENERAL REVIEWS

BENDER M.L. –1960– Mechanisms of catalysis of nucleophilic reactions of carboxylic acid derivatives, in *Chem. Rev.* **60**, 53–113.

JENCKS W.P. –1975– Binding energy, specificity, and enzyme catalysis: the CIRCE effect, in *Adv. Enzymol.* **43**, 219–410.

SPECIALISED ARTICLES

BASH P.A., FIELD M.J., DAVENPORT R.C., PETSKO G.A., RINGE D. & KARPLUS M. –1991– *Biochemistry* **30**, 5826.

BECHET J.J. & YON J. –1964– *Biochim. Biophys. Acta* **89**, 117.

BELL R.P. & KUHN A.T. –1963– *Trans. Farad. Soc.* **59**, 1789.

BENDER M.L., BEGUE CANTOR M.L., BLAKELEY R.L., BRUBACHER L.J., FEDER J., GUNTER C.R., KEZDY F.J., KILLHEFER J.V., MARSHALL T.L., MILLER C.G., ROESKE R.W. & STOOPS J.K. –1966– *J. Am. Chem. Soc.* **88**, 5890.

BENDER M.L., KEZDY F.J. & GUNTER C.R. –1964– *J. Am. Chem. Soc.* **86**, 3714.

BRUICE T.C. & BRADBURY W.C. –1965– *J. Am. Chem. Soc.* **87**, 4846.

BRÜNGER A.T., BROOKS C.L. & KARPLUS M. –1985– *Proc. Natl Acad. Sci. USA* **82**, 8458.

BUNTON C.A. & SHINER V.J. –1961– *J. Am. Chem. Soc.* **83**, 42.

CANNON AND BERKOVIC –1998– *J. Biol. Chem.* **273**, 26257.

CANNON W.R., SINGLETON S.S. & BENKOVIC S.T. –1998– *Nat. Struct. Biol.* **3**, 821.

CANNON W.R. & BENKOVIC S.T. –1998– *J. Biol. Chem.* **273**, 26257.

CHEVALLIER J. & YON J.M. –1966– *Biochim. Biophys. Acta*. **122**, 116.

CLELAND W.W., FREY P.A. & GERTL J.A. –1998– *J. Biol. Chem.* **273**, 25529.

COLLINS K.D. & STARK G.R.–1969– *J. Biol. Chem.* **244**, 1869.

CUI Q. & KARPLUS M. –2001– *J. Am. Chem. Soc.* **123**, 2284.

DEWAR M.J.S. & STORCH D.M. –1985– *Proc. Natl Acad. Sci. USA* **82**, 2225.

FERSHT A.R. & KIRBY A.J. –1967– *J. Am. Chem. Soc.* **89**, 4853 and 4857.

FERSHT A.R. & KIRBY A.J. –1968– *J. Am. Chem. Soc.* **90**, 5818, 5826 and 5833.

FERSHT A.R., LEATHERBARROW R.J. & WELLS T.N.C. –1986– *Nature* **332**, 284.

FERSHT A.R., SHI J.P., KNILL-JONES J., LOWE D.M., WILKINSON A.J., BLOW D.M., BRICK P., CARTER P., YWAYE M.M. & WINTER G. –1985– *Nature* **314**, 235.

FOUCAULT G., SEYDOUX F. & YON J.M. –1974– *Eur. J. Biochem.* **47**, 295.

HOL W.G.–1985– *Adv. Biophys.* **19**, 133.

JENCKS W. & GILCHRIST M. –1968– *J. Am. Chem. Soc.* **90**, 2622.

KOSHLAND D.E. –1962– *J. Theor. Biol.* **2**, 75.

KOSHLAND D.E., NÉMÉTHY G., & FILMER D. –1966– *Biochemistry* **5**, 365.

LEATHERBARROW R.J., FERSHT A.R. & WINTER G. –1985– *Proc. Natl Acad. Sci. USA* **82**, 7840.

LIPSCOMB W.N., HARTSUCK J.A., QUIOCHO F.A. & REEKE G.N. –1969– *Proc. Natl Acad. Sci. USA* **64**, 28.

LOEW G.H. & THOMAS D.D.–1972– *J. Theoret. Biol.* **36**, 89.

LOWE D.M., FERSHT A.N. & WILKINSON A.J. –1985– *Biochemistry* **24**, 5106.

MARX J.L. –1986– *Sci. Res. News* **234**, 141.

MONNOT M., WEINTRAUB H. & YON J.M. –1966– *Bull. Soc. Chim. Biol.* **48**, 959.

NORTHROP D.B. –1975– *Biochemistry* **14**, 2644.

O'KEEFE S.J. & KNOWLES J.R. –1986– *Biochemistry* **25**, 6077.

PAGE M.I. –1973– *Chem. Soc. Rev.* **2**, 295.

POLLACK S.J., JACOBS J.W. & SCHULTZ P.G. –1986– *Science* **234**, 1570.

POMPON D. & LEDERER F. –1985– *Eur. J. Biochem.* **148**, 145.

SEYDOUX F. & YON J.M. –1971– *Biochem. Biophys. Res. Commun.* **44**, 745.

STEARN A.E. –1938– *Ergeb. Enzymforschung* **7**, 1.

STORM D.R. & KOSHLAND D.E. –1970– *Proc. Natl Acad. Sci. USA* **66**, 445.

STORM D.R. & KOSHLAND D.E. –1972– *J. Am. Chem. Soc.* **94**, 5805.

SWAIN C.G. & BROWN J.F. –1952– *J. Am. Chem. Soc.* **74**, 2534.

TAPIA O., BRÁNDÉN C.I. & ARMBRUSTER A.M. –1982– in *Quantum Theory of Chemical Reactions*, Vol. III, R. DAUDEL, A. PULLMAN, L. SALEM & A. VIELLARD eds, Reidel Pub. and Co, 97–123.

TAPIA O. & EKLUND H. –1986– *Enzyme* **36**, 101.

TRAMONTANO A., JARDA K.D. & LERNER R.M. –1986– *Science* **234**, 1566.

URBAN P., ALLIEL P.M. & LEDERER F. –1983– *Eur. J. Biochem.* **134**, 275.

URBAN P. & LEDERER F. –1986– *J. Biol. Chem.* **260**, 11115.

WARSHEL A. –1978– *Proc. Natl Acad. Sci. USA* **75**, 5250.

WARSHEL A. –1981a– *Accounts Chem. Res.* **14**, 284–290.

WARSHEL A. –1981b– *Biochemistry* **20**, 3167.

WARSHEL A. –1998– *J. Biol. Chem.* **273**, 27035.

WARSHEL A., AQVIST J. & CREIGHTON S. –1989– *Proc. Natl Acad. Sci. USA* **86**, 5820.

WARSHEL A. & LEVITT M. –1976– *J. Mol. Biol.* **103**, 227.

WARSHEL A. & WEISS R.M. –1980– *J. Am. Chem. Soc.* **102**, 6218.

WELLS T.N.C. & FERSHT A.N. –1986– *Biochemistry* **25**, 1881.

WINTER G., FERSHT A.N., WILKINSON A.J., ZOLLER M. & SUNTH M. –1982– *Nature* **299**, 756.

YON J.M. –1976– *Biochimie* **58**, 61.

12 – EXAMPLES OF STRUCTURE-FUNCTION RELATIONSHIPS IN ENZYMATIC SYSTEMS

As was shown in the preceding chapter, there is a great variety of enzymes catalysing very diverse chemical reactions. Some enzymes among the best known on the structural and functional level have been chosen in the present chapter to illustrate each of the reaction types. Enzymatic systems for which the three-dimensional structures are known both for the free enzyme and for the enzyme complexed to a substrate or to an analog are preferentially presented. Indeed, a precise knowledge of the position of substrate atoms with respect to enzyme catalytic groups is necessary for understanding reaction mechanisms; otherwise the amount of speculation remains too significant. However, even for particularly well described systems, the transition state and its interactions with the enzyme often remain hypothetical.

Group transfer reactions are treated by taking the examples of acyl transferases, serine proteases and more particularly chymotrypsin which is the paradigm, thiol proteases which include papain, acid proteases like pepsin, and a metallopeptidase, carboxypeptidase. Tyrosine tRNA synthetase and a kinase were chosen for phosphoryl group transfer reactions and lysozyme for glycosyl group transfer reactions. Two examples were retained for oxydoreduction reactions, alcohol dehydrogenase and flavocytochrome b2. Isomerisation reactions are illustrated by triose phosphate isomerase. Aspartate amino transferase, a pyridoxal phosphate enzyme, offers a well documented example of amino group transfer *via* the formation of a SCHIFF base. Aldolases, which induce the formation of a SCHIFF base over the course of the reaction that they catalyse, were retained to illustrate reactions of carbon-carbon bond breaking or formation.

Some of these enzymatic systems were already described under diverse aspects throughout this text. In this chapter, catalytic mechanisms are considered for each enzyme in close relation with its structure. At the risk of redundancy, it seemed useful to us to regroup all the significant relative data to a particular enzyme. Mechanisms of reactions catalysed by allosteric enzymes are presented in Part V in the chapter treating non-covalent regulations.

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_13, © Springer-Verlag Berlin Heidelberg 2010

12.1. PROTEASES

Proteases can be classified according to their activities and the nature of their functional groups. Most serine proteases are endopeptidases. With a few exceptions, thiol proteases are also endopeptidases; they possess a cysteine as the catalytic group which plays the same role as serine. These two types of proteases possess an optimal activity at neutral pH. Acid proteases or aspartyl proteases have carboxylic groups on two aspartyl residues at the catalytic site. Most of them are active at low pH; conversely, renin is active at neutral pH. Metalloproteases and Zn^{++} proteases are active at neutral pH. Among them, carboxypeptidases are exopeptidases and thermolysine is an endopeptidase.

12.1.1. SERINE PROTEASES

12.1.1.1. STRUCTURAL ASPECTS

There is in mammals a large number of serine proteases: trypsin, chymotrypsin, elastase, proteases of the blood coagulation cascade and among them thrombin. They are secreted as precursors or zymogens, their activation resulting from proteolytic cleavage. Many serine proteases of microorganismes are currently known, like subtilisine, α-lytic protease of *Myxobacter 495*, and that of *Streptococcus griseus*. Pancreatic proteases possess great similarities in sequence, from 40 to 45% identity and 50 to 55% homology. Structural analogs are even greater and extend to bacterial proteases of which the three-dimensional structure is known; however, these have a shorter polypeptide chain and only share 21% identity with the preceding ones.

The three-dimensional structure of several serine proteases and their zymogen is known at atomic resolution. Among the first resolved were α-chymotrypsin (BLOW, 1970; BIKTOFT & BLOW, 1972), γ-chymotrypsin (COHEN et al., 1969), chymotrypsinogen (KRAUT, 1971), pig elastase (SHOTTON & WATSON, 1970), trypsin (STROUD et al., 1974; FELDHAMMER & BODE, 1975), trypsinogen (BODE et al., 1976), protease B of *Streptomyces griseus* (DELBAERE et al., 1975), BPN' subtilisine (WRIGHT et al., 1970), Novo subtilisine (DRENTH et al., 1972) and α-lytic protease of *Myxobacter 495*. Since then, other protease structures have been determined and the list increases regularly. All these proteases possess the same catalytic groups. Those of the trypsin family are folded into two distinct domains stabilised by disulphide bridges; each of these domains is made of antiparallel β segments forming a β barrel (Fig. 12.1 opposite). They comprise a short helical segment at the C-terminal end. The principal structural differences are situated in the external loops. These structural similarities suggest that these diverse proteases derive from the same ancestor protein and that they diverged over the course of evolution by preserving their principal structural characteristics but developed different specificities. Subtilisines, although possessing the same catalytic groups, are not homologues of the preceding and present a different folding comprising an important α helix

proportion. Carboxypeptidase II from wheat presents an even different structure; carboxypeptidase Y from yeast and two plant proteases belong to the same family.

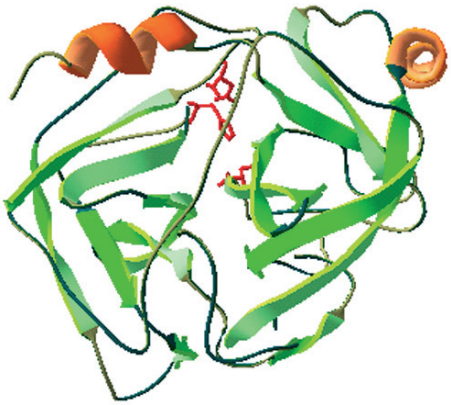


Fig. 12.1 Structure of α-chymotrypsin (from PDB: 3CHA, with SwissPDB viewer software, as well as all those referenced by their PDB identification) *The catalytic groups are indicated* in red

12.1.1.2. ACTIVATION OF ZYMOGENS

Most mammalian serine proteases are synthesised as prezymogens transformed into zymogens by elimination of the signal peptide when crossing the rough endothelial reticulum membrane and secreted under this form (see Chap. 8). Pancreatic proteases are found thus in the pancreas as inactive precursors; this situation constitutes a protection mechanism. Trypsinogen is activated into trypsin by enterokinase which cleaves the Lys6–Ile7 peptide bond, thus releasing the N-terminal hexapeptide. The activation of chymotrypsinogen is more complex. The zymogen is transformed into chymotrypsin by trypsin-catalysed breaking of the Arg15–Ile16 bond which gives rise to π-chymotrypsin. A second cleavage by active chymotrypsin thus formed is produced between Ile13 and Ser14 generating δ-chymotrypsin. Then in a third step, two other cleavages occur, those of Tyr146–Thr147 and Asn148–Ala149 bonds that give rise to α-chymotrypsin. Chymotrypsinogen itself can undergo proteolysis by chymotrypsin, giving rise to neochymotrypsinogens (Fig. 12.2 below). Proelastase is also activated by trypsin which cleaves the Arg15–Val16 bond (the numbering of amino acids is that of chymotrypsinogen).

The breaking of peptide bonds during zymogen activation drives the establishment of a salt bridge between the carboxylate of Asp194 adjacent to the reactive serine (Ser195) and the amino group of Ile16 (or Val16 in elastase); the formation of this salt bridge and its conformational and functional consequences were presented in Chap. 11. Although zymogens are inactive with respect to their physiological substrates that are peptide bonds, they present a weak activity for some specific esters such as p.nitrophenyl esters of different dipeptides as was shown by a study of the group of NEURATH [LONSDALE-ECCLES et al., 1978].

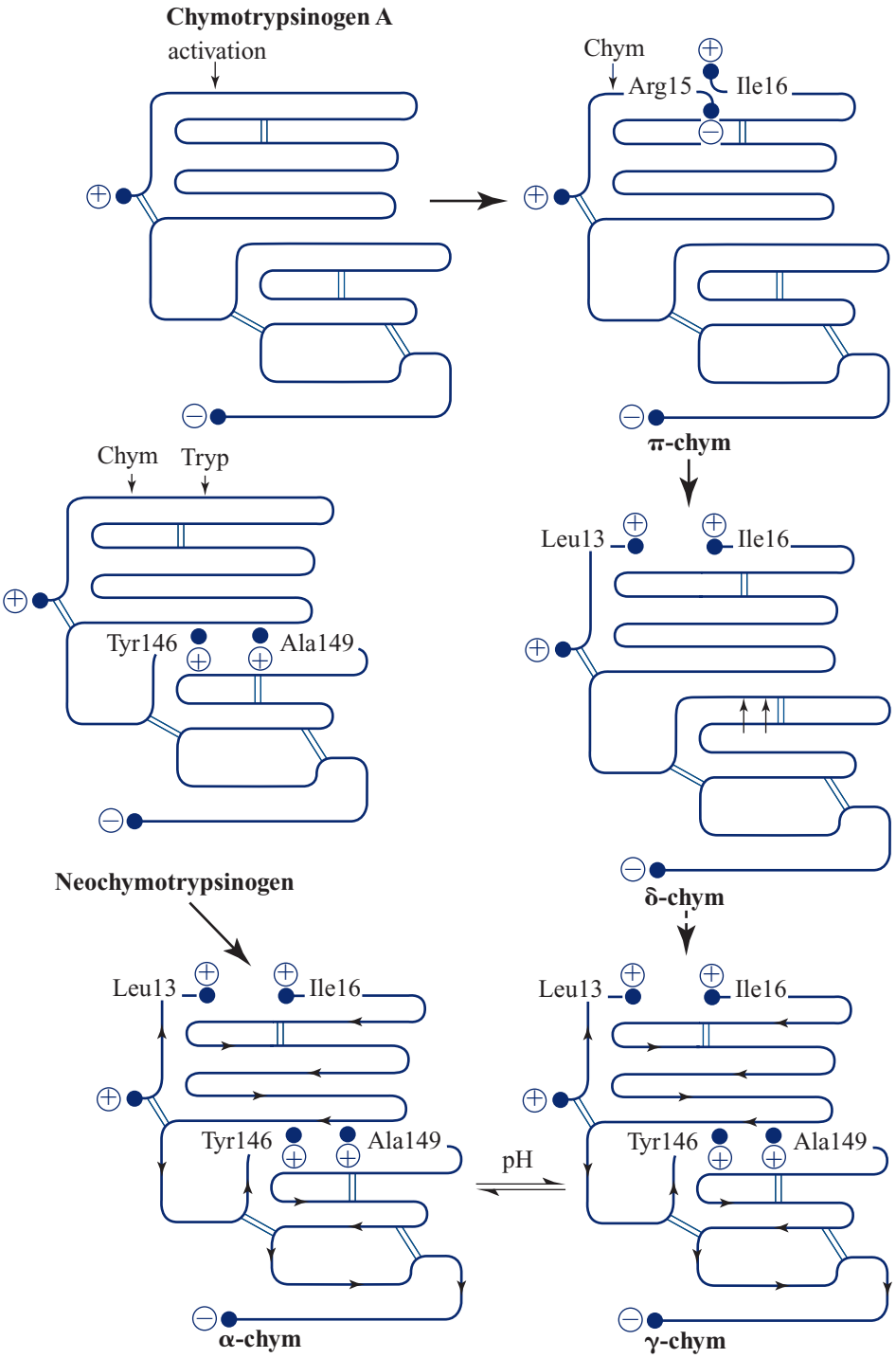


Fig. 12.2 Activation of chymotrypsinogen

Table 12.1 indicates the values of the kinetic parameters obtained for the hydrolysis of diverse substrates by trypsin, chymotrypsin and their respective zymogens. The best substrate of chymotrypsinogen is tertiobutyl oxycarbonyl alanine p.nitrophenyl ester (Boc-Ala-ONP) and that of trypsinogen is benzyl oxycarbonyl hydroxyproline glycine p.nitrophenyl ester (Z-Gly-Hyp-Gly-ONP). Examination of the kinetic parameters obtained for diverse substrates clearly indicates that zymogens present a weak but significant catalytic activity. The comparison of values obtained for each en-zyme and the corresponding zymogen shows that the increase in the K_m value can reach two orders of magnitude whereas k_{cat} is decreased by a factor of 10 000 in the zymogen. On the basis of these results the authors attribute the weak activity of zymogens to a distortion in the primary substrate binding site. In addition, the same authors (LONSDALE-ECCLES et al., 1979) showed that strong ionic strength increases the activity of trypsinogen and chymotrypsinogen, indicating that in the absence of a salt bridge the functional structure of the protein can be stabilised by strong ionic strength. **These differences in efficiency between zymogen and enzyme were interpreted by GHÉLIS and YON (1979) on the basis of energy in terms of conformational coupling between the two domains of the enzyme, the coupling being optimised upon salt bridge formation.**

Table 12.1 Hydrolysis of paranitrophenyl esters by zymogens and the corresponding enzymes (Reprinted with permission from *Biochemistry*, **17**, LONSDALE-ECCLES J.D. et al., 2805. © (1978) American Chemical Society)

	$k_{cat}(s^{-1})$	$K_m(M \infty 10^4)$	$k_{cat}/K_m(M^{-1}.s^{-1})$
Trypsinogen			
Boc-Gly-ONP	6.1×10^{-3}	26.4	2.3
Z-Ser-Gly-Gly-ONP	4.0×10^{-3}	6.3	6.4
Z-Gly-Hyp-Gly-ONP	7.6×10^{-3}	1.8	41.5
Trypsin			
Boc-Gly-ONP	0.22	5.2	0.43×10^{3}
Z-Ser-Gly-Gly-ONP	0.47	5.8	0.81×10^{3}
Z-Gly-Hyp-Gly-ONP	0.99	5.0	1.97×10^{3}
Chymotrypsinogen			
Boc-Gly-ONP	6.6×10^{-3}	6.3	10.5
Z-Ser-Gly-Gly-ONP	3.7×10^{-3}	7.1	5.2
Z-Gly-Hyp-Gly-ONP	9.3×10^{-3}	3.9	23.8
Boc-Ala-ONP	28.1×10^{-3}	4	72
Chymotrypsin			
Boc-Gly-ONP	0.52	0.22	26.6×10^{3}
Z-Ser-Gly-Gly-ONP	0.32	0.47	6.8×10^{3}
Z-Gly-Hyp-Gly-ONP	1.88	1.75	10.7×10^{3}
Boc-Ala-ONP	2.3	21	1.1×10^{3}

The crystallographic data confirmed the existence of a salt bridge previously deduced by chemical, enzymatic and conformational studies (see Chap. 9). This led to the identification of Asp194 (Fig. 12.3). In spite of local rearrangements, the overall structure of chymotrypsinogen and the enzyme are very similar.

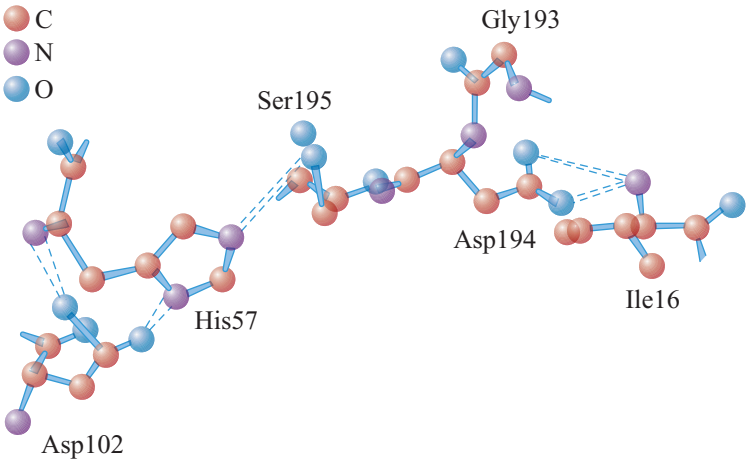


Fig. 12.3 The salt bridge in chymotrypsin according to the crystallographic data
(Reprinted from *The Enzymes*, 3rd ed., Vol. III, BLOW DM., The structure of chymotrypsin, 196. © (1971) Academic Press, with permission from Elsevier)

- Average displacements of α-carbon atoms from zymogen to enzyme are on the order of 1.8 Å. Rearrangements resulting from the activation of the zymogen are very localised. Some segments undergo displacements of a larger amplitude (Table 12.2). Segments I, III, and IV correspond to loops situated at the exterior of the molecule. The movement of segment II accompanies the formation of the salt bridge. Isoleucine 16 undergoes a displacement of 11.3 Å to be placed at the interior of the molecule by interacting with Asp194; this movement takes place by a rotation of the main chain of around 180°. Asp194 is localised in the interior of the molecule in the enzyme like in the zymogen.

Table 12.2 Residues for which the position of Cα differs by more than 3.6 Å between the structures of chymotrypsinogen and α-chymotrypsin
[Reprinted from *The Enzymes*, 3rd ed., Vol. III, KRAUT J., Chymotrypsinogen: X-ray Structure, 165. © (1971a) Academic press, with permission from Elsevier]

Segment	*Residue*	*Displacement in Å*
I	Gln7	4.8
	Pro8	10.0
II	Ile16	11.3
	Val17	6.6
III	Thr37	4.0
	Gly38	6.6

Segment	*Residue*	*Displacement in Å*
IV	Asp72	5.6
	Gln73	9.6
	Gly74	9.1
	Ser75	6.2
	Ser76	10.1
	Ser77	5.6
V	Thr144	5.9
	Arg145	8.7
	Tyr146	4.6
	Ala149	4.7
	Asn150	6.7
	Thr151	4.7
	Pro152	4.6
VI	Met192	8.4
	Gly193	6.6

In segment V, the side chain of Arg145 is displaced by 9 Å and becomes completely accessible to the solvent. In segment VI, Met192 and Gly193 undergo an important displacement that, with Ile16, leads to the setting up of a binding pocket for specific substrates. Met192, deeply buried in the zymogen, comes to the surface during activation permitting access of the substrate to the active centre. These crystallographic data are in agreement with results from chemical studies carried out by the group of NEURATH which showed that, in the enzyme, Met192 is very reactive towards iodoacetate whereas it practically does not react in the zymogen. The residue Gly193 is placed in proximity to His40 and, by the oxygen of the carbonyl and its peptide bond, establishes a hydrogen bond with $N^{\varepsilon 2}$ of His40 that, in the zymogen, forms a hydrogen bond with the oxygen of Asp194. The conformational changes implicating Ile16, Met192, Gly193 and Asp194 induce the formation of the specific substrate binding site during the activation of the zymogen. ◢

It is remarkable that the three catalytic groups Ser195, His57 and Asp102, thought to be linked by hydrogen bonds are found in exactly the same spatial arrangement in chymotrypsin and chymotrypsinogen. The catalytic triad is equally conserved in the same position upon trypsinogen activation into trypsin as was shown by FELHAMMER et al. (1975). This activation is provoked, in this case, by weak differences in the flexibility of regions well localised into the molecule. These regions become more rigid in the active enzyme. Trypsinogen adopts the conformation of trypsin in the presence of the Ile-Val dipeptide provided that it is covalently bound to the p.guanido benzoate group.

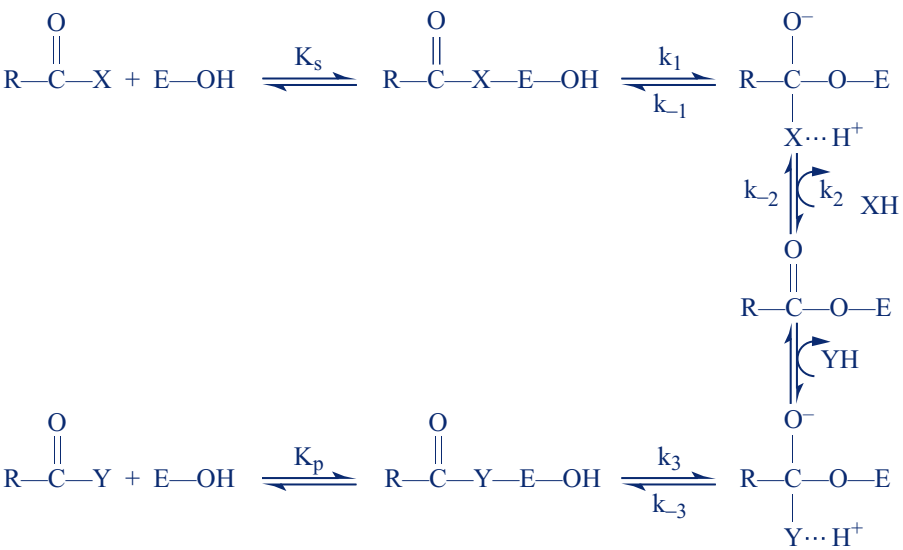
Numerous chemical labelling experiments associated with kinetic studies had shown the catalytic role of Ser195 and His57 before crystallographic studies were carried out (see Chap. 9). The structural analysis confirmed the existence of these two groups

localised into the cleft and interacting by hydrogen bond, forming the active centre near the surface.

◤ The existence of a hydrogen bond between the $N^{\varepsilon 2}$ of His57 and the O^{γ} of Ser195 is based on the interatomic distance between these groups which is of 2.8 to 3 Å; it is 3 Å in trypsin (BODE et al., 1983) and 2.9 Å in elastase for which the binding of an inhibitor induces the alignment of serine and histidine at this distance. Moreover, crystallographic studies revealed that Asp102 is bound by hydrogen bond to a nitrogen proton of His57, which had driven crystallographers to propose the existence of a charge relay system between Ser195-His57-Asp102, and to propose a catalytic mechanism which was invalidated by more recent experiments; the role of Asp102 is discussed in Sect. 12.1.1.6. The intervention of this group in catalysis had not been revealed by chemical analysis. ◢

12.1.1.3. THE REACTION PATHWAY

Reactions catalysed by serine proteases proceed by a sequence of reaction steps common to enzymes of the chymotrypsin family and those of the subtilisine family that possess the same catalytic groups as mentioned above. Most experimental data were obtained on chymotrypsin. The reaction sequence, of which the kinetic evidence was presented in Chap. 5, implicates the formation of the non-covalent MICHAELIS complex and the formation of an unstable tetrahedral intermediate in which a covalent bond is established between the O^{γ} atom of the reactive serine and the substrate, the proton of serine being partially transferred to the leaving group. The tetrahedral intermediate gives rise to the acyl-enzyme and releases the leaving group after the breaking of the C—X bond (see scheme below). Deacylation is the inverse process in which the leaving group is replaced by water or by another nucleophilic compound on which the acyl transfer takes place. The complete reaction scheme is represented below:



R-CO is either a polypeptide, an amino acid or an acyl group, XH is a peptide, an amine or an alcohol; YH can be water or any other acceptor.

As was shown in Part II (Chap. 5), the existence of the acyl-enzyme was proved experimentally from kinetic studies in different conditions. The tetrahedral intermediate is too unstable to be revealed in this manner. Its existence was deduced from indirect arguments, either by analysis of substitution effects on the acylation rate during hydrolysis of synthetic substrates, or by the study of isotope effects. A more direct argument was provided by spectrophotometry data during the hydrolysis of chromophore substrates like N-furyl acryloyl-L-tryptophan amide and the corresponding methyl ester; an additional intermediate situated between the MICHAELIS complex and the acyl-enzyme was observed. FINK and MEEHAN (1979) detected by spectrophotometry the presence of a tetrahedral intermediate upon hydrolysis by elastase of specific di- and tripeptides in carrying out kinetic studies at low temperature (–39°C) as a function of pH. It was assumed that this tetrahedral intermediate is very similar to the transition state.

The crystallographic data revealed the geometry of most of these intermediates and the interactions that stabilise them and contributed to elucidating the catalytic mechanisms. However, the structural data does not permit unambiguously the establishment of a mechanism; it is important to provide complementary data by means of all methods in enzymology. Among serine proteases, chymotrypsin was the focus of the most complete studies; it is therefore this enzyme which is used as an example to present the mechanism of action of these proteases.

12.1.1.4. THE BINDING SITE OF SUBSTRATES AND THE MICHAELIS COMPLEX

Physiological substrates of proteases being polypeptides, they can establish numerous interactions with their specific enzyme. In order to determine their importance, it is useful to divide these types of substrates as a function of the amino acid residues they contain by the notations P_1, P_2, P_3 upstream of the peptide bond to break, that is in the acyl group region, and P'_1, P'_2, P'_3 downstream, that is in the region of the leaving group. The corresponding sub-sites of the enzyme are designated by S_1, S_2, S_3 and S'_1, S'_2, S'_3 (Fig. 12.4 below), the sub-site S_1 defining the protease specificity. Different crystallographic data were obtained in the presence of substrate analogs or specific inhibitors, in particular the complex formed by chymotrypsin and N-formyl-L-tryptophan and the complex of γ-chymotrypsin with polypeptides possessing at the C-terminal end an L-phenylalanine chloromethyl ketone. These studies showed that the binding site of the substrate acyl part is constituted by a region of the polypeptide chain in a β structure containing the amino acids Ser214-Trp215-Gly216. This region is represented in Fig. 12.5 below which corresponds to the complex of chymotrypsin with N-formyl-L-tryptophan. The amido NH group of formyl tryptophan points towards the carbonyl of Ser214. The carboxyl group has a carbon atom in contact with the O^{γ} of Ser195 and a C—O bond practically parallel to the hydrogen bond between the O^{γ} of Ser195 and the $N^{\varepsilon 2}$ of His57.

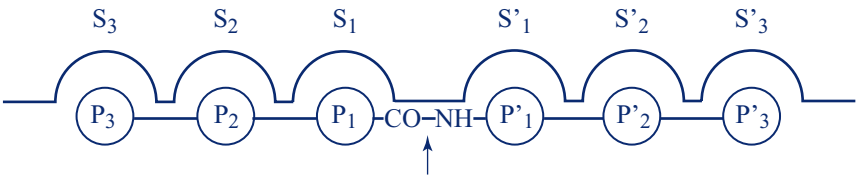


Fig. 12.4 Binding sub-sites of protease substrates

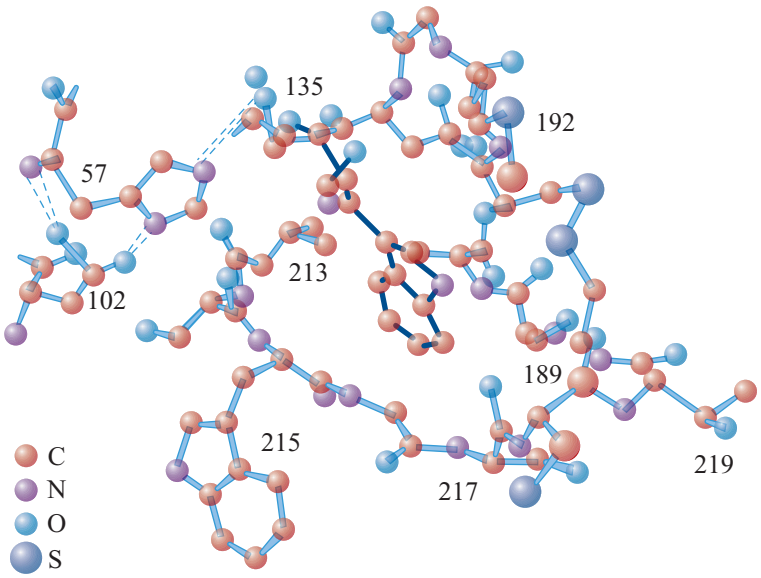


Fig. 12.5 Binding site of N-formyl tryptophan at the active centre of chymotrypsin
(Reprinted from *The Enzymes*, 3rd ed., Vol. III, BLOW DM., The structure of chymotrypsin, 199. © (1971) Academic Press, with permission from Elsevier)

◤ If one considers the different sub-sites, three hydrogen bonds can be formed between a substrate and the enzyme in the region of the acyl group, one with the carbonyl oxygen of S_1 and the amido nitrogen of P_1, another between the amido nitrogen of the peptide backbone of S_3 and the carbonyl oxygen of P_3 and a third between the carbonyl oxygen of S_3 and the amido nitrogen of P_3. Only this last bond can be formed with N-formyl tryptophan which only interacts with the S_1 site. ◢

The indolyl group of tryptophan is placed in the pocket of specificity, the planes of neighboring peptide bonds being practically parallel to the plane of the indol ring. This cavity comprises Ser189, Val213, Ser217; Met192 constitutes the hydrophobic and flexible opening of the cavity. The site of specificity of chymotrypsin accommodates aromatic residues whereas trypsin accepts substrates that possess an L-arginine or an L-lysine. In the crevasse of specificity of this last enzyme, Ser189 is replaced by Arg189 which contracts an electrostatic interaction with the positive charge of the side chain of substrates. The structure of elastase and of its complex with formyl alanine shows that the two glycines which are at the entrance of the

substrate binding pocket in chymotrypsin are replaced by Val216 and Thr226; these residues hinder the entry of encumbering side chains, limiting substrate binding to the dimensions of the Ala residue (SHOTTON & WATSON, 1970).

The formation of the MICHAELIS complex is not accompanied by any significant conformational change. No displacement of the peptide backbone greater than 0.2 Å, which represents the limit of detection of the method, was observed.

The geometry of enzyme interactions with the leaving group is less well established. This results from the fact that, on one hand this interaction is much weaker, and on the other hand it is not easy to have inhibitors that bind strongly and in a productive manner on this part of the active centre. Kinetic studies of nucleophilic competition showed however that chymotrypsin possesses an S'_1 site capable of binding a P'_1 residue provided that the latter does not have a free carboxyl group. In trypsin, analogous studies using aliphatic alcohols of different chain sizes as nucleophile reagents revealed the hydrophobic character of the binding site of the leaving group (SEYDOUX et al., 1969). The hydrophobic character of the sub-site S'_1 was confirmed by crystallography studies of complexes of trypsin with two of its inhibitors, bovine pancreatic inhibitor (HUBER et al., 1974) and soybean inhibitor (BLOW et al., 1974). In elastase the sub-sites S'_1 and S'_2 are also hydrophobic (ATLAS, 1975).

12.1.1.5. THE TETRAHEDRAL COMPLEX AND THE BINDING SITE OF THE OXONIUM ION

The most important particularity of enzymatic catalysis resides in the complementarity of the enzyme for the transition state of the substrate (see Chap. 11). Analogs of the transition state of protease substrates that form tetrahedral compounds must therefore possess a high affinity for the enzyme. The tetrahedral addition compound yields an oxyanion formed by the carbonyl oxygen of the substrate. On the basis of crystallographic data obtained on an acyl-enzyme, indole acryloyl chymotrypsin, HENDERSON (1970) suggested that the oxyanion could contract a pair of hydrogen bonds, one with the NH of the main chain of Ser195, the other with that of Gly193; these hydrogen bonds could be stronger than those which would be formed by the carbonyl oxygen of the acyl-enzyme.

In complexes formed by trypsin with its protein inhibitors, the carbonyl group of the inhibitor is tetrahedral, its distance to the O^{γ} of the reactive serine being too long to establish a VAN DER WAALS interaction. In addition, this carbon has the same geometry in the complex that the inhibitor forms with the modified trypsin in which the serine has been transformed by chemical modification into dihydroalanine. Magnetic resonance studies of the proton indicate that the extra proton would be bound to His57 in the complex formed by pig β-trypsin and bovine pancreatic inhibitor.

The crystal structure of two aromatic derivatives of boronic acid bound to subtilisine was resolved by MATTHEWS et al. (1975). These authors concluded that these systems represent good models for determining the geometry of the transition state. An analogous study was carried out by TULINSKY and BLÉVINS (1987) on complexes formed by chymotrypsin and phenyl ethane boronic acid. This compound is associated

to chymotrypsin with a very good affinity and forms a stable tetrahedral complex with the O^{γ} of Ser195 which occupies one of the tetrahedron positions. The phenyl ethyl boronate molecule occupies a part of the specificity site. The formation of the tetrahedral complex brings about a rotation of –14° in the angle χ_1 of the reactive serine.

12.1.1.6. THE CATALYTIC TRIAD AND THE MECHANISM OF ACYLATION

The catalytic triad is common to all serine proteases which are defined by the presence of a single reactive serine forming an acyl-enzyme with the substrates. Crystallographic data indicated the invariance in the respective position of Asp102, His57 and Ser195 in proteases of the chymotrypsin family. Although the global structure differs noticeably, the same geometry of the catalytic triad is found in subtilisines with the groups Asp32, His64 and Ser221. The distance between these groups is such that they can be connected by hydrogen bonds, which led crystallographers to propose a catalytic mechanism (Fig. 12.6).

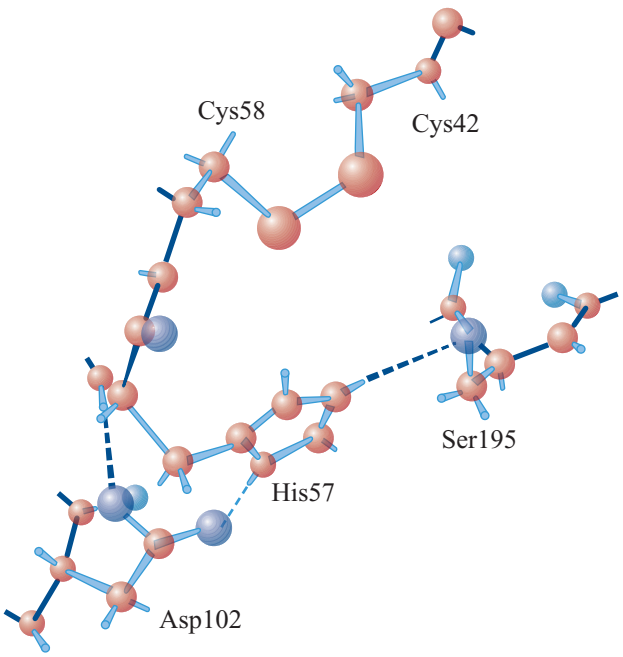


Fig. 12.6 Residue positions in the catalytic triad

(Reprinted from *The Enzymes*, 3rd ed., Vol. III, BLOW DM., The structure of chymotrypsin, 202. © (1971) Academic Press, with permission from Elsevier)

According to their hypothesis, the polarisation of the system would be carried out *via* a ***charge relay system*** in which the negative charge of the carboxylate would drive the increase in the nucleophilic ability of the serine and consequently would make it very reactive. At low pH where the protein crystallises, the system would remain inactive, but at pH 8 which is the pH of optimal activity of the enzyme, the

charge relay would become functional. **However, the hypothesis of the charge relay system proposed by crystallographers was the subject of numerous controversies. Nuclear magnetic resonance studies suggested indeed that the catalytic base is histidine 57 and not aspartate 102, which was ultimately proved.**

One of the limitations of structural studies by X-ray diffraction resides in the fact that this method does not permit localising hydrogen atoms. In contrast, neutron diffraction constitutes an appropriate method to directly determine their exact position. An analysis by neutron diffraction at a resolution of 2.2 Å was carried out in 1981 by KOSSIAKOFF and SPENCER on trypsin covalently bound to an analog of the transition state, the monoisopropyl phosphoryl group. By this method, the authors determined the protonation states of Asp102 and His57. They concluded that the catalytic base is His57 and not Asp102; the serine proton is not transferred to Asp102 as was suggested by the hypothesis of charge relay system. Asp102 would intervene in the transition intermediate by modifying the dielectric environment of His57 of which the pK is usually 7, allowing attainment of a value higher than 9.5; the histidine would remain protonated in the transition state (IT) forming with Asp102 an ionic interaction (Asp^- His^+ IT^-). Figure 12.7 represents the mechanism proposed for the deacylation process which is similar to acylation, the water molecule replacing the catalytic serine. This result is in agreement with the theoretical studies of UMEYAMA et al. (1981) who, using quantum mechanics methods, attributed to Asp102 a purely electrostatic role which has the effect of lowering the energy barrier of proton transfer from serine to histidine, but in which the aspartate does not accept the histidine proton.

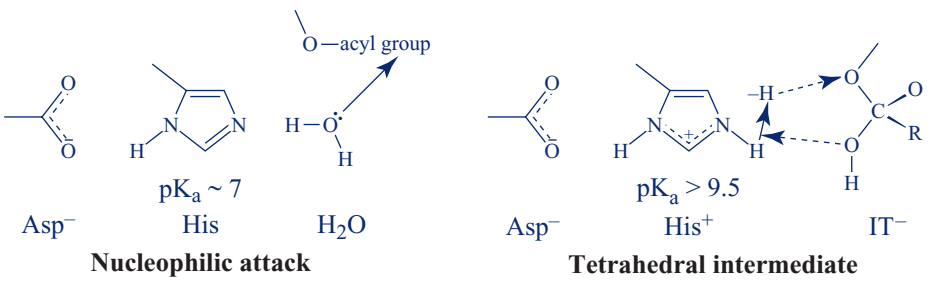


Fig. 12.7 Mechanism of action of serine proteases in the deacylation step

This clearly shows that crystallography, if it is necessary, is insufficient to permit by itself the establishment of a catalytic mechanism.

12.1.1.7. THE ACYL-ENZYME AND THE DEACYLATION STEP

Among the first crystallographic studies, several were carried out on proteases bound covalently and irreversibly by serine 195 to inhibitors in order to avoid autolysis. The structures of tosyl α-chymotrypsin, tosyl γ-chymotrypsin, tosyl elastase, and phenyl methane sulfonyl subtilisine BPN' were resolved. The structure of a more specific acyl-enzyme, indole acryloyl α-chymotrypsin was determined by HENDERSON (1970).

This acyl-enzyme had been previously characterised by kinetic and spectrophotometric studies by the group of BERNHARD et al. (1970). It has a lifespan 10 000 times longer than that of acyl-enzymes of specific substrates and is deacylated more slowly.

◤ Crystallographic studies clearly show that the indolyl part is placed in the site of hydrophobic specificity at a depth of 0.5 to 1 Å, as is the aromatic part of N-formyl tryptophan and of N-formyl-phenylalanine. The carbonyl part of the indole acryloyl group occupies a position close to that of the sulfonyl group in tosyl chymotrypsin. The O^{γ} of serine 195 moves by 2.5 Å upon acylation, which can correspond to a rotation of 120° around the Cα—Cβ bond. A movement of the imidazole of His57 of 0.3 Å towards the solvent region is observed. A water molecule is found localised in a position such that it can form a hydrogen bond with the imidazole $N^{\varepsilon 2}$ of His57 and the carbonyl oxygen of the acyl-enzyme (Fig. 12.8). This water molecule plays the role of the nucleophilic group in the process of deacylation; Met192 is displaced by 1 Å and the Cys191-Cys220 disulphide bridge by 0.35 Å. These movements are produced during acylation but do not bring about important displacements of the peptide backbone. Thus chymotrypsin as well as the ensemble of serine proteases from this family, can be considered as a relatively rigid enzyme, the most important movements accompanying the processes of acylation and deacylation.

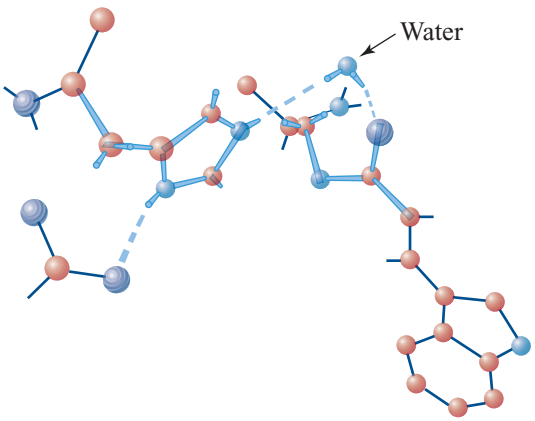


Fig. 12.8 Position of the water molecule with regard to the ester bond of serine 195 in indolacryloyl chymotrypsin ◢

On the basis of structural data obtained for indolacryloyl chymotrypsin, HENDERSON proposed a reactive acyl-enzyme model (Fig. 12.9 opposite). In this model, the N-H amide group very likely forms a hydrogen bond with Ser214. The orientation of the carbonyl oxygen is different from that observed in indolacryloyl chymotrypsin; it is correctly oriented with regard to the water molecule. Thus, deacylation can proceed rapidly by the extraction of a water proton in a general base catalysis, the formation of a tetrahedral intermediate and the appearance of the product.

A problem often discussed is that of the structural identity of a protein in a crystal and in solution. It is indeed very important to resolve this to know if one can transpose directly the conclusions from crystallographic studies to catalytic events that are produced in solution. BERNHARD and ROSSI succeeded in directly measuring

in the crystal the rate of deacylation of indolacryloyl chymotrypsin. They found it equal to that determined in solution in the same conditions of pH and temperature.

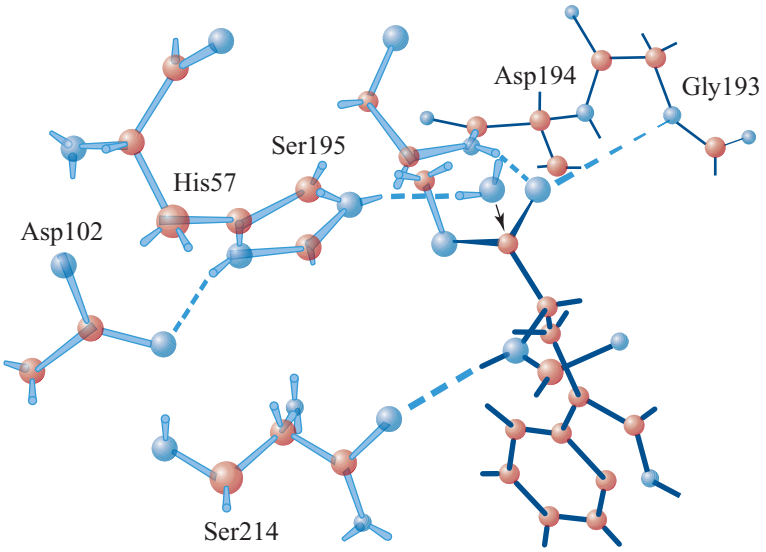


Fig. 12.9 Hypothetical model of the position of atoms in a reacting acyl-enzyme
(Reprinted from *J. Mol. Biol.*, **54**, HENDERSON R., Structure of crystalline α-chymotrypsin: IV. The structure of indoleacryloyl-α-chymotrypsin and its relevance to the hydrolytic mechanism of the enzyme, 351. © (1970) with permission from Elsevier)

The ensemble of structural data from the crystal and the functional and structural studies in solution, completed by theoretical approaches, has therefore allowed to specify the chemical mechanisms implicated in catalysis by serine proteases.

12.1.2. THIOL PROTEASES

Thiol proteases constitute a large family of enzymes which includes papain that was identified first, actinidine, chymopapain, ficine, bromelain, streptococcal protease, clostrinopeptidase B and cathepsins. These enzymes exist in different animal and plant species and in microorganisms. The best known are of plant origin; papain has been particularly well studied. Cathepsins are present in animal tissues. Table 12.3 below indicates the known thiol proteases and their origin.

Thiol proteases are generally monomeric proteins of which the molecular weight is situated between 25 000 and 28 000 with some exceptions. Some are glycoproteins like bromelaine, chymopapain and most cathepsins. All thiol proteases of plant origin are endopeptidases as well as most cathepsins with the exception of cathepsin B that is primarily a carboxypeptidase.

The biological role of most thiol proteases is not known. Enzymes of plant origin have a very broad specificity. It has been suggested that their role is to protect the

fruit against attack from insects or mold. In mammals, cathepsins account for a large part of the proteolytic activity of lysosomes. Some even have specific known functions, for example cathepsin N which has a collagenolytic function and cathepsin P which converts proinsulin into insulin.

Table 12.3 Thiol proteases of different origins

(From The thiol proteases: structure and mechanisms in *Biological macromolecules and assemblies*, **3**, BACKER E.N. & DRENTH J., ed. by F.A. JURNAK & A.M. MCPHERSON, 313-368. © (1984 John Wiley and Sons). This material is reproduced with permission of John Wiley & Sons, Inc.)

Enzyme	*Origin*	*Molecular mass*
Plant enzymes		
Papain	*Carica papaya*	23 350
Chymopapain*	*Carica papaya*	35 000
Peptidase A	*Carica papaya*	26-28 000
Actinidine	*Actinidia chinensis*	24 000
Ficine	*Ficus*	23 500
Bromelain (stem)*	*Ananas comosus*	24-26 000
Bromelain (fruit)*	*Ananas comosus*	33 500
Asclepain A	*Asclepias syriaca*	31 000
Asclepain B	*Asclepias syriaca*	23 000
Calotropine DI	*Calotropis gigantea*	21 000
Calotropine FI*	*Calotropis gigantea*	23 400
Mammalian enzymes		
Cathepsin B*	Rat liver	23-26 000
	Human liver	24-27 500
Cathepsin B_2	Rabbit lung	52 000
Cathepsin H	Rat liver	28 000
	Human liver	28 000
Cathepsin L	Rat liver	23-24 000
	Bovine spleen	23-25 000
Cathepsin I	Rabbit lung	26 000
Cathepsin N	Bovine spleen	18-20 000
	Human placenta	34 600
Cathepsin T	Rat liver	33-35 000
Cathepsin P	Rat islets	31 500
Calpain	Rat liver	90 000
Bacterial enzymes		
Streptococcal proteinase	*Streptococci hemolyticus*	32 000
Clostripain	*Clostridium histolyticum*	55 000

* *glycoproteins*

12.1.2.1. STRUCTURAL ASPECTS

Among all the thiol proteases, papain is the best known from the structural and functional points of view. Its complete sequence was determined in 1978 by CARNE and

MOORE, and its three-dimensional structure is known and has been refined to 1.65 Å resolution (KAMPHUIS et al., 1985). It is constituted of a single polypeptide chain of 212 amino acids folded into two structural domains. The N-terminal domain is essentially made of antiparallel β segments. The C-terminal domain is composed of α helices: a large helix A (residues 24–42) which extends at the interface between the two domains and two shorter helices, helix B (50–57) and helix C (67–78) (Fig. 12.10). The ensemble of the structure is stabilised by three disulphide bridges, two in the N-terminal part and the third in the C-terminal part. The interface between the two domains presents a predominantly polar character comprising four charged residues (Lys17, Glu35, Glu50 and Lys174), and encloses a network of 9 water molecules. Numerous hydrogen bonds are established between the two domains. As in serine proteases, the active centre is localised between the two domains with one of the two catalytic residues, Cys25, in the N-terminal domain, and the other, His159, in the C-terminal domain.

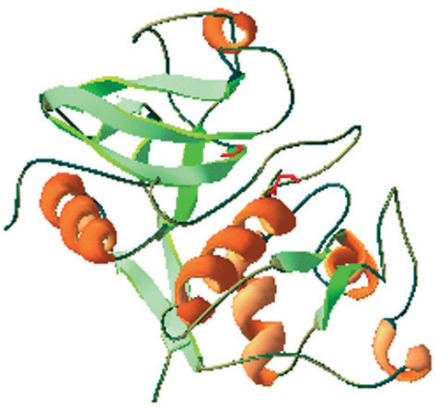


Fig. 12.10 Schematic representation of the three-dimensional structure of papain (PDB: 6PDA)
Catalytic residues Cys25 on the main helix and His159 are represented in red

The structure of other thiol proteases has been determined. Actinidine possesses a sequence and a three-dimensional structure very similar to that of papain, with a small number of insertions and deletions (Fig. 12.11 below). Of the 212 amino acids in papain, 98 are identical in actinidine. The differences are unequally distributed. The sequences forming the helices A and B of the first domain and the β structures of the second domain present 69% homology; 81% of the residues situated at the interface between the two domains are conserved. In contrast, the residues present in the hydrophobic core of the protein present only 41% identity. In the median part of the chain made up of residues 82 to 116, one notes only 11% identity; the helix C presents 17% homology. Despite some differences between the sequences of papain and actinidine, these proteins possess a very high similarity in their three-dimensional structure.

The complete sequences of two mammalian proteases, cathepsins H and B, are known; the other thiol protease sequences have been partially determined. Cathepsin H presents a very high degree of homology with papain and actinidine. The residues situated around the catalytic groups are so well conserved in different thiol

proteases that whether they come from plants or mammals, they probably share a common origin (Table 12.4). The largest difference is observed in the staphylococcal enzyme which does not possess disulphide bridges. Either this enzyme has considerably diverged, or its structure results from a convergent evolution selecting the essential catalytic groups. Thus, there is a great similarity in the three-dimensional structure of known thiol proteases, the most significant differences being localised in the regions that comprise insertions or deletions. Figure 12.11 below indicates the approximate positions of deletions and insertions in some of these enzymes.

Table 12.4 Amino acids around catalytic residues of thiol proteases

	Sequence around the reactive cysteine																				
			20					**25**					30					35			
Papain	N	Q	G	S	C	G	S	C	W	A	F	S	A	V	V	T	I	E	G	I	I	K
Actinidine	S	Q	G	E	C	G	G	**C**	W	A	F	S	A	I	A	T	V	E	G	I	N	K
Bromelain (stem)	N	Q	N	P	C	G	A	**C**	W	A	F	G	A	I	A	T	V	E	S	V	A	S
Bromelain (fruit)	N	Q	N	P	C	G	A	**C**														
Ficine	Q	Q	G	Q	C	G	S	**C**	W													
Chymopapain B	R	V	P	A	S	G	E	**C**	Y													
Chymopapain	N	Q	G	S	C	G	S	**C**	W	A	F	S	T	I	A	T	V	E	G	I	N	K
Cathepsin H	N	Q	G	A	C	G	S	**C**	W	T	F	S	T	T	G	A	L	E	S	A	V	A
Cathepsin L	Y	Q	G	A	C	G	S	**C**	W	A	F	S	A	V	V	L	A	Q				
Cathepsin B	D	Q	G	S	C	G	S	**C**	W	A	F	G	A	V	E	A	M	S	D	R	I	C
Streptococcal proteinase	G	Q	A	A	T	G	H	**C**	V	A	T	A	T	A	Q	I	M	K	Y	H	N	Y
	Sequence around the reactive histidine																					
	155				**159**					165												
Papain	N	K	V	D	**H**	A	V	A	A	V	G	Y	G									
Actinidine	T	A	V	D	**H**	A	I	V	I	V	G	Y	G									
Bromelain (stem)	D	K	L	N	**H**	A	V	T	A	I	G	Y	N									
Ficine	T	S	L	D	**H**	A	V	A	L													
Cathepsin H	D	K	V	N	**H**	A	V	L	A	V	G	Y	G									
Cathepsin B	V	M	G	G	**H**	A	I	R	L	G	W	G										
Streptococcal proteinase	K	V	G	G	**H**	A	F	V	I	D	D	G	A									

12.1.2.2. ACTIVATION OF THIOL PROTEASES

The mechanism of activation of thiol proteases differs from that of serine proteases for which exists a zymogen that must undergo a proteolytic cleavage. Thiol proteases require the presence of an activator to expose the reactive sulfhydril group and permit the expression of the activity. They can be activated by cyanide or reducing agents like cysteine and reduced glutathione as well as 2,3-dimercaptopropanol. Papain is very sensitive to oxidation. The thiol group of the active centre easily participates in exchange reactions with disulphide bridges.

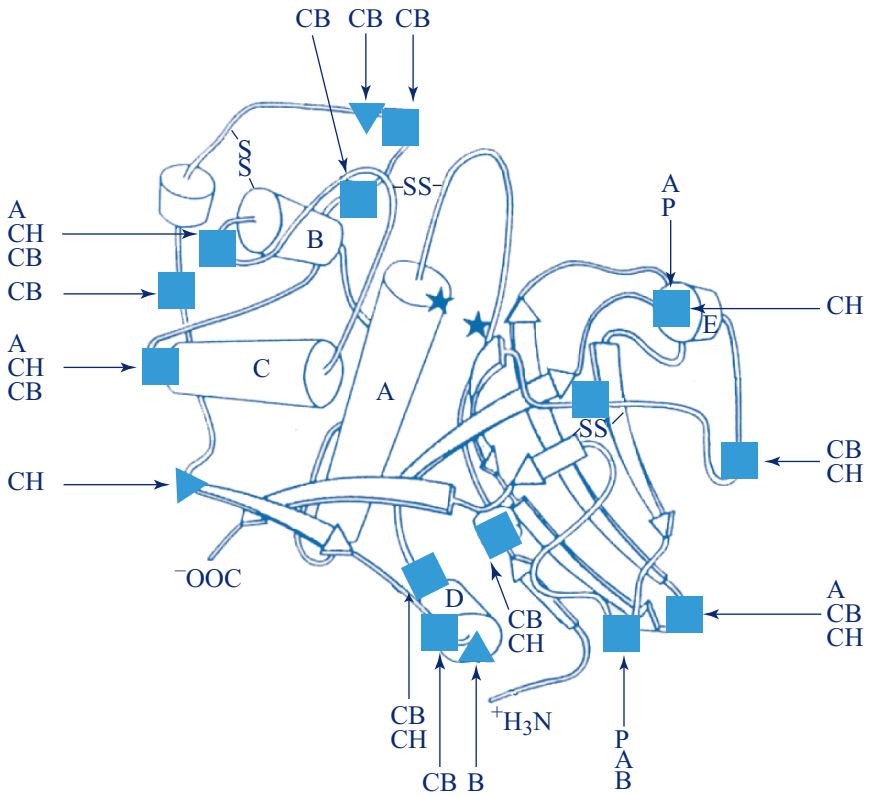


Fig. 12.11 Approximate positions of insertions and deletions in different thiol proteases: papain (P), actinidine (A), bromelain (B), cathepsin H (CH) and cathepsin B (CB), deduced from crystallography studies for papain and actinidine, and from sequence alignments for cathepsins H and B and bromelain

(From The thiol proteases: structure and mechanisms in *Biological macromolecules and assemblies*, **3**, BACKER E.N. & DRENTH J., ed. by F.A. JURNAK & A.M. MCPHERSON, 313-368. © (1984 John Wiley and Sons). This material is reproduced with permission of John Wiley & Sons, Inc.)

12.1.2.3. THE ACTIVE CENTRE

BERGER and SCHECHTER (1970) had shown that papain contains seven binding sub-sites of the substrate; four of them S_1, S_2, S_3 and S_4 are situated in the acyl part of the substrate scissile peptide bond, the three S'_1, S'_2 and S'_3 in the amino part of this. The specificity of papain and enzymes of this family is relatively broad. It requires the presence of a hydrophobic residue, preferentially phenylalanine in the sub-site S_2. Crystallographic studies carried out on complexes of papain with a series of inhibitors, chloromethyl ketones of di-, tri- and tetrapeptides permitted the study of the interactions, and to visualise the association of natural substrates to the binding site as shown in Fig. 12.12 below; hydrogen bonds between the substrate and the enzyme are represented. These studies, as well as NMR experiments, showed that the conformation of the substrate is not distorted in the complex.

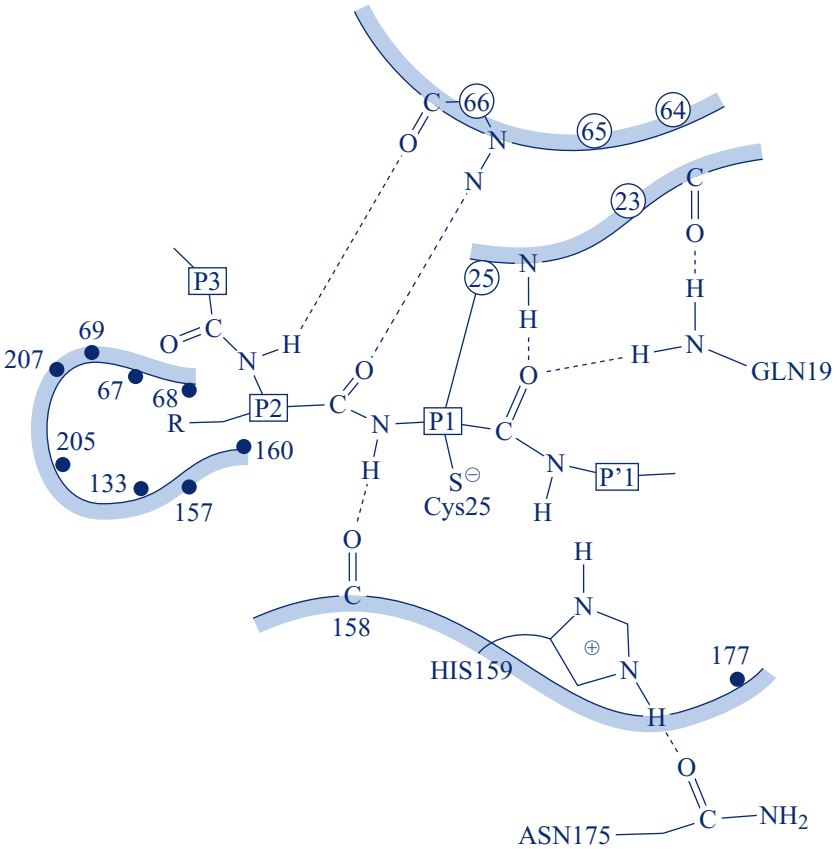


Fig. 12.12 Enzyme-substrate interactions in papain

(From The thiol proteases: structure and mechanisms in *Biological macromolecules and assemblies*, **3**, BACKER E.N. & DRENTH J., ed. by F.A. JURNAK & A.M. MCPHERSON, 313-368. © (1984 John Wiley and Sons). This material is reproduced with permission of John Wiley & Sons, Inc.)

Catalytic groups Cys25 and His159 which constitute the catalytic pair are found in a favorable position relative to the peptide bond to break. In addition, Asn175 interacts by hydrogen bond with the catalytic histidine. One finds again a catalytic triad similar to that of serine proteases (Fig. 12.13). The hydrogen bond His…Asn and the C^{β}—C^{γ} bond of His159 are approximately collinear. The electron density map revealed differences in the orientation of the cycle of the imidazole of His159 in the non-active oxidised papain in which a cysteinyl group is bound to the S^{γ} of Cys25. Following the respective orientation of His159 and Asn175, the rotation of histidine around the C^{β}—C^{γ} bond can be carried out without breaking of the hydrogen bond His…Asn.

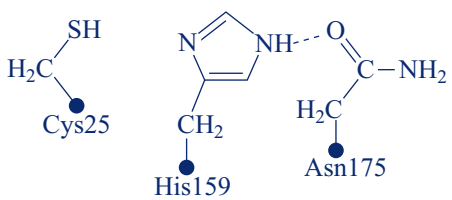


Fig. 12.13 The catalytic triad in papain
(from BACKER & DRENTH, 1984)

The amplitude of internal movements of the protein, evaluated from B factors by crystallographers, indicates that the active centre has a relatively rigid structure. In particular, the movements of the polypeptide chain carrying Cys25 and His159 are of a weak amplitude. They occur in regular structures; Cys25 is on the helix A and His159 on a β segment. In contrast, the corresponding side chains present a great mobility. The residues present at the active centre, principally Trp177 and residues 157–159, 22–24 and 64–67 deviate upon binding of an inhibitor and very likely that of a substrate, enlarging the binding site. No relative movement of domains was revealed; the active centre seems relatively rigid but can enlarge during substrate binding.

12.1.2.4. CATALYTIC MECHANISM

The catalytic mechanism implicates a nucleophilic attack by cysteine of the active centre to form a tetrahedral addition compound in which the oxyanion would be stabilised by a charge of the enzyme. The two groups NH_2 of Gln19 and NH of Cys25 would participate at the site of the oxyanion, interacting by hydrogen bond with oxygen (see Fig. 12.13). The breaking of the bond would occur by a proton transfer from a donor group to the nitrogen of the peptide bond, which would release the leaving group and would form the covalent intermediate, water intervening as a nucleophile in the final step of the reaction. The thiol group under the SH form only possesses a very weak nucleophilic power; in contrast, the S^- form is an excellent nucleophile. Numerous experimental arguments obtained by spectroscopy and NMR indicate that the active form of the enzyme corresponds to the existence of an ion pair between S^- and the protonated form of the imidazole of histidine:

$$S^-\cdots{}^+HIm$$

which is predominant between pH 4 and 8.5. At lower pH the enzyme is inactive, the catalytic groups being under the form $SH\cdots{}^+HIm$; at pH higher than 8.5, the predominant form $S^-\cdots Im$ is inactive. The existence of significant electrostatic fields created by the dipole of the N-terminal helix would have the effect of stabilising the ion pair forming the catalytic site. The role of Asn175 would consist therefore of properly orienting the imidazole nucleus of histidine.

12.1.3. ACID PROTEASES OR ASPARTYL PROTEASES

Acid proteases or aspartyl proteases constitute another class of proteases that includes gastric proteases like pepsin, gastricin, chymosin or rennet. Pepsin from pig has been studied for a long time; as early as 1930, NORTHROP succeeded with the crystallisation. It is secreted as a zymogen made of 370 amino acids. Other enzymes of this family like rennin intervene in some processes of biological control. One also finds acid proteases in some molds, penicillinopepsin of *Penicillium janthinellum* of which the molecule is made of 330 amino acids, pepsin of *Rhizopus chinensis* and of *Endothia parasitica*. No zymogen precursor was detected for these enzymes.

There are also aspartyl-proteases in retroviruses, in particular in the virus HIV 1 responsible for human immunodeficiency (AIDS). This protease plays an important role in the maturation of the virus; by limited proteolysis it produces proteins implicated in the replication of the virus.

Aspartyl-proteases catalyse the hydrolysis of peptide bonds; they are active between pH 1 and pH 7. Gastric enzymes are implicated in the degradation of alimentary proteins. Among the others, rennin has a very narrow specificity in the hydrolysis of the N-terminal peptide of angiotensinogen; this enzyme cleaves only the Leu10-Leu11 bond giving rise to angiotensin which is involved in the regulation of blood pressure.

Diverse approaches showed that the catalytic groups of these enzymes are two aspartic acids; however the catalytic mechanism was the subject of controversies, and different models have been proposed. They will be discussed in the light of structural data.

12.1.3.1. ACTIVATION OF ZYMOGENS

The activation of pepsinogen in pepsin occurs in acidic conditions (pH 3). In a first step, there is the breaking of the Leu16-Ile17 peptide bond which brings about pseudo-pepsin. The activation continues by breaking of the Leu44-Ile45 bond giving rise to an additional peptide which produces active pepsin. The protonation of the zymogen that precedes the proteolytic cleavage at pH 3 is accompanied by a conformational change. At low pH, the activation is essentially intramolecular. For strong zymogen concentrations and at pH 4, the process becomes autocatalytic; the activated molecules hydrolysing the zymogen molecules, the reaction is accelerated. Pepsinogen possesses 11 lysine residues that, with the exception of one (Lys364), are eliminated during activation. Pepsin once formed is susceptible to autolysis in acidic solution.

The group of SIELECKI (1991) solved the structure of pig pepsinogen at 1.8 Å. The results confirm conclusions from studies carried out in solution showing an important conformational change upon zymogen conversion into the active enzyme. The 44 N-terminal amino acids of the pro-segment form a long β strand followed by two approximately orthogonal helices. There are strong electrostatic and hydrophobic interactions between the pro-segment and the active centre of the enzyme. In particular, the interaction between Lys36 of the pro-segment and Asp215 of the enzyme, the hydrogen bonds between Tyr37 of the pro-segment and Asp215 of the enzyme and between Tyr9 and Asp32 of the enzyme block the access of the substrates to the active site in the zymogen. Upon activation Ile1 undergoes an ample movement of 44 Å and these interactions are broken in opening the entry of the active site.

12.1.3.2. STRUCTURAL ASPECTS

The first sequence determined entirely was that of pig pepsin; it comprises 327 amino acids and three disulphide bridges. Among mold acid proteases, only the primary structure of penicillinopepsin was entirely characterised.

The three-dimensional structures of several aspartyl proteases have been solved, including those of penicillopepsin to 1.8 Å, pig pepsin to 2.3 Å, then to 1.8 Å; those of rhizopuspepsin and endothiapepsin were obtained at medium resolution. The structure of aspartyl protease of HIV 1 was determined by the group of NAVIA (NAVIA et al., 1989). It possesses most characteristics of other aspartyl proteases with which it presents great structural analogies, the only difference being that it is a homodimer.

All acid proteases belong to the class of β proteins; besides some very short helical segments, they are essentially made of β segments (Fig. 12.14). They are folded into two structural domains; in pig pepsin, the N-terminal domain consists of 176 amino acids and the C-terminal domain 147. There is an approximate 2-fold symmetry axis between these domains; each of them is itself constituted of portions of similar structure connected also by an approximate 2-fold symmetry axis. This organisation suggested that evolution of acid proteases would result in the product of a quadruple gene formed from four copies of an ancestral unit of around 45 residues organised in a β sheet made of 6 antiparallel β segments. Diverse mutations would have generated the loops that are present today in these proteases. The sheet formed from 6 antiparallel β segments is a structure common to all acid proteases. The interdomain 2-fold symmetry axis is situated between the two central sheets as shown in Fig. 12.15 below that specifies the localisation of hydrogen bonds.

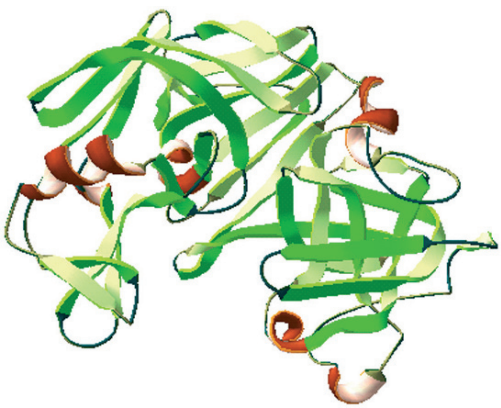


Fig. 12.14 Ribbon diagram of the structure of penicillopepsin (PDB: 3APP)

HIV 1 protease is a dimer consisting of two identical subunits of 99 amino acids which present between them a 2-fold symmetry axis. Each protomer contains the Asp, Thr, Gly triad present in all aspartyl proteases, with a 2-fold symmetry axis. Each protomer corresponds to a domain of other aspartyl proteases.

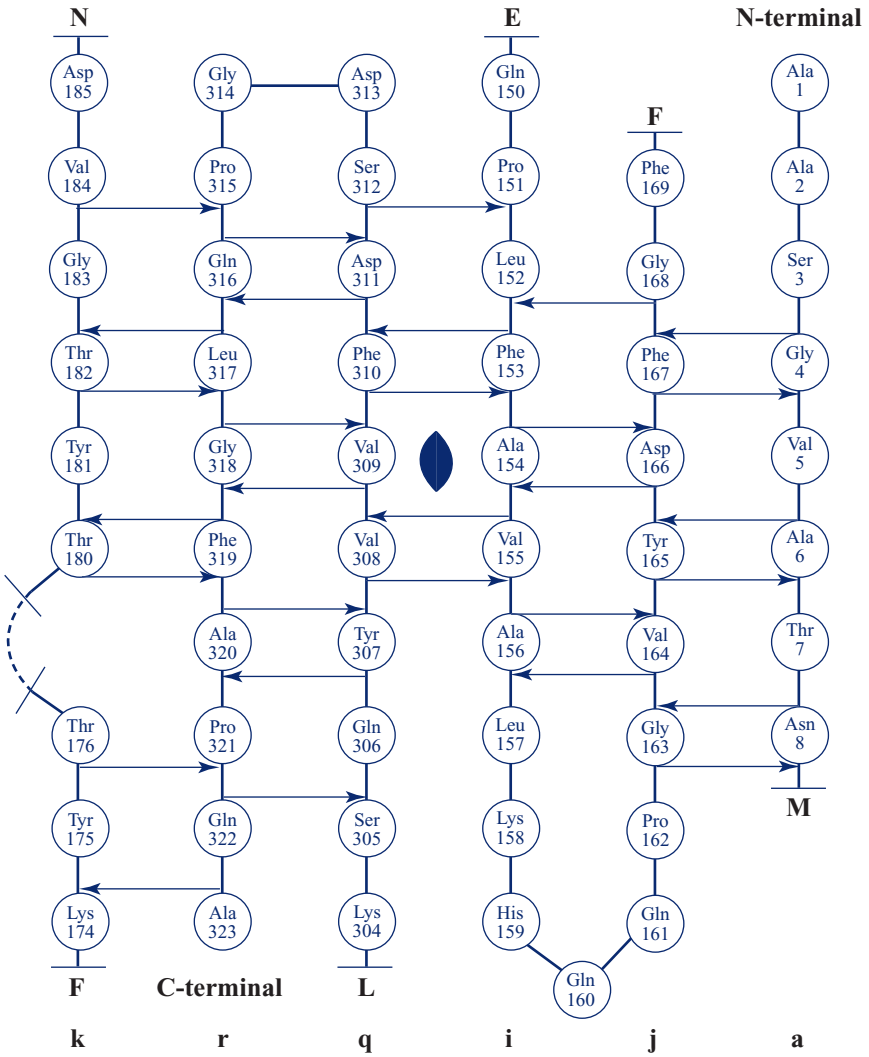


Fig. 12.15 β sheets of acid proteases with their symmetry axis

The position of hydrogen bonds is indicated by arrows

(From Aspartic proteinases and their catalytic pathway in *Biological macromolecules and assemblies*, **3**, JAMES M.N.G. & SIELECKI A.R., ed. by F.A. JURNAK & A.M. MCPHERSON, 431. © (1987 John Wiley and Sons). This material is reproduced with permission of John Wiley & Sons, Inc.)

12.1.3.3. ENZYME-SUBSTRATE ASSOCIATION

Pepsin hydrolyses peptide bonds with a specificity which implicates hydrophobic amino acids in position P_1 and P'_1. The analysis of catalytic constants and principally of the ratio k_{cat}/K_m for a series of synthetic polypeptide substrates suggests that the binding site of the substrate is stretched and can accommodate up to seven amino acids. The lengthening of the polypeptide chain, either near the N-terminal or the C-terminal extremity, increases the value of k_{cat} without modifying the value

of K_m (FRUTON, 1970). These results were confirmed with other aspartyl proteases. From the data as a whole, it has been suggested that a conformational change of the enzyme associated with constraints in the substrate could bring an important contribution to the catalytic efficiency.

Crystallographic data were obtained on complexes formed by acid proteases and inhibitors, such as pepstatin and its analogs. Pepstatin is a powerful natural inhibitor of aspartyl proteases. Its composition is the following: isovaleryl(Iva)-Val-Val-Sta-Ala-StaOH, Sta being the statin residue (acid 4S, 3S-4-amido-3-hydroxyl-6-methyl heptanoic acid). Figure 12.16 gives the formula of this compound. The affinity of this inhibitor for aspartyl proteases is very high ($K_d = 4.6 \times 10^{-11}$ for pig pepsin and 1.5×10^{-10} for penicillinopepsin). This suggested that the central statin residue is an analog of the tetrahedral transition state of true substrates. Diverse analogs such as Iva-Val-Val-StaOEt as well as pentapeptide AcPro-Ala-Pro-Ala-PheOH were used for the crystallographic study of the complexes.

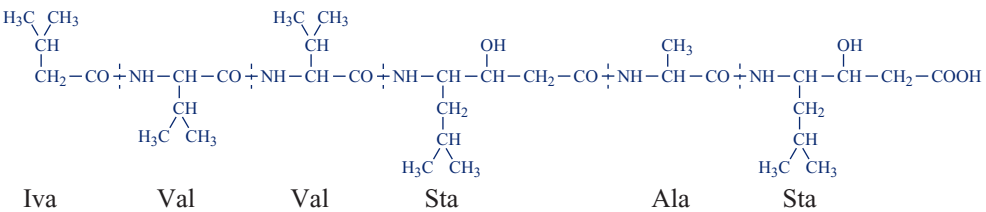


Fig. 12.16 The pepstatin molecule

The structural data clearly indicate that the inhibitor approximately adopts a β type conformation, the side chains alternating left and right of the main chain. Thus, the binding site of the enzyme can accommodate a polypeptide of 7 to 8 amino acids. The association of pepstatin and its analogs to penicillinopepsin is accompanied by a large conformational change of the loop formed by residues Ile73-Ser80 (*flap*). No displacement was observed in this region during binding of the pentapeptide. This conformational change permits the establishment of hydrogen bonds between the enzyme and the inhibitor, and it drives a proper orientation of catalytic residues Asp33 and Asp213 towards the peptide bond to be broken.

These studies also allowed the localisation of the substrate binding sub-sites, in particular S_1 and S'_1 as well as S_4, S_3 and S_2 and to provide by molecular modelling, a representation of the association of a polypeptide substrate to the enzyme (Fig. 12.17 below). In the complex, the carbonyl oxygen of the peptide bond to be cleaved points towards Asp33 and Asp213 displacing a solvent molecule. In addition, the oxygen of a water molecule (O284) is found in an ideal position for a nucleophilic attack of the carbon atom of carbonyl.

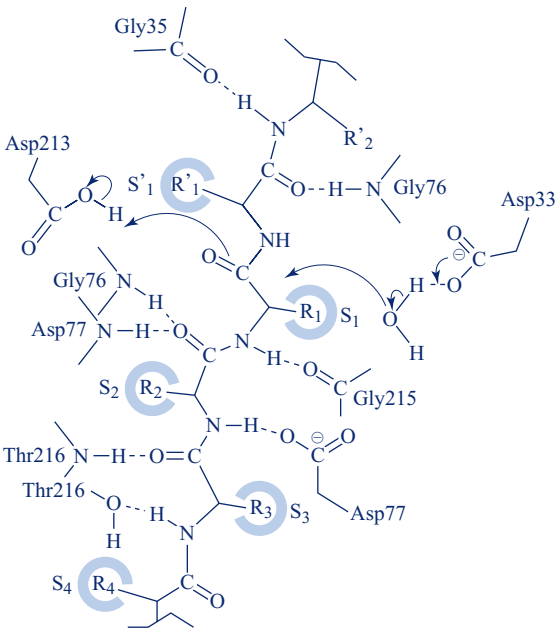


Fig. 12.17 Association of a peptide substrate to penicillopepsin
(From Aspartic proteinases and their catalytic pathway in *Biological macromolecules and assemblies*, **3**, JAMES M.N.G. & SIELECKI A.R., ed. by F.A. JURNAK & A.M. MCPHERSON, 469. © (1987 John Wiley and Sons). This material is reproduced with permission of John Wiley & Sons, Inc.)

12.1.3.4. THE CATALYTIC SITE

The two catalytic groups Asp33 and Asp213 in penicillopepsin, Asp32 and Asp215 in pig pepsin, are present in all aspartyl proteases. Several experimental results had permitted the identification of the catalytic role of two carboxyls before it was found by crystallographic studies. Kinetic studies as a function of pH indicated that the reaction rate depends on the ionisation of two groups of pK 1.2 and 4.7, respectively. Chemical modifications by affinity labelling with diazo and epoxyde compounds had allowed the identification of aspartates. The structural data confirmed these conclusions and showed that Asp33(32) and Asp213(215) are situated at the centre of the substrate binding site; the interdomain 2-fold symmetry axis goes between these two residues which are in a hydrophobic environment. Several solvent molecules were identified at the active site; some of them are displaced by specific inhibitor association with the enzyme.

The active site of the HIV 1 aspartyl protease is made up of Asp25 of one protomer and Asp25' of the other; it possesses an organisation very similar to the active site of pepsin. This active site is created by the symmetrical structure of the assembly of protomers; this represents "an elegant method for creating an active enzyme while encoding a minimal amount of genetic information" (NAVIA et al., 1989). The resolution of the structure of this aspartyl protease marked an important step in the development of therapies against AIDS, permitting the conception and synthesis of inhibitors directed against the active site.

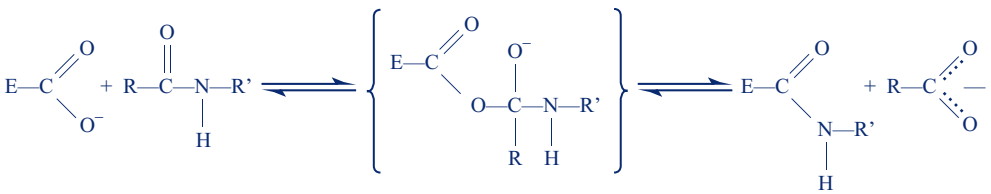
12.1.3.5. FORMATION OF THE TETRAHEDRAL INTERMEDIATE

The catalytic mechanism of acid proteases was the subject of controversies and has not yet been rigorously established. All mechanisms proposed implicate in a first step the formation of a tetrahedral addition intermediate. An electrophilic group polarises the carbonyl, to facilitate the nucleophilic attack of the carbon of the peptide bond and to stabilise the negative charge of the oxyanion. However, it is not yet established if the proton is yielded directly by one of the two groups Asp213 or Asp33 following an electrophilic mechanism or *via* a water molecule (O39) following a mechanism of general acid catalysis.

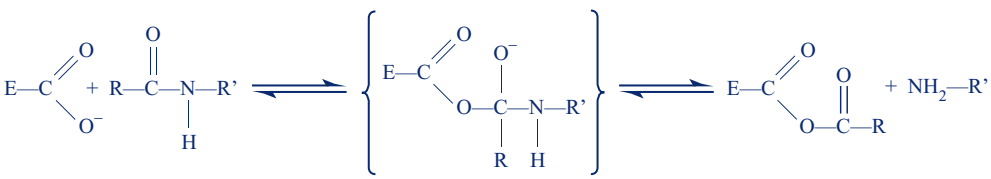
In the absence of proof, the first hypothesis illustrated in Fig. 12.17 is generally accepted (JAMES & SIELECKI, 1987). The catalysis would start by the protonation of the carbonyl oxygen by the residue Asp213. The electronic rearrangements resulting in this protonation would permit the nucleophilic attack of the carbon by the oxygen of a water molecule (O284) assisted by the COO^- group of Asp33 which would play the role of the base catalyst.

12.1.3.6. BREAKING OF THE TETRAHEDRAL INTERMEDIATE

The first hypotheses proposed the formation of a covalent intermediate between one of the two catalytic carboxyls and a part of the substrate, but the nature of this intermediate, amino-enzyme or acid anhydride, was the subject of discussions. According to the first hypothesis, an amino-enzyme is formed with the release of the acid part of the substrate according to the reaction:



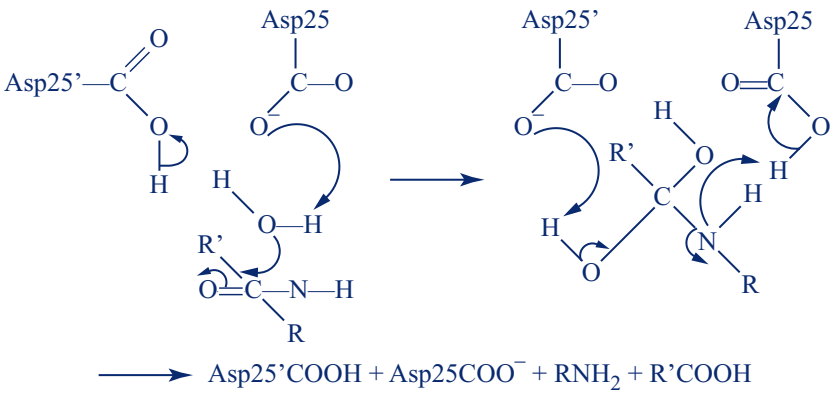
Objections based on arguments of steric order led to the rejection of this mechanism. According to another proposal, the tetrahedral intermediate would give rise to an acyl-enzyme anhydride whereas the amine part would be released according to the scheme:



The acid anhydride would be then decomposed *via* a water molecule. No attempt to detect the covalent intermediate succeeded, not even cryoenzymology experiments performed at –60°C (DUNN & FINK, 1984).

FRUTON in 1976 suggested for the first time that reactions catalysed by aspartyl proteases do not involve a covalent intermediate. Experiments of isotope exchange with ^{18}O brought arguments in favor of a mechanism implicating a direct breaking of the tetrahedral intermediate (ANTONOV et al., 1978). The hypotheses relative to the existence of a covalent intermediate had been formulated before the knowledge of the structure of the enzyme-substrate complex. Although the mechanism is not yet definitely established, the structural data permits to exclude the formation of a covalent intermediate. The breaking of the transition intermediate to give the products implicates first the protonation of the nitrogen of the NH_2 leaving group. This protonation could take place either from the solvent or by a transfer *via* Asp213; no data permits to distinguish between these two mechanisms.

Many theoretical approaches tried to elucidate the reaction mechanism of aspartyl proteases. One of them combining quantum mechanics and classical molecular dynamics was applied by the group of VAN GUNSTEREN (1996) to study the catalytic mechanism of the HIV aspartyl protease. The authors start from the principle that only one of the two Asp is protonated, which is in agreement with the difference in pK between these two groups, although it is not possible to distinguish them. The results of the simulations are in agreement with a mechanism of general acid-base catalysis in which Asp25' is protonated, according to the following scheme:



The proton transfers are predominant in all the proposed mechanisms. However, the X-ray data do not permit the localisation of the protons; it is not possible to define the limiting step of the reaction. The chemical experiments of the solvent isotope effect are difficult to interpret The method of neutron diffraction at high resolution should permit this determination.

12.1.4. METALLOPROTEASES: CARBOXYPEPTIDASE A

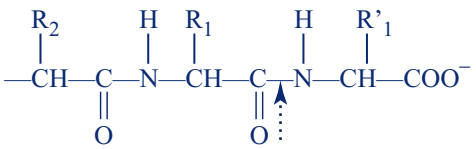
Zn^{++} metalloproteases form a very wide category of enzymes including pancreatic carboxypeptidases, the enzyme for the conversion of angiotensin very important in the regulation of blood pressure, and thermolysine that is a bacterial endopeptidase.

Pancreatic carboxypeptidases are exopeptidases which release the amino acid in the C-terminal position of a polypeptide chain. There are other carboxypeptidases from plants or microorganisms which are not metalloenzymes, in particular carboxypeptidase Y from yeast, dipeptidyl carboxypeptidase from *E. coli*, and D-D carboxypeptidases from *Streptomyces*.

Two carboxypeptidases, carboxypeptidases A and B, are secreted by the pancreas as zymogens or procarboxypeptidases. Carboxypeptidase A was the most studied and has been well-known for a long time; discovered in 1929 by WALDSCHMIDT-LEITZ, it was crystallised for the first time by ANSON in 1935. The presence of Zn^{++} in the molecule was evidenced in 1954; the enzyme as well as the zymogen contains one Zn^{++} ion per molecule. All the carboxypeptidase species resulting from the activation of the procarboxypeptidase contain Zn^{++}.

Carboxypeptidase A was the object of numerous studies, some carried out on the specificity of the enzyme, others on the role of the metal and on the nature of groups implicated in the chelation of this metal, others on groups of the active site, and on the catalytic mechanisms. The sum of works carried out on this enzyme was already impressive enough before the publication of results from crystallographic studies, which furnished solid bases for further studies.

The enzyme specificity was particularly well studied by the group of BERGMANN and FRUTON using synthetic substrates. Carboxypeptidase hydrolyses the last peptide bond at the C-terminal extremity of a polypeptide:



with the condition that the COO^- group is free and that the amino acid R'_1 is in the L configuration. Dipeptides having a free amino group are hydrolysed more slowly than those for which this group is blocked.

Although the specificity of pancreatic carboxypeptidases is relatively large, that of carboxypeptidase A implicates that the C-terminal amino acid is hydrophobic, with preference for an aromatic or branched aliphatic group, tryptophan, phenylalanine, tyrosine, isoleucine, alanine, and glycine; the presence of a proline diminishes or prevents hydrolysis. The residue R_1 can be a tryptophan, a glycine, an alanine, a methionine or a glutamate. Carboxypeptidase B is specific for charged residues (R'_1) lysine or arginine but does not recognise histidine. These enzymes are used to determine the N-terminal sequences of proteins as well as carboxypeptidase Y which releases most C-terminal amino acids including proline, but only slowly releases lysine, arginine and histidine.

12.1.4.1. ZYMOGEN ACTIVATION

According to the species, the precursor of the enzyme procarboxypeptidase A secreted by the pancreas appears either as a monomer or associated non-covalently to other zymogens. In mammals, procarboxypeptidase A forms a ternary complex with chymotrypsinogen C and proproteinase E (CHAPUS et al., 1987). The three-dimensional structure of this heterotrimer was resolved by the group of HUBER (GOMIS-RÜTH et al., 1997) to 2.35 Å by use of X-ray synchrotron. In the complex, procarboxypeptidase A occupies a central position; it is flanked by prochymotrypsinogen C and proproteinase E. Its site of activation is buried at the centre of the heterotrimer whereas the activation sites of chymotrypsinogen C and proproteinase E are localised at the surface and therefore are accessible. This organisation suggests a sequence of activation by trypsin in which chymotrypsinogen C and proproteinase E are activated first, the structure of the complex being very likely maintained and carboxypeptidase staying inactive. Next, the primary cleavage and then the secondary cleavage of procarboxypeptidase are produced bringing about the dissociation of the complex and the appearance of the activity of carboxypeptidase. It was suggested that this coordination in the appearance of proteolytic activities would have the effect of protecting carboxypeptidase from inactivation by the acidity of the gastric juice at the entrance of the duodenum. Procarboxypeptidase A is activated by trypsin cleavage at the N-terminal extremity of a peptide composed of 95 amino acids.

The three-dimensional structure of procarboxypeptidase A was determined by the group of HUBER (GUASH et al., 1992). It shows that the structure of the N-terminal peptide is made of three α helices and four antiparallel β segments forming a structural domain in the main part of the molecule (Fig. 12.18a opposite). This domain appears as an autonomous folding unit. It is suggested that this peptide is involved in the complexes that procarboxypeptidase A forms with chymotrypsinogen C or the proproteinase. In particular three aromatic residues exposed in the zymogen (Phe50, Phe58 and Phe94) would be implicated in this association. The activation of chymotrypsinogen is a slow process, the activation peptide staying associated to the enzyme until there occurs a new proteolytic cleavage at position 74. This peptide exerts an inhibitory effect on the enzyme. From structural considerations, the authors conclude that the segment connecting the activation domain to the enzyme contributes mainly to the inhibition of carboxypeptidase activity by positioning the activation domain on the substrate binding site. In fact, the activation gives rise to three carboxypeptidases of slightly different compositions, the trypsin proteolysis giving rise to α-carboxypeptidase which possesses 307 amino acids, the N-terminal extremity being an alanine. β-carboxypeptidase has 305 amino acids and a serine in the N-terminal position. γ-carboxypeptidase with 300 amino acids possesses an asparagine in the N-terminal position. The C-terminal residue is an asparagine for the three enzyme species.

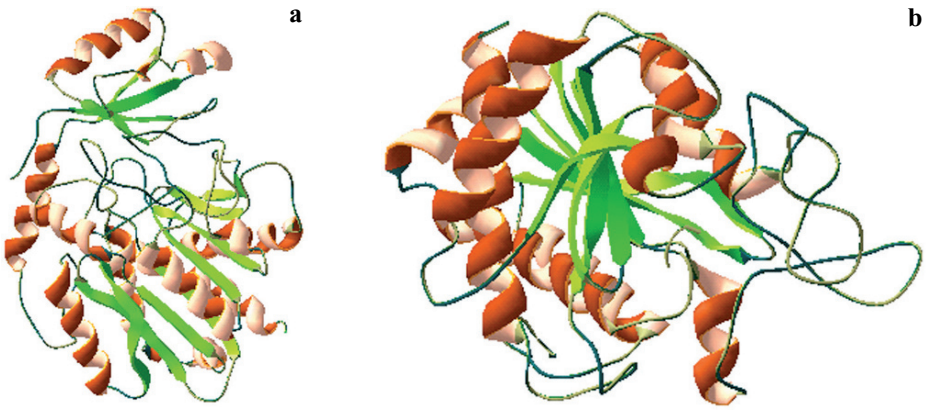


Fig. 12.18 (a) structure of procarboxypeptidase A (PDB: 1PCA)
(b) structure of carboxypeptidase A (PDB: 3CPA)

Carboxypeptidase A with molecular weight 34 000 consists of a single chain. The zymogen has a significant catalytic activity. It catalyses the hydrolysis of esters and also of peptides. It releases the C-terminal leucine of lysozyme with a rate only seven times lower than the enzyme. The binding site of substrates preexists in the zymogen. The difference in activity between the enzyme and the zymogen resides in part in the variations in pK of the carboxylate group of Glu270 which is two units higher in the enzyme (pK 7) than in the zymogen (pK 5). Such a difference suggested that the activation of the zymogen brings about a conformational change that modifies the hydrogen bonds contracted by the carboxylate of Glu270.

Carboxypeptidase B of molecular weight 34 600 is also synthesised as a zymogen which is activated by an enzymatic process. The sequences of the two carboxypeptidases A and B present 49% identity and their three-dimensional structures are very similar. The structure of procarboxypeptidase B was also resolved by the group of HUBER (COLL et al., 1991).

12.1.4.2. STRUCTURE OF CARBOXYPEPTIDASE A

The structure of carboxypeptidase A (form a) was resolved by the group of LIPSCOMB at Harvard at successive resolutions of 6, 2.8 and 1.5 Å (REES et al., 1983). This is the first metalloenzyme for which the three-dimensional structure was determined. Radiocrystallographic studies had been undertaken before the sequence was entirely determined by the group of NEURATH. The two structures were specified simultaneously by a confrontation of results obtained from each of the approaches. The free enzyme has a relatively open structure, the active centre forming a large crevasse on the surface of the molecule. Carboxypeptidase comprises large regions of β structures forming an internal sheet with parallel and antiparallel chains; 45 amino acids (about 15% of the molecule) are implicated in the β structures. The helical parts (25%) are located near the exterior of the molecule (see Fig. 12.18b). β turns of type I, II and III were characterised. Table 12.5 below indicates the parts of carboxypeptidase

organised in regular structures. Water molecules interacting with the protein (195 water molecules per molecule of carboxypeptidase) were localised. Most are found at the exterior of the protein covering 62% of the surface and in tight contact (4 Å) with the atoms of the enzyme; 24 water molecules are buried in the interior, some of them associated by hydrogen bonds. The examination of temperature factors shows that the molecule is subjected to movements of a large amplitude, varying by 5 Å from the centre of the protein to 30 Å from the surface with an average of 8 Å. The localisation of Zn^{++} and its interactions with the side chains of carboxypeptidase were only known from crystallographic studies. It is the same for the amino acid residues which constitute the active centre of the enzyme (LIPSCOMB, 1980).

Table 12.5 Secondary structure of carboxypeptidase A

[Reprinted from *J. Mol. Biol.*, **168**, REES D.C. et al., Crystal structure of carboxypeptidase a at 1.54 Å resolution, 367. © (1983) with permission from Elsevier]

α *helices*	β *segments*	β *turns (type)*
14–28	32–39	3–6 (III)
73–90	47–53	29–32 (I)
93–101	60–66	41–44 (I)
112–122	104–108	69–72 (I)
173–187	191–196	72–75 (III)
215–231	200–204	89–92 (II)
254–261	239–241	99–102 (III)
285–306	267–271	123–126 (I)
		142–145 (II)
		150–153 (II)
		162–165 (I)
		169–172 (II)
		213–216 (I)
		243–246 (III)
		259–262 (I)
		277–280 (II)
		282–285 (I)

The structure of carboxypeptidase B was resolved a little later to 2.8 Å (SCHMIDT & HERRIOT, 1976). The two carboxypeptidases present great similarities; the only differences observed in the refolding of the main chain are situated at the level of external loops.

12.1.4.3. LOCALISATION AND ROLE OF ZN^{++}

Previous studies had shown that Zn^{++} is intimately associated to the activity of the enzyme. It can only be separated from the protein by powerful chelating reagents such as 1,10 phenanthroline, 8-hydroxyquinoleine and dithizone. Other compounds

like cysteine, β-mercaptoethanol and thioglycollate also permit the elimination of Zn^{++}. The loss in activity is directly proportional to the removal of the metal. Apocarboxypeptidase, free of metal, is an inactive but stable protein. Its molecular characteristics are identical to those of the enzyme. The apoenzyme is still capable of forming a complex with a specific dipeptide. Zn^{++} therefore does not have a conformational role. It is possible to reconstitute the enzyme by addition of metal to the apoenzyme; the activity reappears proportionally to the concentration of the metal until one mole per mole of the enzyme is bound. The reconstituted enzyme can be distinguished from the native enzyme neither by its enzymatic properties, nor by its chemical properties, nor by its three-dimensional structure examined by X-rays.

◤ The activity of carboxypeptidase can be restored by other metals than Zn^{++}, such as Co^{++}, Ni^{++}, Mn^{++}, and Fe^{++}; all the metals of the first transition period restore both the peptidase and esterase activities of the enzyme. In the presence of Co^{++}, the activity is even greater than that of the enzyme with Zn^{++}. When Zn^{++} is replaced by a group II metal such as Cd^{++}, Hg^{++} or Pb^{++}, the metalloenzyme obtained still conserves an esterase activity but loses its peptidase activity. A model was proposed to account for differences between the enzyme with Zn^{++} and the enzymes in which Zn^{++} was replaced by Hg^{++} or Cd^{++}. Figure 12.19 shows the attack of the peptide bond by acid AH and base B: groups assisted by Zn^{++} in the peptidase activity of the enzyme. When this is replaced by Cd^{++} or Hg^{++}, the steric hindrance of the metal prevents the action of the basic group of the enzyme, and the amidase activity disappears. In contrast, the breaking of a more labile ester bond can use OH^- as a nucleophile group and therefore the reaction can proceed.

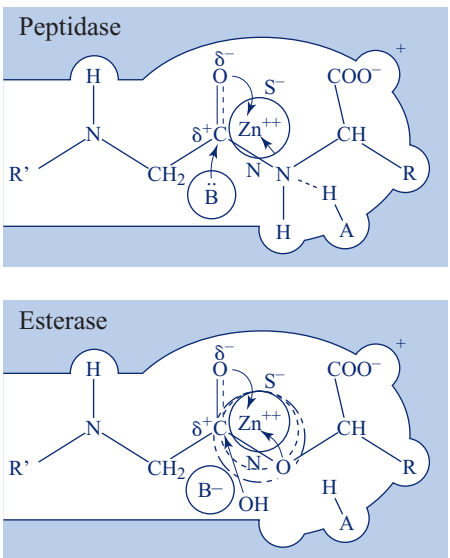


Fig. 12.19 Proposed model accounting for the esterase activity of carboxypeptidase when the Zn^{++} has been replaced by Hg^{++}

The constants of stability of carboxypeptidase complexes with diverse metals follow the following classification:

$$Hg^{++} > Cd^{++} > Zn^{++} > Cu^{++} > Ni^{++}\ Co^{++} > Mn^{++}$$

The dissociation constant has a value of 5×10^{-9} M for Zn^{++} and 1.5×10^{-6} M for Co^{++}. The analysis of stability constants as well as chemical labelling experiments had driven the group of VALLEE to suggest that the metal is chelated by a sulfur atom and a nitrogen atom. An SH group had been titrated by PCMB in the apoenzyme in the absence of metal but was masked in the presence of metal. The possible participation of imidazole had been turned down from experiments of photooxidation in the presence of methylene blue. Nitrogen had been attributed to the N-terminal α-amino group of the protein. ◢

Crystallographic studies showed that Zn^{++} is localised 25 Å from the N-terminal extremity of the enzyme. The groups that participate in chelation of the metal are the $N^{\varepsilon 1}$ nitrogens of residues His69 and His196 and the $O^{\delta 1}$ oxygen of residue Glu72, as well as a water molecule (W571). There are no cysteine residues in the molecule, but a disulphide bridge (Cys138-Cys161) far from the Zn^{++} region. It probably had been reduced in the chemical labelling experiments. His196 had first been identified with some ambiguity by crystallographic studies as a Gln or Lys residue; the authors had opted for lysine on the basis of experiments comparing the reaction of difluorobenzene with apoacetyl-carboxypeptidase and acetyl-carboxypeptidase. The resolution of the sequence established without ambiguity that residue 196 is a histidine (NEURATH et al., 1967). Furthermore, the sequence communicated by NEURATH in 1967 identified residue 69 as an isoleucine. In fact, the electron density map indicated that it was necessary to inverse the sequence 68–69 first identified as Ile-His, a result which was verified by a novel sequence study. The stability constant of the metal-protein complex obtained experimentally is perfectly compatible with the chelation by two nitrogens and an oxygen.

12.1.4.4. THE ACTIVE CENTRE

Before the crystallographic structure was established, a great number of studies using chemical labelling and enzymatic kinetics had been carried out with the goal of determining the catalytic groups of carboxypeptidase. The group of VALLEE in particular had deduced that one (or two) tyrosine residue(s) were implicated in the enzymatic activity. According to variations in the catalytic constant as a function of pH, these authors had suggested that a histidine was able to act as a general base catalyst. However, no chemical method had permitted to identify an essential histidine and it was difficult to choose which group, histidine or tyrosine, played the base catalyst role. Crystallographic studies were carried out by the group of LIPSCOMB not only with the free enzyme, but also with complexes of the carboxypeptidase with either a slightly reactive substrate or even inhibitors like p-phenyl propionate and L-lysyl-tyrosinamide. Data analysis drove to propose a catalytic role for residues Tyr248, Glu270 and Arg145. Nevertheless, later site-directed mutagenesis experiments producing the mutant Tyr248 ⟶ Phe led to a revision of the role attributed to tyrosine (see Sect. 12.1.4.6).

12.1.4.5. ENZYME-SUBSTRATE ASSOCIATION

In the absence of an effector (substrate or analog), the active centre is presented as a deep pocket or by the metaphor of the mouth of a spider having long feelers. Upon substrate binding, one of these feelers moves. The conformational change consists of a coordinated movement of residues Arg145 and Tyr248 that draw near the substrate. Residue Tyr248 is displaced by 12 Å and is drawn to a distance of 2.7 Å from the nitrogen of the peptide bond to break and 3.5 Å from the amino group of lysyl-tyrosine. This movement implicates a rotation of about 120° around the C^{α}—C^{β} bond of this tyrosine. The residue Glu270 goes away from Zn^{++} about 2 Å by a rotation around the C^{α}—C^{β} and C^{β}—C^{γ} carbons. In the free enzyme, the residue Arg145 is interacting with carbonyl 155 of the peptide backbone and residue Tyr248 with residue Glu249. These interactions are suppressed in the complex and the catalytic residues interact with the substrate. On the other hand, the cavity filled with water in the free enzyme becomes a particularly hydrophobic region when the substrate is bound. The water molecule bound to Zn^{++} is chased by the substrate; the metal interacts with the carbonyl oxygen of the bond to break and polarises it. The movement of Tyr248 has the effect of closing the cavity. The ensemble of these movements, driving an adaptation of the enzyme to its substrate, represents an example particularly suggestive of induced fit. An interaction model of a tetrapeptide with carboxypeptidase was proposed by LIPSCOMB (Fig. 12.20).

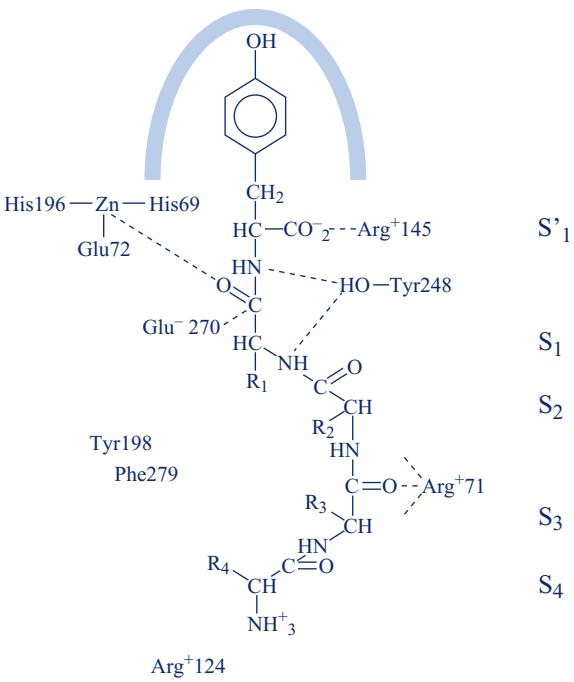


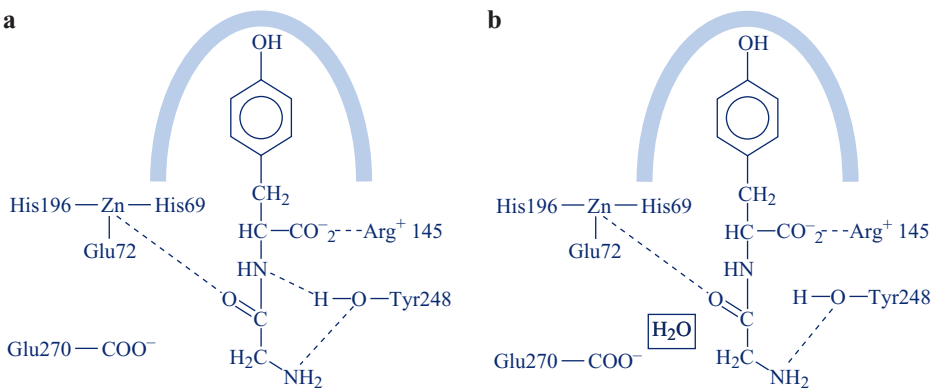
Fig. 12.20 Association of a peptide substrate to carboxypeptidase
(Reprinted from *The Enzymes*, 3rd ed., Vol. III, HARTSUCK J.A. & LIPSCOMB W.N., 1 Carboxypeptidase A, 45. © (1971) Academic press, with permission from Elsevier)

The mode of association of a ketonic substrate analog of ester, 2-benzyl-3-p-methoxypropionic acid, is similar to that of glycyl-tyrosine (LIPSCOMB, 1980). In the complex, the carboxy terminal of the substrate at site S'_1 is bound by two hydrogen bonds to residue Arg145. Residue Arg127 at site S_1 is 4 Å from residue Arg145. At site S_2 the residue Arg71 is found 4 Å from Phe279 and Tyr198 (S_2 to S_3).

The active centre of carboxypeptidase B is very similar to that of carboxypeptidase A. The Zn^{++} ligands are identical (His69, His196 and Glu72). Residues Arg145 and Glu270 are present as well as residue Tyr248 of which the orientation around the C^{α}—C^{β} bond is undetermined. The presence of a residue Asp255 at the centre of the substrate binding pocket explains the difference in specificity between these two enzymes.

12.1.4.6. CATALYTIC MECHANISMS

The breaking of the peptide bond in the COOH terminal position of the substrate implicates a nucleophilic attack of the carbonyl and an electrophilic attack on the amide group of this bond. According to the position of the different atoms (Fig. 12.21a), two hypotheses had been formulated to explain the catalytic mechanism. According to one of them, residue Glu270 would function as a general base catalyst *via* a water molecule (Fig. 12.21b). According to the other hypothesis, a nucleophilic attack would be produced by the carboxylate of Glu270 to generate an anhydride intermediate. Without being possible to choose between these two hypotheses, some steric considerations, in particular the distances and orientations of the groups of reacting atoms, seemed rather in favor of the second. Tyr248 was assumed to act by yielding a proton to the amide group of the peptide bond. In this model, Arg145 would not have a catalytic role but would be indispensable to the stabilisation of the enzyme-substrate complex forming an electrostatic interaction with the carboxy terminal of the substrate.



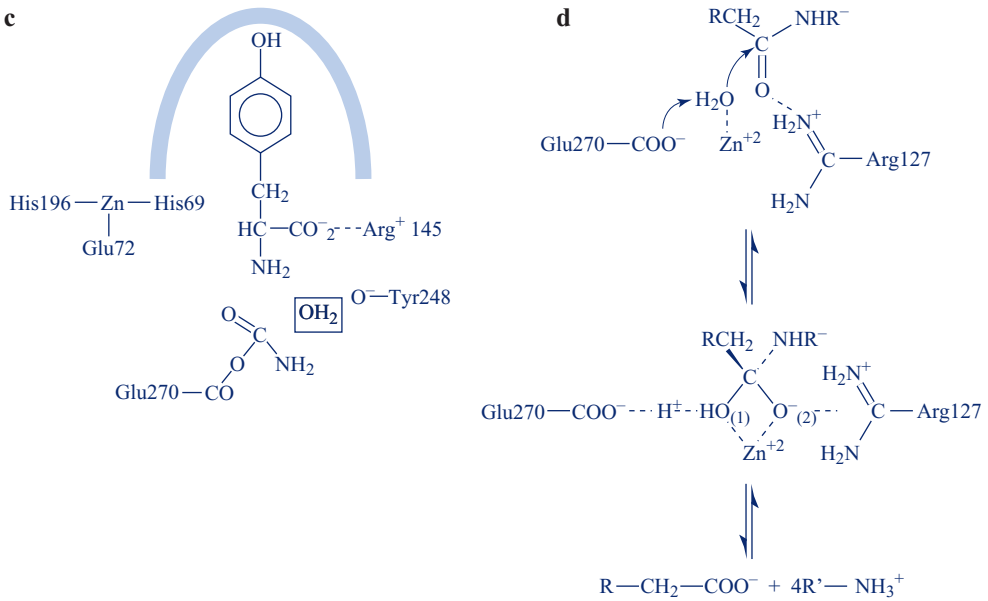


Fig. 12.21 Different catalytic mechanisms proposed for carboxypeptidase
*(**a**) position of enzyme groups with regard to the substrate – (**b**) proposed mechanism according to which Glu270 acts as a base catalyst – (**c**) hypothesis according to which Glu270 intervenes as a nucleophile to form an anhydride intermediate with the substrate (**d**) proposed mechanism in light of more recent data*
(12.21a, b, c – Reprinted from *The Enzymes*, 3rd ed., Vol. III, HARTSUCK J.A. & LIPSCOMB W.N., 1 Carboxypeptidase A, 49. © (1971) Academic press, with permission from Elsevier; 12.21d – Reprinted from *J. Mol. Biol.*, **163**, KUO L.C. et al., Catalytic conformation of carboxypeptidase A: the structure of a true reaction intermediate stabilized at subzero temperatures, 63. © (1983) with permission from Elsevier)

For a long time, it was assumed that the mechanisms implicated in hydrolysis of esters and peptides by carboxypeptidase were different. On the one hand the Tyr248 residue did not seem to participate in the reaction; the acetylation of this group was shown to have no effect on the hydrolysis of esters. In fact, Tyr248 was localised to different positions in the molecule by radiocrystallographic studies. Its localisation in the free enzyme and in the enzyme-substrate complex differs by 12 Å; in the complex, this residue contracts a hydrogen bond with the NH group of the peptide bond of the Gly-Tyr dipeptide. The interaction of tyrosine 248 with Zn^{++} was observed in an arsanilazo derivative of the enzyme; finally it was found bound to the carboxylate of a ketonic substrate. On the other hand, during ester hydrolysis, a water molecule bound to Zn^{++} would seem responsible for the hydrolysis of the anhydride intermediate suggesting that, upon their association to the enzyme, these substrates would not chase completely the water molecule bound to Zn^{++}. However, the binding site S′1 seems common to the two types of substrates. The results of the group of CRAIK (GARDELL et al., 1985) invalidate the catalytic participation of Tyr248 in the hydrolysis of peptides. The authors obtained by site-directed mutagenesis the point mutant Tyr248 ⟶ Phe and evaluated the catalytic parameters for the hydrolysis of some peptide and ester substrates. Table 12.6 below gives the results compared to those obtained with the wild type enzyme. The modified enzyme conserves its catalytic

activity; the most affected parameter is the MICHAELIS constant. The observed increase in the K_m values is compatible with the loss of two hydrogen bonds in the MICHAELIS complex. The authors concluded that the role of Tyr248 is not catalytic, but that this group is involved in the binding of substrates. Crystallographic data shows that residue Tyr248, by its oxygen, interacts tightly with the carboxylate group of the substrate; this situation suggested that its role would be to assist the release of the amino acid product of the reaction, from the active centre of the enzyme.

Table 12.6 Kinetic parameters of hydrolysis of several substrates by wild type carboxypeptidase and the mutant Tyr248 ⟶ Phe (Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **317**, GARDELL S.J. et al., 551. © (1985))

Substrate	*Kinetic parameters*	*Wild type carboxypeptidase*	*Mutant Tyr248 ⟶ Phe*
Bz-Gly-Phe	$k_{cat}(s^{-1})$	17.7 ± 0.7	21.2 ± 0.2
	$K_m(\mu M)$	39.2 ± 5.1	199 ± 4
	$k_{cat}/K_m(M^{-1} . s^{-1})$	4.52×10^2	1.07×10^2
CbZ-Gly-Gly-Phe	$k_{cat}(s^{-1})$	51.6 ± 1.9	23.1 ± 1.6
	$K_m(\mu M)$	27.1 ± 3.4	194 ± 21
	$k_{cat}/K_m(M^{-1} . s^{-1})$	19×10^5	1.3×10^5
Bz-Gly-OPhe	$k_{cat}(s^{-1})$	1194 ± 62	1138 ± 41
	$Km(\mu M)$	95.6 ± 7.7	136 ± 13
	$k_{cat}/K_m(M^{-1} . s^{-1})$	125×10^5	83.8×10^5

Kinetic studies of the reaction in solution and at low temperatures were carried out with Zn^{++} carboxypeptidase and the enzyme in which the metal was substituted by Co^{++}, in order to specify the chemical nature of reaction intermediates and the catalytic conformation of the enzyme (KUO et al., 1983; GALDEN et al., 1986). According to KUO et al. (1983), the hydrolysis of esters and amides proceeds by the same mechanism, but the formation of a mixed anhydride is the limiting step in the hydrolysis of amides, and the deacylation is limiting in the hydrolysis of esters. In addition, the rotation around the C—O bond in the tetrahedral addition compound is required for breaking this bond. However, GALDEN et al. (1986) arrive at different conclusions and, according to their results, suggest that the intermediate formed is different during hydrolysis of esters and amides. ◢

The hypothesis formulated by PHILLIPS in 1990 is represented in Fig. 12.21d. The mechanism proposed relies on crystallographic data, site-directed mutagenesis, enzymatic studies and theoretical simulations. It involves a nucleophilic attack by the water molecule bound to Zn^{++}, assisted by the carboxylate of residue Glu270 that acts as a base catalyst. The oxyanion of the tetrahedral intermediate formed in the transition state would be stabilised by the positive charge of Arg127. The role attributed to Arg127 is supported by structural data of the LIPSCOMB's group who showed that this residue forms a hydrogen bond with the tetrahedral oxygen of the transition state analogs. In addition, the replacement of this group by site-directed

mutagenesis by a methionine or an alanine brings about an important decrease of the catalytic constant. The ratio k_{cat}/K_m is lowered by a factor on the order of 10^4. The association energy of a transition state analog to the enzyme is decreased by 5.4 kcal . mol^{-1} whereas the interaction energy of an inhibitor in the fundamental state is only decreased by 1.7 kcal . mol^{-1}. In agreement with these experimental results, theoretical calculations carried out on the electrostatic contribution permitted the prediction that the positive charge of Arg127 stabilises the transition state from 6 to 8 kcal . mol^{-1}. In a further experiment (PHILLIPS et al., 1990), the authors obtained a partial restoration of the enzymatic activity by the addition of guanidine to mutant Arg127 ⟶ Ala. The first step which drives the formation of the tetrahedral intermediate is the limiting step of the reaction. The hydrolysis of ester substrates like that of peptide substrates depends also on the contribution of Arg127 to catalysis. This mechanism which is currently the best supported by an ensemble of structural arguments and functional studies, seems the most plausible.

12.2. PHOSPHORYL TRANSFER ENZYMES

As was underlined in Sect. 11.3.1.2, the transfer of phosphoryl groups catalysed by enzymes can take place by a nucleophilic attack either on the α phosphorus of ATP or on the γ phosphorus, the attack on the β phosphorus being less frequent. The first type of mechanism is illustrated by reactions that catalyse amino-acyl tRNA synthetases, the second by reactions that catalyse kinases.

12.2.1. AMINO-ACYL tRNA SYNTHETASES

Amino-acyl tRNA synthetases catalyse two successive reaction steps; the first is the formation of an amino-acyl adenylate which consists of a nucleophilic attack of the α phosphorus by the amino acid carboxylate. This first reaction proceeds through the inversion of the α phosphate, without formation of a covalent intermediate. It produces an in-line attack driving a transition state in which phosphorus is pentacoordinated, followed by the release of the pyrophosphate. This mechanism was presented in a detailed manner in Sects. 11.3 and 11.4) in the case of tyrosyl tRNA synthetase, the enzyme for which the conjunction of structural studies, molecular modelling, and site-directed mutagenesis associated with enzymology studies, drove a very precise analysis of the reaction. The reader is deferred to this chapter.

The second reaction step, the charge of the amino-acyl adenylate on the tRNA is produced by a nucleophilic attack on the carbon of the ester phosphoric bond by the oxygen of the hydroxyl group in position 2′ with the release of AMP and formation of the amino-acyl tRNA. The migration from the position 2′ to 3′ is then carried out very rapidly.

12.2.2. KINASES – PHOSPHOGLYCERATE KINASE

Kinases reversibly catalyse the transfer of the γ phosphate group of ATP to a nucleophile acceptor. For some enzymes, other purine or pyrimidine nucleoside triphosphates can replace ATP. The presence of a metallic cation, generally Mg^{++}, is necessary for the reaction. In fact, in most cases the true substrate is Mg-ATP. The role of metal is to facilitate the catalysis; it acts as a super acid or an electrophilic catalyst, probably by stabilising the transition state. There are different types of kinases that one can distinguish according to the nature of the nucleophilic substrate. For hexokinase and phosphofructokinase the nucleophilic substrate is of the ROH type. For phosphoglycerate kinase it is of the R—COO^- type, for pyruvate kinase of the R—CO—COO^- type, for arginine kinase and creatine kinase of the RNH_2 type, and for protein kinases of the protein-serine OH type. In most of these cases, there occurs a direct transfer of γ phosphate on the nucleophilic substrate. For some protein kinases, the reaction would pass by a phosphoryl-enzyme intermediate, which is not the case for most other kinases.

Phosphoglycerate kinase is the first enzyme of the glycolytic pathway which catalyses the production of ATP by the transfer of a phosphoryl group from 1,3-diphosphoglycerate (1,3 diPG) to ADP with formation of 3-phosphoglycerate (3PG) according to the reaction:

$$\text{Mg-ADP} + \text{1,3-diPG} \rightleftharpoons \text{Mg-ATP} + \text{3PG}$$

The equilibrium constant of this reaction is in favor of the production of ATP:

$$K_{eq} = \frac{(\text{Mg-ATP})(\text{3PG})}{(\text{Mg-ADP})(\text{1,3-diPG})} = 3.4 \times 10^3 \text{ at } 25°\text{C}$$

In physiological conditions, this reaction is coupled to that which is catalysed by the preceding enzyme of the glycolytic pathway, glyceraldehyde-3-phosphate dehydrogenase:

$$\text{G3P} + NAD^+ + P_i \rightleftharpoons \text{3PG} + \text{NADH}$$

G3P being glyceraldehyde-3-phosphate. It follows that the ensemble of the two reactions equilibrates approximately at the physiological concentrations of substrates (SCOPES, 1973).

Phosphoglycerate kinase exists in a great variety of species and tissues, yeasts, microorganisms, muscles and red cells of mammals. It also exists in the plant kingdom where it is implicated in the taking up of carbon from CO_2. Indeed it makes part of the multienzymatic complex made up of ribose phosphate isomerase, phosphoribulokinase, ribulose bisphosphate carboxylate or Rubisco, phosphoglycerate kinase and glyceraldehyde-3-phosphate dehydrogenase.

12.2.2.1. STRUCTURAL PROPERTIES

The sequences of more than thirty phosphoglycerate kinases were determined, by either protein sequencing or gene sequencing. They show that this enzyme was particularly well conserved over the course of evolution. The three-dimensional structures of the enzyme of horse muscle (BANKS et al., 1979), yeast (WATSON et al., 1982), *Bacillus stearothermophilus* (DAVIES et al., 1993), pig (HARLOS et al., 1992) and *Trypanosoma brucei* (BERNSTEIN & HOL, 1998) were solved. Outside of some local differences, these structures present great similarities, particularly in the C-terminal part.

The molecule of molecular weight 44 500 is made of a single polypeptide chain folded into two domains of nearly equal size (Fig. 12.22). The elements of secondary structure are organised in repeating αβ motifs; the structure is of the α/β type. Each domain is made of six parallel β segments forming an internal sheet surrounded by helices. In total, 25% of the chain is implicated in the β structures and 42% in α helices. The twelve last residues of the polypeptide chain are strongly associated to the N-terminal domain. Residues 186–189 which join helix 7 and the segment βF (helix V and βG in the yeast enzyme) form the link between the two domains. The C-terminal domain comprises the ROSSMANN fold, characteristic of the NAD^+ binding domain of dehydrogenases. There is no sequence similarity between the two domains which differ also in the detail of their three-dimensional structure. The free enzyme presents a large groove between the two domains, an open structure in the absence of substrates, as was described for other kinases, in particular hexokinase.

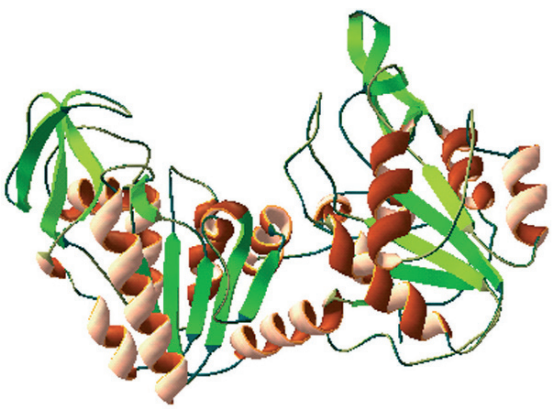


Fig. 12.22 Schematic representation of the phosphoglycerate kinase structure (PDB: 2PGK) showing the two structural domains

12.2.2.2. THE BINDING SITE OF NUCLEOTIDE SUBSTRATES

Radiocrystallographic studies were carried out on complexes formed by the enzyme with either Mg-ADP or MgATP. They reveal only a single binding site for nucleotide substrates; this site, located in the C-terminal domain, is identical for Mg-ATP and Mg-ADP. The enzymes of horse, yeast and *Bacillus stearothermophilus* show analogous characteristics. Figure 12.23 below illustrates the interactions of the horse enzyme with Mg-ATP. Bound to the enzyme, Mg-ATP like Mg-ADP, has a

different configuration than in the ATP crystal. The base is in an *anti* configuration with respect to sugar, the ribose is C′2 *endo* and the C4′-C5′ exocyclic bond is left *trans*; the triphosphate group extends far from the adenine nucleus. This latter is completely buried in a very hydrophobic region of the site, its amino group likely contracting a hydrogen bond with the carbonyl oxygen of the polypeptide chain of Gly237. The adenine binding site is defined by the parts of the main chain consecutive to the segments βG (residues Gly212, Gly213 and Ala214), βH (Gly236, Gly237 and Gly238) and βJ (Val339, Gly340 and Val341). The side chains of Leu313 and Leu256 are in contact with the adenine nucleus. Ribose is localised in a depression close to Pro338, the hydroxyls in 2′ and 3′ forming hydrogen bonds with Gly343. In the triphosphate chain, only the α phosphate interacts with the ε-amino group of Lys219 forming an ion pair. In the native enzyme, this position is occupied by a sulfate or tartrate ion according to the composition of the crystallisation solution. α and β phosphates are situated 5 Å from the N-terminal part of helix 13, probably interacting with the positive dipole of the helix. The Mg-ADP substrate has the same configuration as Mg-ATP and the same interactions with the enzyme. The only difference resides in the position of the metal ion. In Mg-ADP, this metal is located between the α and β phosphates and the carboxylate of Asp374 of helix 13. In Mg-ATP, the situation is less clear; the metal interacts very likely with the β or the γ phosphate or both. Crystallographic studies only revealed a single binding site for the nucleotide substrates, and the existence of a second site was revealed by several other studies in solution.

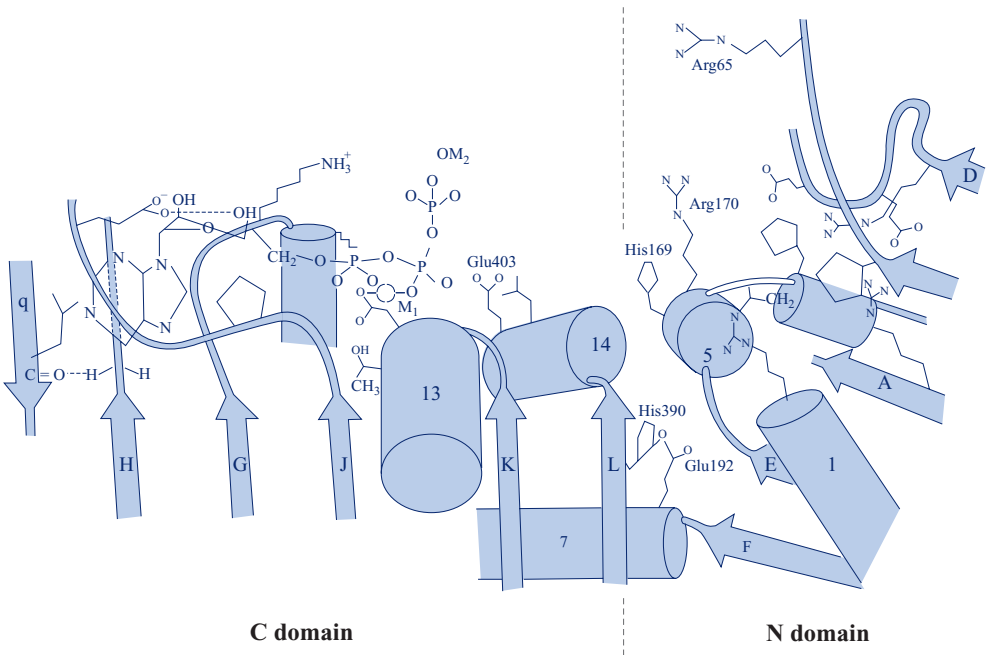


Fig. 12.23 Interactions of phosphoglycerate kinase with Mg-ATP

(Reprinted from *Phil. Trans. Roy. Soc. London*, **B293,** BLAKE C.C. & RICE D.W., Phosphoglycerate Kinase, 98. ©(1981) with permission from The Royal Society)

According to radiocrystallographic studies, nucleotide binding to the enzyme does not seem to be accompanied by significant conformational changes (BLAKE & RICE, 1981). However, studies of the fluorescence anisotropy decay of tryptophan residues of the enzyme reveal a displacement in helix 13 upon binding of Mg-ATP, permitting the free rotation of Trp335 which was impeded in the free enzyme (MOUAWAD et al., 1990).

12.2.2.3. THE BINDING SITE OF PHOSPHOGLYCERATE SUBSTRATES

The localisation of phosphoglycerate substrates was much less clear at the time of determination of the first structures. 1,3-diphosphoglycerate is too unstable to permit a crystallographic study of the complex. The only studies which attempted to localise the site of the second substrate were carried out with 3-phosphoglycerate. The localisation of this substrate on the enzyme remained however hypothetical, no precise structural data concerning the complex having been obtained. BLAKE and RICE (1981) suggested that 3PG is bound in the N-terminal domain between two arginine residues, Arg38 and Arg170 of the horse enzyme (Arg38 and Arg168 in the yeast enzyme); these two residues belong to a cluster of basic groups that include 5 arginines and 3 histidines. Data obtained by NMR showed also an interaction between 3PG and the cluster of basic groups. The substrate would be found thus at a distance of 12 to 15 Å from the γ phosphate. If such a localisation is correct, a relative movement of two domains is necessary for the chemical reaction to take place. From these considerations the group of BLAKE proposed a movement of domains around an axis localised between helix 7 and segment βF (*hinge bending motion*). A rotation of 20–25° would suffice to reduce the distance between the substrates from 12 to 5 Å; this rotation would be assured by a "scissor" movement of helices 7 and 14 (Fig. 12.24).

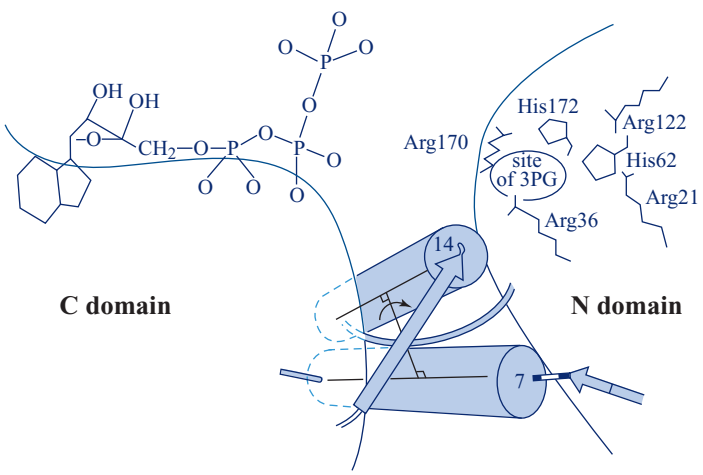


Fig. 12.24 Hinge bending movement which would bring together the two domains
(Reprinted from *Phil. Trans. Roy. Soc. London*, **B293,** BLAKE C.C. & RICE D.W., Phosphoglycerate Kinase, 98. © (1981) with permission from The Royal Society)

It would be induced by the binding of phosphoglycerate substrates as indicated in Fig. 12.25, the nucleotide substrates not bringing about conformational variations from the ensemble of the molecule.

In 1992, the group of BLAKE (HARLOS et al., 1992) succeeded in determining the structure of the complex of the pig muscle enzyme with 3-phosphoglycerate to a resolution of 2 Å. 3-phosphoglycerate is bound altogether by hydrogen bonds and electrostatic interactions with the cluster of basic residues of the N-terminal domain; the carboxylic group of phosphoglycerate points towards the C-terminal domain to which it is bound *via* a water molecule; this interaction involves nitrogen atoms of the main chain of helix 14 (Fig. 12.26 opposite). A rotational movement of 7°7 with respect to the free enzyme brings together the two domains. This value is weaker than that which had been predicted by the authors. Nevertheless, these results are in agreement with the preceding hypotheses of BLAKE.

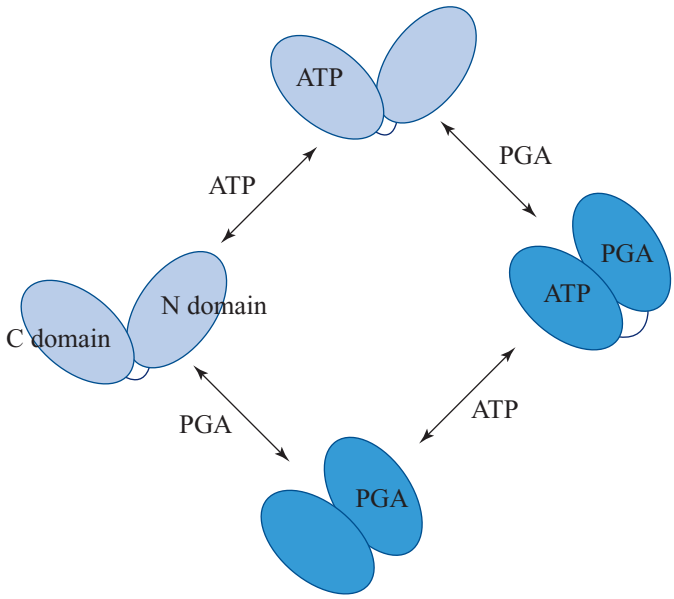


Fig. 12.25 Proposed hypothesis for the transition from the open to the closed structure of phosphoglycerate kinase

12.2.2.4. THE TERNARY COMPLEX AND THE MOVEMENT OF DOMAINS

For a long time, all the attempts to crystallise the ternary complex remained unsuccessful. In 1996, the group of BLAKE obtained the structure of the ternary complex of pig phosphoglycerate kinase with 3-phosphoglycerate and an analog of ATP, Mn-adenylimidodiphosphate. The same year, MCPHILLIPS et al. (1996) published the structure of a mutant of yeast phosphoglycerate kinase complexed with Mg-AMP-PNP and 3-phosphoglycerate. However, the movement of domains observed in the two cases was not more significant than in the binary complex with 3-phosphoglycerate and therefore was insufficient to permit the bringing together of substrates.

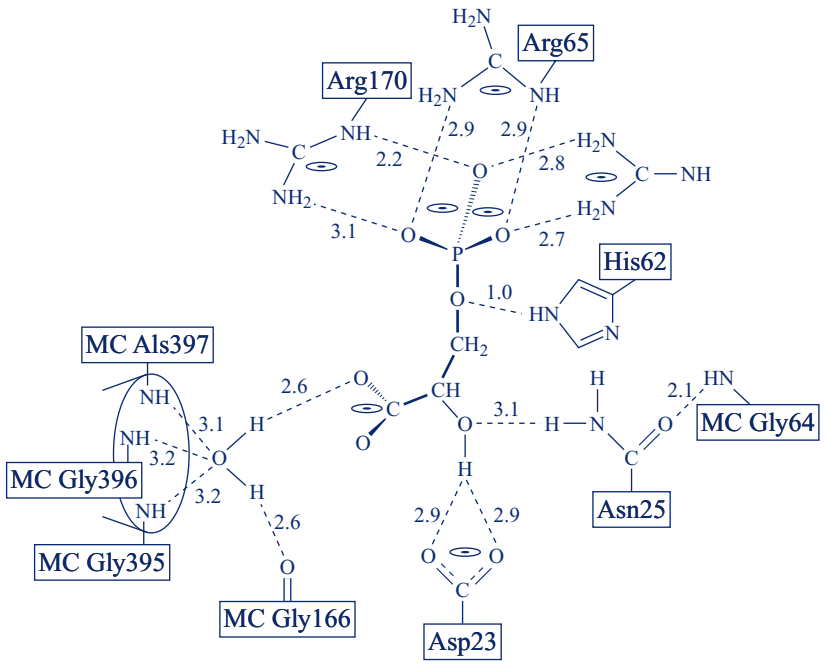


Fig. 12.26 Schematic representation of hydrogen bonds and electrostatic interactions of 3-phosphoglycerate with phosphoglycerate kinase of pig muscle

(From *Proteins: Struct. Funct. Genet.*, **12**, No. 2, HARLOS K. et al., 1992, 133–144. © (1992 Wiley-Liss, Inc.). Reprinted with permission of John Wiley & Sons, Inc.)

In return, several studies in solution had revealed major conformational movements upon binding of the two substrates. In addition, the group of PERAHIA showed by molecular dynamics simulations that the movement of the domains is an intrinsic property of the molecule that, in the absence of substrates, takes place in all directions. The binding of substrates by rigidifying the domains brings about a directionality of movement favouring the closure of the active site and bringing together the substrates (GUILBERT et al., 1995). In 1997, the group of VAN HOL (BERNSTEIN et al., 1997) obtained the structure of the ternary complex of the enzyme of *Trypanosoma brucei* in the presence of 3-phosphoglycerate and ADP; this shows a closure of the active site which brings together the substrates, thus allowing the enzymatic reaction.

12.2.2.5. CATALYTIC MECHANISM

Although phosphoglycerate kinase is a monomeric enzyme, the kinetics of the enzymatic activity do not follow the law of MICHAELIS in conditions of weak ionic strength. In return, at strong ionic strength, one observes a michaëlien behaviour. The deviation from the law of MICHAELIS recalls the behaviour of hysteretic or mnemonic enzymes (see Part V). In the case of phosphoglycerate kinase, this behaviour could result from the relative movement of domains, the return to the open

form of the enzyme upon desorption of reaction products occurring according to slow kinetics, but no experimental proof has been produced. Whatever it may be, kinetic data obtained in conditions where the enzyme follows the law of MICHAELIS, indicate that the association of substrates to the enzyme occurs following a random mechanism to give the active ternary complex.

The catalytic mechanism does not proceed via the formation of a phosphoryl-enzyme intermediate. It consists of an attack in-line on the g phosphate of Mg-ATP with inversion of configuration and direct transfer of this group (WEBB & TRENTHAM, 1980). In the transition intermediate, the γ phosphorus is pentacoordinated. Figure 12.27 represents the most probable catalytic mechanism that involves one enzyme group as a general acid catalyst and another as a general base catalyst. The nature of the residues implicated in catalysis is not yet precisely known. The participation of His169 (His167 in the yeast enzyme), Lys215 (Lys213) and Glu403 (Glu401) was suggested but not really proven. Numerous site-directed mutagenesis experiments were undertaken with the yeast enzyme, in particular by the group of MAS. Most basic residues localised to the active centre were replaced. The results obtained up to now show that only residue Arg68 is critical for the activity. Residues His167, Arg168, Arg21 as well as His62 would be implicated in the binding of the anion activator.

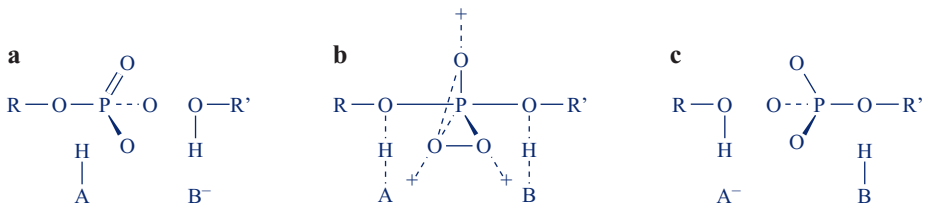


Fig. 12.27 Mechanism of the reaction catalysed by phosphoglycerate kinase
*(**a**) configuration of groups of atoms permitting the nucleophilic attack by phosphoglycerate – (**b**) transition intermediate in which the phosphorus is pentacoordinated – (**c**) products of the reaction*

12.3. GLYCOSYL TRANSFER ENZYMES – LYSOZYME

Glycosidases or glycosyl transferases catalyse the breaking of glycosidic bonds and the transfer of the glycosyl group to an acceptor which may be water. Among these enzymes, chicken egg lysozyme represents the best-known system on the structural and functional level.

The biological function of lysozyme is to hydrolyse bacterial cell walls. Bacteriological properties of egg white were recognised in 1909 by LASCHTCHENKO and lysozyme identified as responsible for these properties by FLEMING in 1922. There are numerous sources of lysozyme; besides egg white this enzyme is found in differ-

ent tissues and secretions, in particular placenta, leukocytes, tears, saliva etc. There also are phage lysozymes. Since its discovery, chicken egg lysozyme was the object of many studies. This is the first enzyme for which the three-dimensional structure had been resolved (BLAKE et al., 1965, 1967; PHILLIPS, 1967). The substrate consists of alternating units of N-acetyl-glucosamine (NAG) and muramic acid (NAM) bound by glycosidic bonds 1 ⟶ 4. Before crystallographic studies, neither the catalytic groups nor the mechanism of action of the enzyme were known. The examination of the structure of the enzyme and its complexes with inhibitors led to attribute a catalytic role to residues Asp52 and Glu35. These data stimulated a great number of following works.

12.3.1. STRUCTURAL PROPERTIES

The sequence of chicken egg lysozyme was established the same year (1963) by the group of JOLLÈS in France and the group of CANFIELD in England. The sequences of lysozyme from other species of vertebrates and bacteriophages were established later. The enzyme of chicken egg white is a small molecule of molecular weight 14 500 made of a polypeptide chain of 129 amino acids stabilised by four disulphide bridges (Cys6-Cys127; Cys64-Cys80; Cys76-Cys94; Cys30-Cys115). The three-dimensional structure (Fig. 12.28) shows that the enzyme is of the type ($\alpha + \beta$) according to the nomenclature of CHOTHIA. The molecule comprises an α structural domain with six helical regions, corresponding to 44% of the molecule, certain parts approaching a $helix_{3,10}$ conformation. Four antiparallel β segments are localised in the other domain. Residues 1-3 and 38-40 associated by two hydrogen bonds form a small antiparallel β sheet.

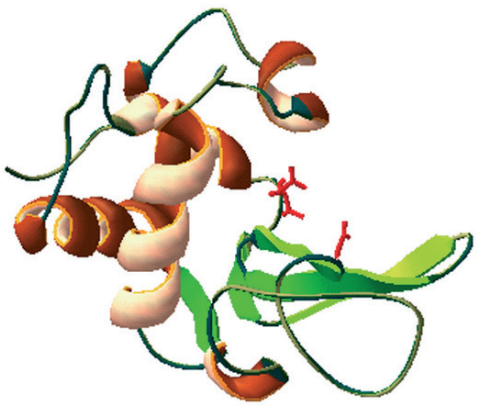


Fig. 12.28 Schematic representation of the chicken egg lysozyme structure (PDB: 6LYZ)

Figure 12.29 below represents the different hydrogen bonds that stabilise the segments of secondary structure. Three water molecules buried in the interior of the molecule were localised by radiocrystallography as well as a great number of water molecules situated at the surface.

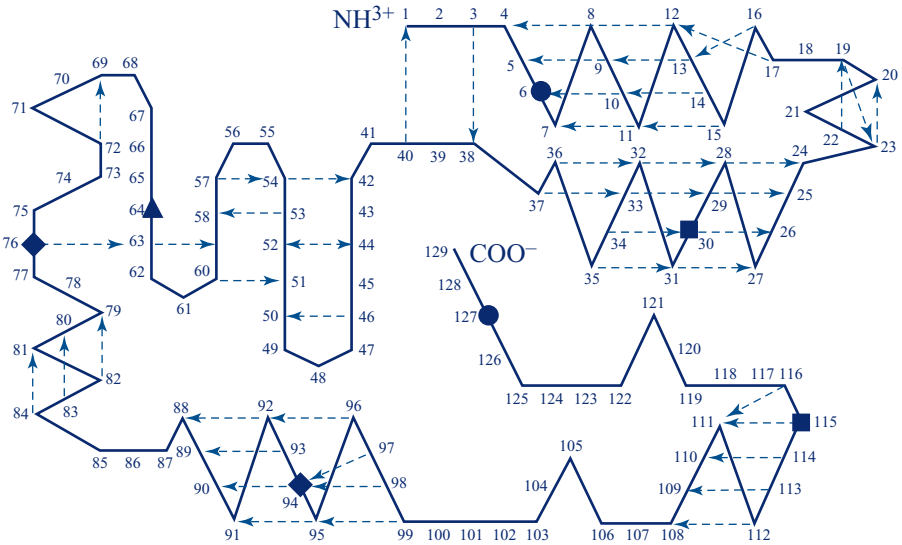


Fig. 12.29 Position of hydrogen bonds in lysozyme

(Reprinted from *The Enzymes*, 3rd ed., Vol. VII, IMOTO T. et al., Vertebrate lysozymes, 666-868. © (1972) Academic Press, with permission from Elsevier)

12.3.2. THE ACTIVE CENTRE

The active centre of the enzyme appears as a deep pocket crossing the molecule (see Fig. 12.28). Structural studies were carried out on complexes of lysozyme with diverse inhibitors, in particular (NAG)3; this latter forms a non-productive complex with the enzyme. Indeed, the minimum length of the substrate is a hexasaccharide, the cleaved bond being found between saccharide residues 4 and 5. Upon association of a trisaccharide to the enzyme, Trp62 undergoes a rotational movement of 0.75 Å towards the inhibitor, a movement which decreases the width of the binding site in the region of sub-site B. These conformational variations had previously been observed by difference spectrophotometry of the enzyme-substrate complex with respect to the free enzyme in solution. In addition, thermal relaxation studies had indicated the existence of an isomerisation of the enzyme-inhibitor complexes.

From the data obtained on enzyme-inhibitor complexes, a model of the hexasaccharide substrate was constructed and the complex was reconstituted by molecular modelling. The binding site of lysozyme substrates is formed from sub-sites A, B, C, D, E and F. Residues NAM are bound only to sites B, D and F, residues NAG to the other sites. It was proposed that non-covalent interactions between the enzyme and the saccharide ring in position D bring about the distortion of this latter which would take a half-chair configuration. Kinetic studies in solution gave an estimation of the interaction energies of different sub-sites with the substrate. Results obtained are given in Table 12.7 opposite; they seem to favor a constraint at the level of site D.

However, the theoretical calculations carried out by LEVITT (1974) have questioned the hypothesis of a substrate distortion in the fundamental state of the complex. It was suggested that the weak association of the substrate to site D results in fact in the displacement of two water molecules bound to the carboxylate of Asp52 in the free enzyme which are displaced upon complex formation. This was later confirmed by crystallographic studies (KELLY et al., 1979).

Table 12.7 Binding energies of substrates to sub-sites of lysozyme

Sub-site	*Bound residue*	ΔG_{ass} *(kcal . mol^{-1})*
A	NAG	–2
B	NAG	–3
	NAM	–4
C	NAG	–5
D	NAM	+3
	NAG	0
E	NAG	–4
F	NAG	–2

The saccharide bond is cleaved between sites D and E. It is localised between the carboxyl groups of Glu35 and Asp52 which are protonated and deprotonated, respectively (PARSON & RAFTERY, 1972).

12.3.3. CATALYTIC MECHANISM

The identification of the two residues Glu35 and Asp52 suggested two possibilities for the catalytic mechanism. One of these mechanisms had been proposed by KOSHLAND in 1953, well before the knowledge of structural data. He assumed a nucleophilic attack by Asp52 on the C_1 carbon of the saccharide ring to give an intermediate ester, a glycosyl-enzyme by analogy with the acyl-enzyme of serine proteases. According to this hypothesis (Fig. 12.30 below, pathway A), the carboxylate of Asp52 acts as a nucleophile to produce the covalent intermediate glycosyl-enzyme in a SN_2 reaction, with inversion of configuration. In a second step, a water molecule would intervene to hydrolyse the glycosyl-enzyme and a second inversion of configuration would restore the initial configuration of the enzyme.

However, this mechanism has been discarded in favor of a mechanism proposed by PHILLIPS on the basis of crystallographic data and modelling. According to PHILLIPS, the reaction would proceed with a retention of configuration according to a SN1 mechanism with the formation of a carbocation stabilised by the negative charge of Asp52 (Fig. 12.30, pathway B). The release of the product would occur following a general acid catalysis by the protonated residue Glu35. Water would intervene in a second step to form the second product of the reaction and to protonate residue Glu35.

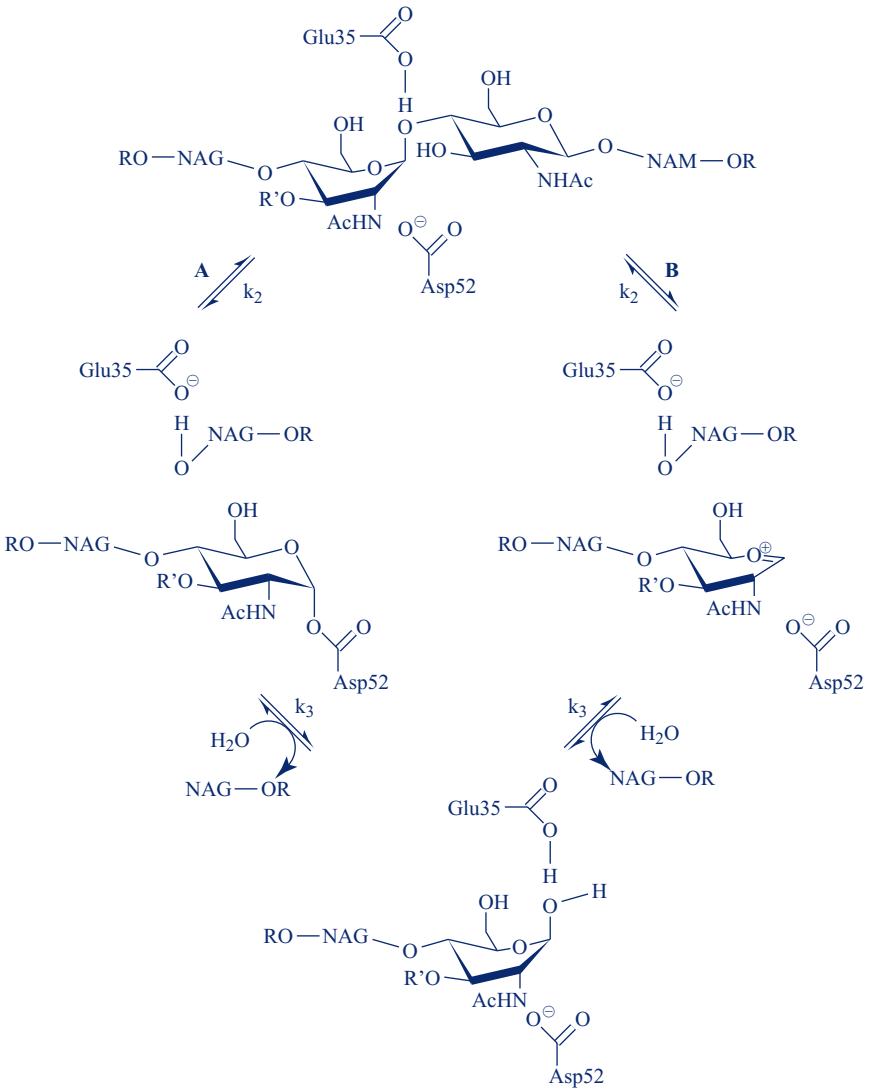


Fig. 12.30 Proposed mechanisms for lysozyme-catalysed reactions
Pathway A: formation of a covalent intermediate; pathway B: formation of a carbocation stabilised by the charge of Asp52

During more than forty years, diverse groups attempted without success to isolate and identify the intermediate. The difficulty resided in obtaining an intermediate which has a sufficiently long life time to permit the study. Recently, the proof of the validity of the first mechanism, that is the formation of a covalent intermediate, was carried out by VOCADIO et al. (2001). These authors succeeded to accumulate the covalent intermediate by combining two approaches, the first being the study of the reaction with a lysozyme mutant in which Glu35 was replaced by a glutamine (Glu35 ⟶ Gln), the second being the use of a doubly fluorinated substrate. The

covalent intermediate had a life span long enough to permit mass spectrometry studies, and to determine the crystal structure of this glycosyl-enzyme.

Such a mechanism seems valid for the ensemble of glycosidases.

This example shows once more that the knowledge of the enzyme structure alone is not sufficient to determine a catalytic mechanism.

12.4. OXYDOREDUCTION ENZYMES

In biological systems, oxidations are frequently coupled to the production of energy, and are carried out by electron transfers (see Chap. 11), which depend on the oxydoreduction potentials of the substrates (see Part I). Two oxydoreduction enzymes were chosen as examples; the first is a NAD^+ dehydrogenase, alcohol dehydrogenase, and the second a flavine enzyme, flavo-cytochrome b_2.

12.4.1. ALCOHOL DEHYDROGENASE

Alcohol dehydrogenase catalyses the oxidation of diverse primary and secondary alcohols in aldehydes and ketones in the presence of a coenzyme, NAD^+. The reaction proceeds through a ternary complex intermediate:

$$\text{enz-NAD}^+ + \text{alcohol} \rightleftharpoons \text{enz-NAD}^+\text{-alcohol} \rightleftharpoons \text{enz-NADH-aldehyde} \rightleftharpoons \text{enz-NADH} + \text{aldehyde}$$

The yeast enzyme and that of mammals differ in their specificity and their catalytic activity. Alcohol dehydrogenase of yeast is more specific for ethanol and acetaldehyde, whereas the mammalian enzymes have a much broader specificity. Among these latter, alcohol dehydrogenase of horse liver was the object of detailed structural and functional studies. Its three-dimensional structure is known with atomic resolution as well as that of enzyme-substrate complexes due to the work of the group of BRÁNDÉN in Uppsala, Sweden. The catalytic mechanism has been well established.

12.4.1.1. STRUCTURAL PROPERTIES

The horse liver enzyme is a dimer of molecular weight 80 000 constituted of two identical subunits. Each subunit possesses a binding site for NAD^+ and two strongly bound Zn^{++} atoms, one being essential to catalysis. The sequence of 374 amino acids of the polypeptide chain was determined by JÖRNVALL in 1970. Since then, other sequences of alcohol dyhydrogenases have been resolved. The human enzyme is a complex system of isoenzymes encoded by at least five genes. The sequence of the corn enzyme was deduced from the sequence of two genes ADH1 and ADH2. The known sequences of alcohol dehydrogenases of mammals, yeast, plants and bacteria permit the identification of conserved regions and differences. There is 87

to 89% homology between the human enzyme and that of horse. The horse enzyme and that of corn contain approximately 50% conserved residues; only 25% are identical between the yeast enzyme and that of mammals. The important residues at the catalytic site are conserved in all species as well as the catalytic Zn^{++} ligands. The residues that bind the coenzyme are less well conserved. Conversely, the entire region distant from the active site, which is not implicated in the association of the subunits, is found in all species. It includes residues 35, 77, 80, 86 and 87. Such a permanence suggests that these residues have an essential role in the folding of the catalytic domain.

Figure 12.31 shows the general organisation of the dimer molecule. This has the form of an ellipse of dimensions $40 \times 55 \times 100$ Å. In the absence of the substrate and the coenzyme, it presents an open structure. Each subunit is itself folded in two structural domains separated by a deep crevasse. One of the domains consisting of 143 residues binds the coenzyme; the other, made of 231 residues, contains the Zn^{++} ligands and most of the residues which are involved in the binding of the substrate. The two subunits are associated by the coenzyme binding domain.

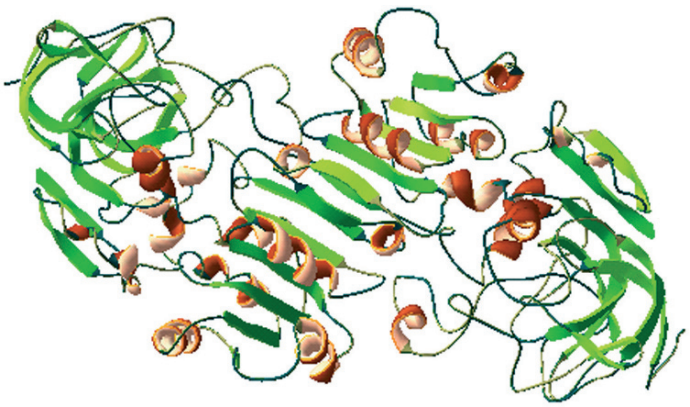


Fig. 12.31 Three-dimensional structure of horse liver alcohol dehydrogenase (PDB: 6ADH)

The NAD^+ binding domain has the typical α/β structure or ROSSMANN fold, with a regular sequence of six $\alpha\beta$ motifs, a central β sheet formed from six parallel β segments surrounded by the helices, the structure observed in most NAD^+ dehydrogenases. The comparison of the structures of malate dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase, lactate dehydrogenase and alcohol dehydrogenase shows a remarkable similarity in the binding domain of NAD^+, whereas there is no sequence homology between these enzymes. In contrast, the catalytic domains totally differ. This observation had suggested to ROSSMANN et al. (1974) a relationship between these enzymes over the course of evolution. The dehydrogenases would have evolved by fusion of genes between ancestral polypeptide chains. The NAD^+ binding domain is common to these enzymes and would have fused with different catalytic domains. The discovery of introns and exons in eukaryotic genes

suggested that the exons correspond to protein modules separated by introns. The intron-exon arrangement could thus constitute a mechanism that would permit an increase in the rate of evolution (EKLUND & BRÁNDÉN, 1984). On the basis of the position of introns in the gene of corn alcohol dehydrogenase, a diagram of the enzyme subunit is represented in Fig. 12.32.

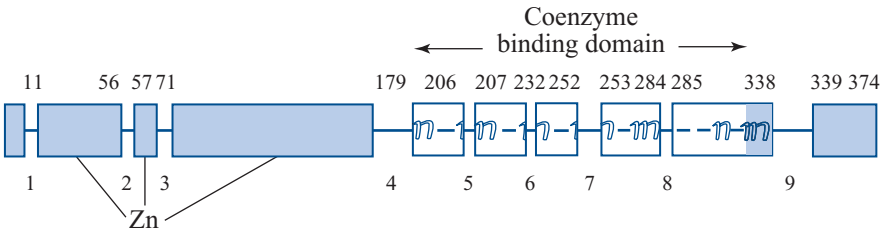


Fig. 12.32 Schematic representation of the polypeptide chain of alcohol dehydrogenase

The dark thin line *(non-framed) represents the position of the introns. The catalytic domain is indicated in* light blue

The catalytic domain is formed from two regions of the polypeptide chain localised between residues 1-174 and 319-374. The two Zn^{++} atoms are bound to ligands situated in this domain, the catalytic Zn^{++} by residues Cys46, His67 and Cys174. Each of these residues is encoded by a different exon. The second Zn^{++} atom is bound by four sulfur atoms belonging to residues Cys97, 100, 103 and 111; it is completely buried in the interior of the molecule. Figure 12.33 shows the arrangements of these two Zn^{++} atoms. The function of the second zinc atom is not known; a conformational role has often been attributed to it.

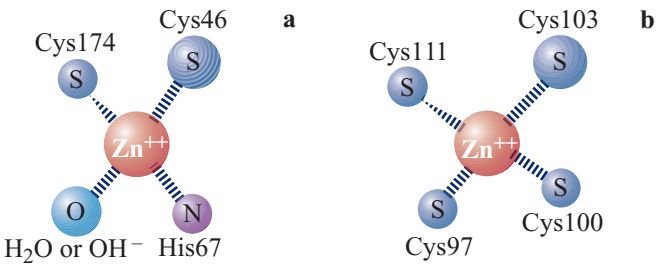


Fig. 12.33 Arrangement of the Zn^{++} atoms in alcohol dehydrogenase
(a) *catalytic Zn^{++} –* ***(b)*** *non-catalytic Zn^{++}*

12.4.1.2. CONFORMATIONAL CHANGE OF THE ENZYME INDUCED BY THE COENZYME BINDING

The binding of the coenzyme to alcohol dehydrogenase (LADH) brings about an important conformational change of the molecule that moves from an open to a closed form. The two binding domains of the coenzyme conserve the same orientation. The major change resides in a rotation of the catalytic domain with respect to the centre

of the dimer (Fig. 12.34). However, the internal structure of the catalytic domains remains identical. This conformational change is therefore a rotational movement of one domain with respect to the other; this recalls the *hinge bending movement* described for phosphoglycerate kinase. Consequently, the interaction areas between the domains are modified.

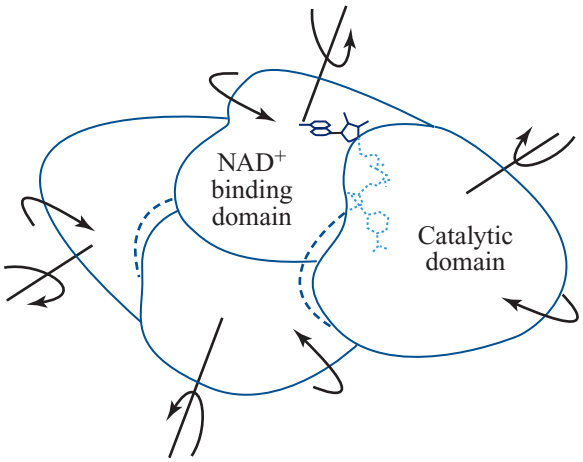


Fig. 12.34 Rotational movement of domains in alcohol dehydrogenase upon NAD+ binding (From Alcohol deshydrogenase in *Biological macromolecules and assemblies*, **3**, EKLUND H. & BRANDEN C.I., ed. by F.A. JURNAK & A.M. MCPHERSON, 73–142. © (1984 John Wiley and Sons). This material is reproduced with permission of John Wiley & Sons, Inc.)

The movements of the greatest amplitude are observed in this region. In particular, the loop comprising residues 292–298 of the binding domain of the coenzyme moves, as well as residues 51–58 of the catalytic domain (Fig. 12.35).

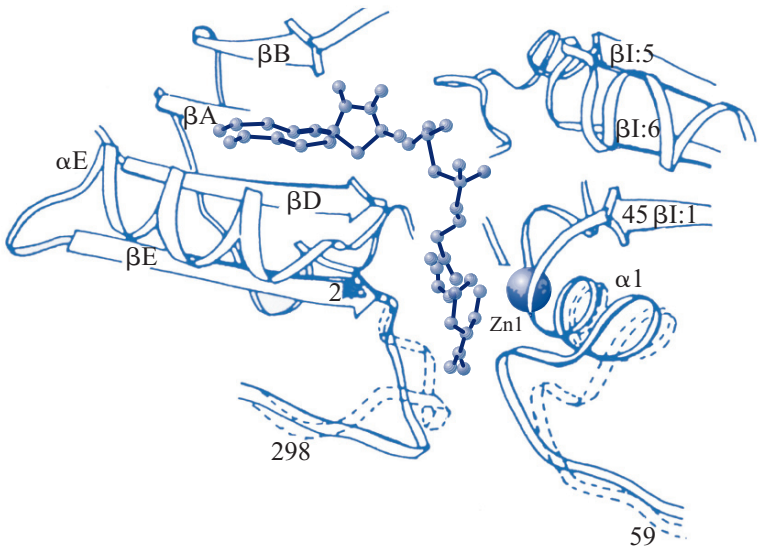


Fig. 12.35 Detail of alcohol dehydrogenase conformational variations upon NAD+ binding (From Alcohol deshydrogenase in *Biological macromolecules and assemblies*, **3**, EKLUND H. & BRANDEN C.I., ed. by F.A. JURNAK & A.M. MCPHERSON, 73–142. © (1984 John Wiley and Sons). This material is reproduced with permission of John Wiley & Sons, Inc.)

This movement has the effect of burying the coenzyme more deeply into the enzyme molecule. The integrity of the coenzyme is necessary to induce the movement; the ADP ribose portion alone does not produce any change. The dissociation of the coenzyme, necessary for the return to the open structure of the molecule, constitutes the limiting step of the enzymatic reaction.

12.4.1.3. BINDING OF THE COENZYME

The structure of the enzyme was resolved in the presence of the coenzyme, the substrates and diverse inhibitors. The interactions of the enzyme with its different ligands are therefore well determined (Fig. 12.36).

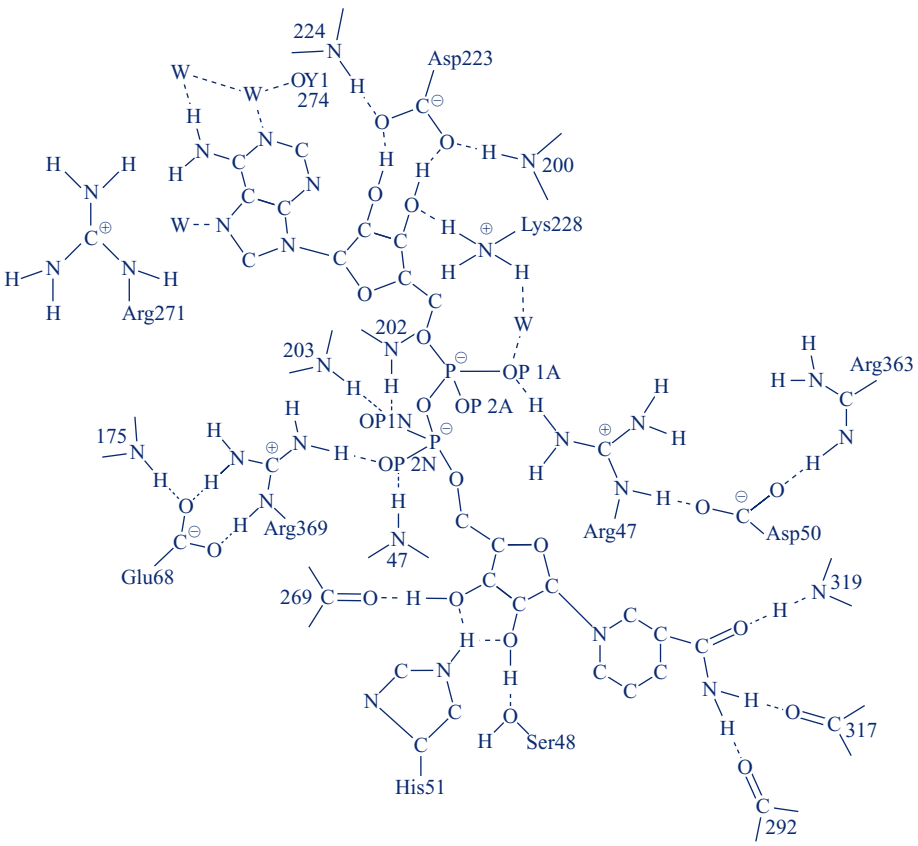


Fig. 12.36 Detail of interactions between NAD$^+$ and the protein groups at the binding site

Hydrogen bonds are indicated in dotted lines. *They correspond to distances smaller than or equal to 3.5 Å between donor and acceptor groups*

The adenosine part is bound in the coenzyme binding pocket, and the adenine nucleus is in VAN DER WAALS contact with the side chains of Ile224 and Ile269. Nitrogen 6 is in contact with Asp271. Ribose is interacting with Asp223 and Lys228. The

other parts of the coenzyme are bound in the region situated between the two domains of the same subunit. Pyrophosphate is localised in the central part of the crevasse between the two domains forming hydrogen bonds with nitrogens of the main chain and also with the side chains of Arg47 and Arg369. The dipole moments of the helices render this site very positive. The nicotinamide ring is in the crevasse at the interior of the protein 15 Å from the solvent. It interacts through one of its faces with Thr178, Val203 and Val294; the other face is directed towards the active site in proximity of the catalytic Zn^{++}. In ternary complexes, it interacts with substrates or inhibitors. The conformation of NAD^+ in the enzyme is represented in Fig. 12.35.

12.4.1.4. BINDING OF SUBSTRATES

The binding site of alcohol dehydrogenase substrates is larger and more hydrophobic than that of other dehydrogenases, in agreement with the larger specificity of LADH. The catalytic zinc is situated deeply in the interior of the subunit near the covalent bond which joins the two domains. It is found at the bottom of the substrate binding site. Zn^{++} ligands and Ser48 are the only polar groups of this region whose inner walls are made essentially of hydrophobic residues Leu57, Phe93, Phe110, Leu116, Phe140, Leu141, Leu294 and Ile318, as well as Met306 and Leu309 of the other subunit.

The mode of association of the substrate was studied by radiocrystallography in three different complexes, with trifluoroethanol, dimethylaminocinnamaldehyde, and a complex in equilibrium with bromobenzyl alcohol and bromobenzyl aldehyde. The substrate is bound to the catalytic Zn^{++} atom. It replaces a water molecule which, in the free enzyme, constitutes the fourth Zn^{++} ligand. The oxygen of alcohol interacts by hydrogen bond with Ser48 which itself is linked by hydrogen bond to His51. The system of hydrogen bonds extends to Ile269. Carbon 1 of the alcohol is near the nicotinamide ring. The other parts of the alcohol interact with the hydrophobic region of the site.

12.4.1.5. CATALYTIC MECHANISM

The enzymatic reaction proceeds following an ordered mechanism, the binding of the coenzyme preceding that of the alcohol [DALZIEL, 1975]. The transition of an open structure to a closed structure changes the properties of the active centre favoring the binding of the substrate. The mechanism of oxidation of alcohols is essentially an electrophilic catalysis by the Zn^{++}. One of the important points is the ionisation state of the alcohol in the complex. The proton cannot be localised by radiocrystallography. However, studies carried out in solution at various pH, using several substituted alcohols, led to establish that the substrate binds to the enzyme as an alcoholate ion. The system of hydrogen bonds including Ser48 and His51 would constitute a sort of shuttle for the migration of the proton. Zn^{++} would have the role of stabilising the alcoholate ion and orienting the substrate in a convenient position with regard to NAD^+. In the reverse reaction, Zn^{++} would act as an electron

attractor increasing the electrophilic character of the aldehyde, which would facilitate the transfer of the hydride ion to the aldehyde (Fig. 12.37).

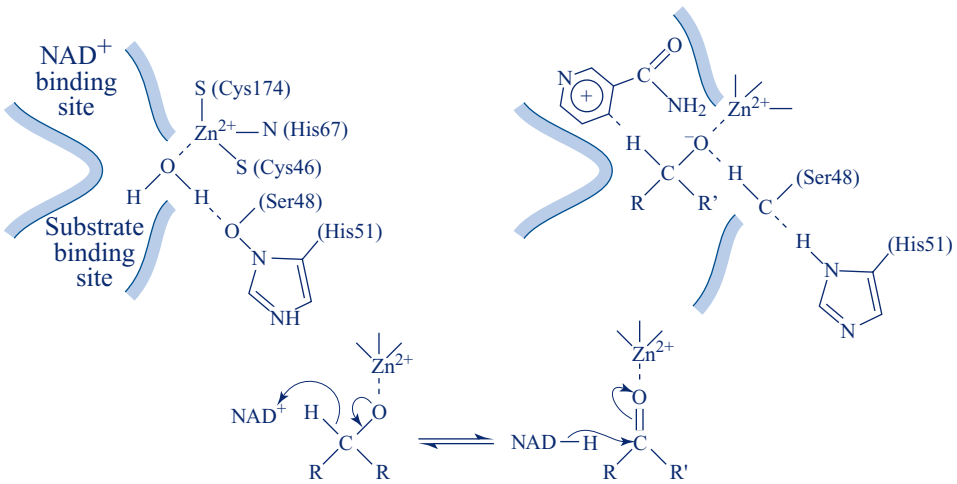


Fig. 12.37 Mechanism of action of alcohol dehydrogenase

Crystallographic data does not permit to establish a mechanism implicating a transfer of hydride. This was demonstrated by studies of isotope effects. Moreover, by molecular modelling it was shown that from very weak rearrangements in the position of bromobenzyl such as observed in the crystal, a correct geometry is assured for transferring the hydride between the substrate and the coenzyme. In addition, theoretical studies confirmed this mechanism.

12.4.2. FLAVOCYTOCHROME B_2

Flavocytochrome b_2 is a lactate ferricytochrome c oxydoreductase. This enzyme, localised in the intermembrane space of yeast mitochondria, catalyses the oxidation of lactate followed by the transfer of electrons to cytochrome c. It contains a heme, a flavine and a flavine mononucleotide or FMN, associated to the enzyme in a non-covalent manner. The amino acid sequence of the enzyme was established in France by the group of LEDERER in 1985; the three-dimensional structure was obtained to 3, then to 2.4 Å resolution by the group of MATTHEWS in the United States in 1987.

12.4.2.1. STRUCTURAL PROPERTIES

The enzyme is a tetramer of molecular weight 230 000 constituted of four identical subunits. Each chain contains 511 amino acids. The first hundred form a domain resistant to trypsin hydrolysis; this is the binding domain of heme or "cytochrome domain" whose sequence is homologous to that of mitochondrial cytochrome b_5. The four subunits are assembled according to a 4-fold symmetry axis (Fig. 12.38 below). Each subunit is itself organised into two domains, the heme binding domain

(residues 1–100), the flavine binding domain (residues 101–486), and then a C-terminal part (487–511) containing a short portion of a helix that extends to the centre of the tetramer in contact with the three other subunits. Only two of the four heme binding domains are visible on the electron density map; the two others seem to present a disordered structure (Fig. 12.39).

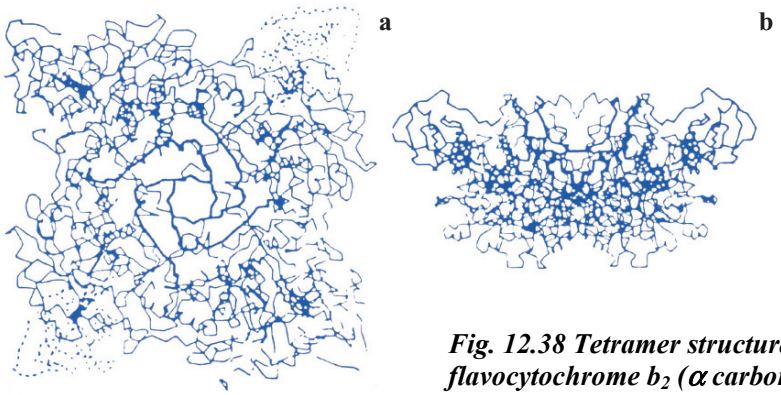


Fig. 12.38 Tetramer structure of flavocytochrome b_2 (α carbon tracing)

*(**a**) view according to the 4-fold symmetry axis of the molecule – (**b**) view according to the vertical 4-fold symmetry axis.* (Reprinted from *Flavins and flavoproteins*, MATTHEWS F.S. & ZIA Z.X., 123. © (1987) Walter DE GRUYTER and C°, Berlin)

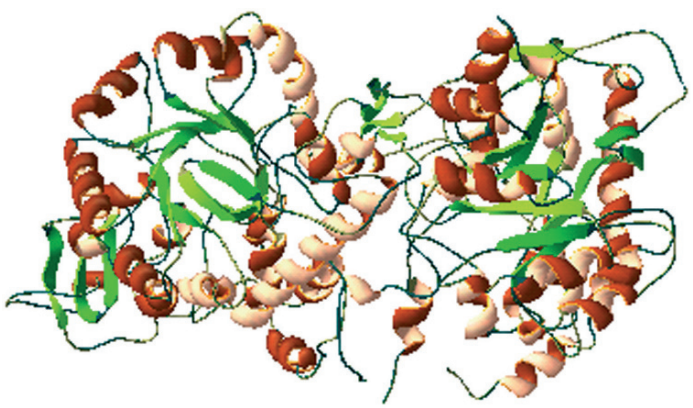


Fig. 12.39 Structure of flavocytochrome b_2 (PDB: 1FBC) showing the secondary structures

The cytochrome domains point towards the exterior. Each of them is folded similarly to the corresponding domain in cytochrome b_5; it is oriented such that the heme crevasse faces the flavine binding domain (Fig. 12.40 opposite). In the cytochrome domain, the heme crevasse has at the bottom a β sheet constituted of six β segments and the inner walls of two pairs of antiparallel helices, each pair being localised in each side of the crevasse. The flavine binding domain is constituted by the portion of the chain going from residue 192 to residue 465; it is folded into a structure of type α/β with eight αβ motifs. It comprised moreover a small segment of the N-terminal

part (101–191) with three short helices. This organisation is similar to that observed in the flavine binding domain of other enzymes, in particular glycolate oxydase and trimethylamine dehydrogenase (TMADH), as is illustrated in Fig. 12.41 below. The interface between the two domains seems to be inaccessible to the solvent.

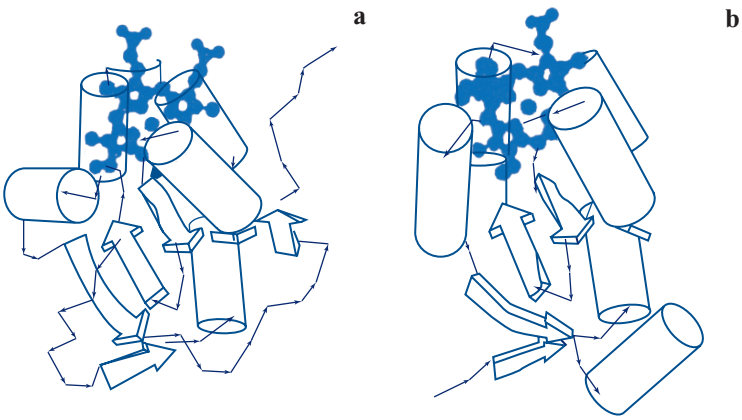


Fig. 12.40 Cytochrome domain
*(**a**) in flavocytochrome b_2 – (**b**) in cytochrome b_5*
(Reprinted from *Flavins and flavoproteins*, MATTHEWS F.S. & ZIA Z.X., 123. © (1987) Walter DE GRUYTER and C°, Berlin)

12.4.2.2. BINDING OF HEME AND FLAVINE

The heme is situated at the interface between the two domains which fits into the hydrophobic part of the N-terminal domain. The iron ligands of heme are the N^{ε} of histidines 43 and 66. The propionate groups extend towards the centre of the molecule and the flavine. Each contracts a hydrogen bond with a tyrosine, one with Tyr97 of the cytochrome domain, the other with Tyr143 of the flavine binding domain.

Flavine lies at the C-terminal extremity of the central sheet. The isoalloxazine nucleus is inclined and its plane is approximately parallel to segment β1. The pyrimidine nucleus is near the axis of the β barrel and the benzene ring near the surface. The ribityl chain is situated between segments β7 and β8. The phosphate group is near the N-terminal extremity of a short helix situated just after segment β8. FMN establishes contacts with the main chain and the side chains of residues situated in six of the eight β segments. The phosphate group contracts salt bridges with Arg413 and Arg433 (in β8) and hydrogen bonds with peptide nitrogens of residues 432 and 433 (in β8). The O^3 oxygen of the ribityl group interacts with Asp409 (in β7), the O^2 oxygen with Ser195 (in β1), Lys349 (in β5) and the carbonyl oxygen of residue 196 (in β1). The N^1 nitrogen and the O^2 oxygen of flavine interact with Lys349 (in β5) and Thr280 (in β4), respectively. The N^3 nitrogen interacts with Glu252 (in β3) and the O^4 oxygen with Ser228 (in β2). The N^5 nitrogen interacts with the peptide nitrogen of residue 198 (in β1). The plane of flavine forms with that of the heme an angle around 17°. The distance of the heme iron to the centre of the isoalloxazine nucleus is around 16 Å.

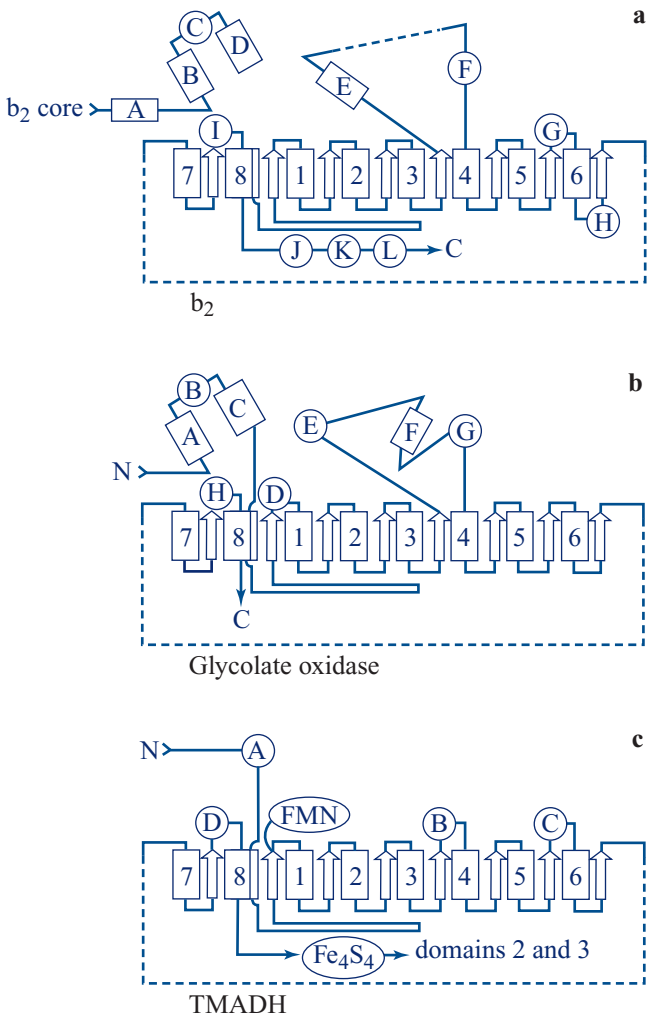


Fig. 12.41 Topology of the flavine binding domain in three FMN enzymes
*(**a**) flavocytochrome b_2 – (**b**) glycolate oxydase – (**c**) trimethylamine dehydrogenase*
(Reprinted from *Flavins and flavoproteins*, MATTHEWS F.S. & ZIA Z.X., 123. © (1987) Walter DE GRUYTER and C°, Berlin)

12.4.2.3. BINDING OF THE SUBSTRATE

An isolated zone of electron density, localised in one of the subunits where the cytochrome domain is disordered, was attributed to pyruvate. The orientation of pyruvate was deduced from this. The carboxylate group of the substrate would interact with residues Arg376 and Tyr143 whereas the ketonic oxygen would be near residues Tyr254 and His373. The substrate is found thus in a position such that its carboxylate is near the benzene ring of the cofactor, its plane being practically parallel to that of flavine (Fig. 12.42 opposite).

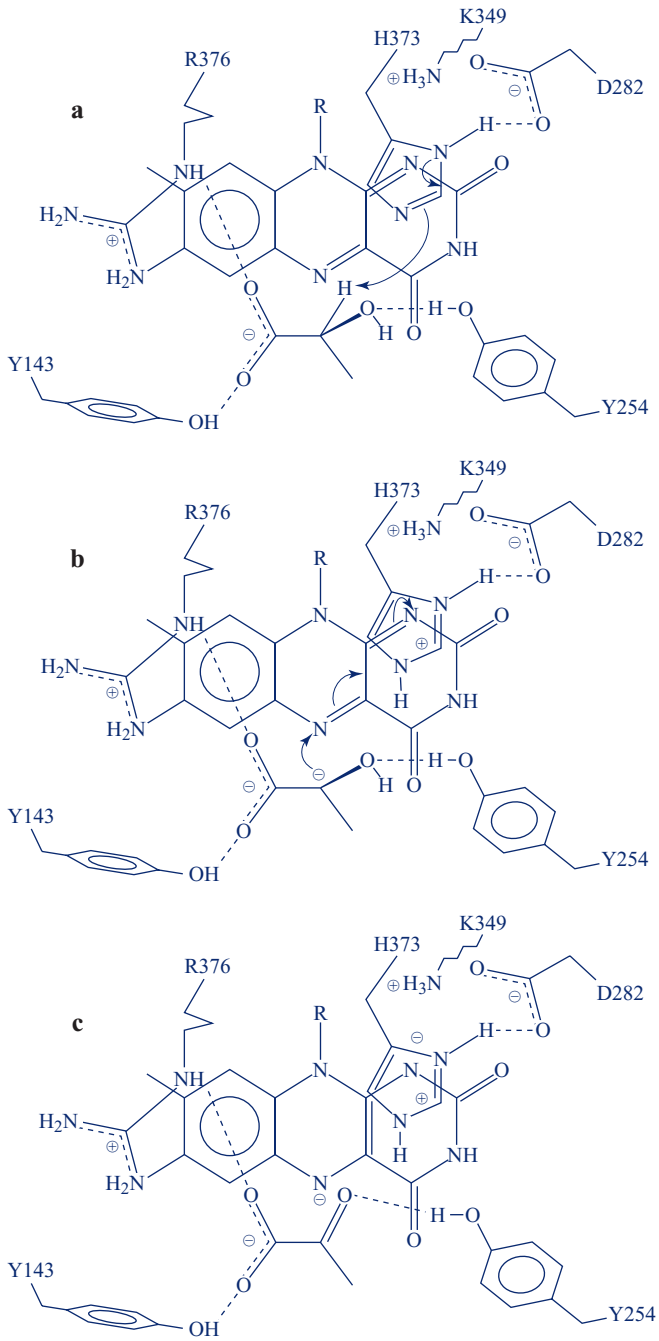


Fig. 12.42 Interactions of the substrate at the active centre of flavocytochrome b_2 and reaction mechanism *(courtesy of Florence LEDERER)*

12.4.2.4. REACTION MECHANISM

Flavocytochrome b_2 or L-lactate ferricytochrome c oxydoreductase catalyses the electron transfer from lactate to cytochrome c. The first step is the dehydrogenation of lactate to give pyruvate with reduction of flavine; it occurs according to a $2e^-$ transfer step process. Reoxidation of flavine takes place by a twice $1e^-$ transfer step process, the transfer being carried out from flavine to the heme, then to cytochrome c. Detailed kinetic studies permitted the establishment of diverse steps of the reaction and the determination of their rate constants (CAPEILLÈRE et al., 1975). Figure 12.43 illustrates the reaction pathway.

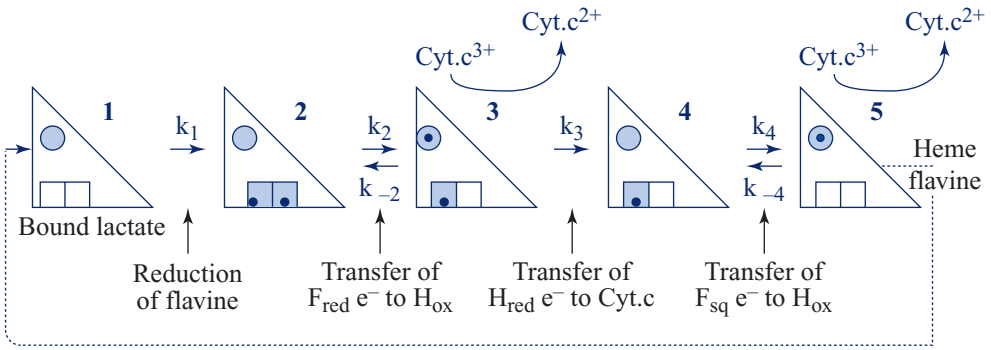


Fig. 12.43 Reaction pathway of the electron transfers from lactate to cytochrome c
(Reprinted from *Eur. J. Biochem.*, **54,** CAPEILLIAIRE-BLANDIN C. et al., Flavocytochrome b_2: Kinetic Studies by Absorbance and Electron-Paramagnetic-Resonance Spectroscopy of Electron Distribution among Prosthetic Groups, 549. © (1975) with permission from Blackwell Publishing)

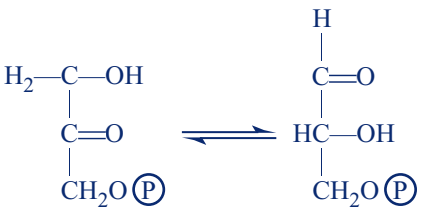
The first step of the reaction is the abstraction of a hydrogen in α. From an ensemble of experiments including the use of halogenated substrates such as bromolactate the group of LEDERER proposed a mechanism which implicates the extraction of a proton and the formation of a carbanion intermediate (see Chap. 11). The structural data confirmed this reaction mechanism and permitted specifying the details as well as the role of different enzyme groups in the reaction (see Fig. 12.42). This would implicate the following events: the binding of the carboxylate of the substrate by an electrostatic interaction with Arg376 and a hydrogen bond with Tyr143, a general acid-base catalysis on the C_2 by His373. Tyr254 stabilises the transition state. The entry of electrons in the cofactor is facilitated by Lys349 of which the positive charge stabilises the anionic character of the reduced flavine.

The transfer of electrons between flavine and heme is carried out certainly at the interior of a same subunit. Indeed, the distance between the edges of heme and flavine being 10 Å and their respective orientation being near coplanar, such a situation favors a rapid transfer of electrons as is revealed by kinetic studies (75–500 s^{-1}). The distance between the cofactors belonging to two different subunits is at least 45 Å and therefore an inter-subunit transfer would be much too slow.

The intermolecular transfer of electrons between the heme of flavocytochrome b_2 and cytochrome c is the last step of the reaction. These two proteins are capable of associating to form a complex with a stoichiometry of a heme c by heme b_2 facilitating thus the electron transfers.

12.5. TRIOSE PHOSPHATE ISOMERASE

Triose phosphate isomerase is an enzyme of the glycolytic pathway which catalyses the interconversion of dihydroxyacetone phosphate into D-glyceraldehyde-3-phosphate according to the reaction:



In the metabolic pathway, the reaction is favored in the forward direction, glyceraldehyde-3-phosphate being rapidly transformed. This enzyme is an aldose-ketose isomerase. The structures of several triose phosphate isomerases have been determined and many functional studies were reported.

12.5.1. STRUCTURE OF THE ENZYME

The amino acid sequences of triose phosphate isomerases (TIM) were determined in a great variety of prokaryotic and eukaryotic organisms. The first three-dimensional structures known for TIM were those of the chicken muscle enzyme (BANNER et al., 1975) and yeast (ALBER et al., 1981) solved to 2.5 and 3 Å, respectively. Since then, the structures of other triose phosphate isomerases were resolved: those of yeast (1.90 Å), of *Leishmania mexicana*, and of *Vibrio marinus*. In these different species the structures are very similar. The enzyme is a symmetrical dimer of molecular weight 53 000. The polypeptide chain of the chicken enzyme contains 247 amino acids; that of the yeast enzyme 248.

The three-dimensional structure (Fig. 12.44 below) is very typical with alternating β segments (22%) and α helices (55%) organised into a cyclic structure of type $(\alpha\beta)_8$, referred to as the *TIM barrel*. The practically parallel β segments are internal to the molecule, the α helices being situated at the periphery. These regular structures are practically superposable in the yeast and chicken enzymes, the differences residing in the position of irregular external loops.

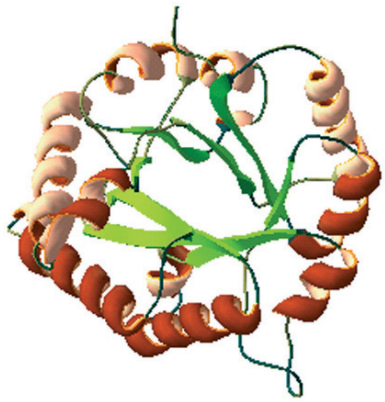


Fig. 12.44 Three-dimensional structure of a triose phosphate isomerase subunit (PDB: 8TIM) showing the $(\alpha\beta)_8$ or TIM barrel structure

12.5.2. STRUCTURE OF THE ENZYME-SUBSTRATE COMPLEX

The structure of the complex formed by triose phosphate isomerase and dihydroxyacetone phosphate was solved to 3.5 Å for the yeast enzyme and to 6 Å for the chicken enzyme. Although in the last case the resolution is not excellent, the structure of the complex formed by the yeast enzyme with a competitive inhibitor, phosphoglycohydroxamate was determined to 1.9 Å resolution. Previously, chemical labelling studies with affinity labels had permitted to attribute a catalytic role to Glu165. The structural data revealed that the substrate is effectively juxtaposed to residue Glu165. Residue His95 situated at the N-terminal extremity of the helix formed by residues 95–103 is in a position allowing its interaction with the oxygens in C_1 and C_2. It can also form a hydrogen bond by its nitrogen atom with the NH in the main chain of residue Glu97. Residue Lys12 of the yeast enzyme (Lys13 in the chicken enzyme) can also interact with the oxygen in C_2 or with that of the phosphoric ester. Residues 168 to 177 which belong to a mobile external loop, have their position modified and their mobility decreased in the complex.

12.5.3. REACTION MECHANISM

The dimer is the functional unit, however the enzymatic reaction follows the MICHAELIS law and does not show any cooperativity. **This enzyme approaches "catalytic perfection", the value of k_{cat}/K_m being on the order of magnitude of the diffusion rate of reagents ($k_{cat}/K_m = 2 \times 10^8$ mol^{-1} . s^{-1}).**

The reaction proceeds through the formation of a cis-enediol intermediate [FERSHT, 1985]. Its existence was evidenced by isotope exchange experiments showing that the reaction proceeds by proton exchange with the milieu, excluding a direct transfer. Isotope effect experiments with tritium demonstrated the intramolecular transfer of the proton between C_1 and C_2. It had been deduced that this transfer is catalysed by a basic group of the enzyme; the proton is transferred on the same face of the substrate driving the formation of a cis-enediol. The conjunction of results from

kinetic and spectroscopic studies had suggested the involvement of a residue susceptible to act as an acid catalyst according to the scheme given in Fig. 12.45.

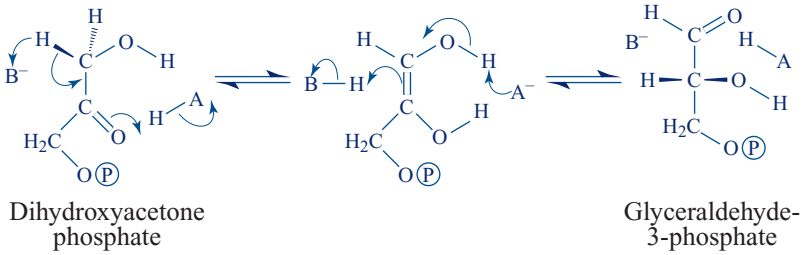


Fig. 12.45 Proposed mechanism for the reaction catalysed by triose phosphate isomerase mediated by an acid catalyst

Isotope exchange data (NICKBARG & KNOWLES, 1988) clearly showed that a rapid proton exchange with the solvent takes place at the level of the cis-enediol intermediate associated to the enzyme (Fig. 12.46).

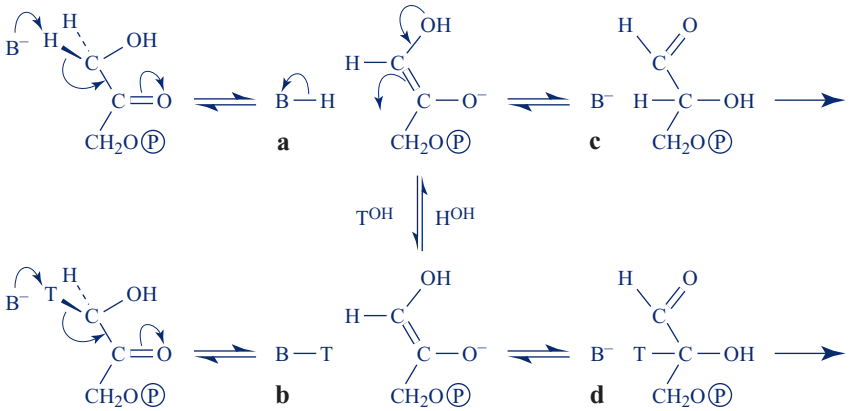


Fig. 12.46 Proton exchange with the solvent in the cis-enediol intermediate

(Reprinted with permission from *Biochemistry*, **27,** NICKGARD E.B. & KNOWLES J.R., 5939. © (1988) American Chemical Society)

A reaction mechanism was proposed on the basis of the crystallographic data; it involves a basic group of the enzyme, Glu165 which catalyses the extraction of the proton from C_1 and an acid group attributable to protonated His95 or maybe Lys12. NMR experiments showed indeed that His95 is not titrated between pH 5.4 and 9; however no formal attribution of non-titratable histidine residues has been carried out. The results from crystallographic studies are compatible with the establishment of a hydrogen bond between residue His95 and the hydrogen of the nitrogen amide of Glu97. On these bases, a mechanism was proposed in which the non-protonated histidine would catalyse the enolisation of the substrate by electrostatic stabilisation. In this case, there would not be a proton transfer from His95 to the oxygen. This alternative is presented in Fig. 12.47 below (a and b).

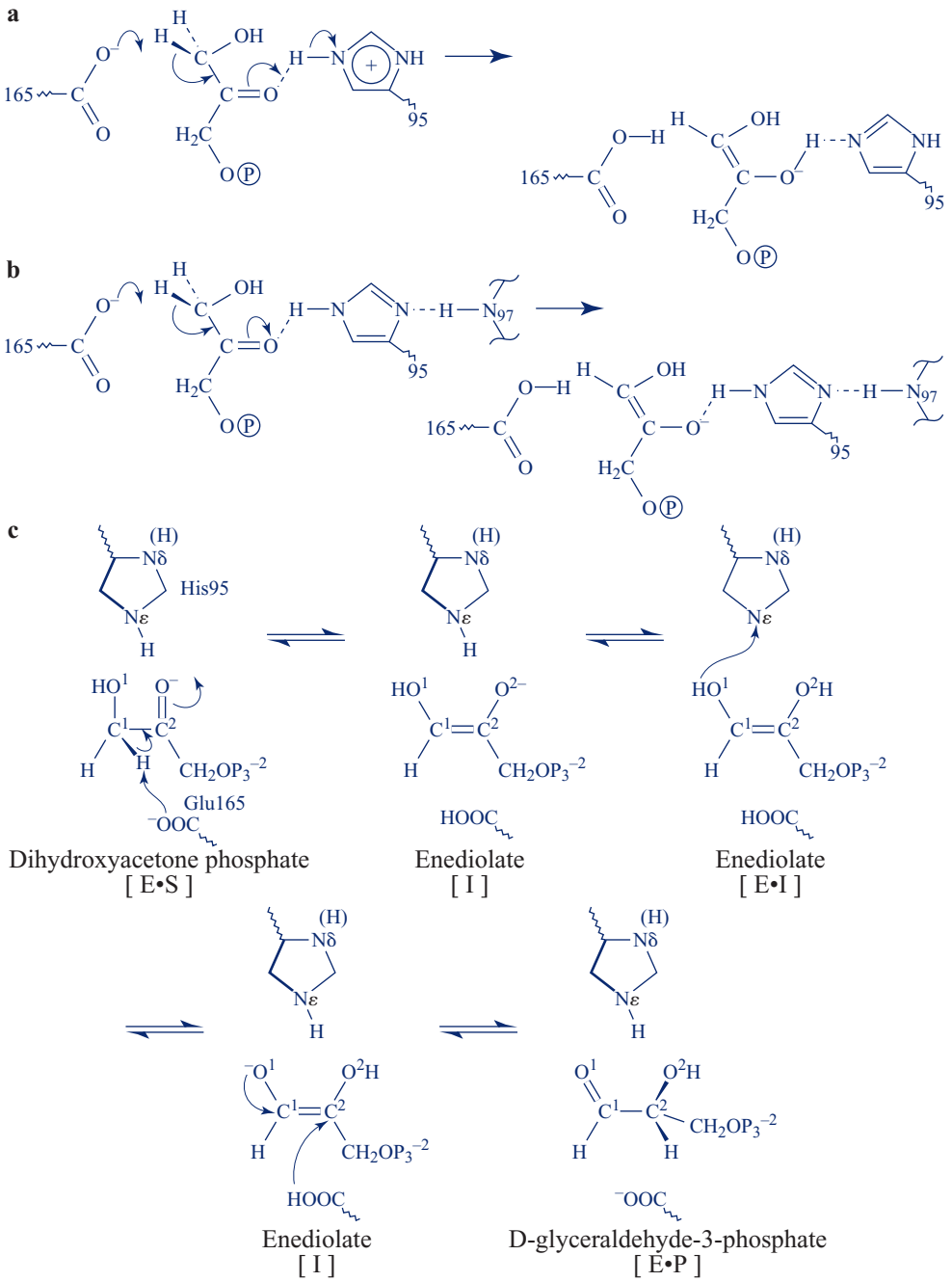


Fig. 12.47 Reaction schemes proposed from both crystallographic studies and NMR data *(**a**) histidine 95 would play the role of an acid catalyst – (**b**) histidine 95 would play the role of an electrostatic stabiliser – (**c**) histidine 95 under its neutral form would play the role of an acid catalyst* (Reprinted with permission from *Biochemistry*, **27**, NICKBARG E.B. et al., 5948. © (1988) American Chemical Society)

In order to specify the role of His95, an approach using site-directed mutagenesis (NICKBARG et al., 1988) consisted of replacing this histidine by a glutamine (mutant H95Q) in the yeast enzyme. This mutation brings about a decrease in the specific activity by a factor of 400. The inhibition constants determined for substrate analogs, 3-phosphoglycohydroxyacetate and 2-phosphoglycolate, indicate that these analogs have a respective affinity of 8 and 35 times weaker for the mutant than for the wild type enzyme. The analysis of isotope effects shows that there is practically no exchange between the intermediate and the solvent neither for the forward reaction, nor for the reverse reaction. These results suggested that the role of His95 would be to facilitate, by electrophilic assistance, the extraction by Glu65 of the proton in C_1 from dihydroxyacetone phosphate bound to the wild type enzyme. In the mutant, Glu165 would achieve this extraction without the catalytic assistance of His95 and therefore less efficiently, and the enediolate intermediate would be less stable.

The catalytic mechanism remained uncertain however. In 1991 KOMIVES et al., carried the proof of electrophilic catalysis by histidine 95. The experimental approach was driven by FOURIER transform infrared spectroscopy and the crystallographic analysis of mutants His95 $\longrightarrow$ Asn and His95 $\longrightarrow$ Gln. The decrease in the stretching frequency of dihydroxyacetone phosphate bound to the wild type enzyme revealed an electrophilic attack by the enzyme, which polarises the carbonyl group of the substrate towards the transition state for enolisation. No perturbation having been observed in the case of the two mutants, the electrophile catalyst can only be His95. Moreover, the crystallographic data showed that Glu165 is the only residue significantly displaced in the mutant His95 $\longrightarrow$ Gln, its new position permitting the proton transfer from C_1 to C_2. This role would be assured by this histidine in the wild type enzyme. Consequently, the authors conclude that residue His95 would play the double role of the electrophilic group and the general acid-base catalyst in the isomerisation of triose phosphate.

In 1991, DAVENPORT et al. determined at 1.9 Å resolution the structure of triose phosphate isomerase complexed to an analog of the reactive intermediate, phosphoglycolohydroxamate. The analysis of the refined structure permitted specification of the geometry of the active site residues and the interactions with the inhibitor and, by analogy, with the substrates. The study was completed by a theoretical simulation of the reaction pathway by the group of KARPLUS (1991). The data were in agreement with a mechanism of general acid-base catalysis in which the carboxylate of Glu165 extracts a proton from the carbon of the substrate during which His95, under its neutral form rather than its imidazolium form, gives a proton to oxygen to generate the enediol intermediate. The implication of this histidine under its neutral form was finally verified by NMR. Glu165 undergoes, in the complex, a movement of 2 to 3 Å with regard to its position in the free enzyme which brings it to a non-aqueous environment, increasing thus its basic character.

This mechanism is presented in Fig. 12.47c opposite, which also shows the enediolate intermediates. A theoretical analysis indicates that this mechanism seems

the most energetically favorable (CUI & KARPLUS, 2001). Lys12 would orient or polarise the substrate oxygens by hydrogen bond or electrostatic interaction, stabilising the charged transition state. Asn10 would play a similar role.

The energy profile of the reaction was determined by an ensemble of studies. It corresponds to the following scheme:

$$E + D \rightleftharpoons ED \rightleftharpoons EX \rightleftharpoons EG \rightleftharpoons E + G$$

D, X, and G being respectively dihydroxyacetone phosphate, enediol intermediate and glyceraldehyde-3-phosphate. Figure 12.48 reproduces the free energy profile of the reaction catalysed by the yeast enzyme. This profile shows clearly that the steps of proton transfer are not limiting. The highest transition state corresponds to the binding step of the least stable substrate, glyceraldehyde-3-phosphate, which is the limiting step of the reaction; the rate 2×10^8 $mol^{-1} . s^{-1}$, corresponds to a diffusion process. The most stable state is that of the enzyme-dihydroxyacetone phosphate complex. The free energy profiles of the chicken enzyme and of the yeast enzyme are very similar despite the distance of the species in the phylogenetic tree and despite the sequence difference (50%). Triose phosphate isomerase seems therefore to have developed its catalytic potential very early over the course of evolution.

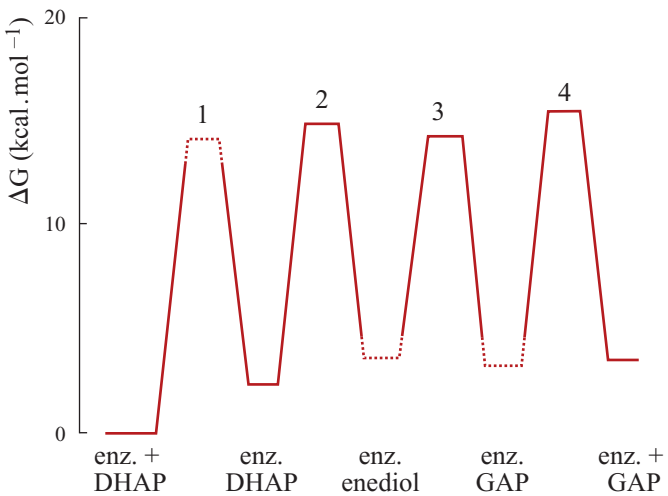


Fig. 12.48 Energy profile of the reaction catalysed by triose phosphate isomerase
(Reprinted with permission from *Biochemistry*, **30,** BASH P.A. et al., 5827. © (1991) American Chemical Society)

12.6. ASPARTATE AMINOTRANSFERASE

Aminotransferases are enzymes involved in the metabolism of amino acids. The metabolic scheme (Fig. 12.49 opposite) shows the role of aspartate aminotransferase with regard to the KREBS cycle and to the cycle of urea (from COOPER & MEISTER, 1989).

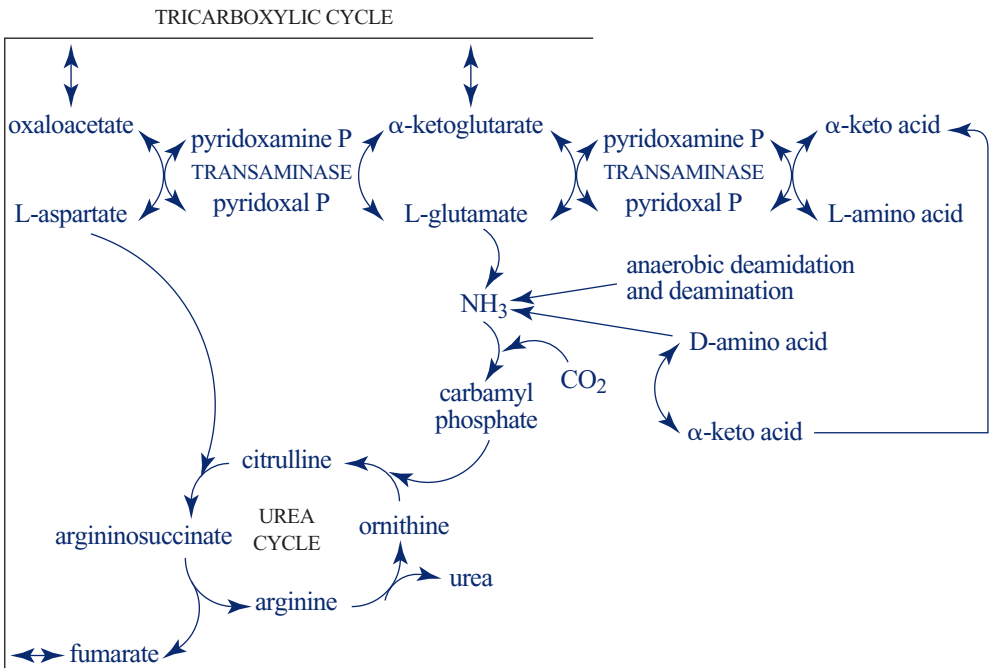
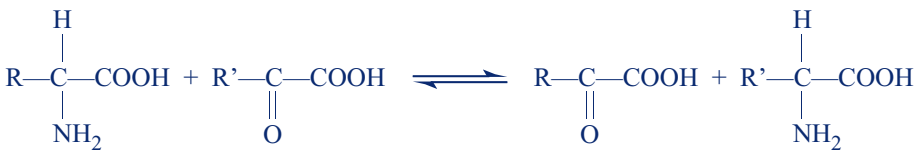
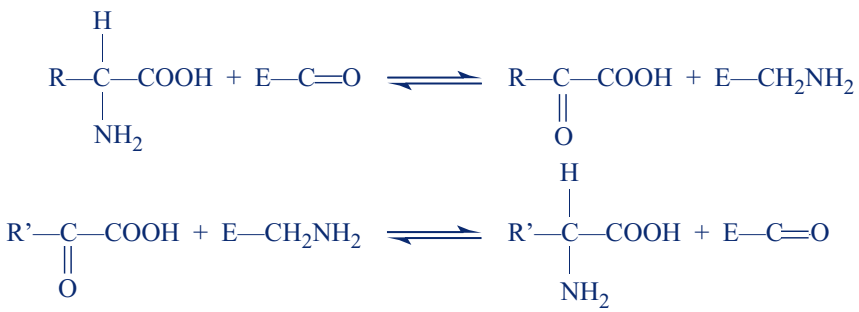


Fig. 12.49 Position of aspartate aminotransferase in cellular metabolism

These enzymes reversibly catalyse the exchange of the amino group between two amino acids through the intermediate of the corresponding keto acids by labilising the bond between the α carbon and the nitrogen of the amino group, according to the reaction:



This is carried out by the intermediate of a coenzyme, pyridoxal phosphate or vitamin B_6, and is decomposed in two half reactions:



in which the coenzyme is transformed from pyridoxal (E—C═O) to pyridoxamine (E—CH_2NH_2) (Fig. 12.50 below).

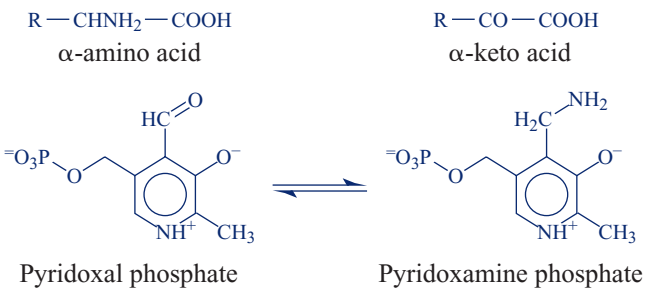


Fig. 12.50 Mechanism of transamination by pyridoxal phosphate

Transamination reactions were characterised by BRAUNSTEIN and collaborators in 1937, then by SNELL in 1945. These authors as well as the groups of METZLER and JENKINS carried a contribution decisive to the study of enzymatic transamination mechanisms. The coenzyme alone is capable of catalysing deamination reactions although with a much weaker efficiency; this property permitted its use as a model in the study of chemical mechanisms implicated in the reaction.

Each transaminase is specific for a couple of amino acids. Among these enzymes, aspartate aminotransferase catalyses the transfer of the amino group from L-aspartate to α-ketoglutarate (αKG) to give L-glutamate and oxaloacetate (OAA):

$$\text{L-Asp} + \alpha\text{-KG} \rightleftharpoons \text{L-Glu} + \text{OAA}$$

The kinetics of the reaction proceeds according to a Bi Bi ping-pong mechanism (see Chap. 5).

Aspartate aminotransferase of pig heart was the first pyridoxal phosphate enzyme which was purified (JENKINS et al., 1959). From this time, numerous enzymatic and physico-chemical studies of this enzyme were reported. Among these, one must cite the contributions of the groups of BRAUNSTEIN, KARPEISKY, OUVCHNINIKOV, FASELLA, YON, CHRISTEN, METZLER, MARTINEZ-CARRION and MEISTER. Since then, aspartate aminotransferases of other species were isolated and studied. In higher animals, there are two isoenzymes coded by two nuclear genes, a cytosolic form (cAAT) and a mitochondrial form (mAAT). This latter is synthesised as a precursor and imported in mitochondria where the signal peptide is released (see Chap. 7). The three-dimensional structure of the enzyme from several origins was obtained at high resolution. From these data, it was possible to specify the catalytic mechanisms.

12.6.1. STRUCTURAL PROPERTIES

Aspartate aminotransferase is a dimer of molecular weight 90 000 made of two identical protomers. Each protomer contains a molecule of pyridoxal phosphate (PLP) bound to the apoenzyme as a SCHIFF base (aldimine) with a lysine residue, Lys258, in the pig heart enzyme.

The complete sequence of several aspartate aminotransferases was established, including those of cytosolic and mitochondrial enzymes of pig heart (c and mAAT, respectively), of chicken (c and mAAT), as well as the rat enzyme (mAAt) and that of *E. coli*. Cytosolic enzymes consist of 411 or 412 amino acids per polypeptide chain; the mitochondrial enzyme is made of 401 amino acids and the *E. coli* enzyme of 396. Isoenzymes corresponding to different species present 85% identity whereas isoenzymes of a same species have only 46 to 48% identity. The *E. coli* enzyme is equally distant from the two types of isoenzymes presenting 40% identity with mAAT and cAAT of other species.

Radiocrystallography studies of the enzyme from several species were carried out by different groups of researchers, the cytosolic enzymes of chicken by the group of BORISOV, of pig by the group of ARNONE and METZLER, of yeast by the group of PETSKO, and the mitochondrial enzyme of chicken by the group of JANSONIUS. The enzyme structures of several microorganisms including *E. coli*, *Thermus thermophilus* and *Paracoccus denitrificans* were determined by the group of KAGAMIYAMA. These molecules present great structural similarities. The three-dimensional structure of the apoenzyme, the enzyme in the presence of PLP or pyridoxamine, as well as that of the enzyme in the presence of different substrate analogs like methyl aspartate, or dicarboxylic acids, competitive inhibitors, was obtained with a good resolution, as well as that of several mutants. One currently has access to 92 structures of aspartate aminotransferase from different species including the complexes and several mutants in the Protein Data Bank.

The structure of the mitochondrial enzyme of chicken (mAAT) bound to pyridoxal phosphate was obtained at 1.92 Å resolution (Fig. 12.51 below). The dimer molecule presents a 2-fold symmetry axis. Each subunit of identical structure has an ellipsoid form. It is folded in three parts, a stretched N-terminal region (residues 3 to 14), a small discontinued domain (15–47 and 326–410) and a large domain comprising residues 48 to 325 (the numeration used is that of the pig heart cytosolic enzyme). This latter is the domain of the coenzyme which is bound to Lys258 as an aldimine. Chicken mAAT, which only possesses 401 amino acids, presents deletions with respect to cAAT of pig heart in the following positions: 1, 2, 65, 127, 128, 131, 132, 153, 407 and 412. The interface between the two subunits is established *via* the binding domains of the coenzyme. The N-terminal part of the chain interacts with the base of the binding domain of the coenzyme of each subunit, contributing thus to the stability of the dimer. Each subunit comprises 16 helices including a very long one of 50 Å, helix 13 (Fig. 12.51), and a large β sheet composed of six parallel β segments and an antiparallel β segment, as well as two β sheets each constituted of two β segments, parallel for one and antiparallel for the other. The global structure comprises 50% of α helix, 14% of β structure and 14% of loops. The large β sheet is found in the coenzyme binding domain. The interface between the two domains is flat. This permits, over the course of the enzymatic reaction, a sliding movement of the small domain on the surface of the coenzyme binding domain (to see later).

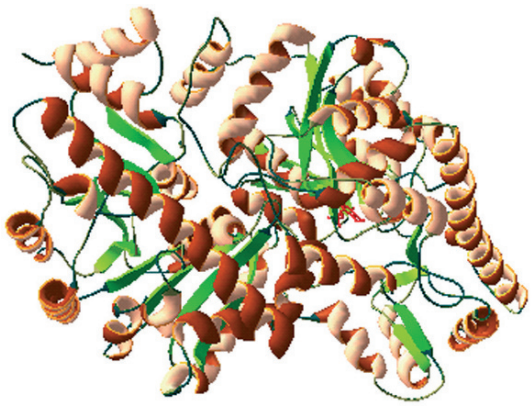


***Fig. 12.51 Three-dimensional structure of the aspartate aminotransferase dimer (PDB: 1TAR); PLP is indicated* in red**

The structure of the enzyme in the presence of pyridoxamine phosphate is slightly different. The only variations are localised to the immediate vicinity of the coenzyme. The inorganic phosphate occupies the same place as the phosphate group of the coenzyme. This structural analogy suggests that the binding of the coenzyme takes place without constraints.

12.6.2. BINDING OF THE COENZYME

The coenzyme PLP is bound to the enzyme under the aldimine form with lysine 258 (Fig. 12.52 opposite). Its face A is oriented towards the protein. The pyridine ring is surrounded by residues Ala224 and Tyr225 which contract a hydrogen bond with the hydroxyl group in 3′, as well as by residues Phe360, Asn194, Asp222 and Trp140. The $O^{3'}$ oxygen contracts a hydrogen bond with Asn194, the N^{1} nitrogen forms a salt bridge with protonated Asp222. The phosphate group is associated to the enzyme by nine hydrogen bonds with, for OP^{3} the OH groups of Ser107 and Ser255, the NH group of the main chain of Gly108, for OP^{4} the NH group of the main chain and the OH of Thr109, the $N^{\eta 1}$ of Arg266, for OP^{2} the $N^{\eta 2}$ of Arg266 and the phenolic group of Tyr70 of the other subunit. The role of three histidines 143, 189 and 193, situated under the coenzyme does not appear clearly yet; it could be related to the charge dissipation during the catalysis. According to the crystallographic data, the torsion angle of the phosphate group with respect to the pyridine ring is about–37°, i.e. smaller than what had been initially assumed.

When the coenzyme is under the pyridoxamine form, variations are observed in its orientation. The phosphate group remains localised and oriented as in PLP bound to the enzyme, but the aromatic ring is inclined 15° towards the front. The hydrogen bond with Asp222 remains. Rotations of weak amplitude between phosphorus and C^{5} are sufficient for the change from one form to the other. This implies that the torsion angle C^{4}-C^{5}-$C^{5'}$-OP^{1} passes through 0°, since in PMP, the OP^{1} oxygen is behind face A of the pyridine ring with a torsion angle around 24°. Since PMP does not form a covalent bond with the enzyme, it can be inclined even more in front as that occurs in aldimine formed with methylaspartate.

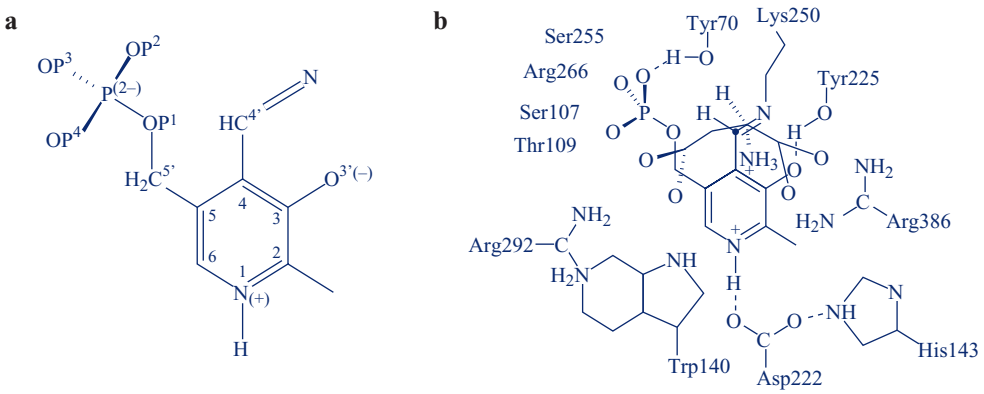


Fig. 12.52 (a) pyridoxal phosphate bound to the enzyme as a SCHIFF base (b) environment of the SCHIFF base and position of a dicarboxylic acid at the active centre of the enzyme

The inclination of the coenzyme is accompanied by movements of weak amplitude of the side chains of Tyr225 and Trp140, although the hydrogen bond between Tyr225 and $O^{3'}$ is maintained. The amino group of the coenzyme turns behind the plane of the pyridine ring with a torsion angle C^3-C^4-$C^{4'}$-$N^{4'}$ of –117°. The freed side chain of Lys258 extends towards the front, the NH_3^+ group establishing hydrogen bonds with the nitrogen of the amino group of the coenzyme and the carbonyl oxygen of Gly38.

12.6.3. BINDING OF THE SUBSTRATE: CONFORMATIONAL CHANGE OF THE ENZYME

Radiocrystallographic study of the enzyme in the presence of an inhibitor like maleate or the substrate analog, methylaspartate, showed that the binding of these ligands induces the closure of the enzyme structure. This movement corresponds to a rotation of 13° of the small domain towards the active site. It provokes the straightening of helix 13. During this conformational change, a sliding of the small domain occurs on the binding surface of the coenzyme, modifying the interface between the two domains. This involves residues 161–166, 192–199, 228–231 and 323–325 of the coenzyme binding domain, and residues 326–328, 348–364, 386–390 of the small domain. The residues of the coenzyme binding domain only undergo weak displacements, with the exception of the side chain of Phe228. The single hydrogen bond between the two domains (NH of Ile357 and O of Gly197) remains. In contrast, His352 undergoes a rotation of nearly 90° and Cys166 becomes accessible resulting in the "syncatalytic" increase in the reactivity of this cysteine, as was also observed in studies carried out in solution. The loop 37–39 undergoes a rearrangement. The largest movement affects helix 1 (residues 16–25) which orients towards the coenzyme, residues 15–18 blocking the entrance to the active site and enclosing the

inhibitor or the substrate. The side chains undergo significant displacements. However, no change in the regular structures, helices or β segments, occurs.

12.6.4. THE ACTIVE SITE

The residues of the active site, His193, Phe360 and Arg386, are included in the interface between the two domains. There are important differences in the active site between the open and closed structures. In the closed structure whose formation is induced by maleate, the pyridine ring is reoriented following a rotation around N1-C4 and becomes coplanar with the double bond of aldimine; the torsion angle C5-C5′ is reduced to –16°. Following the movement of the coenzyme, the Trp140 nucleus is reoriented. Residues 37–39 undergo a movement which drives Val37 to establish a VAN DER WAALS interaction with Tyr70; the peptide bond 37–38 is oriented such that it contracts a hydrogen bond with the carbonyl of Gly38 and the εNH_3^+ group of Lys258 after transaldimination. The NH group of Gly38 forms then a hydrogen bond with the carboxyl which interacts with Arg386. The guanidinium group of Arg392 is reoriented towards the active site and contracts two hydrogen bonds with the carboxylate of maleate, whereas in the open structure it forms a salt bridge with residue Asp15.

12.6.5. CATALYTIC MECHANISM

The first step of the enzymatic reaction is the association of the aspartate substrate to the enzyme-PLP. In the coenzyme bound to Lys258 as an aldimine, the pyridine nitrogen N^1 is protonated and the 3′OH group deprotonated. When the aspartate is associated with the enzyme, it is oriented in the active site by electrostatic attraction of the group NH_3^+ towards $O^{3'-}$ of the coenzyme and by interaction of the α-carboxylate with Arg386 and of the β-carboxylate with Arg292. These charges are situated in such a way that the active centre of the enzyme can only accommodate L-amino acids. The β-carboxylate turns to a left orientation relative to the α-carboxylate, thus optimising the interactions. This is accompanied by the release of some water molecules and the transition towards the closed form of the molecule.

The non-polar environment created around the substrate increases the interactions of the charges, and the proton of the amino group of the substrate is yielded to aldimine which, once protonated, presents a characteristic absorption spectrum with a maximum at 430 nm (Fig. 12.53 opposite). The protonation of aldimine induces a rotation of the pyridine ring which turns –16° around the C^5—$C^{5'}$ bond and becomes practically coplanar with the double bond of aldimine. The α-carboxylate of the substrate remains implicated in a hydrogen bond with Arg386 and contracts another one with Asn194 ($N^{\delta 2}$). The remaining protons of the α-amino group are reoriented towards the carboxylates to form internal hydrogen bonds. It results that the pair of nitrogen electrons is oriented towards the $C^{4'}$ of the coenzyme permitting its nucleophilic attack. The first step of transaldimination is the transient formation of a

diamine intermediate presenting an absorption spectrum centred at 340 nm. Then the non-protonated Lys258 is freed; the plane of the pyridine ring is inclined by 30° with respect to its position in the MICHAELIS complex and is clustered with Trp140 at a distance of 3.5 Å. It remains linked by hydrogen bond to Asp122. Tyr225, remaining bound by hydrogen bond to O′3, follows the movement of the coenzyme. The C^{α}—H bond has an optimal orientation, practically orthogonal to the plane of the pyridine ring. The NH_2 group of Lys258 hangs above the plane of the coenzyme; in this position, it can establish hydrogen bonds with OP^1, Tyr70, Tyr225 (O^{η}) and Gly38 (O). Hydrogen bonds with the last two oxygens place this group in proximity to C^{α}, permitting the transfer of the α proton to Lys258, which brings about the formation of a quinonoid intermediate. Indeed, the nitrogen and the C^{α} of the substrate are found then in the plane of the pyridine ring, forming a system of π electrons which extend to the α-carboxylate (intermediate III of Fig. 12.53).

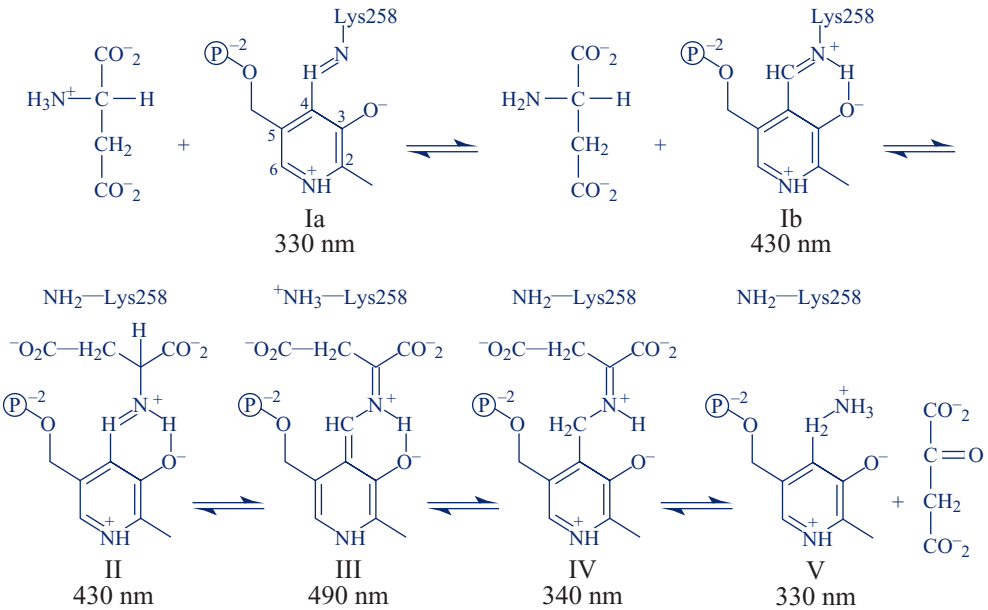


Fig. 12.53 Reaction intermediates during catalysis by aspartate aminotransferase

The transition of the external aldimine to this hyperconjugated form implicates a small rotation (< 10°) around the C^5—$C^{5'}$ bond. Trp140 follows this movement by a slight rotation which deletes its hydrogen bond with the β-carboxylate. The N^1 nitrogen in the sp^3 hybridation is uncharged; it can establish a strong hydrogen bond with the two oxygens of Asp222. This chemically unstable intermediate is thus stabilised by its interactions with the enzyme. It presents an absorption spectrum with a maximum at 490 nm (intermediate III).

The water molecule in position 7 can then enter the top of the active site, displacing Lys258 and protonating the $C^{4'}$ to give the ketimine intermediate (IV) which absorbs at 340 nm. This protonation splits the double bond between atoms N and C^{α}

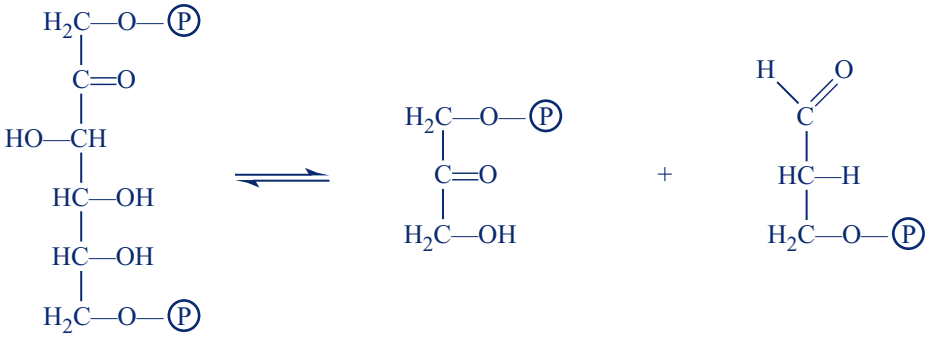
of the π electrons system in the pyridine ring. This latter is relaxed by a rotation towards the back which strengthens the hydrogen bond between Asp222 and the charged N^1 nitrogen. The part of the substrate which is coplanar with $C^{4'}$ optimises its position by establishing an internal hydrogen bond between the nitrogen of aldimine and the $O^{3'}$ oxygen, and another hydrogen bond between the β-carboxylate and Trp140 ($N^{\varepsilon 1}$). Lysine is displaced from its preceding position. It captures the proton of a water molecule situated in the vicinity. The OH group thus formed is oriented towards the C^{α} and produces a nucleophilic attack on this. The emergence of the sp^3 hybridation around the C^{α} repels the nitrogen that acquires a supplementary proton, probably originating from Lys258 before it is freed. The coenzyme turns towards the back. The nitrogen having taken back a second proton becomes oriented between the two carboxylates in order to optimise the hydrogen bonds. Lysine takes back a proton. The C^{α}—N bond thus broken produces oxaloacetate and the coenzyme as PMP. The transfer of the amino group to the α-ketoglutarate is carried out according to the reverse process.

On the basis of structural data, many questions concerning the catalytic mechanism of aspartate amino-transferase were resolved. The conception of mutants and their analysis permitted the specification of some aspects. CRONIN and KIRSCH in 1988 replacing Arg292 by an aspartate confirmed the key role of this arginine in the specificity of the enzyme for the substrate. The mutant Arg292 ⟶ Asp has an activity (k_{cat}/K_m) decreased by five orders of magnitude for the transformation of its physiological substrates. Conversely, it presents an activity 9 to 16 times greater than that of the wild type enzyme for using cationic amino acids. The replacement of tyrosine in position 70 (Tyr70Phe) permitted the specification of the role of this residue. The OH group of Tyr70 stabilises the transition state by 2 kcal . mol^{-1}. In addition, the benzene ring is essential to the selectivity of the enzyme for L-glutamate and 2-oxoglutarate. His143 is not necessary for the catalysis but is involved in the formation of the enzyme-substrate complex. Studying mutants in which either Arg292 or Arg386 was replaced by a leucine, MIZAGUSHI et al. (2001) concluded that the constraints are more important than the electrostatic interactions in the control of the SCHIFF base pK.

Thus, a wide ensemble of work involving structural and enzymatic studies and the application of genetic techniques permitted to establish in detail the catalytic mechanisms implicated in the activity of aspartate amino-transferase.

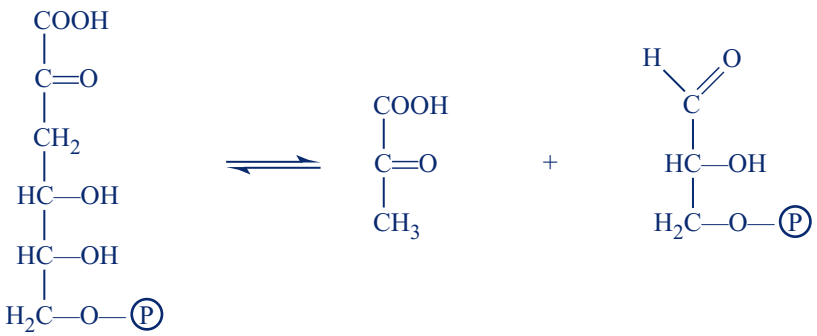
12.7. ALDOLASES

Aldolase of muscle was discovered in 1934 by MEYERHOF and LOHMAN who observed the reversible cleavage of hexose diphosphates catalysed by an enzyme which they name zymohexase. Later, MEYERHOF demonstrated that aldolase catalyses the transformation of fructose-1,6-diphosphate in dihydroxyacetone phosphate and 3-phosphoglycerate, according to the reaction:



Fructose 1,6-diphosphate Dihydroxyacetone phosphate 3-phosphoglycerate

Aldolase is a highly ubiquitous enzyme. It is found in animal and plant tissues and in most microorganisms. Aldolases can be divided into two distinct classes. Aldolases of class I are found in animals and higher plants; they do not require the presence of a metal ion and form a SCHIFF base intermediate with their substrate. They are inactivated upon reduction by sodium borohydride which blocks the intermediate. In vertebrate tissues, aldolase is tetrameric and there are different isoenzymes. Type A aldolase is present in skeletal muscle; that of type C in cardiac muscle. Liver aldolase is of type B. Aldolases of type A and C have a preferred catalytic specificity for fructose 1,6-diphosphate. Aldolase B plays a major role in the catabolism of fructose-1-phosphate. Aldolases of class II are met in bacteria and molds; their activity requires the presence of a metal ion. Their catalytic mechanism differs from that of aldolases of class I; they do not form a SCHIFF base with their substrate. However a microorganism, *Pseudomonas putida,* possesses an aldolase of class I specific for 2-keto-3-deoxy-6-phosphogluconate or KDPG aldolase. This enzyme catalyses the following reaction:



2-keto-3-deoxy-phosphogluconate Pyruvate Glyceraldehyde-3-phosphate

Only aldolases of class I are considered in this chapter. The primary structures of many of these aldolases are currenly known, including the human and the rabbit muscle enzymes, KDPG aldolase of *Pseudomonas putida* as well as the aldolases of some parasites. The first three-dimensional structures resolved were that of KDPG aldolase to 2 Å (MAVRIDIS et al., 1982), that of rabbit muscle to 2.7 Å (SYGUSH et al., 1987), and that of human muscle to 2 Å (GAMBLIN et al., 1991). Today PDB

contains more than a hundred aldolase structures of diverse organisms, some resolved to less than 2 Å. The sequences of human pathological mutants were determined. Some specific mutations provoke intolerance to fructose or hemolytic anemia. The mutation sites are generally localised to the vicinity of the active site or at the interfaces between subunits. The analysis of aldolase sequences of parasites responsible for malaria (*Plasmodium falciparum*) and sleeping sickness (*Trypanosoma brucei*) reveals some differences in the C-terminal part of the protein.

12.7.1. STRUCTURAL PROPERTIES

Rabbit muscle and human aldolases are tetramers made of four identical subunits of molecular mass 36 000. Each subunit comprises 365 amino acids. The two enzymes exhibit a high degree of sequence homology and also a great analogy in their three-dimensional structure. KDPG aldolase, contrary to mammalian aldolases, is a trimeric molecule made of identical protomers. Each subunit, shorter than those of mammalian aldolases, is constituted of a polypeptide chain of 225 amino acids. It does not present any sequence homology with fructose 1,6-diphosphate aldolase of muscle.

Structural data obtained by radiocrystallography shows that each protomer, for mammalian aldolases as well as for KDPG aldolase, presents a structure of type α/β similar to that of triose phosphate isomerase and pyruvate kinase. The motif is made of eight αβ units organised through an 8-fold pseudo symmetry axis (Figs. 12.54 and 12.55 opposite). The α helices are approximately antiparallel to the β segments. As compared to related proteins, muscle aldolases contain four additional helices, three of them being located after segments a, b, and c. KDPG aldolase is shorter, but contains however an additional N-terminal helix like mammalian aldolases.

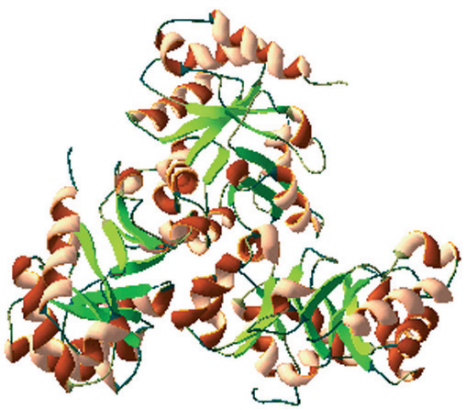


Fig. 12.54 Structure of the KDPG aldolase trimer (PDB: 1EUN)

The β barrel forms the centre of the protein (Fig. 12.55 opposite). It is constituted not only of hydrophobic amino acids but also of polar residues. In the muscle enzymes, residues Asp33, Lys107, Lys146, Glu187 and Lys229 are located at the interior of the cavity. The potentially charged parts of these residues are collinear,

the acid and basic residues being alternatively arranged, a situation which can assure charge neutralisation and permits the refolding of these residues towards the interior of the β barrel. In KDPG aldolase, there are twelve pairs of ions important to the stability of the structure. Some charged groups are situated in the central cavity of the molecule (Glu56, Arg141, Asp166), the others at the interface between the subunits. Some of them form interprotomer salt bridges (Arg75-Asp223′; Asp207-Arg210′). These interactions contribute to maintaining the trimer structure.

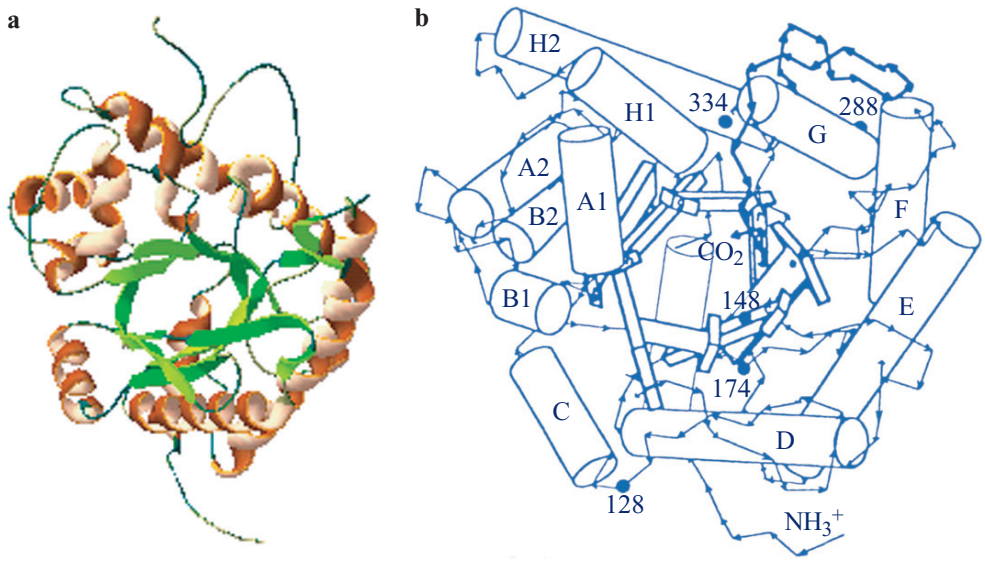


Fig. 12.55 (a) the protomer of rabbit muscle aldolase (PDB: 1ADO) (b) that of the human enzyme

(12.55b – Reprinted from *J. Mol. Biol.*, **219,** GAMBLIN S.J. et al., Activity and specificity of human aldolases, 174. © (1991) with permission from Elsevier)

In rabbit muscle aldolase, the C-terminal extremity presented an apparent disorder in the region 346–354. This region was identified in the human enzyme. The first nine residues (345–353) form a flexible loop at the surface of the enzyme which extends from the C-terminal extremity to helix H2 at the amino-terminal end of helix G. The side chain of serine 355 contracts a hydrogen bond with the nitrogen of the main chain at the extremity of the helix. This interaction orients the last C-terminal part that crosses the αβ motif; the three last residues are localised to the centre of the barrel (see Fig. 12.55b). This region of the protein seems to play a role in the catalytic activity. Indeed, its suppression by proteolysis reduces the enzymatic activity. It was suggested that the C-terminal peptide could modulate the enzymatic activity by assuring the binding and the alignment of the substrate during catalysis. This part of the molecule does not exist in KDPG aldolase. In contrast, it is present in enzymes of the parasites plasmodium and trypanosome.

In the quaternary structure of muscle enzyme, the subunits are arranged in a tetrahedral configuration. The regions of contact are essentially hydrophobic. The largest

of the two interfaces is established by the interaction of side chains of helices E and F of adjacent subunits. The arrangement of subunits is such that the two pairs of helices bound by symmetry are approximately antiparallel. Helix E interacts with its equivalent E′ for practically all its length. Helices F and F′ enter into contact only by their C-terminal extremity. The interface with the weakest contact between subunits consists of the loop connecting the βc strand to helices C and D of the two subunits.

12.7.2. THE ACTIVE SITE

The residues implicated in the catalytic reaction are located in a pocket that extends from the surface of the subunit to the centre of the β barrel and includes the region where the charged groups are located. The localisation of the active site differs from that observed in enzymes that present a similar structure, like triose phosphate isomerase where the active site is situated in the region of loops at the COOH extremity of the β strands of the β barrel. In muscle aldolase, lysine 229 which is implicated in the formation of the SCHIFF base with C^2 of the substrate extends from the centre of the β barrel towards the middle of segment βf. Lysine 107 which interacts with phosphate in C^6 is at the surface of the substrate binding pocket. Residues Lys107, Lys146 and Lys229 form an alignment which seems to favor the binding of the triphosphate part of ATP or inositol triphosphate, powerful inhibitors of aldolase.

The inner walls of the active centre pocket are lined with hydrophobic residues coming from the β segments. The closest of the ε-amino group of Lys229 are Ile185, Ile77 and Phe144 near the N-terminal extremity of the β barrel. Residue Arg148, a probable candidate for the binding of phosphate in C^1, is situated at the surface of the substrate binding pocket on the side opposite to Lys107. The distance between the binding sites of the two phosphate groups is 11 Å, which is compatible with the interatomic distances of C^1 and C^6 in the non-cyclic configuration.

From the complete structure of the human enzyme including the C-terminal part of the protein, the group of WATSON carried out molecular modelling of the substrate 1,6-fructose diphosphate at the interior of the active centre. Figure 12.56 opposite illustrates the different enzyme-substrate interactions. The phosphate site in C^1 consists of residues Arg148 and Lys146. The phosphate in C^6 interacts with residues Lys41, Arg42 of helix A1 and Arg303 situated on the loop between the segment βh and helix H1. His361 is in close contact with Arg148 and at a distance permitting the establishment of a hydrogen bond with Tyr363. It could play a role in the positioning of this tyrosine during catalysis. In addition, Tyr363 is at a distance compatible with the establishment of a hydrogen bond with the protonated SCHIFF base made by lysine 229 with the C^3 of fructose 1,6-diphosphate. Site-directed mutagenesis experiments showed that Asp33 situated in the proximity of Lys229 is a residue essential to catalysis. It could act as a base catalyst to extract the proton of the hydroxyl group in C^4 [MORRIS & TOLAN, 1993].

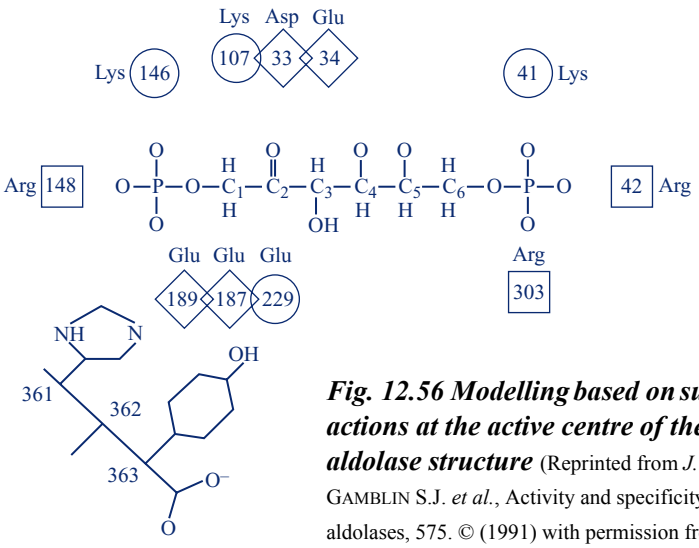


Fig. 12.56 Modelling based on substrate interactions at the active centre of the human aldolase structure (Reprinted from *J. Mol. Biol.*, **219**, GAMBLIN S.J. *et al.*, Activity and specificity of human aldolases, 575. © (1991) with permission from Elsevier)

All the amino acid residues of the active centre including the potentially charged groups of the β barrel were conserved over the course of evolution in all aldolases of class I. By comparison the residues in the vicinity of Lys144 of KDPG aldolase implicated in the formation of the SCHIFF base are indicated in Fig. 12.57.

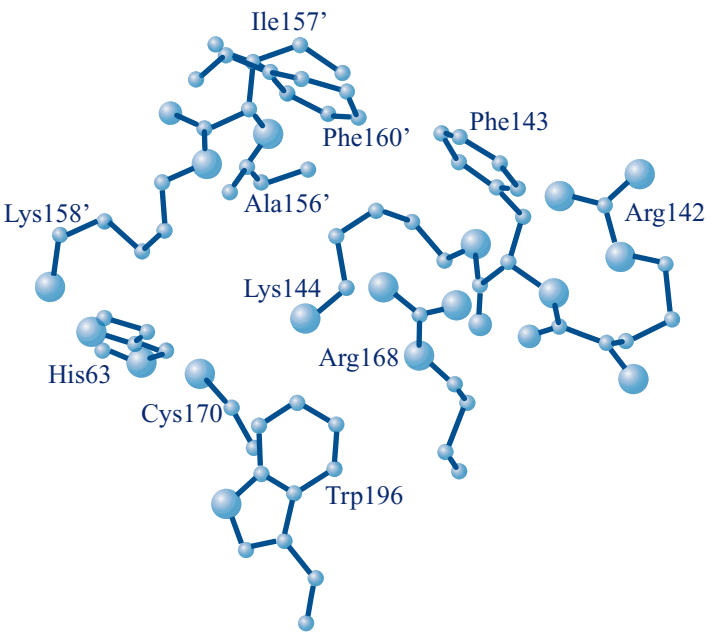


Fig. 12.57 Amino acid residues present in the vicinity of the catalytic lysine (Lys144) of KDPG aldolase (Reprinted from *J. Mol. Biol.*, **162**, MAVRIDIS I.M. et al., Structure of 2-keto-3-deoxy-6-phosphogluconate aldolase at 2.8 Å resolution, 441. © (1982) with permission from Elsevier)

This residue lies in a depression at the surface of the subunit. The entrance to the cavity is bordered by residues Leu145, Pro147, Phe169 and Pro171 of one subunit, and by residues Gly152′, Gly153′ and Ala155′ of another subunit. Like in the muscle enzyme, Lys144 is surrounded by a cluster of positively charged groups (Arg142, Arg168, Lys158′ and His63) and hydrophobic residues.

Although the residues around the active site are well conserved, only 4 of the 22 residues of the C-terminal region are found in the different aldolases, suggesting that this part of the molecule can be implicated in the variation in specificity and activity of these enzymes towards their substrates.

12.7.3. CATALYTIC MECHANISM

Mammalian aldolases are tetramers. But contrary to most oligomeric enzymes, the isolated monomer presents a significant enzymatic activity. The experiments of CHAN (1970, 1972) had shown that the subunits isolated from the rabbit enzyme bound to a solid matrix conserved 27% of the specific activity of the tetrameric aldolase. More recently, BEERMINK and TOLAN (1996) obtained a monomeric aldolase by a double mutation Q125D/E224A which destabilises the tetramer. This monomer presented 72% of the specific activity of the wild type aldolase.

A catalytic mechanism was proposed by HORECKER in 1972 for rabbit muscle fructose 1,6-diphosphate aldolase. The general mechanism involves a SCHIFF base intermediate. The structural data obtained since then, which concerns the reduced SCHIFF base intermediate formed by the transaldolase of *E. coli* and dihydroxyacetone phosphate, allowed to propose a catalytic mechanism. In the complex, the SCHIFF base formed with Lys132 (corresponding to Lys129 of the human enzyme) interacts by hydrogen bond with several residues as shown in Fig. 12.58a opposite. The active site is filled with water molecules; one of them is very close to the ε-amino group of the catalytic lysine and of the dihydroxyacetone part and forms hydrogen bonds with Glu96 and Thr156. This water molecule very likely plays a role in the catalytic mechanism. On the basis of these data a mechanism was proposed by JIA et al. (1997). It is represented in Fig. 12.58b. The first step of the reaction is a nucleophilic attack by the ε-amino group of lysine 132 on the C^2 carbon of carbonyl; the water molecule could act in the proton transfer from Lys132 to Glu96. The negative charge of oxygen in C^2 in the transition state could be stabilised by hydrogen bonds with the water molecule and the side chain of Thr156. The departure of a water molecule from carbaminolamine drives the formation of an imine with the enzyme. The cleavage of the imine and the release of the first product, glyceraldehyde-3-phosphate, are facilitated by the deprotonation of the hydroxyl group in C^4 which would be assisted by Asp17. The resulting SCHIFF base is stabilised by resonance until an aldose is bound to the active site. The carbanion of dihydroxyketone reacts afterwards with the carbonyl of the aldose in the reverse order from what occurs during transaldolisation. The proposed roles for Asp17 and Glu96 are based on structural data.

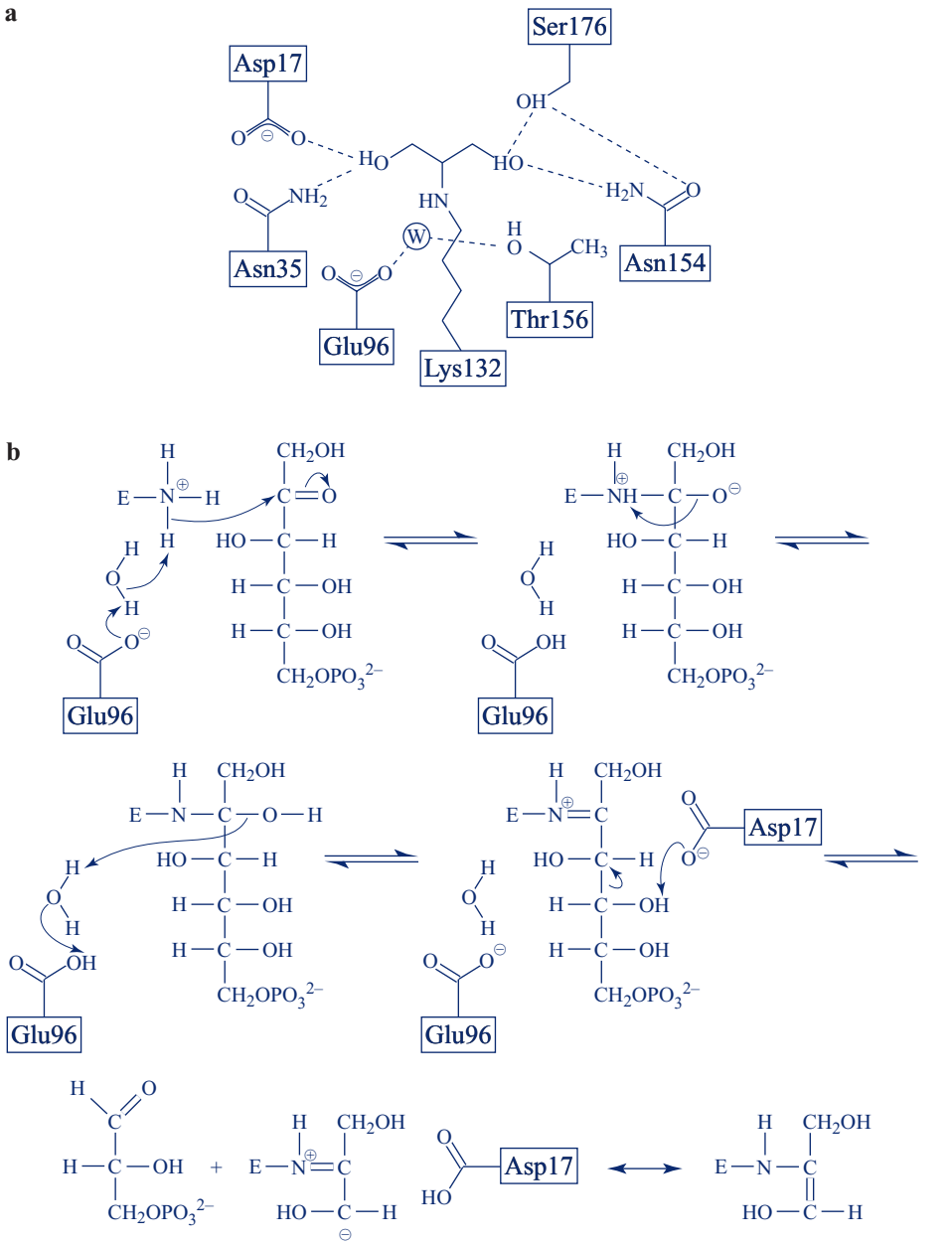


Fig. 12.58 (a) interactions of the reduced SCHIFF base intermediate at the active site of E. coli transaldolase – (b) catalytic mechanism proposed by JIA et al., 1997

(From *Protein Sci.*, **6**, No. 1, JIA J. et al., 1997, 119–124. © (1997 The Protein Society). Reprinted with permission of John Wiley & Sons, Inc.)

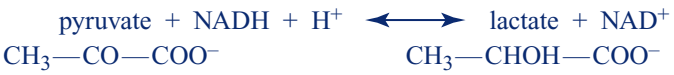
Despite the ensemble of structural and functional data obtained on these diverse enzymatic systems taken as examples since they are among the best known, many points still remain to be specified.

12.8. CONCLUSIONS AND PERSPECTIVES

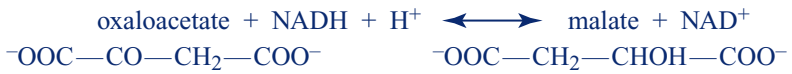
The examples presented illustrate particularly well the contribution of precise structural data in the understanding of catalytic mechanisms implicated in enzymatic reactions. They show also their limits since, on the one hand these data remain static, and on the other hand the X-ray diffraction diagrams do not permit the localisation of protons. However the improvement of techniques that are currently developed in the domain of radiocrystallography, particularly the method of LAUE, should permit the rapid recording of diffraction data and foresee the possibility of carrying out the kinetic studies, meaning to follow the displacement of atoms during an enzymatic reaction. However, the application of this method to the study of enzymatic reactions still poses large problems, principally that of synchronisation of all the reacting molecules. Also it has only been applied to reactions triggered by a "flash" of light. In any case, the use of a plurality of methods is necessary to understand the diverse events which occur in the path of an enzymatic reaction. The use of NMR methods and neutron diffraction remains necessary to localise protons. Kinetic studies in solution and chemical labelling methods or analysis of isotope effects are indispensable. **It is clear that it is only by the use of a variety of methodologies relevant to different disciplines that one succeeds in understanding and mastering the ensemble of factors which cooperate in the catalytic efficiency of enzymes.** Despite all the progress made, there still are numerous enzymes for which the mechanism of action is not known, and even for well described systems, many problems still remain unresolved.

Throughout this chapter, the reader has realised the contribution of chemical labelling and more recently of site-directed mutagenesis in the study of catalytic mechanisms. The possibility of selectively labelling an amino acid residue or replacing this amino acid by another considerably improved our knowledge. But these methods do not only permit knowing, but also intervening and modifying the catalyst. Consequently they are used in new strategies with the goal of modulating or modifying the specificity and catalytic efficiency of an enzyme. The example of mutations carried out on trypsin showed that it was possible, by changing a single amino acid in a well determined position, to increase the specificity of the enzyme for peptide bonds near arginine or lysine. In 1988, WILKS et al. transformed lactate dehydrogenase of *B. stearothermophilus* in malate dehydrogenase by site-directed mutagenesis. Although these two enzymes which function with NAD^+ had practically

no sequence homology, they present great structural analogies, in particular in the binding domain of NAD^+ binding domain. Lactate dehydrogenase reversibly catalyses the reduction of pyruvate in lactate by NADH:



and malate dehydrogenase, that of oxaloacetate in malate by NADH :



Taking into account the position of atoms in the ternary complexes at the active centre of the two enzymes, mutants Asp197 ⟶ Asn, Thr246 ⟶ Gly and Gln102 ⟶ Arg of lactate dehydrogenase were constructed and analysed. Figure 12.59 shows the geometry of the active site in the ternary complex LDH-NADH-oxamate. Table 12.8 gives the values of the kinetic parameters of mutants and the wild type enzyme. The last mutation Gln102 ⟶ Arg gives rise to an enzyme very specific for malate with an excellent catalytic efficiency. The maximal rate obtained is twice larger than that of malate dehydrogenase of the same organism. Thus, engineered malate dehydrogenase, more active than the natural enzyme, was obtained by remodelling of another enzyme.

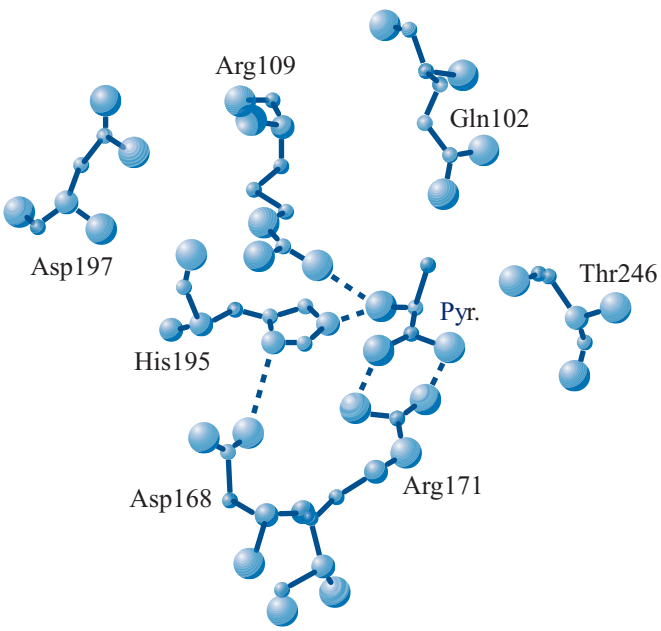


Fig. 12.59 Geometry of the active centre of lactate dehydrogenase (LDH) in the ternary complex LDH-NADH-oxamate (From *Science*, **242**, WILKS H.M. et al., A specific, highly active malate dehydrogenase by redesign of a lactate dehydrogenase Framework, 1541. © (1988) reprinted with permission from American Association for the Advancement of Sciences)

Table 12.8 Kinetic parameters of lactate dehydrogenase and its mutants

Enzyme	*Substrate*	K_m(mM)	$k_{cat}(s^{-1})$	$k_{cat}/K_m(M^{-1}.s^{-1})$
	Pyruvate			
Wild type		0.060	250	4.2×10^6
Asp197 ⟶ Asn		0.66	90	1.3×10^5
Thr246 ⟶ Gly		13	16.0	1.3×10^3
Gln102 ⟶ Arg		1.8	0.9	5.0×10^2
	Oxaloacetate			
Wild type		1.5	6.0	4.0×10^3
Asp197 ⟶ Asn		0.15	0.50	3.0×10^3
Thr246 ⟶ Gly		0.20	0.94	4.7×10^3
Gln102 ⟶ Arg		0.06	250	4.2×10^6

We have cited this example which was among the first well documented ones. Since then, other works have resulted in the modification of specificity obtained by mutation of several amino acids at the binding site of substrates after consideration of the structure.

Taking into account the successes of these approaches, research in molecular enzymology is oriented more and more towards the concept of new catalysts. Several strategies are currently being developed. The first uses the very strict specificity of abzymes (see Chap. 11). Results show that these catalysts are generally very selective. However, their catalytic efficiency, although higher than that of chemical catalysts, remains relatively weak compared to that of enzymes.

Another strategy, oriented by molecular modelling, utilises the genetic or even chemical approach (hemisynthesis) for the remodelling of enzymes or the conception of new enzymes. The goal of devising new proteins (*protein design)* is to obtain *de novo* polypeptides having a predetermined activity. The conception step consists of choosing a reaction mechanism, chemical groups and a spatial arrangement susceptible to giving rise to the researched activity. Then, one must imagine a polypep tide scaffold that will support these active groups in the correct orientation. Finally, there remains to be determined an amino acid sequence which will be conveniently folded to give a stable three-dimensional structure. Several approaches are possible. One consists of taking as a model a protein of known structure having properties close to those of the researched catalyst and locally remodelling it to create the active centre desired. This is the **local conception**. The other approach, or the **global conception,** conceives, starting from zero, a structure having one of the classical folds. In this case, the remodelling can be made by analogy using the data bank of known crystallographic structures. The last step consists of synthesising the corresponding gene coding for the sequence thus conceived.

Finally, the development of genetic engineering techniques now permit obtaining very numerous variants of a protein by methods of **in vitro *evolution*** associated with screens permitting the selection of the desired protein.

Despite the empirical success which was obtained so far, it is important to underline that significant progress will only be made when the folding code of proteins will be deciphered. Indeed, it is easy today to modify a protein sequence, even to conceive new ones, but to obtain an active enzyme, it is still necessary that this protein can acquire an adequate three-dimensional structure. The knowledge of mechanisms which drive a polypeptide chain to an operational spatial structure is an indispensable preliminary to the rational conception of new enzymes and represents an important challenge for biotechnology.

Whatever it may be, it is clear that molecular enzymology is long from falling obsolete; due to the utilisation of new tools, it has acquired a new youth and currently represents a domain in full development by both its fundamental aspects and its applications. With the success obtained in the sequencing of genomes, a large field of study is dedicated today to proteins and their structural and functional properties.

BIBLIOGRAPHY

BOOKS

CHRISTEN P. & METZLER D.E. eds –1985– *Transaminases,* John Wiley & Sons, New York.

FERSHT A. –1985– *Enzyme, structure and mechanisms*, 2nd ed., Freeman and Co, New York.

PELMONT J. –1996– *Enzymes : catalyseurs du vivant*, Collection Grenoble Sciences, EDP Sciences, Paris.

GENERAL REVIEWS

BACKER E.N. & DRENTH J. –1984– The thiol proteases: structure and mechanisms, in *Biological macromolecules and assemblies*, Vol. 3, F.A. JURNAK & A.M. MCPHERSON eds, John wiley & Sons, New York, 313–368.

BLOW D.M. –1971– The structure of chymotrypsin, in *The Enzymes*, 3rd ed., Vol. III, P.D. BOYER ed., Acad. Press, New York, 185–212.

BRAUNSTEIN A.E. –1985– Introduction: an historical survey in transamination and transaminases, in *Transaminases*, P. CHRISTEN & D.E. METZLER eds, John Wiley & Sons, New York, 1–35.

DALZIEL K. –1975– Kinetics and mechanism of nicotinamide nucleotide linked dehydrogenases, in *The Enzymes*, 3rd ed., Vol. III, P.D. BOYER ed., Acad. Press, New York, 1–60.

EKLUND H. & BRÁNDÉN C.I. –1984– Alcohol dehydrogenase, in *Biological macromolecules and assemblies*, Vol. 3, F.A. JURNAK & A.M. MCPHERSON eds, John wiley & Sons, New York, 73–142.

FRUTON J.S. –1976– The mechanism of the catalytic action of pepsin and related acid proteinases, in *Adv. Enzymol.* **44**, 1–36.

HARTLEY B.S. & SHOTTON D.D. –1971– Pancreatic elastase, in *The Enzymes*, 3rd ed., Vol. III, P.D. BOYER ed., Acad. Press, New York, 323–373.

HARTSUCK J.A. & LIPSCOMB W.N. –1971– Carboxypeptidase, in *The Enzymes*, 3rd ed., Vol. III, P.D. BOYER ed., Acad. Press, New York, 1–79.

HORECKER B.L., TSOLAS O. & LAI C.Y. –1972– Aldolases, in *The Enzymes*, 3rd ed., Vol. VII, P.D. BOYER ed., Acad. Press, New York, 213–258.

IMOTO T., JOHNSON L.N., NORTH A.C.T., PHILLIPS D.C. & RUPLEY J.A. –1972– Vertebrate lysozymes, in *The Enzymes*, 3rd ed., Vol. VII, P.D. BOYER ed., Acad. Press, New York, 666–868.

IVANOV V.I. & YA KARPEISKY M. –1969– Dynamic three-dimensional model for enzymic transamination, in *Adv. Enzymol.* **32**, 21–53.

JAMES M.N.G. & SIELECKI A.R. –1987– Aspartic proteinases and their catalytic pathway, in *Biological macromolecules and assemblies*, Vol. 3, F.A. JURNAK & A. MCPHERSON eds, John wiley & Sons, New York, 413–482.

JANSONIUS J.N. & VINCENT M.G. –1987– Structural basis for catalysis by aspartate aminotransferase, in *Biological macromolecules and assemblies*, Vol. 3, F.A. JURNAK & A. MCPHERSON eds, John wiley & Sons, New York, 187–285.

KRAUT J. –1971a– Chymotrypsinogen: X-ray structure, in *The Enzymes*, 3rd ed., Vol. III, P.D. BOYER ed., Acad. Press, New York, 165–183.

KRAUT J. –1971b– Subtilisin: X-ray structure, in *The Enzymes*, 3rd ed., Vol. III, P.D. BOYER ed., Acad. Press, New York, 547–560.

KRAUT J. –1977– Serine proteases: structure and mechanism of catalysis, in *Annu. Rev. Biochem.* **46**, 331–358.

MARTINEZ-CARRION M., HUBERT E., IRIARTE A., MATTINGLY J.R. & ZITO S.W. –1985– Mechanism of aminotransferase action, in *Transaminases*, P. CHRISTEN & D.E. METZLER eds, John Wiley & Sons, New York, 307–316.

METZLER D.E. –1979– Tautomerism in pyridoxal phosphate and in enzyme catalysis, in *Adv. Enzymol.* **50**, 1–40.

SCOPES R.K. –1973– 3-Phosphoglycerate kinase, in *The Enzymes*, 3rd ed., Vol. VIII, P.D. BOYER ed., Acad. Press, New York, 335–351.

SPECIALISED ARTICLES

ALBER T., BANNER P.W., BLOOMER A.C., PETSKO G.A., PHILLIPS D.C., RIVER P.S. & WILSON I.A. –1981– *Phil. Trans. Roy. Soc. London* **B293**, 159.

ANTONOV V.K., GINODMAN L.M., KAPITANNIKOV Y.V., BARSHEVSKAYA T.N., GUROVA A.G. & RUMSH L.D. –1978– *FEBS Lett.* **88**, 87.

ATLAS D. –1975– *J. Mol. Biol.* **93**, 39.

BANKS R.D., BLAKE C.C.F., HASER P.R., RICE D.W., HARDY G.W., MERRETT M. & PHILLIPS A.W. –1979– *Nature* **279**, 773.

BANNER D.W., BLOOMER A.C., PETSKO G.A., PHILLIPS D.C., PEGSON C.L., WILSON I.A., CONNAN P.H., FURTH A.J., MILMAN J.D., OFFORD R.D., PRIDDLE J.D. & WALEY S.G. – 1975– *Nature* **255**, 609, London.

BASH P.A., FIELD M.J., DAVENPORT R.C., PETSKO G.A., RINGE D. & KARPLUS M. –1991– *Biochemistry* **30**, 5826.

BEERNINK P.T. & TOLAN D.R. –1996– *Proc. Natl Acad. Sci. USA* **93**, 5374.

BERGER A. & SCHECHTER I. –1970– *Phil. Trans. R. Soc. London B Biol. Ser.* **259**, 249.

BERNHARD S. ET AL. –1970– *J. Mol. Biol.* **49**, 85.

BERNSTEIN B.E. & HOL W.G.J. –1998– *Biochemistry* **37**, 4429.

BERNSTEIN B.E., MICHELS P.A. & HOL W.G.J. –1997– *Nature* **385**, 275.

BIRKOFT J.J. & BLOW D.M. –1972– *J. Mol. Biol.* **68**, 187.

BLAKE C.C.F., KOENIG P.F., MAIR G.A., NORTH A.C.T., PHILLIPS D.C. & SARMA V.P. –1965– *Nature* **206**, 757.

BLAKE C.C., JOHNSON L.N., MAIR G.A., NORTH A.C., PHILLIPS D.C. & SARMA D.R. –1967– *Proc. R. Soc. London B Biol. Sci.* **169**, 378.

BLAKE C.C.F.&.RICE D.W –1981– *Phil. Trans. Roy. Soc. London* **B293**, 93.

BLOW D.M., JANIN J. & SWEET R.M. –1974– *Nature* **249**, 54.

BODE W., FELDHAMMER H. & HUBER R. –1976– *J. Mol. Biol.* **106**, 325.

CANFIELD R.E. –1963– *J. Biol. Chem.* **238**, 2691 and 2698.

CAPEILLIAIRE-BLANDIN C., BRAY R.C., IWATSUBO M. & LABEYRIE F. –1975– *Eur. J. Biochem.* **54**, 549.

CHAN W.W. –1970– *Biochim. Biophys. Res. Com.* **41**, 1198.

CHAN W.W. –1972– *Arch. Biochem. Biophys.* **14**, 136.

CHAPUS C., KERFELEC B., FOGLIZZO E. & BONICEL J. –1987– *Eur. J. Biochem.* **166**, 379.

COHEN G.H., SILVERSTON E.W., MATTHEWS B.W., BRAXTON H. & DAVIES D.R. –1969– *J. Mol. Biol.* **44**, 129.

COLL M., GUASH A., AVILES F.X. & HUBER R. –1991– *Embo. J.* **10**, 1.

COOPER A.J. & MEISTER A. –1989– *Biochimie* **71**, 387.

CUI Q. & KARPLUS M. –2001– *J. Am. Chem. Soc.* **123**, 2284.

DAVENPORT R.C., BASH P.A., SEATON B.A., KARPLUS M, PETSKO G.A. & RINGE D. –1991– *Biochemistry* **30**, 5821 and 5826.

DAVIES G.J., GAMBLIN S.J., LITTLECHILD J. & WATSON H.C. –1993– *Proteins* **15**, 283.

DELBAERE L.J.T., HUTCHEON W.L.B., JAMES M.N.G. & THIESSEN W.E. –1975– *Nature* **257**, 758.

DRENTH J., HOL W.G., JANSONIUS J.N. & KOEKOEK R. –1972– *Eur. J. Biochem.* **26**, 177.

DUNN B.M. & FINK A.–1984– *Biochemistry* **23**, 5241.

EKLUND H., PLAPP B., SAMAMA J.M. & BRÁNDÉN C.I. –1982– *J. Biol. Chem.* **257**, 14349.

FELDHAMMER H. & BODE W. –1975– *J. Mol. Biol.* **98**, 683.

FINK A. & MEEHAN P. –1979– *Proc. Natl. Acad. Sci. USA* **76**, 1566.

FLEMING A. –1922– *Proc. Roy. Soc. London* **B93**, 306.

FRUTON J.S.–1970– *Adv. Enzymol. Rel. Areas Mol. Biol.* **33**, 401.

GALDEN A., AULD D.S. & VALLEE B.L. –1986– *Biochemistry* **25**, 646.

GAMBLIN S.J., DAVIES G.J., GRIMES J.M., JACKSON R.M., LITTLECHILD J.A. & WATSON H.C. –1991– *J. Mol. Biol.* **219**, 573.

GARDELL S.J., CRAIK C.S., HILVERT D., URDEA M.S. & RUTTER W.J. –1985– *Nature* **317**, 551.

GERTLER A., WALSH K.A. & NEURATH H. –1974– *Biochemistry* **13**, 1302.

GHELIS C. & YON J.M. –1979– *C.R. Acad. Sci. Paris* **D289**, 197.

GOMIS-RÜTH F.X., GOMEZ-ORTIZ M., VENDRELL J., VENTURA S., BODE W., HUBER R. & AVILES F.X. –1997– *J. Mol. Biol.* **269**, 861.

GUASH A., COLL M., AVILES F.X. & HUBER R –1992– *J. Mol. Biol.* **224**, 141.

GUIARD B. –1985– *Embo. J.* **4**, 3265.

GUIARD B., GROUDINSKI O. & LEDERER F. –1974– *Proc. Natl Acad. Sci. USA* **84**, 2629.

GUILBERT C., PERAHIA D. & MOUAWAD L. –1995– *Comput. Phys. Commun.* **91**, 263.

HARLOS K., VAS M. & BLAKE C.C. –1992– *Proteins* **12**, 133.

HARUYAMA K., NABAI T., MIYAHARA I., HIROTSU K., MIZYGUSHI H. & KAGAMIYAMA H. –2001– *Biochemistry* **40**, 4632.

HENDERSON R. –1970– *J. Mol. Biol.* **54**, 341.

HUBER R, KUKLA D., BODE D., SCHWAGER P., BARTELS P., DEISENHOFER J. & STEIGEMANN W. –1974– *J. Mol. Biol.* **89**, 73.

INOUE K., KARAMITSU S., OKAMOTO A., HIROTSU K., HIGUSHI T. & KAGAMIYAMA H. –1991– *Biochemistry* **30**, 7796.

JENKINS W.J., YPHANTIS D.A. & SINGER I.W. –1959– *J. Biol. Chem.* **234**, 51.

JIA J., SCHÖRKEN U., LINDVIST Y., SPENGER G.A. & SCHNEIDER G. –1997– Crystal structure of the reduced Schiff-base intermediate complex of transaldolase B from *Escherichia coli*: mechanistic implications for class I aldolases, in *Protein Sci.* **6**, 119.

JOLLES J., JAUREGUI-ADELL J., BERNIER I. & JOLLES P. –1963–*Biochim. Biophys. Acta* **78**, 668.

KAMPHUIS I.G., DRENTH J. & BAKES E.N. –1985– *J. Mol. Biol.* **182**, 317.

KAMPHUIS I.G., KALK K.H., SWART M.B. & DRENTH J. –1985– *J. Mol. Biol.* **179**, 233.

KARPEISKY M.Y., KHOMUTOV R.M., SEVERIN E.S. & BRENSOV Y.N. –1963– in *Chemical and biological aspects of pyridoxal catalysis*, A.E. BRAUNSTEIN, E.S. SEVERIN & Y.M. TORCHINSKY eds, Pergamon Press, Oxford, 323.

KELLY J.A., SIELECKI A.R., SYKES B.D., JAMES M.N.G. & PHILLIPS D.C. –1979– *Nature* **282**, 875.

KERR M.A., WALSH K.A. & NEURATH H. –1976– *Biochemistry* **15**, 5566.

KIRBY A.J. –2001– *Nat. Struct. Biol.* **8**, 737.

KLINMAN J.P. –1981– *C.R.C. Crit. Rev. Biochem.* **10**, 39.

KNOWLES J.R. –1970– *Phil. Trans. Roy. Soc. London* **B257**, 135.

KOMIVES E.A., CHANG L.C., LOLIS E., TILTON R.F., PETSKO G.A. & KNOWLES J.R. –1991– *Biochemistry* **30**, 3011.

KOSHLAND D.E. –1953– *Biol. Rev.* **28**, 416.

KOSSIAKOFF A.A. & SPENCER S.A. –1981– *Biochemistry* **20**, 6462.

KUO L.C., FUKUYAMA J.M. & MAKINEN M.W. –1983– *J. Mol. Biol.* **163**, 63.

LAI C.Y. –1975– *Arch. Biochim. Biophys.* **166**, 358.

LASHTCHENKO P. –1909– *Z. Hug. Infektionkrankh.* **64**, 419.

LEDERER F. –1987– in *Flavins and flavoproteins*, Walter de Gruyter and Co, Berlin, 513.

LEDERER F. & MATTHEWS F.S. –1987– in *Flavins and flavoproteins*, Walter DE GRUYTER and Co, Berlin, 133.

LEVITT M. –1974– in *Peptides, polypeptides and proteins*, E. BLOUT, F. BOVEY, M. GOODMAN & N. LOTAN eds, John Wiley pub., New York, 99.

LIPSCOMB W.N. –1970– *Accounts for Chem. Res.* **3**, 81.

LIPSCOMB W.N. –1980– *Proc. Natl Acad. Sci. USA* **77**, 3875.

LONSDALE-ECCLES J.D., NEURATH H. & WALSH K. –1978– *Biochemistry* **17**, 2805.

LONSDALE-ECCLES J.D., KERR M.A., NEURATH H. & WALSH K. –1979– *FEBS Lett.* **100**, 157.

MAKINEN M.W., YAMAMURA K. & KAYSER E.T. –1976– *Proc. Natl Acad. Sci. USA* **68**, 3882.

MATTHEWS D.A., ALDEN R.A., BIRKTOFT J.J. & KRAUT J. –1975– *J. Biol. Chem.* **250**, 7120.

MATTHEWS F.S. & ZIA Z.X. –1987– in *Flavins and flavoproteins*, Walter DE GRUYTER and Co, Berlin, 123.

MAVRIDIS I.M., HATADA M.H., TULINSKI A. & LEBRODA L. –1982– *J. Mol. Biol.* **162**, 419.

MAY A., VAS M., HARLOS K. & BLAKE C.C.F. –1996– *Proteins: Struct. Funct. Genet.* **24**, 292.

MCLACHLAN A.D. & SHOTTON D.H. –1971– *Nature* **229**, 202.

MCPHILLIPS T.M., HSU B.T., SHERMAN M.A., MAS M.T. & REES D.C. –1996– *Biochemistry* **35**, 4118.

MIZYGUCHI H., HAYASHI H., OKADA K., MIYAHARA I., HIROTSU K. & KAGAMIYAMA H. –2001– *Biochemistry* **40**, 353.

MORRIS A.J. & TOLAN D.R. –1993– *J. Biol. Chem.* **268**, 1095.

MOUAWAD L., DESMADRIL M., PERAHIA D., YON J.M. & BROCHON J.C. –1990– *Biopolymers* **30**, 1151.

NAVIA M.A., FITZGERALD P.M.D., MCKEEVER B.M., LEU C.T., HEIMBACH J.C., HERBER W.K., SIGAL I.S., DARKE P.L. & SPRINGER J.P. –1989– *Nature* **337**, 615.

NEURATH H., WALSH K.A. & WINTER W.P. –1967– *Science* **158**, 1638.

NICKBARG E.B. & KNOWLES J.R. –1988– *Biochemistry* **27**, 5939.

NICKBARG E.B., DAVENPORT R.C., PETSKO G.A. & KNOWLES J.R. –1988– *Biochemistry* **27**, 5948.

OUCHNINIKOV Y.A., EGOROV C.A., ALDANOVA N.A., FEIGINA M.Y., LIPKIN V.M., ABDULAEV N.G., GRISHIN E.V., KISELEV A.P., MODYANOV N.M., BRAUENSTEIN A.E., POLYANOVSKI O.L. & NISONOV V.V. –1973– *FEBS Lett.* **29**, 31.

PARSON S.M. & RAFTERY M.A. –1972– *Biochemistry* **11**, 1623.

PHILLIPS D.C., FLETTERICK R. & RUTTER W.J. –1990– *J. Biol. Chem.* **265**, 20692.

ROSSMANN M.G., MORAS D. & OLSEN K.W. –1974– *Nature* **250**, 194.

REES D.C., LEWIS M. & LIPSCOMB W.N. –1983– *J. Mol. Biol.* **168**, 367.

SCHINDLER M., ASSAF Y., SHARON N. & CHIPMAN D.M. –1977– *Biochemistry* **16**, 423.

SCHMIDT M.F. & HERRIOT J.R. –1976– *J. Mol. Biol.* **103**, 175.

SEYDOUX F., NÉMÉTHY G. & YON J.M. –1969– *Biochim. Biophys. Acta* **171**, 145.

SIELECKI A.R., FEDOROV A., BOODHOO A., ANDREVA N.S. & JAMES N.G. –1990– *J. Mol. Biol.* **214**, 143.

SIELECKI A.R., FUJINAGA M., READ R.J. & JAMES N.G. –1991– *J. Mol. Biol.* **219**, 671.

SHOTTON D.M. & WATSON H.C. –1970– *Nature* **225**, 811.

SNELL E.E. –1945– *J. Am. Chem. Soc.* **67**, 194.

STROUD R.M., KAY L.M. & DICKERSON R.E. –1974– *J. Mol. Biol.* **83**, 185.

SUZUKI N. & WOOD W.A. –1980– *J. Biol. Chem.* **255**, 3427.

SYGUSH J., BRANDY J., & ALLAIRE M. –1987– *Proc. Natl. Acad. Sci. USA* **84**, 7846.

TONEY M.D. & KIRSCH J.F. –1991– *Biochemistry* **30**, 7456.

TULINSKY A. & BLEVINS R.A. –1987– *J. Biol. Chem.* **262**, 7737.

UMEYAMA H., NAGAWA S. & KUDO T. –1981– *J. Mol. Biol.* **150**, 409.

URBAN P. & LEDERER F. –1985– *J. Biol. Chem.* **260**, 11115.

VALLEE B.L. & NEURATH H. –1955– *J. Biol. Chem.* **217**, 253.

VOCADIO D.J., DAVIES G.J., LAINE R. & WITHERS S.G. –2001– *Nature* **412**, 835.

WALTER J. & BODE W.–1983– *Hoppe Seylers Z. Physiol. Chem.* **364**, 949.

WATSON H.C., WALKER N.P.C., SHAW P.J., BRYANT T.N., WENDELL P.L., FOTHERGILL L.A., PERKINS R.E., CONROY S.C., DOBSON M.J., TUITE M.F., KINGSMAN A.J. & KINGSMAN S.M. –1982– *Embo. J.* **1**, 1635.

WEBB M.R. & TRENTHAM D.R. –1980– *J. Biol. Chem.* **255**, 1775.

WILKS H.M., HART K.W., FEENEY R., DUNN C.R., MUIRHEAD H., CHIA W.N., BARSTOW D.A., ATKINSON T., CLARKE A.R. & HOLBROOK J.J. –1988– *Science* **242**, 1541.

WRIGHT C.S., ALDEN R.A. & KRAUT J. –1970– *Nature* **221**, 235.

XIA Z.X., SHAMALA H., BETHGE P.H., LIM L.W., BELLAMY H.D., XUONG N.H., LEDERER F. & MATHEWS F.S. –1987– *Proc. Natl Acad. Sci. USA* **84**, 2629.

PART V

REGULATION OF ENZYME ACTIVITY

INTRODUCTION

If we consider cellular metabolism as the integrated ensemble of a few thousand chemical reactions in microorganisms, a few tens of thousands in superior organisms, it is fairly easy to understand that at least some of these reactions must be regulated. The aim of regulation is not to maintain these reactions at a constant level, but, on the contrary, to vary their rate in order to cope with cellular metabolism's requirements and to allow it to adapt to variations in the environment.

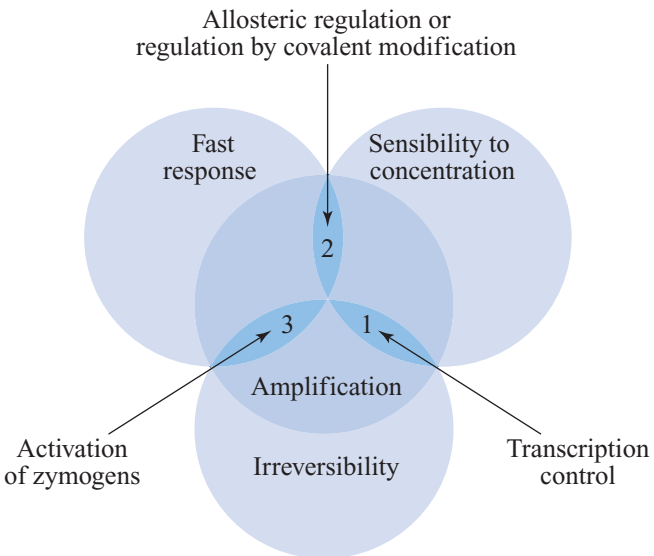


Diagram showing the inter-relation between the fundamental aspects of metabolic regulation by means of the three main mecanisms

1: transcription control; 2: allosteric regulation or regulation by covalent modification; 3: activation of zymogens. Each of the three mecanisms implies a combination of three of the four aspects noted in the circles; the combination is represented by the overlapping section of the circles (from NEURATH, 1977).

The regulation of a chemical reaction rate in the cell at a given time can be the result of a variation in intracellular concentration of the enzyme which catalyses the reaction, either due to an accrued biosynthesis, or to the activation of an inactive precursor, or, in the opposite direction, to the reversible or irreversible inactivation of the enzyme. For metabolic requirements, the catalytic power of an enzyme can also be

regulated either by the covalent binding of some chemical groups, or by non-covalent binding of certain metabolites or even by the reversible association of the enzyme with another protein. As shown in the diagram, these different types of regulation differ in properties such as reversibility, the rate of the response or its amplifying nature. The differences allow the cell to cope with very diverse metabolic requirements.

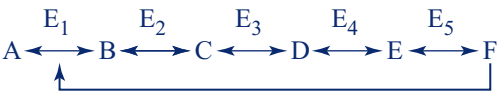
We shall consider the different types of regulation, except that of the *de novo* synthesis of enzymes which is beyond the scope of this book.

13 – REGULATION BY NON-COVALENT INTERACTIONS

The regulation of enzymatic activity through non-covalent interactions allows the cell to react quickly and in a reversible manner.

13.1. ALLOSTERIC REGULATION

So-called "allosteric" enzymes are responsible for the phenomenon of **regulation through retroinhibition** which is observed in most metabolic pathways. The phenomenon arises when one of the final metabolites inhibits the activity of the first enzyme that leads to its biosynthesis:



If we consider the transformation of metabolite A into metabolite F, which involves five steps each catalysed by a specific enzyme (E_1 to E_5), retroinhibition consists in the fact that the final metabolite F inhibits the activity of the enzyme E_1, limiting in this manner its own production. The described schematic case is the simplest one. Frequently, regulatory enzymes are also subject to an **activation** process by a metabolite which belongs to another metabolic pathway, which leads to crossed regulation phenomena, and we shall see examples of this. Several antagonisms of the type described contribute, for example, to equilibrate the biosynthesis of purine and pyrimidine nucleotides in the cell. The enzymes which are the subject of this type of regulation are called **allosteric enzymes** (MONOD et al., 1963). The term was chosen to express the fact that the regulating metabolite (F) whose structure is different than that of the substrate, binds to a stereospecific site (regulatory site), which is distinct from the catalytic site. The hypothesis was verified in all the cases of allosteric enzymes whose structure has been determined with sufficient precision.

The majority of allosteric enzymes are oligomers which show cooperative effects in substrate binding. In the simplest case, the binding of the first substrate molecules favours the binding of the following ones. In other cases, it is the catalytic efficiency which increases with the proportion of catalytic sites occupied by the substrate.

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_14, © Springer-Verlag Berlin Heidelberg 2010

Later on, we shall return to the distinction which should be made between **allostery and cooperativity**. The phenomenon of cooperativity between catalytic sites of an oligomeric enzyme also allows it to modulate in a positive or negative manner the affinity of its substrate sites as a function of the substrate concentration.

13.2. PHENOMENOLOGICAL ASPECT OF COOPERATIVITY

When the apparent affinity of an oligomeric enzyme for the substrate increases with the degree of occupation of the catalytic sites, the curve of the variation of v, the reaction rate, as a function of substrate concentration [s], is no longer a hyperbola but displays a sigmoidal shape (Fig. 13.1 opposite). The consequence is that the saturation curve cannot be linearised as was shown earlier (Chap. 2). Figure 13.1b and c show the curves obtained from the LINEWEAVER-BURK and EADIE-HOFSTEE plots, respectively.

In the case of the enzyme not displaying a cooperative phenomenon (enzyme displaying Michaelian behaviour), we can, on the basis of the MICHAELIS equation, define the function describing the saturation by the substrate using the equation:

$$Y_s = \frac{v}{V_m} = \frac{s / K_m}{1 + (s / K_m)}$$

s/K_m is called the **reduced substrate concentration.**

As described in Chap. 5, the reaction rate varies with substrate concentration following a hyperbolic law and attains the maximum value V_m when the enzyme is saturated by the substrate. When v varies from 0 to V_m, the saturation function Y_s varies from 0 to 1. If the considered reaction follows the MICHAELIS-MENTEN law, the ratio of substrate concentrations which correspond to 90% ($s_{0.9}$) and 10% ($s_{0.1}$) of enzyme saturation has a defined value: $s_{0.9}/s_{0.1} = 81$. In the case of an enzyme displaying cooperative phenomenon, the ratio is inferior to 81. On the contrary, this ratio is superior to 81 in the case of enzymes displaying **anticooperativity**, in which case the binding of the first substrate molecules de-favours the binding of the subsequent ones (Fig. 13.1a).

Differences in behaviour are also evident in the variation of reaction rate as a function of the logarithm of substrate concentration. In the case of a Michaelian enzyme, the curve is symmetric on either side of the half-saturation point $s_{0.5}$ and $(s_{0.9})(s_{0.1})/(s_{0.5})^2 = 1$. The symmetry of the curve is also observed in non-Michaelian systems, in particular in the case where the saturation law presents cooperative phenomena. When there is anticooperativity however, the curve is no longer symmetrical.

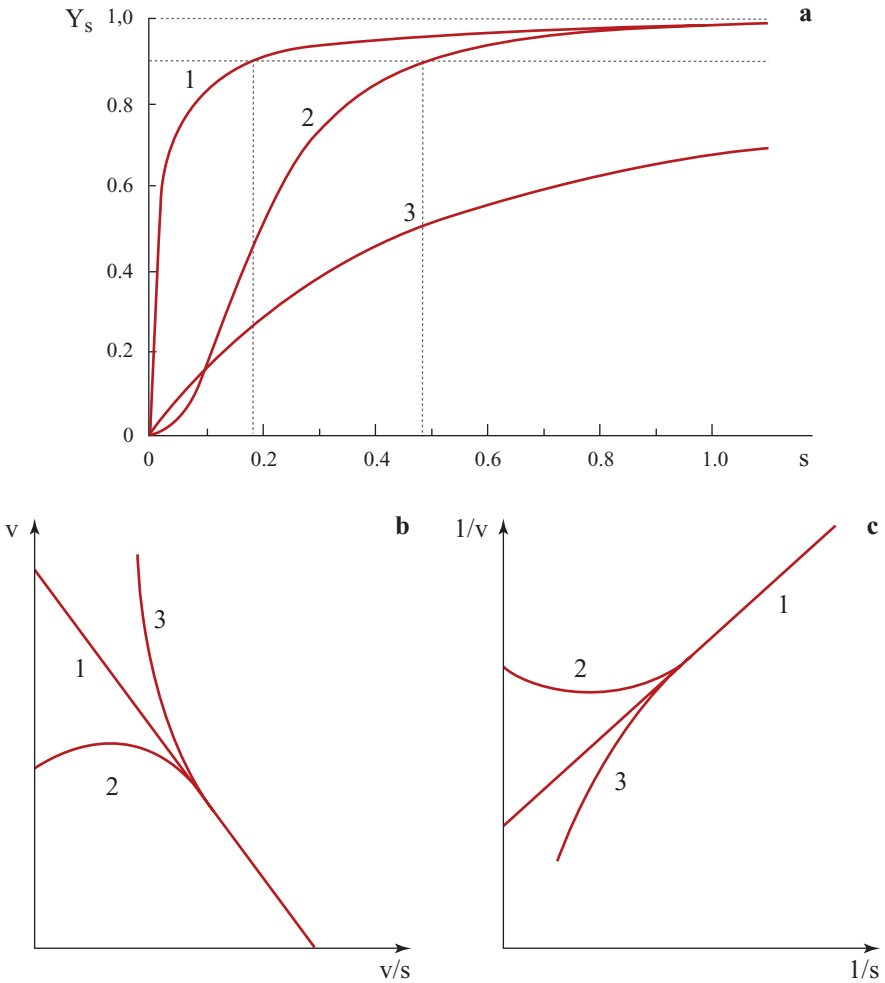


Fig. 13.1 (a) variation of the saturation function by the substrate, (1) for a Michaelian enzyme; (2) for an enzyme showing cooperative effects; (3) for an enzyme showing anti-cooperative effects – (b) shape of the curves in the EADIE-HOFSTEE plot – (c) shape of the curves in the LINEWEAVER-BURK plot

Allosteric enzymes are inhibited or activated by allosteric effectors. Phenomenologically, if the substrate saturation curve is a sigmoid (cooperative effects), the presence of an inhibitor accentuates the sigmoidal nature of the curve. On the contrary, the presence of an activator lessens it: when the activator concentration becomes sufficiently high, the curve tends towards a hyperbola (Fig. 13.2 below). Everything occurs as if the allosteric inhibitor increases the apparent K_m value without changing the maximum rate, ***however it is not competitive inhibition***. The activator reduces the value of K_m without changing the maximum rate.

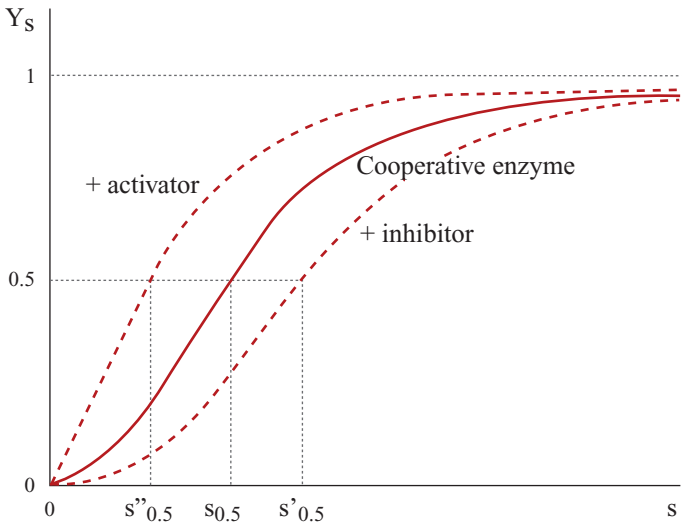


Fig. 13.2 Saturation by the substrate of an enzyme displaying cooperative effects
In the absence of an effector, in the presence of an activator, and in the presence of an inhibitor

Other allosteric enzymes can display different behaviour with respect to effectors. Their kinetics obeys the MICHAELIS law in relation to the substrate saturation of the enzyme. In the presence of an inhibitor, the saturation curve of the enzyme by the substrate stays Michaelian and the same is true in the presence of an activator. The inhibitor reduces the apparent maximal rate; the activator increases it. However, it is ***not a classical non-competitive mechanism***. The law describing the satur-ation of the enzyme either by the inhibitor or the activator is a sigmoid, and non-Michaelian. When studying an inhibition or activation phenomenon, the test is essential before deciding on the nature of the model.

How can these behaviours be explained? The study of different models that have been proposed, allows us to better understand the other function of enzymes, the **regulating function**.

13.3. PHENOMENOLOGICAL MODELS

13.3.1. THE HILL EQUATION

The first formalisms to describe cooperativity phenomena were proposed by HILL, (1910) to explain oxygen binding to haemoglobin (see Chap. 2).

In the case of an enzymatic reaction, the HILL relation can be written in the follow-ing form:

$$Y_s = \frac{s^{nH}}{K_m + s^{nH}}$$

or:

$$\frac{Y_s}{1 - Y_s} = \frac{s^{nH}}{K_m}$$

The HILL equation is:

$$\log Y_s/[1 - Y_s] = \log K' + n_H \log [s]$$

The HILL proposition consists in representing the variation of $\log [Y_s/1 - Y_s]$ as a function of $\log [s]$. In the case of an enzyme the variation of $\log [v/V_m - v]$ as a function of $\log [s]$ is represented (Fig. 13.3). The slope at the half-saturation point $s_{0.5}$ is n_H, the cooperativity index or HILL number. In HILL's formalism, it can be shown that the numerical value of n_H is between 1 and n, the number of subunits of the enzyme. Today, HILL's plot is mainly used to furnish a phenomenological index, which indicates the degree of cooperativity of analysed saturation curves, without taking into account the underlying molecular mechanisms. It has been shown by WYMAN (1964) that at extreme levels of substrate concentration the slope of HILL's curve tends towards 1. The extrapolation of the extreme slopes theoretically allows the determination of the association constants K_1 and K_n, corresponding to the binding of the substrate to the first and the nth site, respectively (Fig. 13.3).

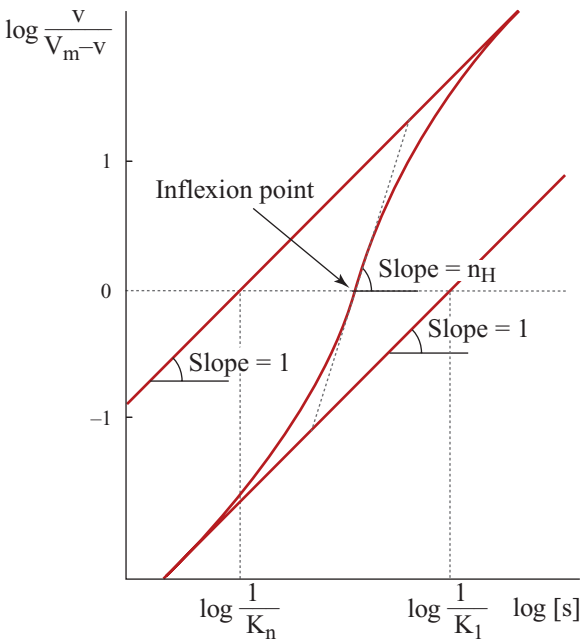


Fig. 13.3 Substrate saturation curve of an enzyme displaying cooperative effects, according to HILL's plot

WYMAN (1964) defined the interaction energy between subunits as a function of the two association constants through the relation:

$$\Delta G_I = -RT \ln K_n/K_1$$

The apparent interaction energy is therefore defined as the amount of energy that must be furnished to change the affinity from K_1 to K_n. On the basis of this definition, WYMAN calculated that the apparent interaction energy is equal to –3 kcal. mol^{-1} per tetramer in the case of haemoglobin. The energy is distributed over the different interfaces between subunits.

In the case of a dimer: $n_H = 2/\left(1+\sqrt{K_1/K_2}\right)$, where K_1 and K_2 represent the intrinsic affinity constants for the first and second sites, respectively. Additionally, as was previously defined:

$$\Delta G_I = -RT \ln K_2/K_1$$

The combination of the two preceding equations establishes a relationship between HILL's index and the interaction energy between subunits:

$$\Delta G_I = +2RT \ln (2 - n_H)/n_H$$

The interaction energy is the difference between the free energy of ligand binding to the second site and the free energy of ligand binding to the first site.

$$\Delta G_I = (\Delta G_1 - \Delta G_2)/2 - \Delta G_1$$

or:

$$\Delta G_I = \tfrac{1}{2} (\Delta G_2 - \Delta G_1)$$

The interaction energy is therefore zero for a Michaelian oligomeric enzyme ($\Delta G_2 = \Delta G_1$). It is negative in the case of an enzyme displaying cooperativity and positive in the case of an enzyme displaying anticooperativity [LEVITZKI, 1978].

13.3.2. THE ADAIR EQUATION

In 1925, after showing that haemoglobin is a tetramer, ADAIR proposed an explanation for the sigmoidal nature of the protein's saturation curve by oxygen as a function of four association constants. As for the preceding model, the thermodynamical model makes no hypothesis on the structural mechanism involved in the variation of the constants. The ADAIR equation in its most general form is constructed in the following way. If we consider a protein of n identical subunits, each with a ligand binding site, and suppose that the ligand binding does not modify the degree of oligomerisation, the ligand binding can be described by the following equilibria:

$$E + S \longleftrightarrow ES \qquad \Phi_1 = \frac{[ES]}{[E][S]}$$

$$E + 2S \longleftrightarrow ES_2 \qquad \Phi_2 = \frac{[ES_2]}{[E][S]^2}$$

$$E + 3S \longleftrightarrow ES_3 \qquad \Phi_3 = \frac{[ES_3]}{[E][S]^3}$$

$$E + iS \longleftrightarrow ESi \qquad \Phi_i = \frac{[ES_i]}{[E][S]^i}$$

$$E + nS \longleftrightarrow ES_n \qquad \Phi_n = \frac{[ES_n]}{[E][S]^n}$$

where:

$\Phi_1 = K_1$, $\Phi_2 = K_1K_2$, $\Phi_3 = K_1K_2K_3$, $\Phi_i = K_1K_2K_3\ldots K_i$ and $\Phi_n = K_1K_2K_3\ldots K_i\ldots K_n$

The average number of ligand molecules bound per enzyme molecule is given by:

$$N_s = \frac{[ES] + 2[ES_2] + 3[ES_3] + ... + i[ES_i] + ... + n[ES_n]}{[E] + [ES] + [ES_2] + [ES_3] + ... + [ES_i] + ... + [ES_n]}$$

or:

$$N_s = \frac{\Phi_1[S] + 2\,\Phi_2[S]^2 + 3\,\Phi_3[S]^3 + ... + i\,\Phi_i[S]^i + ... + n\,\Phi_n[S]^n}{1 + \Phi_1[S] + \Phi_2[S]^2 + \Phi_3[S]^3 + ... + \Phi_i[S]^i + ... + \Phi_n[S]^n}$$

However, the expressions described above remained at a phenomenological level, and did not offer any interpretation of the change in intrinsic constant. In 1935, PAULING was the first who tried to relate the variations of affinity constants to the geometry of the protein. He supposed that the four haemoglobin sites are equivalent and therefore possess the same intrinsic affinity constant for oxygen. Furthermore, he introduced an interaction factor depending on the arrangement and conformation of the four subunits. The explanation of the change in the interactions between hemes upon oxygen binding was based upon conformational variations in the protein. Crystallographic studies brought the experimental proof of PAULING's interpretation (1960). Various models were developed thirty years later, all of which explained the observed effects in terms of protein conformational changes.

13.4. THE CONCERTED MODEL (MONOD, WYMAN & CHANGEUX, 1965)

13.4.1. DEFINITION

The concerted model developed by MONOD, WYMAN and CHANGEUX is based on the following postulates:

- In the absence of any ligand, a protein which displays cooperative effects exists as an equilibrium of two conformations:

$$R_0 \overset{L_0}{\longleftrightarrow} T_0$$

The equilibrium is defined by the **allosteric constant** L_0, such that $L_0 = (T_0)/(R_0)$. The two states differ by the interaction free energy between the subunits, the

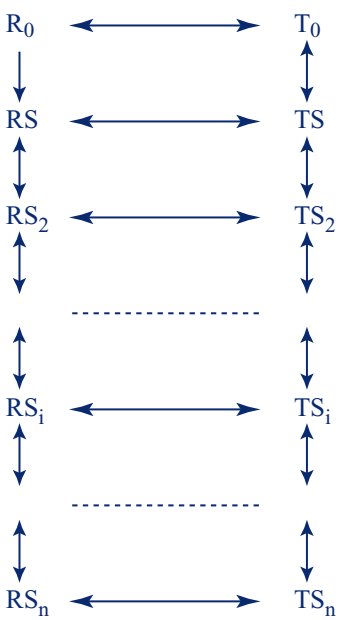
interaction energy being stronger in the T state, in which the subunits are subjected to stronger conformational constraints.

- All the protomers in the oligomer are identical and display at least one symmetry axis between them, in both the R and T states. The concept of **symmetry** assumes a great importance in the model. It follows that the transition from T to R is concerted, without formation of hybrid intermediate species, the symmetry is conserved on the passage from one form to the other.
- Each protomer possesses one substrate binding site, these sites are equivalent and independent (see Chap. 2). The sites are related to one another by the same symmetry axis as the protomers in the oligomer.
- Each protomer can possess one or several regulatory sites capable of binding an allosteric inhibitor or activator, which are distinct from the catalytic site. The regulatory sites possess the same symmetry axis as the protomers in the oligomer.
- The model distinguishes **homotropic effects** between sites binding the same ligand and **heterotropic effects** between sites which bind different ligands (substrates and effectors for example).

From the described postulates, two types of system were proposed, K systems and V systems.

13.4.2. K SYSTEMS

In K systems, R and T states display respectively strong and weak affinities for the substrate S, by definition ($K_R < K_T$, K_R and K_T are the intrinsic substrate dissociation constants for R and T forms). They both possess the same catalytic constants however ($k_{catR} = k_{catT}$). The consequence is that, when the substrate concentration increases, the equilibrium is shifted towards the R state, although the binding sites are not yet saturated. If the protein has n substrate binding sites, the following equilibria exist:



In the model, all the sites are considered equivalent and independent for both R and T states. Therefore, there are multiple equilibria of n equivalent and independent sites (vertical equilibria) linked by conformational equilibria (horizontal equilibria). The model allows the protein's **state function** to be defined, that is, the fraction

of molecules in the R state, and the **saturation function**, fraction of sites occupied by the substrate. The preceding definitions imply that, as the concentration of S increases, the state function pre-empts the saturation function.

In the following we shall consider an oligomeric enzyme formed of identical protomers each with a catalytic site and the following definitions:

- n = the number of protomers;
- K_R and K_T, the substrate microscopic dissociation constants for the R and T states, respectively;
- $\alpha = [s]/K_R$ the reduced substrate concentration;
- $c = K_R/K_T$ the ratio of the microscopic dissociation constants. The ratio is named the non-exclusion coefficient because it takes into account the fact that S can also bind to the T state, which is the opposite of what occurs in a so-called exclusive system. In the case of an exclusive system, $c \longrightarrow 0$.

As a function of the above parameters, the saturation function is written as:

$$Y_s = \frac{\text{(sum of all molecules that have bound substrate)}}{\text{(total protein)}}$$

$$Y_s = \frac{\sum_{i=1}^{n} RS_i + \sum_{i=1}^{n} TS_i}{\sum_{i=0}^{n} RS_i + \sum_{i=0}^{n} TS_i}$$

or:

$$Y_s = \frac{L_0 c\alpha(1+c\alpha)^{n-1} + \alpha(1+\alpha)^{n-1}}{L_0(1+c\alpha)^n + (1+\alpha)^n}$$

The above expression has two terms; the term $\alpha(1 + \alpha)^n$/[denominator] expresses the R state saturation and $L_0c\alpha(1 + c\alpha)^{n-1}$/[denominator] expresses the T state saturation.

The state function is given by the relation:

$$R = \frac{\sum_{i=0}^{n} RS_i}{\sum_{i=0}^{n} RS_i + \sum_{i=0}^{n} TS_i}$$

that is:

$$R = \frac{(1+\alpha)^n}{L_0(1+c\alpha)^n + (1+\alpha)^n}$$

It is useful to express Y_s as a function of R:

$$Y_s = R\frac{\alpha}{1+\alpha} + (1-R)\frac{c\alpha}{1+c\alpha}$$

This expression clearly shows that the observed cooperativity is the result of a Michaelian saturation, linked to a transformation function of the protein which varies as a function of the saturation. The first term expresses this fact for the R form, the second for the T form. The difference between the above and the MICHAELIS laws is therefore due to the protein transformation function.

In K systems in the concerted model, the homotropic effects are necessarily cooperative. The cooperativity depends on the values of the parameters L_0 and c. If the substrate displays a higher affinity for the R state, c will be less than 1. If the constant L_0 is large, the T state will be the dominant form in the absence of a ligand. The larger L_0, the smaller c, and the more important the cooperativity.

In the case of exclusive substrate binding in the R state of the enzyme, the expressions given for the saturation and state functions are simplified; they become:

$$Y_s = \frac{\alpha(1+\alpha)^{n-1}}{(1+\alpha)^n + L_0}$$

and:

$$R = \frac{(1+\alpha)^n}{(1+\alpha)^n + L_0}$$

and the relationship between the two:

$$Y_s = R\frac{\alpha}{1+\alpha}$$

Additionally, if the constant L_0 is very small, $L_0 < 50$, the expressions are further simplified and:

$$Y_s = \frac{\alpha}{1+\alpha} \quad \text{and} \quad R = 1$$

which is the expression corresponding to a Michaelian saturation law. The MICHAELIS law can therefore be considered as a limiting case of the allosteric model of MONOD, WYMAN and CHANGEUX.

It is to be noted that if c = 1, that is to say, when both R and T states have the same substrate affinity, which is true in V systems as shall be discussed later, the saturation function is also Michaelian.

If we consider that the two extreme states differ only in their substrate affinity, the reaction rate will be proportional at all times to Ys and the proportionality factor is the catalytic constant k_3.

$$v = k_3 Y_s = k_3[ES]$$

The more complex situation in which the extreme states T and R also differ in their catalytic efficiency will be considered later.

The model was also conceived to take into account **heterotropic effects that imply allosteric inhibition or activation.** MONOD, WYMAN and CHANGEUX considered that an inhibitor binding site and an activator binding site is present on each proto-

mer. The state which has the highest substrate affinity also has the highest activator affinity. The state which has the lowest substrate affinity displays the highest affinity for the inhibitor. The allosteric activation and inhibition are therefore explained by the shift of the conformational equilibrium towards the state which has the highest substrate affinity (activation), and towards the state which has the lowest substrate affinity (inhibition). Everything occurs as if the conformational constant L_0 is modified by the presence of the effectors. As shall be seen later, this postulate is not always in agreement with experimental data.

The same expressions as given previously are valid for the functions Ys and R, where L_0 is replaced by a new allosteric constant L′ such that:

$$L' = L_0 \frac{[(1+\beta)(1+\gamma e)]^n}{[(1+\beta d)(1+\gamma)]^n}$$

with: $\beta = [I]/K_{I,T}$; $\beta d = [I]/K_{I,R}$; $\gamma = [A]/K_{A,R}$ and $\gamma e = [A]/K_{A,T}$

When **the binding is exclusive**, I only binds to the T state and A to the R state, and the expression simplifies to become:

$$L' = L_0 \frac{(1+\beta)^n}{(1+\gamma)^n}$$

L′ varies with the saturation degree of the enzyme for each effector. L′ increases when the inhibitor concentration increases, and decreases with the activator concentration. L′ is an apparent allosteric constant such that:

$$L' = \frac{\sum_0^n T_I}{\sum_0^n R_A}$$

Since the substrate and the activator have a preferential affinity for the R state, when all the enzyme is saturated by the activator, the cooperativity for substrate binding disappears and the Ys function becomes Michaelian. In a reciprocal fashion, when the enzyme is saturated by the substrate, the activation phenomenon disappears. If the activator concentration is very high, Y_s tends towards a maximum value $Y_{s,max}$, and under these conditions:

$$\frac{Y_s}{Y_{s,max} - Y_s} = \frac{R}{1-R} = \frac{(1+\alpha)^n}{L'} = \frac{(1+\alpha)^n (1+\gamma)^n}{L}$$

By plotting the nth square root of the ratio $Y_s/(Y_{s,max} - Y_s)$, which is equivalent to $v/(V_A - v)$, as a function of s for a given concentration of the activator A, a line is obtained. The intersection of the line with the x-axis gives $-K_R$. The intersection with the y-axis gives $\sqrt[n]{1/L'}$. By performing an experiment for several activator concentrations, plotting the determined value of $\sqrt[n]{1/L'}$ as a function of A, and plotting a second curve, the values of K_A and L_0 are obtained (Fig. 13.4 below). The value of the ratio $R/(1 - R)$ can be determined experimentally from the enzymatic

reaction rate. Under the described conditions, the ratio is equal to v/(V_A – v), where V_A is the reaction rate observed when the enzyme is saturated by the activator.

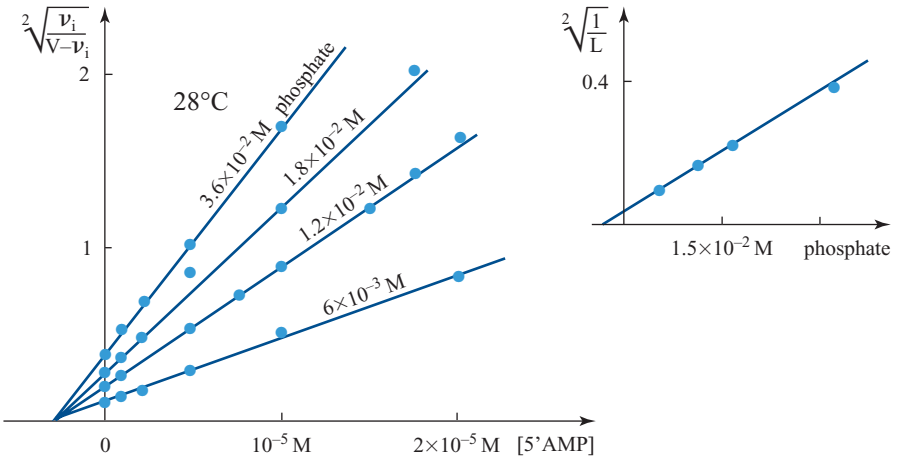


Fig. 13.4 Determination of the dissociation constant of AMP for the R state of phosphorylase b, in the presence of different concentrations of phosphate
The insert corresponds to the secondary plot. (Reprinted from *Biochem. Biophys. Res. Commun.*, **28**, BUC H., On the allosteric interaction between 5′AMP and orthophosphate on phosphorylase b. Quantitative kinetic predictions, 62. © (1967) with permission from Elsevier)

◤ RUBIN and CHANGEUX (1966) expressed the apparent allosteric constant L′ as a function of the substrate concentration at the half-saturation point, $\alpha^{½}$, so that Ys = ½. The expression of L' is therefore:

$$L' = \frac{(\alpha^m - 1)(1+\alpha^m)}{(1-\alpha^m c)(1+\alpha^m c)}$$

When c << 0, log $\alpha^{½}$ is a sigmoidal function of log L′.

The values of Y_s and R have also been analysed by RUBIN and CHANGEUX in the case of non-exclusive binding and for widely varying parameters. The function R varies from $(L' + 1)^{-1}$ when α tends towards zero, to $(L'c^n + 1)^{-1}$, when α tends towards infinity. The authors define the "allosteric domain", Q, by the difference in the two values of R:

$$Q = R_{\alpha\to\infty} - R_{\alpha\to 0} = \frac{L'(1+c^n)}{(L'+1)(L'c^n+1)}$$

the maximum value of Q is obtained when $L = c^{-n/2}$:

$$Q_{max} = \frac{(1 \quad c^{n/2})}{(1 \quad c^{n/2})}$$

Q_{max} tends towards 1 when c tends towards zero, that is, for an exclusive binding, on the condition that L >> 1 (equilibrium displaced towards the T form). Figure 13.5 shows the "allosteric zone" for non-exclusive ligand binding. The state function R is represented as a function of log L for the two extreme values of substrate concentra-

tion ($\alpha = 0$ and $\alpha \longrightarrow \infty$). The "allosteric zone" is represented in Fig. 13.5 for a given value of L_0. The distance between the two curves is indicated by a horizontal line which depends on the relative ligand affinity for the R and T states and on the number of protomers. The distance has the value $\log c^{-n}$, where n is the number of protomers. The curves are given for $n = 2$, $c = 0.1$ and $L_0 = 100$. L′ varies with the different effectors.

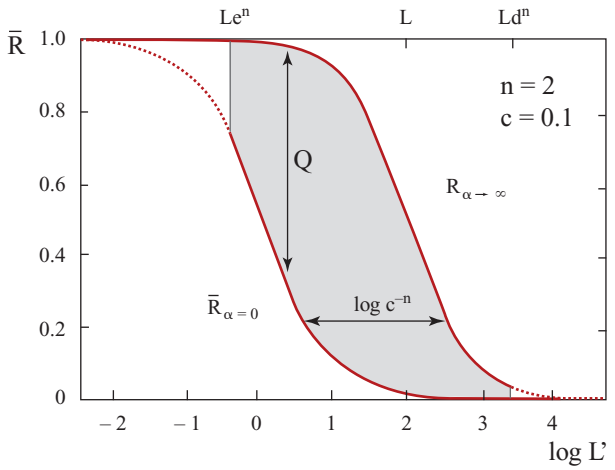


Fig. 13.5 Allosteric transition zone upon non-exclusive ligand binding
The grey region is delimited by the state function R at zero and infinite substrate concentrations. For the diagram, the parameters are $L_0 = 100$, $c = 0.1$ and $e = 0.4$. The "allosteric domain" Q_{max} under these conditions is equal to 0.82 (Reprinted from *J. Mol. Biol.*, **21**, RUBIN M. & CHANGEUX J.P., On the nature of allosteric transitions: implications of non-exclusive ligand binding, 265. © (1966) with permission from Elsevier)

13.4.3. V SYSTEMS IN THE CONCERTED MODEL

In the preceding description, it has been considered that the T and R conformations differ in their intrinsic dissociation constants K_T and K_R in relation to the substrate but display the same catalytic constant, $k_{cat,T} = k_{cat,R}$. The model also allows for V systems where the two extreme conformations display the same affinity for the substrate ($K_R = K_T$), but differ in their catalytic constants ($k_{cat,R} > k_{cat,T}$). It follows that the substrate saturation curve of such an enzyme is a hyperbola and does not display any sign of cooperativity. The enzymes differ however from enzymes displaying Michaelian behaviour if the equilibrium between the two extreme conformations is displaced by allosteric effectors. In V systems, the activator has a higher affinity for the enzyme conformation which has the highest catalytic constant, by definition. On the contrary, the inhibitor has a preferential affinity for the form of the enzyme with the weakest catalytic constant. Therefore, in the presence of both the inhibitor and the activator, the substrate saturation curve stays a hyperbola, although the maximum rate varies with effector concentration (Fig. 13.6a). If the

effectors act strictly on the T ⟷ R equilibrium and if, when saturated in inhibitor, the equilibrium is entirely displaced towards the T state, it becomes possible to determine $k_{cat,T}$. The same is true for $k_{cat,R}$ in the presence of saturating concentrations of activator. From a phenomenological point of view, everything occurs as if it would be a Michaelian enzyme sensitive to the presence of an activator or a non-competitive inhibitor. However, if homotropic interactions are present between regulatory effector sites, differences from the MICHAELIS law are observed. For example, the variation of the maximal rate as a function of inhibitor concentration or allosteric activator has a sigmoidal shape (Fig. 13.6b).

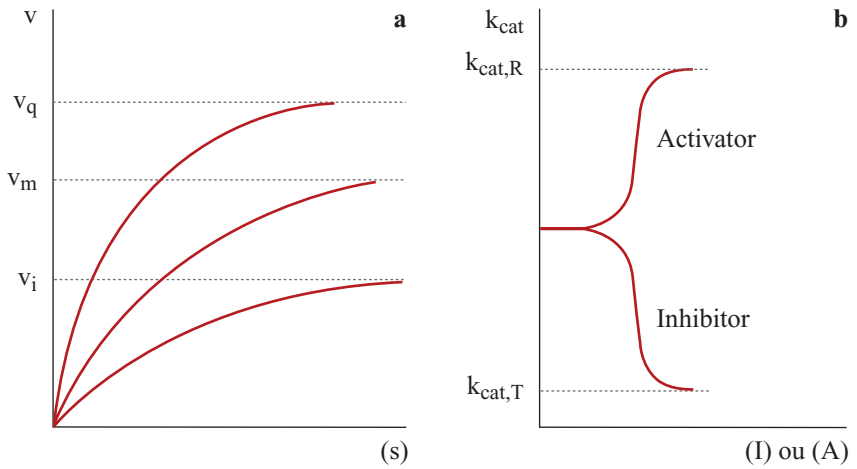


Fig. 13.6 V systems

*(**a**) variations in the rate as a function of substrate concentration, in the absence (middle curve) and in the presence of allosteric effectors: activator (top curve) and inhibitor (bottom curve) – (**b**) variations in the experimental catalytic constant as a function of allosteric inhibitor or activator concentration*

Nothing prevents an enzyme from displaying mixed behaviours, and that its extreme conformations differ both in their substrate dissociation constant and in their catalytic constant. The two types of parameters are extremely sensitive to very discrete variations in the topology of the catalytic site and it is unlikely that a conformational change in the substrate binding site does not influence the catalytic constant.

It is to be noted that the concerted model does not explain anticooperativity since the ligand cannot displace the thermodynamic equilibrium R ⟷ T towards the form for which it has the lowest affinity.

13.5. SEQUENTIAL MODEL [KOSHLAND, NÉMÉTHY & FILMER, 1966]

13.5.1. DEFINITION

The sequential model can explain anticooperative phenomena as well as cooperative ones. This model is based on the induced-fit theory developed previously by KOSHLAND to explain the efficiency of enzymatic reactions based on the flexibility of protein structures (Chap. 10).

The induced-fit theory is based on three postulates:

- the substrate binding to an enzyme provokes a reversible and discrete change in the conformation of the enzyme;
- in order for enzymatic activity to occur, a suitable and very precise orientation of the enzyme catalytic groups in relation to those of the scissile bond of the substrate is necessary;
- the substrate induces its own orientation in relation to the enzyme *via* the change it provokes in the conformation of the enzyme.

During the dynamic process, the substrate "teaches" the enzyme the conformation it must adopt. KOSHLAND thought to apply the principles to oligomeric enzymes to explain differences in their behaviour compared to the MICHAELIS theory. Because of the basic premise, the KOSHLAND model is sometimes described as an **instructive mechanism** in opposition to the **selective mechanism** at play in the concerted model (CITRI, 1973).

The sequential model does not at all take into account the symmetry conservation and assumes the existence of hybrid conformations. It is based on the following postulates:

- the protein exists in one form A only in the absence of a ligand;
- the ligand binding to one subunit induces a conformational change of the subunit towards a conformation B;
- the conformational change of the concerned subunit provokes variations in interactions between subunits;
- the ligand binds preferentially to one of the two conformations. The proposed mechanism confers a particular importance to interactions between subunits, and the formalisms that describe the mechanism depend on the geometry of the oligomeric molecule. If a tetramer is considered, the nature of the interactions depends on the spatial arrangement of the subunits. KOSHLAND also considered and treated different models that had been proposed by PAULING in 1935 (Fig. 13.7 below);
- all the sites are equivalent and independent in the absence of a ligand.

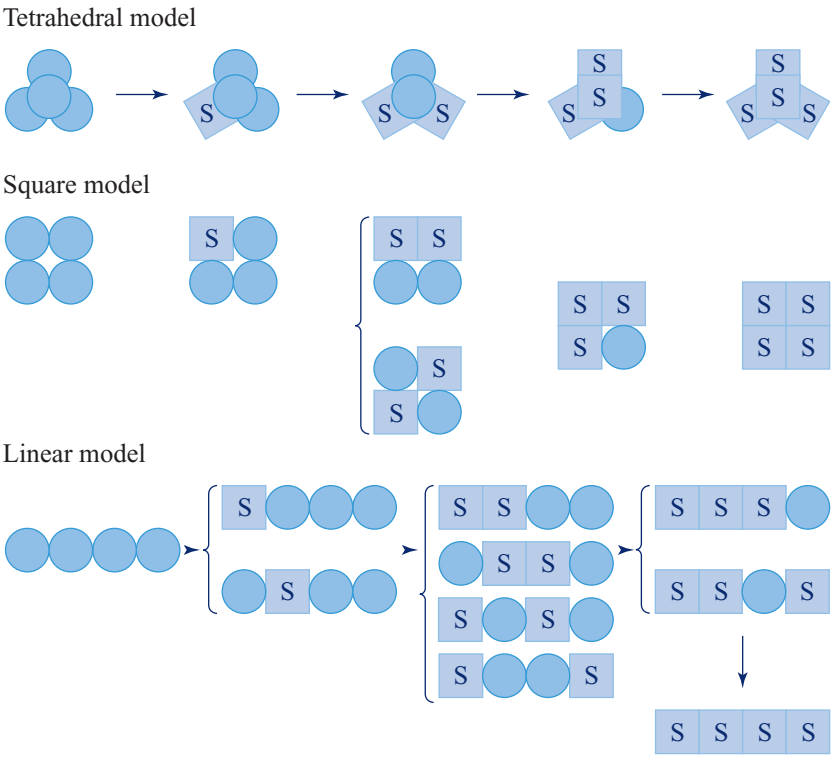


Fig. 13.7 Different geometries of a tetrameric molecule considered in the model of KOSHLAND, NÉMÉTHY and FILMER

In the following formulas, A and B refer to the two conformations each subunit may adopt and the ligand preferentially binds to the B conformation with an affinity constant:

$$K_s = \frac{[BS]}{[B][S]}$$

The equilibrium between the two conformations A and B is given by the ratio: $K_t = [B]/[A]$. The constant does not include the effect of the variation of interactions between subunits. The interaction variation is taken into account using the constants K_{AA}, K_{AB} and K_{BB} such that:

$$K_{AB} = \frac{[AB][A]}{[AA][B]}$$

$$K_{BB} = \frac{[BB][A][A]}{[AA][B][B]}$$

In the above equations [AB] indicates the interacting subunits, and [A] and [B] are non-interacting subunits. To simplify the calculations, it is supposed that $K_{AA} = 1$.

If, in the preceding model, the deviation from the MICHAELIS law is expressed by a conformational equilibrium linked to binding equilibria, in the present model it is expressed in terms of the variation in the interaction between protomers as the sites progressively become occupied.

If $K_{AA} < K_{BA} < K_{BB}$, that is to say, if the interaction between protomers becomes stronger as the sites become occupied, cooperativity is observed. If, on the contrary, $K_{AA} > K_{AB} > K_{BB}$, anticooperativity is observed for the substrate binding. **The cooperative and anticooperative effects depend on the relative value of the constants K_{AA}, K_{AB}, and K_{BB}.** They also depend however on the **geometry of the molecule**, since the number of interactions is a function of the geometry. To prove the preceding, the saturation function for the tetrahedric and square models will be considered.

The fraction of occupied sites is given by:

$$N_s = nY_s = \frac{[\text{bound substrate}]}{[\text{enzyme}]}$$

where N_s represents the average number of substrate molecules bound to an enzyme molecule and n is the number of subunits.

13.5.2. TETRAHEDRAL TETRAMER MODEL

In the case of a tetrahedral tetramer, the number of interactions is 3AB in the species A_3BS, 1BB, 4AB and 1 AA in the species $A_2B_2S_2$, 3BB and 3AB in the species AB_3S_3 and 6BB in the species B_4S_4. The corresponding concentrations of enzyme molecules which have bound a given number of substrate molecules are given by the following expressions:

$$[ES] = [A_3BS] = 4K_{AB}^3(K_sK_t[S])[A_4]$$

$$[ES_2] = [A_2B_2S_2] = 6K_{AB}^4K_{BB}(K_sK_t[S])^2[A_4]$$

$$[ES_3] = [AB_3S_3] = 4K_{AB}^3K_{BB}^3(K_sK_t[S])^3[A_4]$$

$$[ES_4] = [B_4S_4] = K_{BB}^6(K_sK_t[S])^4[A_4]$$

The equation of N_s is given by the following relation:

$$N_s = \frac{4K_{AB}^3(K_sK_t[S]) + 12K_{AB}^4K_{BB}(K_sK_t[S])^2 + 12K_{AB}^3K_{BB}^3(K_sK_t[S])^3 + 4K_{BB}^6(K_sK_t[S])^4}{1 + 4K_{AB}^3(K_sK_t[S]) + 6K_{AB}^4K_{BB}(K_sK_t[S])^2 + 4K_{AB}^3K_{BB}^3(K_sK_t[S])^3 + 4K_{BB}^6(K_sK_t[S])^4}$$

13.5.3. SQUARE TETRAMER MODEL

The concentrations of the species ES, ES2, ES3 and ES4 are evaluated as done in the preceding model:

$$[ES] = A_3BS = 4K^2_{AB}(K_t[S]/K_s)(A_4)$$
$$[ES_2] = A_2B_2S_2 = (4K^2_{AB}K_{BB} + 2K^4_{AB})(K_t[S]/K_s)^2(A_4)$$
$$[ES_3] = AB_3S_3 = 4K^2_{AB}K^2_{BB}(K_t[S]/K_s)^3(A_4)$$
$$[ES_4] = B_4S_4 = K^4_{BB}(K_t[S]/K_s)^4(A_4)$$

The average number of substrate molecules that are bound by an enzyme molecule in the square model is given by:

$$N_s = \frac{4K^2_{AB}(K_sK_t[S]) + 4(K^4_{AB} + 2K^2_{AB})(K_sK_t[S])^2 + 12K^2_{AB}(K_sK_t[S])^3 + 4(K_sK_t[S])^4}{1 + 4K^2_{AB}(K_sK_t[S]) + (2K^4_{AB} + 4K^2_{AB})(K_sK_t[S])^2 + 4K^2_{AB}(K_sK_t[S])^3 + 4(K_sK_t[S])^4}$$

The expression differs to that corresponding to a tetrahedral model, which illustrates the importance of the geometry of the molecule in enzymatic behaviour.

In the square model, the protein's saturation and state functions are the same since the conformational change of each protomer is induced by the substrate binding during saturation. These two functions have four parameters, K_s, K_t, K_{AB} and K_{BB}, since K_{AA} is taken as a reference. The model has one extra parameter compared to the model of MONOD, WYMAN and CHANGEUX.

Figure 13.8 opposite shows the variation of the saturation curves as a function of the molecule's geometry; all the other parameters are kept constant. Figure 13.9 opposite on the following page represents the variations of the saturation function for different parameter values. A few examples of saturation curves for the square model show the role of variations in K_{AB} and K_{BB} (Fig. 13.10 below). The product K_sK_t is arbitrarily set to 100 for all the curves. The curves which correspond to the same K_{AB} value have the same asymptote for small log[s] values; the curves which correspond to the same value of K_{BB} intersect at $N_S = 2$. The curves which have a constant K_{BB}/K_{AB} ratio are asymptotic at high log[s] values.

13.5.4. ANTICOOPERATIVITY

Anticooperativity consists in the fact that, in opposition to what has been described previously, the binding of the first ligand molecules to an oligomeric protein decreases the apparent affinity for the other ligand sites. It can be understood in terms of a conformational change in a subunit on binding the first ligand molecule that induces a modification in a neighbouring subunit which reduces its ligand affinity. The sequential model described previously for cooperative effects can take into account anticooperative phenomena perfectly well. As has been highlighted previously, anticooperativity arises when $K_{AA} > K_{AB} > K_{BB}$.

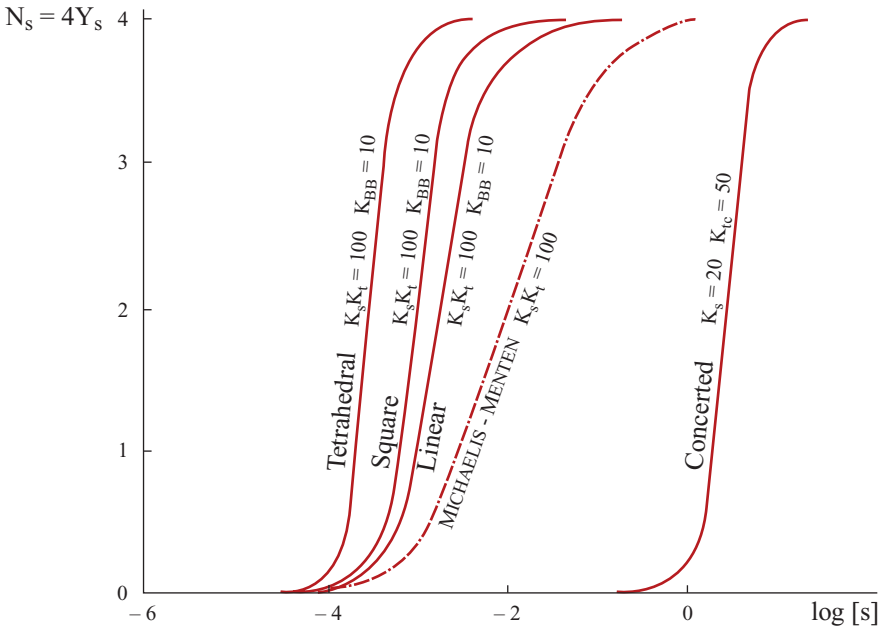


Fig. 13.8 Substrate saturation curves for different geometries of a tetrameric enzyme *For* all curves*, the product K_sK_t is taken equal to 100, $K_{BB} = 10$ and $K_{AB} = 1$* (Reprinted with permission from *Biochemistry*, **5**, KOSHLAND D.E. et al., 366. © (1966) American Chemical Society)

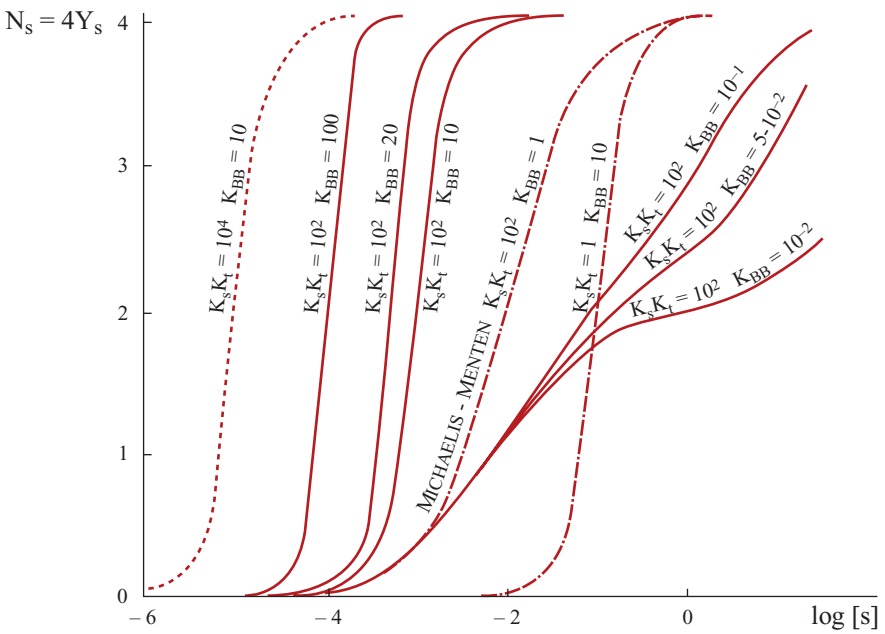


Fig. 13.9 Substrate saturation curves in the square model showing the effect of variations in K_sK_t and K_{BB} *($K_{AB} = 1$ for* all curves*)*
(Reprinted with permission from *Biochemistry*, **5**, KOSHLAND D.E. et al., 365. © (1966) American Chemical Society)

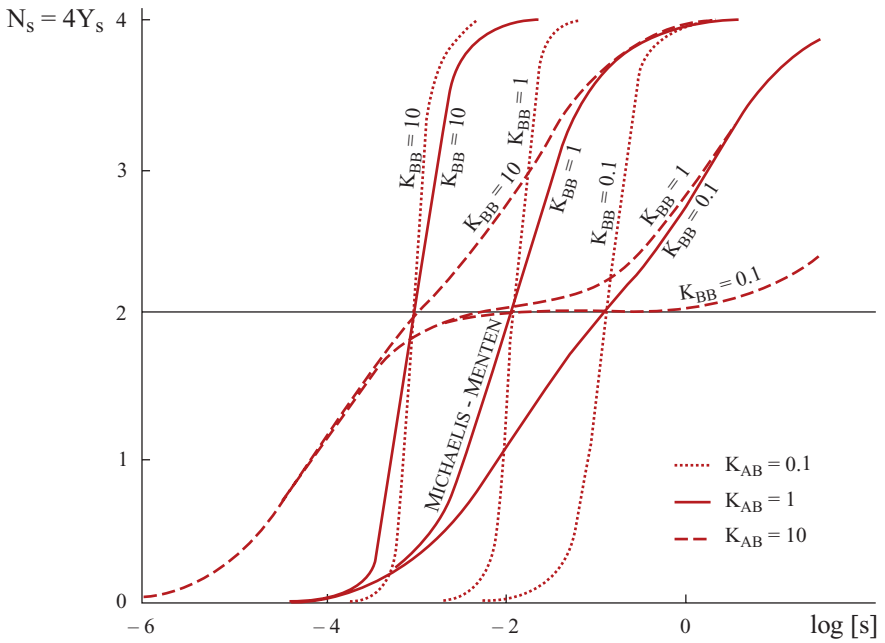


Fig. 13.10 Substrate saturation curves in the square model for different values of K_{AB} and K_{BB} *($K_sK_t = 100$* for all curves*)*

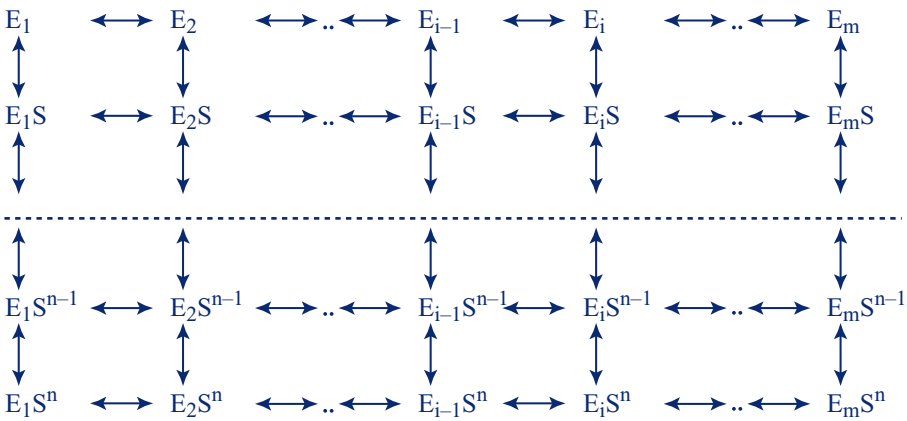
(Reprinted with permission from *Biochemistry*, **5**, KOSHLAND D.E. et al., 365. © (1966) American Chemical Society)

The substrate saturation curve of an enzyme which displays anticooperative phenomena cannot be linearised. Figure 13.1b and c show the general shape of curves obtained in the case of anticooperativity using the LINEWEAVER-BURK and EADIE-HOFSTEE plots. In practice, it is very difficult to distinguish between anticooperativity and heterogeneity of the binding sites. In the first case, in the absence of a ligand, all the binding sites display the same affinity, and it is the conformational change induced by the binding of the first ligand molecules which confers a lower apparent affinity to the non-occupied sites, or even a practically non-existent affinity. In the second case, several site categories exist with different substrate affinities (see Chap. 2), which exist prior to ligand binding.

In certain cases, the anticooperative phenomenon is extreme; the binding of the first ligand molecules induces a structural modification of the protein which confers a practically infinite dissociation constant to the residual sites. The phenomenon is referred to as **half-site reactivity**. Amongst the enzymes which display anticooperative phenomena, glyceraldehyde-3-phosphate dehydrogenase has been particularly well studied (HENIS & LEVITZKI, 1980). A particular aspect of half-site reactivity has been described by LAZDUNSKI et al. (1971): a **flip-flop** mechanism, where each protomer of the dimeric enzyme is alternately phosphorylated by the substrate, while the other protomer is de-phosphorylated, is used to explain the behaviour of *E. coli* alkaline phosphatase.

13.6. THE GENERALISED MODEL

A generalised model was proposed to explain both cooperative and anticooperative effects. In 1965, WEBER proposed an extension of the ADAIR model, taking into account the conformational mobility of the protein. WEBER assumed that in the absence of any ligand, the protein can exist in m conformations, each form having n sites which can bind either a substrate, S, or an effector. The diagram below can take into account the different equilibria:



The model is described by a general equation of the same type as the equation of ADAIR in which the dissociation constants have a complex significance, containing a term corresponding to the dissociation constant between the substrate and its receptor site and a transformation constant of the protomer in the oligomer. The different equilibria in the vertical and horizontal directions are linked. Figure 13.11 below illustrates the model in the case of a tetramer. It is interesting to note that each of the models described previously represents a special case of the generalised model. If only the two outside vertical columns are considered in Fig. 13.11, the model of MONOD, WYMAN and CHANGEUX is described. The diagonal of the diagram represents the KOSHLAND, NÉMÉTHY and FILMER model. If the interaction energy between protomers in the oligomer is larger than the protein-ligand association energy, the system is described by the concerted model. On the contrary, if the protein-ligand interaction energy is larger than the interaction between protomers, the system is well described by the sequential model. It is the relative interaction energies which control the enzymatic system.

13.7. THERMODYNAMICAL COUPLING BETWEEN LIGAND BINDING ENERGY AND SUBUNIT INTERACTION ENERGY

◤ The notion of thermodynamical coupling between ligand binding energy and the subunit interaction energy is implicit in the models described in the preceding sections.

It is also included in the thermodynamical considerations of WYMAN (1964) on the importance of binding potentials. The concept of coupling, stated by NOBLE (1967), was extensively developed by WEBER, who conceived a protein behaviour analysis method for proteins displaying cooperative effects, in terms of variations in free energy, associated with both ligand binding and interactions between protein subunits.

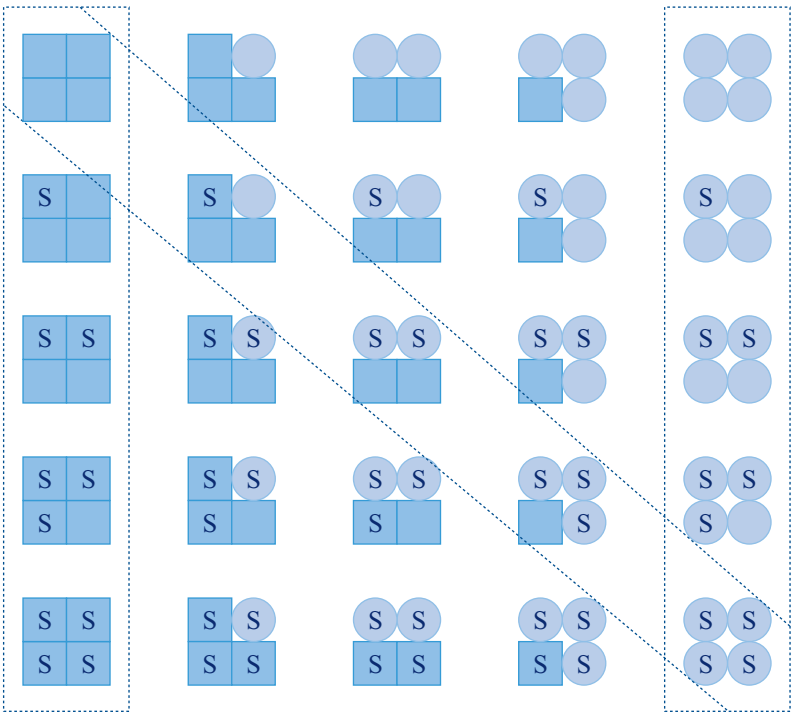
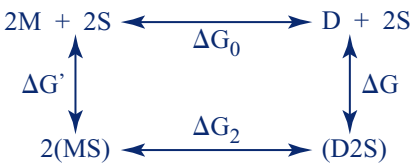


Fig. 13.11 The generalised ADAIR-WEBER model

Let us consider the extreme case of a monomer-dimer equilibrium. If the substrate S has a higher affinity for the dimeric form D than for the monomeric form M, it will shift the equilibrium towards the dimer. Every physico-chemical parameter which favours the dissociation will however disfavour the binding of S. The described situation is characterised by the following cycle:



From a thermodynamical point of view, the different values of ΔG correspond to the variations in free energy, which accompany each reaction. The variation in free energy to go from one state to another is independent of the chosen path, so the sum of the variations around the cycle is zero, and:

$$\Delta G - \Delta G' = \Delta G_2 - \Delta G_0$$

The above means that, to every difference in ligand affinity for the protein's two states, corresponds an equal difference in the subunit interaction energy of the free and ligand associated forms. It is then possible to describe the progressive ligand binding to the protein along the thermodynamical diagram presented in Fig. 13.12.

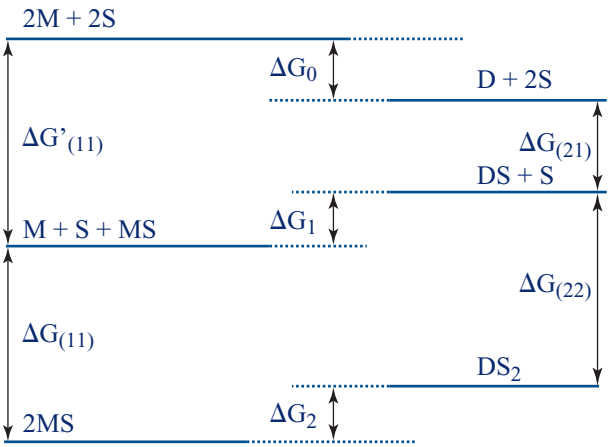


Fig. 13.12 Thermodynamical coupling between dimerisation and substrate binding *The figure represents the case where, despite cooperativity of substrate binding to the dimeric enzyme, the formation of the monomer is favoured* (Reprinted with permission from *Biochemistry*, **11**, WEBER G., Ligand binding and internal equilibiums in proteins, 868. © (1972) American Chemical Society)

ΔG_0 represents the variation in free energy of the monomer-dimer equilibrium in the absence of the ligand, $\Delta G'_{(11)}$ the free energy of ligand binding to the monomer, $\Delta G_{(21)}$ and $\Delta G_{(22)}$, the free energy variation on binding the first and second ligand molecules to the dimer, respectively. If $\Delta G_{(22)} > \Delta G_{(21)}$, the binding is cooperative; if in addition, $|\Delta G_0| > |\Delta G_1|$, the dissociation is favoured despite the cooperativity of ligand binding to the dimer. In these conditions and depending on the relative values of ΔG_0, ΔG_1 and ΔG_2, WEBER defines a **first order** system as one in which it is the binding of the first S molecule which provokes the subunit interaction energy change. In a so-called **second order** system it is the binding of the second S molecule which induces the effect. The **intermediate order** corresponds to the case where the binding of both ligand molecules contributes. Depending on the values of ΔG, the binding of S will be cooperative, anticooperative or Michaelian, and the binding will favour or disfavour dimer formation.

The same reasoning applies to the case where the ligand binding is not linked to the complete dissociation of the protein, but only to a **variation in subunit interaction energy**, which is the most general case where proteins binding their ligand in a cooperative manner are concerned. Considering the different possible cases, WEBER (1972) established the following rules:

- in a first order system, the ligand binding decreases the subunit interaction if the binding of S is cooperative, and increases the interaction if the binding is anticooperative;

- in a second order system, the ligand binding increases the subunit interaction if the binding of S is cooperative, and decreases it if the binding is anticooperative.

13.8. KINETIC COOPERATIVITY: RICARD MODEL

The models described above, with the exception of V systems, involve variations in the affinity of the catalytic sites for substrates and consider dissociation constants only in relation to the substrates, at equilibrium. The descriptions at thermodynamical equilibrium are insufficient to explain what happens in the case of enzymes where the equilibria are immediately broken by the reaction and the appearance of the products. Both WHITEHEAD (1970) and RICARD (1989) developed new formalisms destined to explain in a more complete manner the cooperative phenomena displayed by the enzymes.

RICARD's approach attempts to describe the way in which variation of subunits interactions can change the catalytic constant. In the authors own words "what is important is not to understand how the interactions between subunits and conformational constraints modulate substrate binding, but in a more integrated way, to understand how the interactions and constraints control the rate of product apparition in the stationary phase".

The rate constant of a reaction is related to its activation energy by the relation:

$$k = \frac{k_B T}{h} e^{-\Delta G^{\neq}/RT}$$

in which k_B is the BOLTZMANN constant, h PLANCK's constant, R the gas constant and T the absolute temperature.

The basic idea of RICARD's model is that the interactions between subunits can have two types of effect on the rate of the catalysed reaction:

- the interactions can change the rate of conformational transitions involved in the reaction. The corresponding contribution to the free activation energy is called the **protomer arrangement contribution**, $\sum({}^{\alpha}\Delta G_{int})$;
- the interactions between subunits can induce a distortion of the catalytic site. The corresponding energetic contribution is called the **quaternary constraint contribution**, $\sum({}^{\sigma}\Delta G_{int})$.

As a consequence, the global activation energy, $\Delta G^{\neq}$, of the catalysed reaction is equal to:

$$\Delta G^{\neq} = \Delta G^{\neq *} + \sum({}^{\alpha}\Delta G^{int}) + \sum({}^{\sigma}\Delta G^{int})$$

where $\Delta G^{\neq *}$, called the **intrinsic energy contribution**, corresponds to what the free energy of the reaction would be if the subunits were free from all interactions. As a consequence, the expression of the rate constant is:

$$k_{int} = \frac{k_B T}{h} \exp[-(\Delta G^{\neq *} + {}^{\alpha}\Delta G_{int} + {}^{\sigma}\Delta H_{int})/RT]$$

The theories developed on the basis of this formulation lie on three postulates:

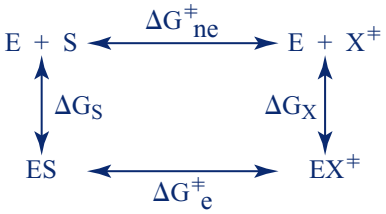
- The theories of enzyme catalysis suggest that the intramolecular constraints are abolished in the transition state. The idea is applied to oligomeric enzymes and the model supposes that the quaternary constraints are suppressed on formation of the transition state. Suppose the following reaction, catalysed or not by an enzyme:
  - catalysed by the enzyme

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-e}}{\overset{k_e}{\rightleftharpoons}} EP \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P$$

  - not catalysed by the enzyme

$$S \underset{k_{-ne}}{\overset{k_{ne}}{\rightleftharpoons}} P$$

  The rate constants k_e and k_{ne} are compared in the following thermodynamic cycle, which takes into account the formation of a transition state X:



$$\Delta G^{\neq}_{ne} - \Delta G^{\neq}_{e} = \Delta G_X - \Delta G_S$$

  Since $\Delta G^{\neq}_{ne}$ is much larger than $\Delta G^{\neq}_{e}$, it follows that $\Delta G^{\neq}_{X} >> \Delta G_S$. The enzyme displays a much larger affinity for the transition state than for the fundamental state of the substrate and in the enzyme-transition state complex the quaternary constraints must be abolished.
- In the absence of quaternary constraints, the subunits may adopt a limited number of conformations. If two conformations A and B are present, substrates and products stabilise the same conformation (for instance B).
- The subunit that has bound the transition state is in the A conformation which is identical to that of the free subunit.

Several important conclusions on the analysis of oligomeric enzyme behaviour follow from the development of the above formalisms:

- a weak interaction between subunits can imply a kinetic cooperativity phenomenon, although the substrate saturation curve is not a sigmoid;
- if the subunit interaction energy is strong, several cases can be distinguished:
  - either the substrate binding cooperativity is positive and the kinetic cooperativity must be positive,
  - or the substrate binding cooperativity is negative and the catalytic cooperativity can be negative or positive.

Thus, even in the case of a strong interaction between subunits, a strong catalytic cooperativity can generate a substrate saturation curve which does not have a sigmoidal shape.

13.9. COOPERATIVITY AND ALLOSTERY

The original and strict definition of allostery is that a metabolite whose structure is different to that of the substrate, on binding to a site, which is not the substrate binding site, induces an alteration of the active centre's properties, in terms of substrate affinity and/or catalytic efficiency (MONOD et al., 1963).

From an epistemological point of view, it is interesting that to explain such a well defined phenomenon, formalisms destined to account for a different phenomenon, that is, cooperativity between catalytic sites, were used. An explanation can be found in the fact that the first regulatory enzymes to be studied at this period (threonine deaminase for example) displayed both phenomena, which were sought to be explained by a unique model.

The information gathered over the past years on the regulatory mechanisms of different enzymes, and in particular aspartate transcarbamylase, have lead to a revision of this view. Indeed, it emerges that, in aspartate transcarbamylase, cooperative effects between catalytic sites, and allosteric effects between regulatory and catalytic sites, operate by different mechanisms which are however coupled. The same may apply to other regulation enzymes. In addition to this, certain enzymes only show one or the other of these phenomena; certain dehydrogenases such as the glyceraldehyde 3-phosphate dehydrogenase show cooperativity between catalytic sites without being sensitive to physiological effectors. (HENIS & LEVITZKI, 1980). Others, on the contrary, only show sensitivity to allosteric effector, without showing cooperativity between catalytic sites. This is the case, for example, of aspartate transcarbamylase from certain bacteria and the yeast *Saccharomyces cerevisiae* (BETHELL & JONES, 1969; PENVERNE & HERVÉ, 1983). It is also the case of ribonucleotide reductases, some of which have a monomeric structure (ERIKSON & SJÖBERG, 1989).

In the present state of our knowledge, and in an attempt to clarify the discussion of these phenomena, it is practical to discuss cooperativity on the one hand, and allostery on the other, even if it turns out that in some cases the two types of phenomena are explained by the same mechanisms (THIRY & HERVÉ, 1978).

We shall now consider a few examples of allosteric enzymes.

13.10. EXAMPLES OF ALLOSTERIC ENZYMES

At present, numerous allosteric enzymes have been studied, and for several of them, structures have been determined by X-ray crystallography. For some enzymes (aspartate transcarbamylase, glycogen phosphorylase, phosphofructokinase), the crystallographic studies have resolved the protein structure in several states, which

are involved in the mechanisms of regulation. In this context, the idea emerges that, to a certain extent, each of these systems possesses its own logic, its own mechanisms. Some of the mechanisms are similar, some are not. Some are in agreement with the theoretical models that have been proposed, some are not. It therefore appears that, **in the field of metabolic regulation, as in other fields of biology, nature has used several solutions to face the same demands of survival.**

13.10.1. GLYCOGEN PHOSPHORYLASE

Glycogen phosphorylase b catalyses the first step of intracellular degradation of glycogen:

$$(\alpha\ 1,4\text{-glycogen})n + P_i \longleftrightarrow (\alpha\ 1,4\text{-glycogen})n-1 + \alpha\ \text{glucose-1-P}$$

13.10.1.1. ALLOSTERIC REGULATION

The enzyme is subjected to allosteric regulation and to a reversible phosphorylation. The regulation process can be explained in terms of several conformational states, from the T state, which shows a weak substrate affinity, to the R state, which shows a strong affinity. The quaternary constraints and the tertiary structure are modulated by the ligand binding.

Phosphorylase a is the phosphorylated form, phosphorylase b, the dephosphorylated one. Physiological activation induces the conversion of inactive phosphorylase b to phosphorylase a, by phosphorylation of the catalytic serine 14 by a kinase, under the influence of hormonal and neuronal action. The phosphorylase b can also be activated by AMP, IMP and strong concentrations of phosphate. Its affinity for phosphate is increased 15-fold in the presence of AMP. It is inhibited by glucose-6-phosphate, glucose, ATP and ADP. The activation by AMP as well as the phosphorylation bring about a change in the oligomeric state of the enzyme; it becomes tetrameric. The tetrameric enzyme possesses only 12 to 33% of its phosphorylase activity, but it can be dissociated by glycogen or oligosaccharides into dimers which are completely active. Figure 13.13 below is a schematic representation of the activation of glycogen phosphorylase.

Phosphorylase a is less sensitive to effectors than phosphorylase b. Phosphorylase a does not show cooperativity for glycogen or phosphate binding (Fig. 13.14 below).

However, in the case of phosphorylase b, a strong cooperativity is observed for glucose-1-phosphate binding, which is the enzyme's substrate in the reverse reaction (Fig. 13.15 below).

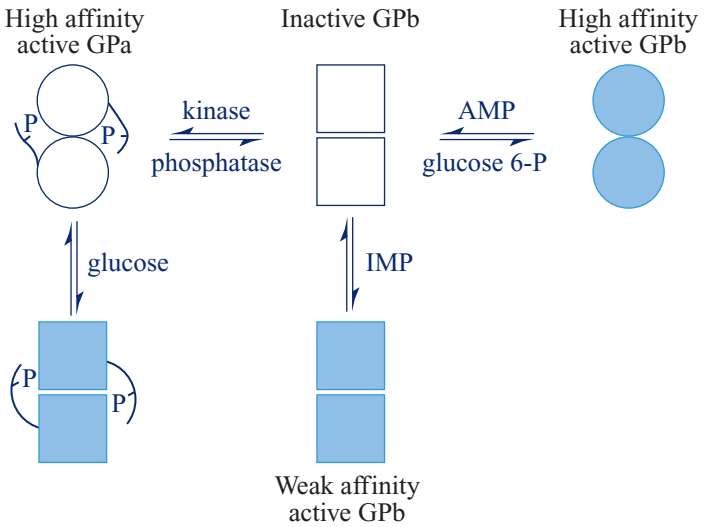


Fig. 13.13 Schematic representation of the allosteric and covalent activation of glycogen phosphorylase

The subunits R and T are symbolised by squares and circles respectively (Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **340**, BARFORD D. & JOHNSON L.N., 609. © (1989))

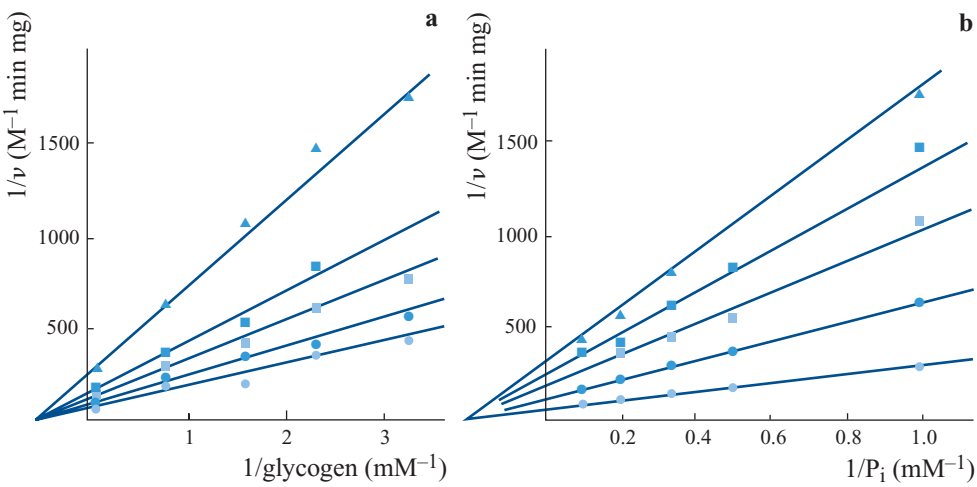


Fig. 13.14 Variation of phosphorylase a activity

(a) *as a function of glycogen concentration (AMP = 0.10 mM), for different concentrations of P_i: 10 mM (○), 5 mM (●), 3 mM (□), 2 mM (■), 1.0 mM (▲) –* ***(b)*** *as a function of P_i concentration (AMP = 0.10 mM) for different concentrations of glycogen: 12.4 mM (○), 1.24 mM (●), 0.617 mM (□), 0.432 mM (■), 0.308 mM (▲)*

(From *J. Biol. Chem.*, **243**, BLACK W.J. & WANG J.H., 5892. © (1968) with permission from The American Society for Biochemistry and Molecular Biology)

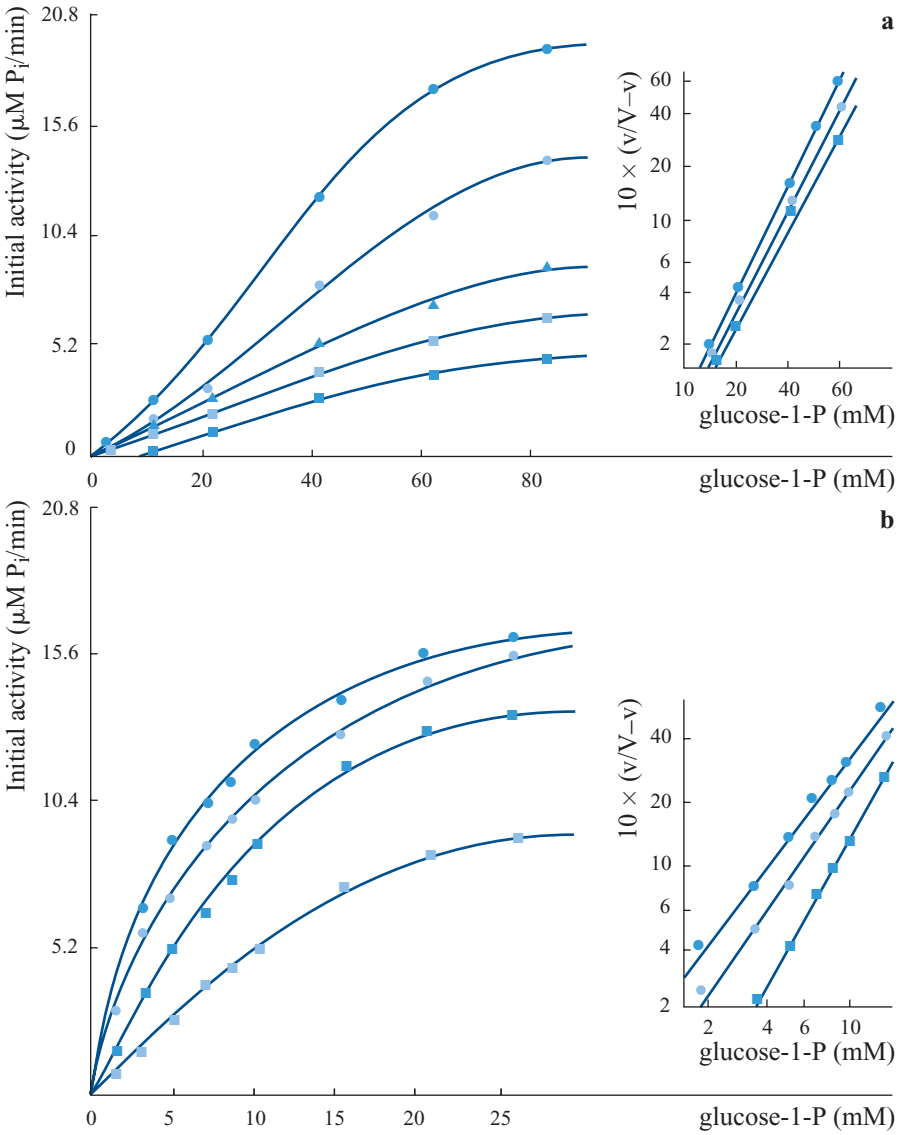


Fig. 13.15 Variation of phosphorylase b activity as a function of the concentration of glucose-1-phosphate

*(**a**) for different concentrations of IMP: 4.16 mM (●), 1.66 mM (○), 0.83 mM (▲), 0.33 mM (■), 0.0833 mM (□) – (**b**) for different concentrations of AMP: 2.08 mM (●), 0.208 mM (○), 0.104 mM (■), 0.052 mM (□). Inset, HILL's plot*

(From *J. Biol. Chem.*, **243**, BLACK W.J. & WANG J.H., 5892. © (1968) with permission from The American Society for Biochemistry and Molecular Biology)

13.10.1.2. PHOSPHORYLASE STRUCTURE

The structures of glycogen phosphorylase a and b have been resolved in the T and R states (Fig. 13.16).

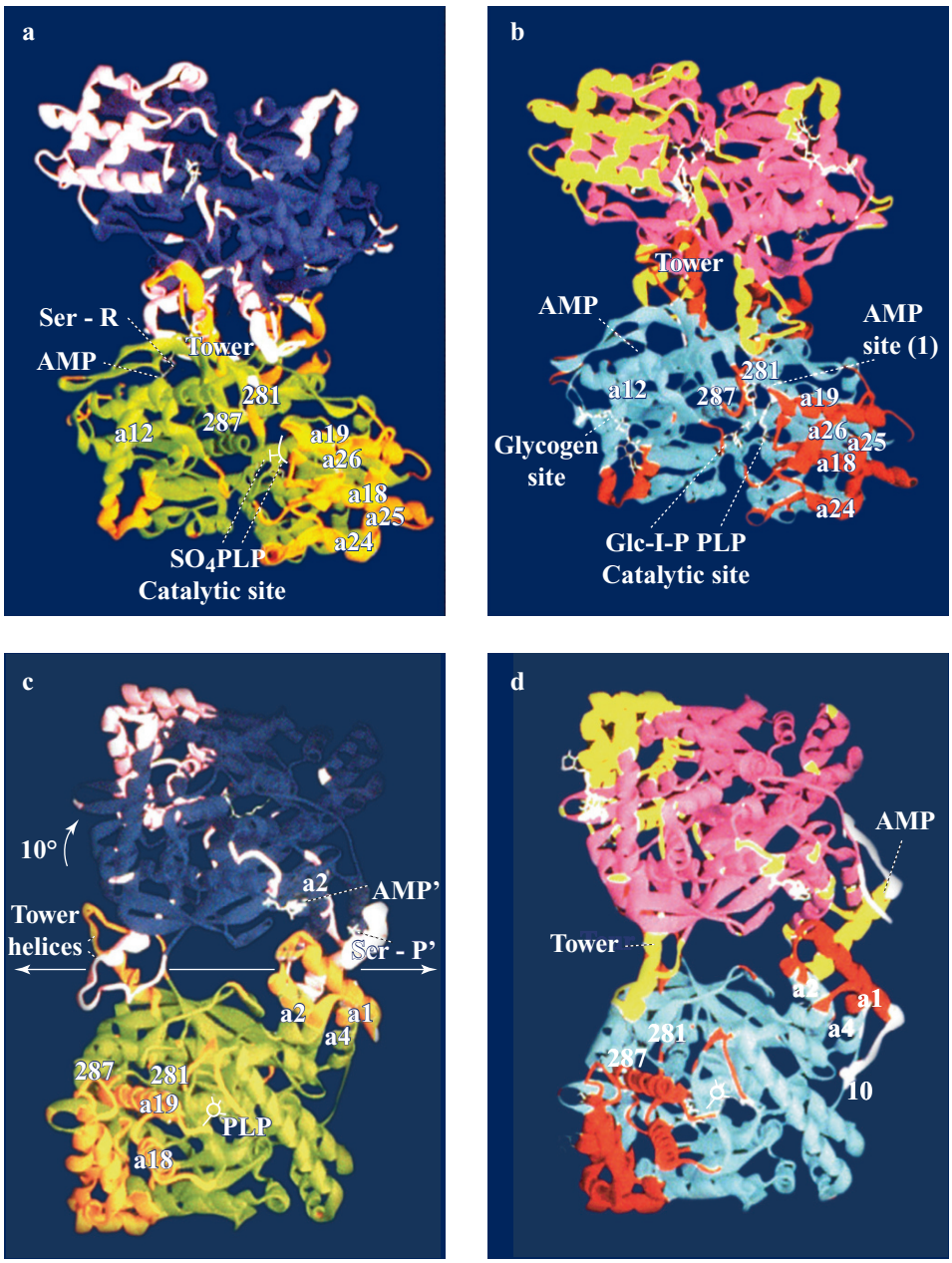


Fig. 13.16 Structures of the R state of the phosphorylase a dimer and the T state of the phosphorylase b dimer

In (a) and (c): in the R state one of the subunits is in green, the other in blue. The regions that differ in C^{α} positions by less than 1 Å between tertiary structures of R and T states are in orange for one subunit and in pink for the other. In (b) and (d): the T state of one subunit is in cyan; the other in purple; the regions which differ in C^{α} positions by more than 1 Å between the R and T states are shown in red and yellow in each of the subunits, respectively. The N-terminal (10–23) and C-terminal (837–852) residues are shown in white. In (a) and (b): the view is shown from the 2-fold axis of the dimer with the catalytic site and the "tower" helices turned towards the front. In (c) and (d): the view is perpendicular to the 2-fold axis of the dimer to show the change in quaternary structure, where one subunit turns by 10° in relation to the other. The view also shows the change in interface between subunits. The ligand binding sites, the catalytic sites and the phosphorylation sites are indicated. (Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **340**, BARFORD D. & JOHNSON L.N., 609. © (1989))

Figure 13.17 below shows the subunit in the T state with its different domains and ligand binding sites that have been identified. The organisation of the secondary structure is the same for the R and T states. The protein is of the α/β type and possesses two structural domains, the N-terminal domain comprising amino acids 1 to 484, and the C-terminal domain the amino acids 485 to 842. The pyridoxal phosphate binding site, the catalytic site C, the AMP binding site N, which is 30 Å away from the catalytic site, the glycogen reserve site G, and the allosteric contact zone with the other subunit are shown in the figure. When the two glycogen phosphorylase dimers associate to form the tetramer, the interface formed by the association contains residues of the catalytic face of the subunit, and access to the catalytic site is hindered. The situation is at the origin of the weak activity of the tetramer in the T state. Another contact zone contains the residues of the glycogen reserve region, in particular a key residue of the site, the residue Glu 433, which forms a hydrogen bond with the interface: this explains why glycogen induces dissociation of the tetramer into dimers.

The transition of the T to the R state is accompanied by small changes in the tertiary structure that are mainly localised in the ligand binding sites and at the interface between subunits, in particular helix α_7 ("tower" helix). The regions of the molecule that undergo the largest displacements are the zones which are important, either for the allosteric response, or for the tetramer formation. The variations in the tertiary structure are coupled to large changes in the quaternary structure; a rotation of 10° between subunits occurs. In the T state, the interface is formed by an antiparallel association of the two "tower" helices. During the transition towards the R state, the arrangement of the two helices is greatly modified, their axes form an angle of –80°. The loop located above the helix also undergoes an important change. The change in quaternary structure directly affects the N site and the phosphorylation site, and transmits a signal to the catalytic site and to the inhibitor site I, *via* the conformational change of the two "tower" helices.

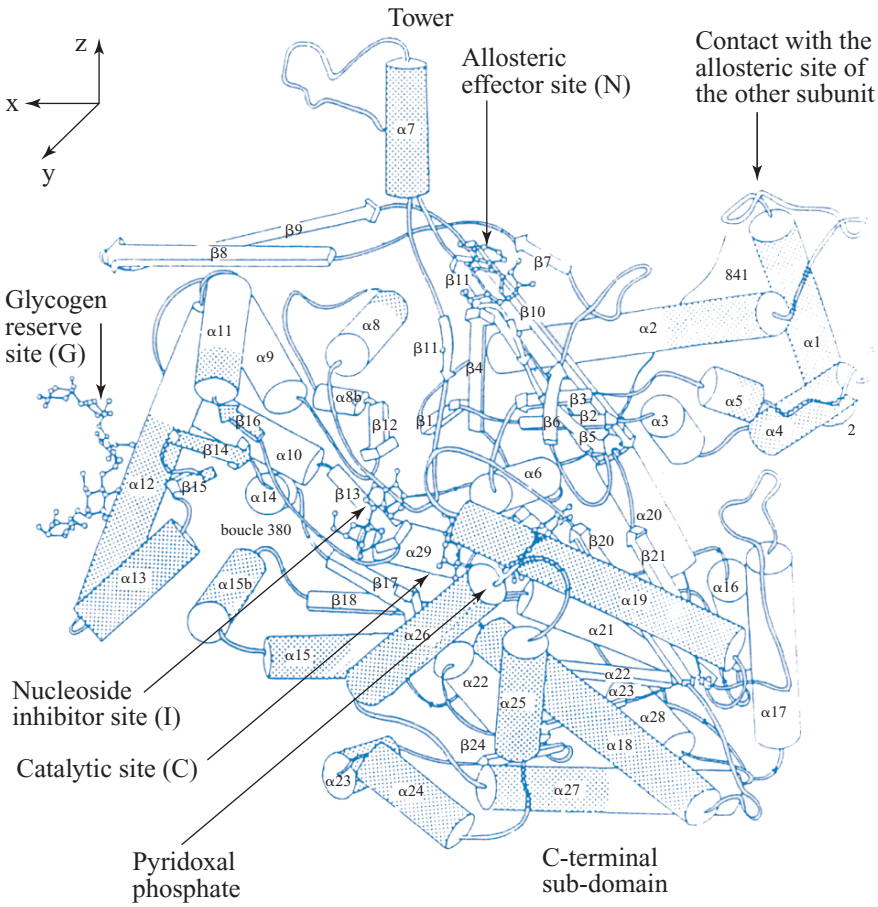


Fig. 13.17 3-dimensional structure of a subunit of glycogen phosphorylase b in the T state, seen along the crystallographic axis

The grey *areas represent regions that undergo changes during the transition to the R state* (From *Protein Sci.*, **1**, No. 4, BARFORD D. & JOHNSON L.N., 472–493. © (1992 The Protein Society). Reprinted with permission of John Wiley & Sons, Inc.)

Glycogen phosphorylase represents one of the first enzymes whose catalytic properties have been described by the concerted model of MONOD, WYMAN and CHANGEUX (BUC, 1967). The structural data appear to agree with the model (JOHNSON, 1992). However, to explain certain experimental results, several authors have been led to complicate the model. To explain the activation by small concentrations of AMP, a three state model has been proposed. Four state models have been suggested to describe the heterotropic interactions between AMP and glucose-1-P, and AMP and glucose-6-P. In glycogen phosphorylase, the situation is complicated, on the one hand by the existence of the dimeric and tetrameric forms, and on the other hand by the existence of phosphorylated forms. The crystal structures of the two R states of glycogen phosphorylase a and b are identical however. The

same is true for the different T states of Fig. 13.16. The crystallographic data do not allow the interpretation of the regulatory properties of the enzyme with more than two states.

13.10.2. PHOSPHOFRUCTOKINASE

Phosphofructokinase is an enzyme of the glycolytic pathway, which catalyses the transfer of the γ-phosphate of ATP to the C_1 hydroxyl of fructose-6-phosphate (Fruc-6-P), to give fructose-1,6-diphosphate (Fruc-1,6-P2), following the reaction:

$$\text{Fruc-6-P} + \text{ATP} \rightleftharpoons \text{Fruc-1,6-P2} + \text{ADP}$$

in the presence of Mg^{++}. The reaction is strongly exergonic and practically irreversible.

The enzyme is very specific to the hexose phosphate, but accepts a wide variety of triphosphate nucleosides as phosphate donors. In vivo, phosphofructokinase is subjected to multiple regulations that confer a key role to the enzyme in the control of glycolysis according to the energetic requirements of the cell. A number of allosteric effectors, inhibitors or activators of phosphofructokinase exist, which differ in different organisms (Fig. 13.18).

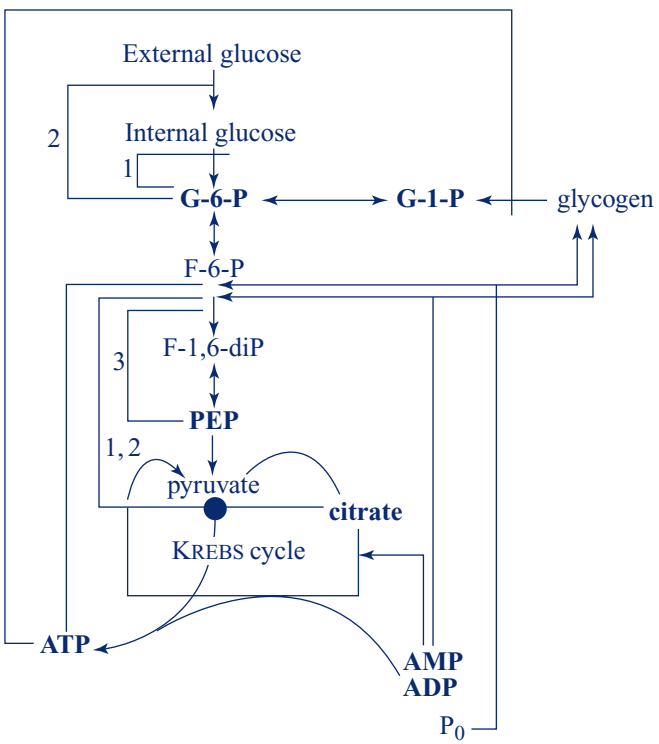


Fig. 13.18 Allosteric regulation of phosphofructokinase, a key enzyme of glycolysis, in different organisms: 1, mammals; 2, yeast; 3, E. coli

13.10.2.1. STRUCTURE OF PHOSPHOFRUCTOKINASE

Phosphofructokinase has a large variety of structures in prokaryotic and eukaryotic organisms. In prokaryotes, such as *E. coli* and *B. stearothermophilus*, the enzyme is a tetramer, formed of identical protomers. The enzyme is generally tetrameric in mammals, but is subject to association-dissociation equilibria that vary according to tissues. The yeast enzyme is octameric, formed of four α and four β subunits. These subunits have been isolated and identified (LAURENT & YON, 1989) and their structural and functional properties found to be different. The experimental results showed that the β subunits, which bind fructose-6-phosphate, are the catalytic subunits, and that the α subunits bear the regulator sites.

The three dimensional structures of the enzymes from *E. coli* and *B. stearothermophilus* were determined by X-ray crystallography by the EVANS group in Cambridge (EVANS, 1992). Figure 13.19 opposite shows the structure of the *B. stearothermophilus* enzyme. The enzyme is made up of four identical protomers of molecular weight 33 900 containing 316 amino acids. The protomers of the oligomer are associated by three orthogonal 2-fold symmetry axes. The centre of the tetramer has a 7 Å diameter cavity filled with water molecules (the cylinder in Fig. 13.19a). Each subunit is in tight contact with only two neighbouring subunits. The cavity separates the pairs of subunits that do not interact. The enzyme belongs to the α/β class of proteins. Each protomer is structured into two domains of approximately equal sizes.

The binding sites of substrates and effectors have been determined (Fig. 13.19b). Fructose-6-phosphate is in interaction with residues of two different subunits, His249 of one, and Arg159 and Arg240 of the other. The ATP binding site is between domains 1 and 2, and that of the allosteric effectors is different. The interactions of ADP, an allosteric activator of the enzyme, involve residues of both subunits. The β-phosphate interacts with Arg151 of one subunit and Arg21 and Arg25 of the other. The ribose forms interactions with His212 and Thr155. The phosphate group of phosphoenolpyruvate seems to bind to the same site as the β-phosphate of ADP.

The structures of two conformations of the enzyme were resolved, one in the presence of the reaction products, the other in the presence of an inhibitor, analogue of phosphoenolpyruvate (SHARAKIHAVA & EVANS, 1988). At this time, it was supposed that these structures corresponded to the R and T conformations, respectively. The main difference between the two forms resides in the rotation by 7° of the AB dimer in relation to the CD dimer (Fig. 13.19a). The movement is coupled to the change in fructose-6-phosphate binding sites. The most important change in tertiary structure concerns the position of the two helices 8 and 9 and the rearrangement of loop 6F. The rearrangement of this loop couples the fructose-6-phosphate and allosteric effector binding to the change in quaternary structure of the molecule. The symmetry is conserved during the transition, which suggests that it is concerted.

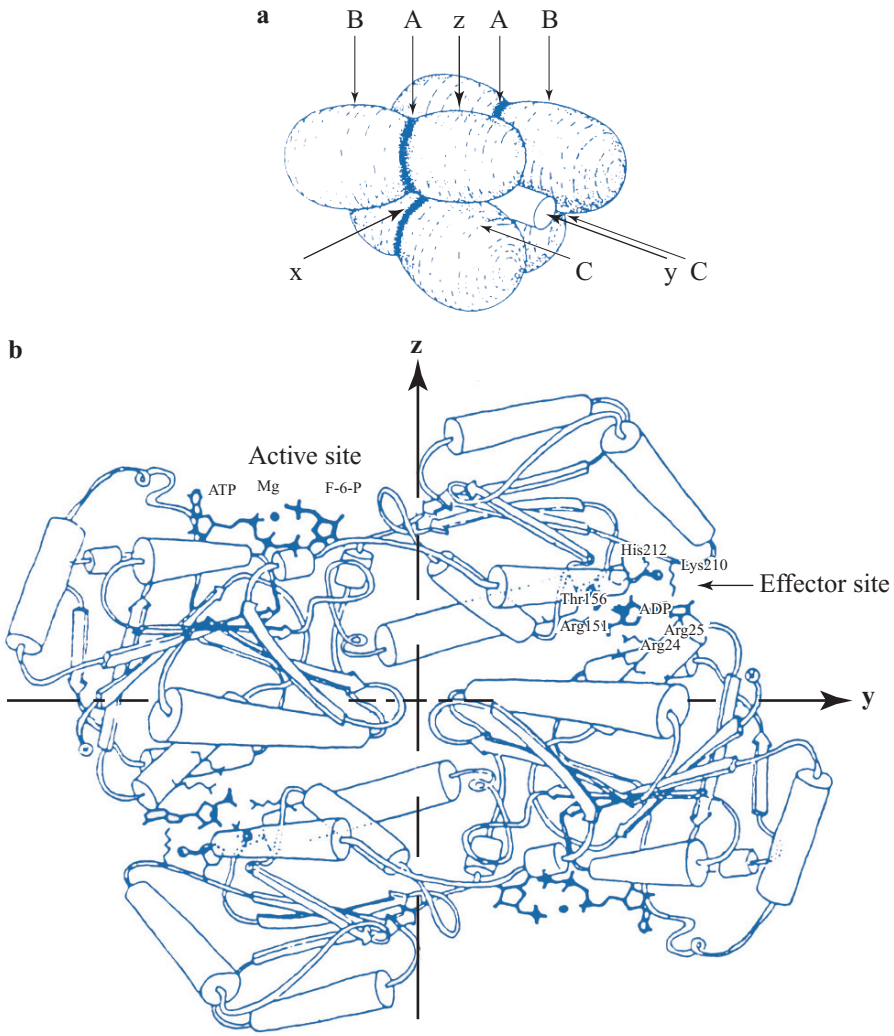


Fig. 13.19 Three-dimensional structure of phosphofructokinase from B. stearothermophilus

(a) *schematic representation of the protomer arrangement in the oligomer showing the tetramer geometry. Each protomer is formed of two domains of approximately equal sizes –* ***(b)*** *representation of two subunits along the x-axis. The positions of substrates and effectors are shown* (Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **279**, EVANS P.R. & HUDSON P.J., 500. © (1979))

13.10.2.2. PHOSPHOFRUCTOKINASE ALLOSTERIC REGULATION

Phosphofructokinases from prokaryotes, in particular those of *E. coli* and *B. stearothermophilus*, which have been well-studied, show kinetic cooperativity for the substrate fructose-6-phosphate, but not for ATP. There is however a difference between the enzymes from *E. coli* and *B. stearothermophilus:* the latter only displays a

cooperative phenomenon for fructose-6-phosphate in the presence of phosphoenolpyruvate. The enzymes are under allosteric control, being activated by ADP and other diphosphonucleosides, and inhibited by phosphoenolpyruvate. Contrary to the eukaryotic enzymes (see below), they are not inhibited by ATP or citrate, and not activated by AMP. The *E. coli* enzyme was one of the first allosteric enzymes to which the concerted model was applied BLANGY et al., 1968. As an illustration, Fig. 13.20 shows the cooperative kinetics as a function of fructose-6-phosphate concentration for different ADP concentrations. Figure 13.21 opposite shows the enzyme's Michaelian behaviour in relation to ATP.

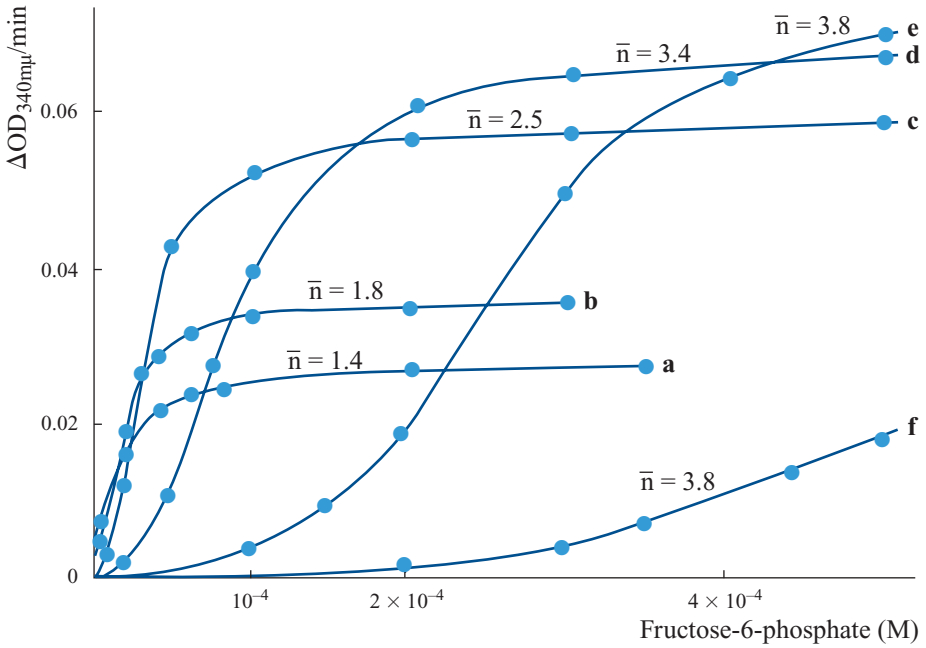


Fig. 13.20 Kinetic cooperativity of the reaction catalysed by E. coli phosphofructokinase as a function of fructose-6-phosphate, in absence and presence of ADP at different concentrations,

a: 0.82 mM; b: 0.52 mM; c: 0.22 mM; d: 70 µM; e: 20 µM; f: without ADP; ATP = 0.10 mM; Mg^{++} = 1 mM. HILL's number $\bar{n}$ is indicated for each curve. (Reprinted from *J. Mol. Biol.*, **31**, BLANGY D. et al., Kinetics of the allosteric interactions of phosphofructokinase from *Escherichia coli*, 13. © (1968) with permission from Elsevier)

For many years, the enzyme was considered a perfect example of an enzyme obeying the concerted model of MONOD et al. (1965). More recently, two unexpected observations have led to this interpretation being fundamentally revised:

- the determination of the enzyme's crystallographic structure in the absence of substrate showed that the structure was identical to that obtained in the presence of substrate, showing that the substrate itself is incapable of provoking the supposed transition between the T and R states (RYPNIEWSKI & EVANS, 1989);

- the direct measurement of substrate binding by fluorescence showed that the binding occurs in a Michaelian fashion without cooperativity (AUZAT et al., 1995).

It is now known that it is the allosteric effector binding that brings about the quaternary structure transition. Cooperativity for fructose-6-phosphate is of catalytic nature. It seems that the cooperativity results in a change in the limiting step of the reaction when the concentration of fructose-6-phosphate increases (AUZAT et al., 1995).

Various site-directed mutagenesis experiments have been performed with the *E. coli* enzyme. The experiments showed that aspartate 127 plays a key role in catalysis. Mutating it to a serine reduces the k_{cat} value 18 000 times. The mutants obtained by replacing threonine 156 by a serine or a glycine, and serine 159 by an asparagine, display Michaelian kinetics and are blocked in the T state. In addition, the T156S or T156G mutants show a hyperbolic response to activation by phosphoenolpyruvate. The main results of the site-directed mutagenesis experiments indicate that the region between residues 156–160 is critical in the transmission of interactions between effector and catalytic sites. In addition, it has been shown that the removal of the C-terminal extremity by subtilisin suppresses the allosteric effects (LE BRAS & GAREL, 1985).

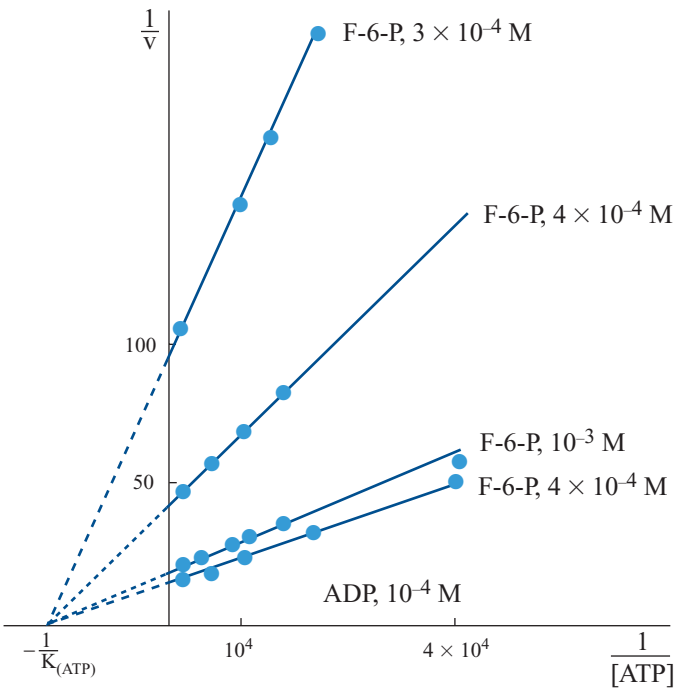


Fig. 13.21 Kinetics of the reaction catalysed by E. coli phosphofructokinase as a function of ATP concentration for different fructose-6-phosphate concentrations

(Reprinted from *J. Mol. Biol.*, **31**, BLANGY D. et al., Kinetics of the allosteric interactions of phosphofructokinase from *Escherichia coli*, 13. © (1968) with permission from Elsevier)

Phosphofructokinases from eukaryotic organisms display different behaviour with respect to allosteric effectors. Contrary to the *E. coli* enzyme, the yeast enzyme has effectors, which are products from reactions that are very close in the metabolic pathway. It is not affected by phosphoenolpyruvate. ADP is only a very weak activator. AMP and ATP are very strong effectors however, AMP is an activator and ATP both substrate and inhibitor. The enzyme's ATP saturation curves are biphasic, showing that two types of site exist: strong affinity sites that correspond to ATP behaving as an allosteric inhibitor ($K_d = 3\ \mu M$) and weak affinity sites ($K_d = 36\ \mu M$), corresponding to ATP substrate binding sites (Fig. 13.22). The 5'AMP activator binding involves only one type of site. The fructose 2,6-biphosphate is also an activator of yeast phosphofructokinase. The enzyme's substrate saturation function is cooperative for fructose-6-phosphate but stays Michaelian for ATP, as is the case for enzymes from prokaryotic organisms.

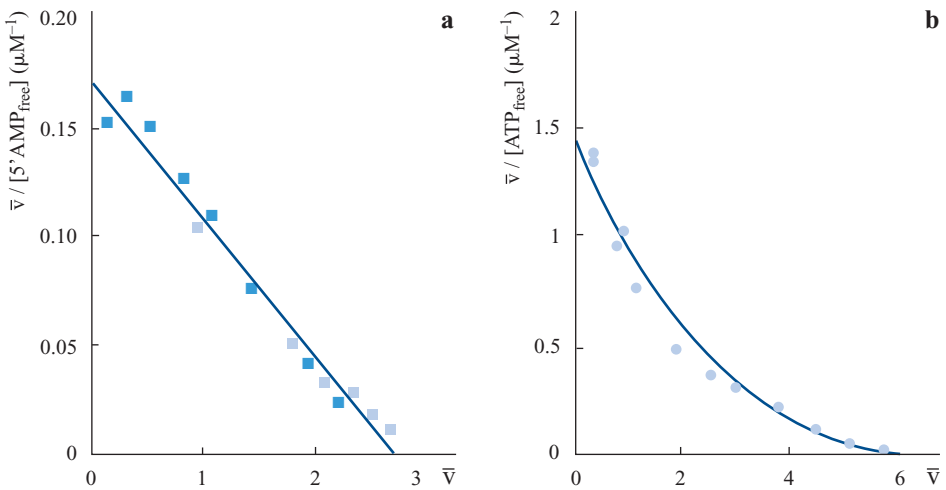
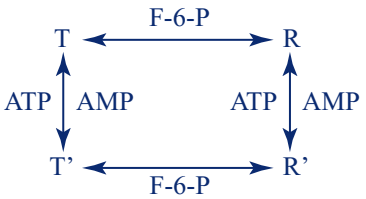


Fig. 13.22 Binding isotherms of 5'AMP (a) and of ATP (b) to yeast phosphofructokinase determined by flow dialysis experiments

SCATCHARD plot (Reprinted from *Biochem. Biophys. Res. Commun.*, **80**, LAURENT M. et al., Binding of nucleotides AMP and ATP to yeast phosphofructokinase: evidence for distinct catalytic and regulatory subunits, 650. © (1978) with permission from Elsevier)

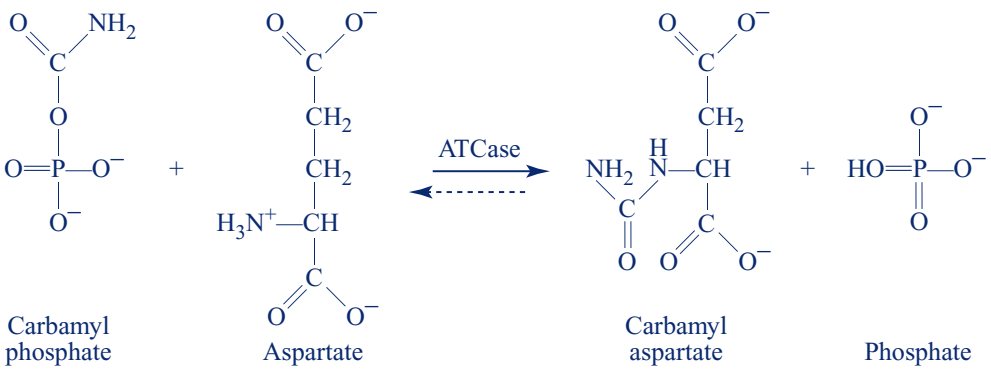
The behaviour of yeast phosphofructokinase cannot be described by the simple two state concerted model however. In fact, inhibition by ATP induces a progressive displacement of the saturation curves by fructose-6-phosphate. For this enzyme, a decoupling between homotropic and heterotropic effects has been observed. ATP binding induces a new conformation of the enzyme T', and the T ⟶ T' transition rate is higher than that of the T ⟶ R transition. In the same way, the binding of the activator AMP brings about a R ⟶ R' transition. To explain all of the experimental data, a model has been proposed:



In this model, the horizontal transitions corresponding to the binding of fructose-6-phosphate are concerted, with a constant L_0 of the order of 3 to 6. The vertical transitions corresponding to heterotropic effects (ATP or AMP binding to the regulatory sites) obey the sequential model and the transition is proportional to the ligand binding. This model in which the four catalytic and four distinct regulator sites are involved explains both the conformational and functional properties of yeast phosphofructokinase. It has also been shown that in this case, it is the allosteric effector binding and not that of the substrates, which brings about a transition in quaternary structure. This shows that the concerted model and the sequential model are too simple to describe the behaviour of the yeast enzyme.

13.10.3. E. COLI ASPARTATE TRANSCARBAMYLASE

Aspartate transcarbamylase (ATCase) catalyses the first specific reaction of the pyrimidine nucleotide biosynthesis pathway, i.e., the carbamylation of the aspartate amino group by the carbamyl phosphate, as is shown below:



In *E. coli*, the reaction proceeds following an ordered mechanism whereby the carbamyl phosphate binds first, followed by aspartate, and the reaction products are released in the order: carbamyl aspartate, then phosphate (PORTER et al., 1969; HSUANYU & WEDLER, 1987). The ATCase of *E. coli* is extensively studied as a model system for the understanding of cooperativity and allostery mechanisms (ALLEWELL, 1989; HERVÉ, 1989; LIPSCOMB, 1992). The enzyme shows the different types of interactions that enzymes regulated by non-covalent modification display: cooperative effects between catalytic sites, positive and negative hetero-heterotropic interactions between regulatory and catalytic sites, as well as cooperative and anti-cooperative effects between catalytic sites.

The ATCase structure shows several levels of organisation. The enzyme is constituted by the association of two trimers of catalytic chains (catalytic subunits of molecular mass 102 kDa), maintained in contact by their association with three dimers of regulatory chains (regulatory subunits of molecular mass 34 kDa). ATCase contains therefore six chains of each type (Fig. 13.23). The catalytic sites are located at the interface between two catalytic chains belonging to the same trimer, and the regulatory sites are located near the interface between the two regulatory chains.

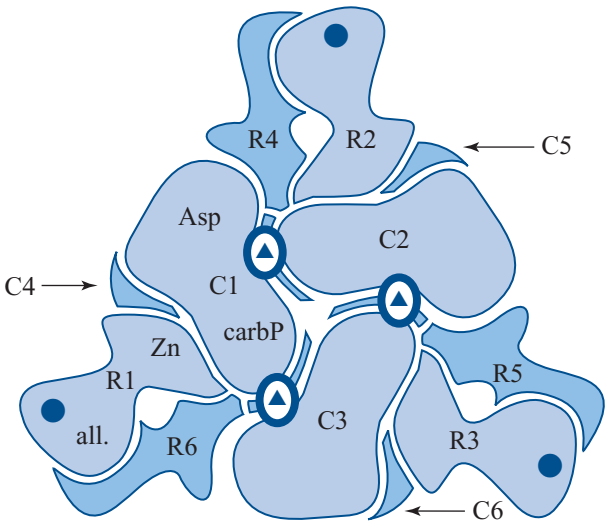


Fig. 13.23 Schematic representation of the quaternary structure of ATCase
(▲) catalytic site region; (●) regulatory site region; Asp: aspartate binding domain in the catalytic subunits; carbP: carbamyl phosphate binding domain in the catalytic subunits; Zn: Zn domain in the regulatory subunits; all.: allosteric domain where the effectors bind in the regulatory subunits. (Reprinted from *J. Mol. Biol.*, **193**, KRAUSE K.L. et al., 2.5 Å structure of aspartate carbamoyltransferase complexed with the bisubstrate analog N-(phosphonacetyl)-aspartate, 528. © (1987) with permission from Elsevier)

The crystallographic structures of several ATCase conformations are known with a 2.5 Å resolution thanks to the work carried out in LIPSCOMB's laboratory. Figure 13.24 opposite shows the structure of a catalytic chain together with a regulatory chain (R1C1 in the Fig. 13.23). Each of the chains is made up of two structural domains: the carbamyl phosphate and the aspartate binding domains in the case of the catalytic chains, the zinc and allosteric domains in the case of the regulatory chains. Each of the latter contain a zinc atom bound to the sulphur atoms of the four cysteine residues found in the C-terminal region of the chains. The reaction of some mercurial agents with the cysteine residues allows the dissociation of ATCase into catalytic and regulatory subunits. The isolated catalytic subunits conserve their capacity to catalyse the carbamylation of aspartate amino group but show no regulatory properties. The regulatory subunits possess no catalytic activity, but conserve their capacity to bind regulatory nucleotides.

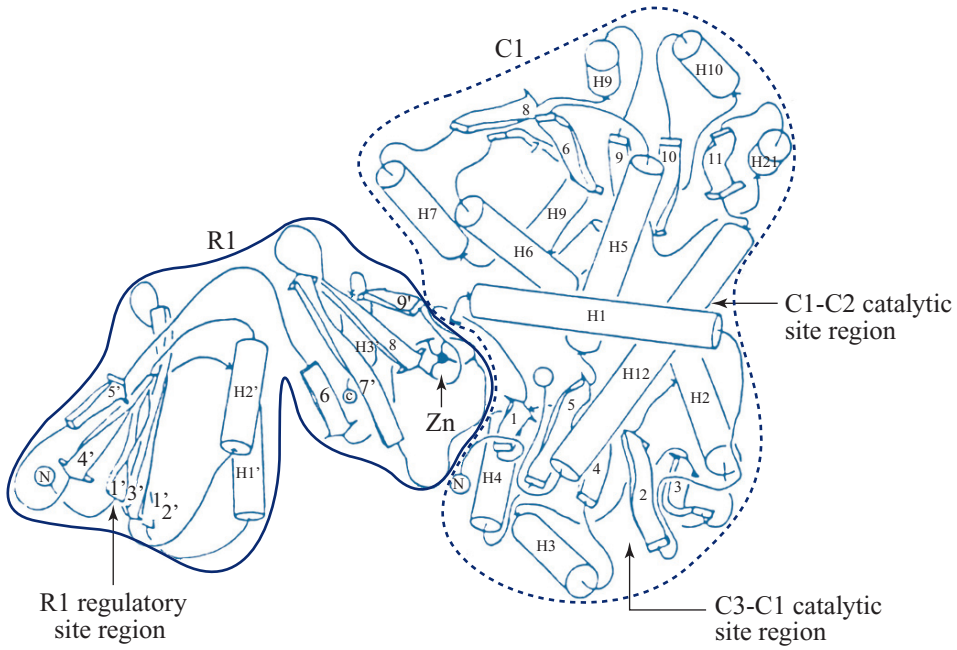


Fig. 13.24 Schematic representation of the structure of a catalytic subunit together with a regulatory subunit

(Reprinted from *J. Mol. Biol.*, **193**, KRAUSE K.L. et al., 2.5 Å structure of aspartate carbamoyltransferase complexed with the bisubstrate analog N-(phosphonacetyl)-aspartate, 527. © (1987) with permission from Elsevier)

13.10.3.1. COOPERATIVE EFFECTS BETWEEN CATALYTIC SITES

ATCase displays cooperative effects between the catalytic sites for aspartate binding. The enzyme saturation curve for aspartate is sigmoidal in the presence of saturating carbamyl phosphate. Succinate, a competitive analogue of aspartate, displays the same phenomenon on binding under the same conditions. The existence of these cooperative effects is explained by the transition of ATCase from one conformation, which displays a weak affinity for aspartate (T state), to a conformation that has a strong affinity for this substrate (R state). The crystallographic structures of the two extreme conformations (Fig. 13.25) are known at 2.5 Å resolution (HONZATKO et al., 1987; KE et al., 1988). The transition from one conform-ation to the other involves three aspects that are coupled: a quaternary structure transition, a tertiary structure transition and a catalytic site conformational transition.

Quaternary structure transition

During the change from the T state to the R state (Fig. 13.25 below), the two catalytic subunits move apart by 11 Å along the 3-fold symmetry axis of the molecule, that is, by a tenth of the diameter of the molecule. This movement is associated with their rotation by 5° in opposite directions around the same axis, and with a reorientation of the regulatory dimers. A variation in the relative position of the polypeptide

loops belonging to the catalytic chains (in particular the loop containing the residue 240) also characterises the motion. The variation in quaternary structure brings about an alteration in the interfaces between subunits. In this way, during the transition, the C1C4 and R1C4 type interfaces are broken.

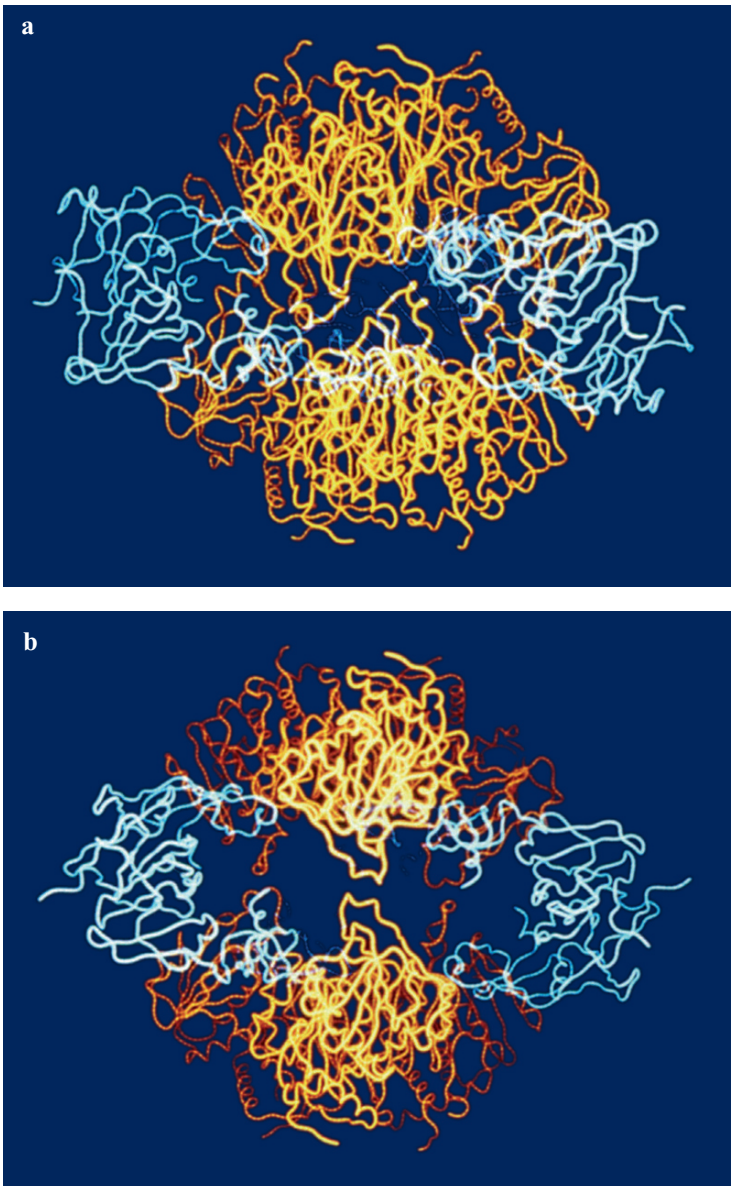


Fig. 13.25 ATCase structure, in the T state (a) and in the R state (b)
(PDB: 1RAB and 8ATC)

The rearrangement is characterised, in particular, by the breaking of a certain number of ionic interactions, both between catalytic chains, and between catalytic and

regulatory chains (Fig. 13.26). The interactions are replaced by new ionic intra-chain interactions which, in some cases, involve the same amino acid side-chains.

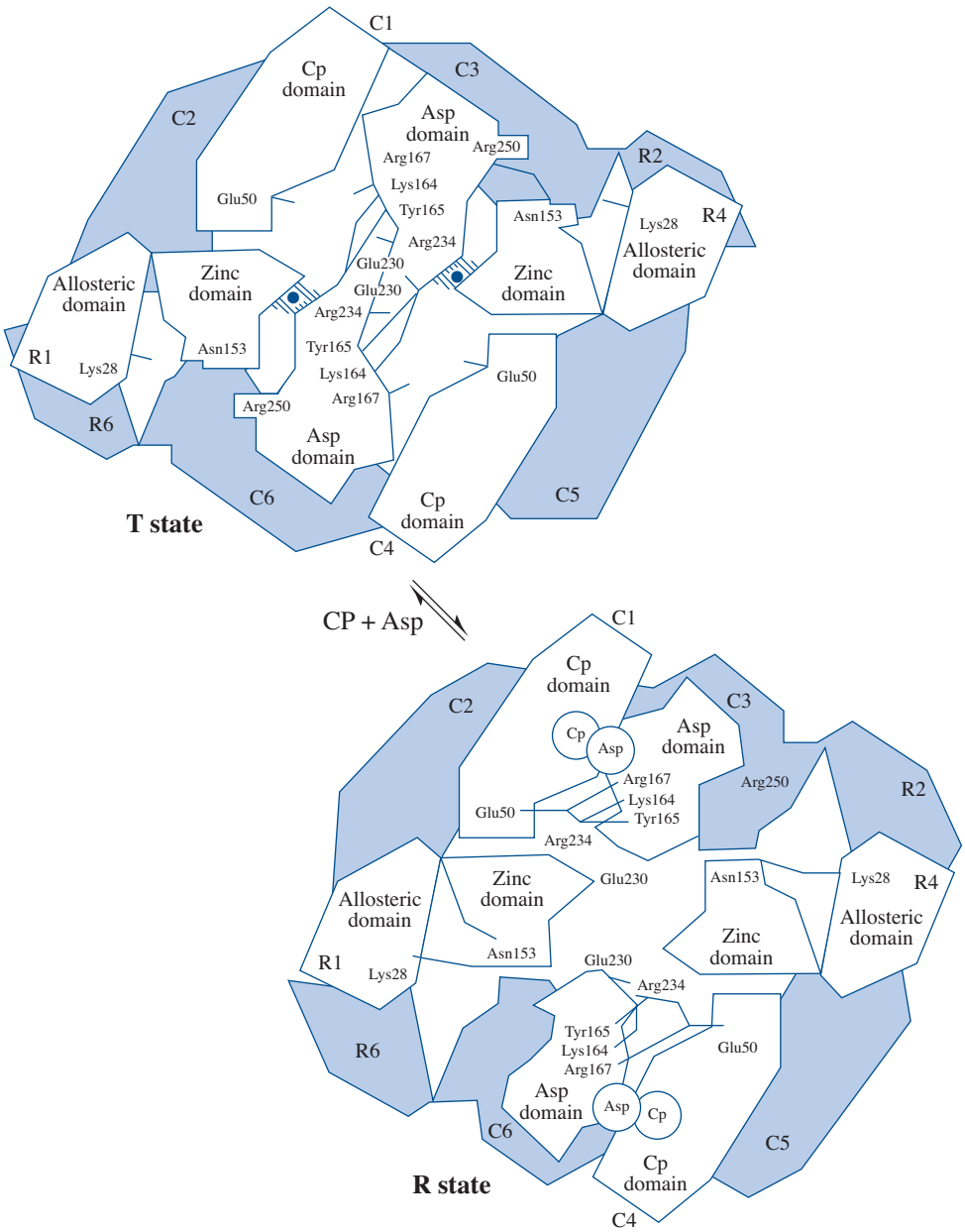


Fig. 13.26 Change in the subunit interactions that occurs during the allosteric transition of aspartate transcarbamylase

(Reprinted from *J. Mol. Biol.*, **125**, THIRY L. & HERVÉ G., The stimulation of Escherichia coli aspartate transcarbamylase activity by adenosine triphosphate: relation with the other regulatory conformational changes; a model, 515. © (1978) with permission from Elsevier)

Tertiary structure transition

A tertiary structure transition of the catalytic chains is associated with the quaternary structure transition. In particular, the transition in tertiary structure concerns two loop regions located in one and the other of the two catalytic chain domains (Fig. 13.27). This variation in tertiary structure results in the two domains of the catalytic chains coming closer together, and in this way tighten more closely to the substrate molecules bound to the catalytic site (Fig. 13.28).

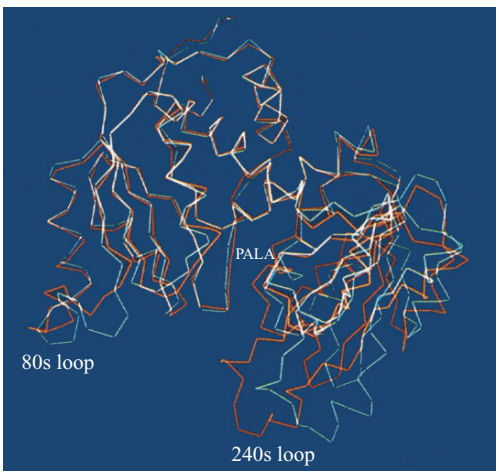


Fig. 13.27 Tertiary structure transition
Red: *R state;* blue: *T state. PALA indicates the position of the catalytic site*

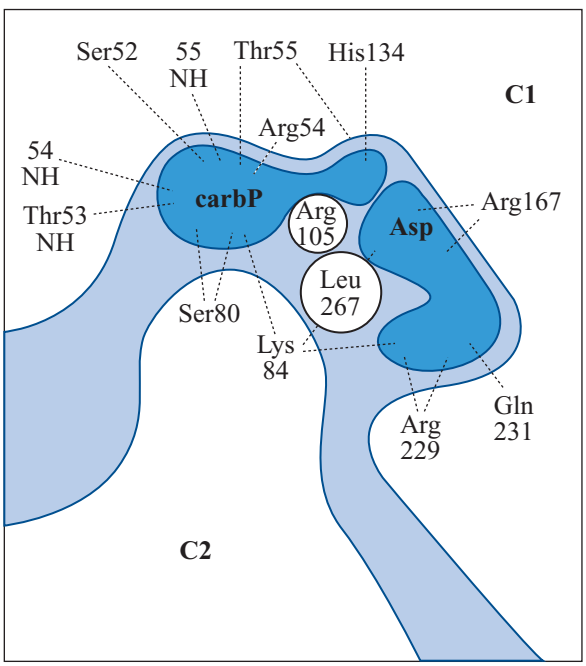


Fig. 13.28 Schematic representation of ATCase's active centre, with the approximate carbamyl phosphate and aspartate positions
The enzyme amino acids involved in substrate binding are shown
(Reprinted from *J. Mol. Biol.*, **204**, KE H. et al., Complex of N-phosphonacetyl-aspartate with aspartate carbamoyltransferase: X-ray refinement, analysis of conformational changes and catalytic and allosteric mechanisms, 725. © (1988) with permission from Elsevier)

Catalytic site conformational transition

The change in tertiary structure of the catalytic chains therefore implies a modification of the catalytic site, a modification that tends to bring the amino acid side-chains that interact with the substrate molecules closer together. The catalytic site rearrangement includes a variation in the pKa of the group(s) involved in aspartate binding and/or in catalysis (Fig. 13.28). In fact, in ATCase, the pH dependence of the catalysed reaction varies considerably during the change from the T to R states. A shift of the activity optimum by one unit of pH is evidence of the important variation. The pH profile of the reaction in the R state is identical to that observed for the isolated catalytic subunits, in which it is insensitive to the substrate concentration. This insensitivity has made it possible to analyse the pH profile. The variation in the enzymatic activity as a function of pH is a result of the protonation-deprotonation of three groups (Fig. 13.29). The binding of the first substrate, carbamyl phosphate, shifts the pKa of a group from 8.2 to 7. The binding of the second substrate, aspartate, involves a group of the protein whose pKa is shifted from 7.2 to 9.4 in the complex. Finally, the catalytic act involves a group whose pKa is 7.2. Site-directed mutagenesis has been used to identify the groups corresponding to the different pKa, in particular the group whose pKa is shifted during the T to R transition. The technique has allowed histidine 134 to be identified as the residue whose deprotonation increases the affinity of the enzyme for the carbamyl phosphate (XI et al., 1990).

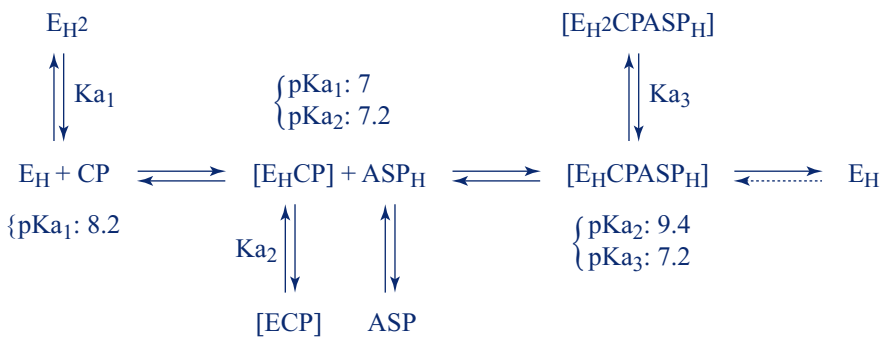


Fig. 13.29 pK variations of the active centre groups in the different steps of the reaction
(Reprinted from *J. Mol. Biol.*, **125**, THIRY L. & HERVÉ G., The stimulation of *Escherichia coli* aspartate transcarbamylase activity by adenosine triphosphate: relation with the other regulatory conformational changes; a model, 515. © (1978) with permission from Elsevier)

Although the extreme conformations, R and T, involved in the cooperative homotropic effects between catalytic sites are now relatively well-known, it is not the case for the mechanism by which the enzyme passes from one state to the other. On the basis of currently available data, it seems that the transition involved is both **concerted and induced**. It is concerted in the sense that the complete displacement towards the R state is observed before the saturation of catalytic sites by aspartate or its analogues (GRIFFIN et al., 1973; FETLER et al., 1995). This transition is induced in the sense that it appears not to be the result of a simple pre-existing thermodynamical

equilibrium, which would be displaced by the exclusive or preferential substrate or analogue binding to the R state. The structure of the substrate or its analogues appears to bring information, and the process therefore probably involves an induced fit (BAILLON et al., 1985; FOOTE et al., 1985; FOOTE & LIPSCOMB, 1981).

13.10.3.2. ALLOSTERY – HETEROTROPIC INTERACTIONS BETWEEN REGULATORY AND CATALYTIC SITES

The activity of ATCase is retro-inhibited in a synergistic manner by the final metabolites CTP and UTP (WILD et al., 1989). On the contrary, the activity is increased by ATP. The resulting antagonism tends, in the cell, to equilibrate the biosynthesis of the pyrimidine and purine nucleotides. The first models that attempted to explain ATCase's regulatory properties supposed that CTP and ATP acted directly on the transition implicated in the homotropic effects between catalytic sites described above. Over the years, experimental evidence was accumulated that showed that the nucleotides work based on different mechanisms.

"Primary-secondary effects" mechanism

The totality of the obtained results concerning the action of CTP and ATP effectors has driven to the conclusion that these effectors do not act directly on the T ⟷ R transition, but alter, through localised conformational changes, the catalytic sites' affinity for aspartate (primary effect). In the presence of a given aspartate concentration, the affinity change modifies the substrate occupation of the catalytic sites, which displaces the T ⟷ R equilibrium (secondary effect). In this way, the effect on the equilibrium is mediated by the substrate and not directly by the nucleotide (THIRY & HERVÉ, 1978; TAUC et al., 1982). The mechanism is presented in a schematic way in Fig. 13.30 opposite. The most simple formalism capable of describing the process associates the equation of the MONOD, WYMAN and CHANGEUX concerted model to the formalism that describes the partial competitive inhibition, the type of inhibition whereby the inhibitor binds to a site distinct from the catalytic site (see Chap. 5).

It has been possible to verify the predictions that follow on from this model, by using small angle X-ray scattering experiments (HERVÉ et al., 1985). In addition, isotope exchange experiments at equilibrium have allowed the verification that the CTP and ATP effectors do not directly affect the equilibrium T ⟷ R (HSUANYU & WEDLER, 1988), but affect the aspartate binding rate constant to the catalytic site.

The fact that the nucleotides do not act only directly on the transition involved in the homotropic cooperative effects between catalytic sites, associated to the fact that in ATCase the transition is induced, has lead to the proposition of a modified version of the mechanism, called "effector modulated quaternary transition" (XI et al., 1991). In this mechanism, the local conformational change provoked by the

nucleotide binding does not in itself induce any change in the affinity of the catalytic site for aspartate. On the contrary, through the modification of the interfaces between catalytic and regulatory chains, whose interaction energy is modified, the nucleotide facilitates (ATP) or on the contrary disfavours (CTP) the quaternary structure transition induced by the substrate. In other words, and for example in the presence of ATP bound to the regulatory sites, it would suffice to bind a lesser number of aspartate molecules to the catalytic sites to provoke the complete transition to the R conformation. The inverse phenomenon would occur in the presence of CTP. In fact, the determination of the crystallographic structures of the R and T states, which have CTP and ATP bound (STEVENS et al., 1990; GOUAUX et al., 1990), as well as the results of site-directed mutagenesis experiments and the use of pseudo-substrates, suggest that the two mechanisms described operate simultaneously. On the basis of the crystallographic data, it was proposed by STEVENS and LIPSCOMB (1992) that the nucleotide binding to the regulatory sites would induce the transmission of a direct and an indirect signal.

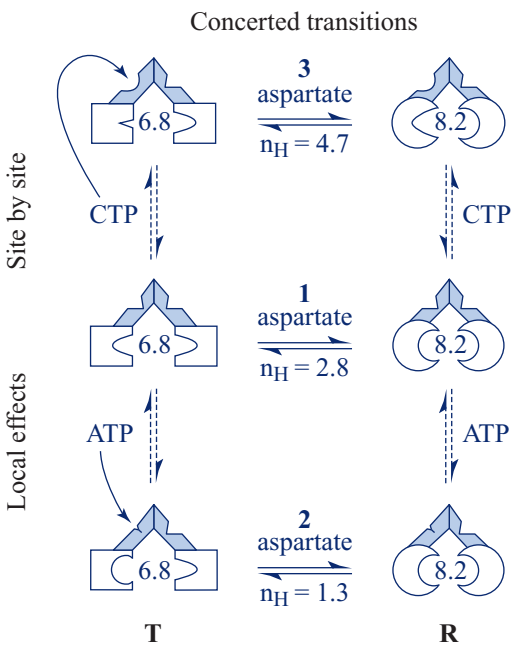
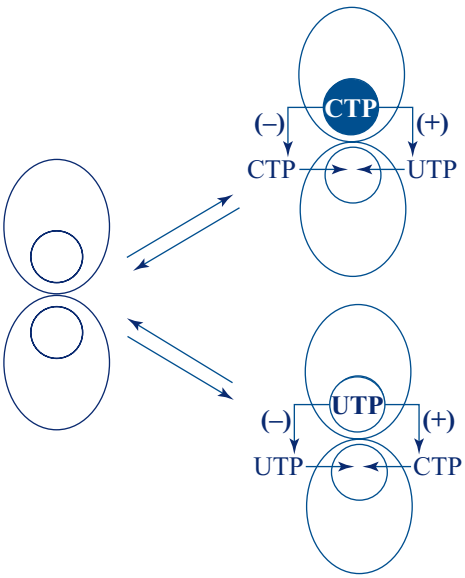


Fig. 13.30 Heterotropic effects in ATCase: the "primary-secondary effects" mechanism
The activity optima pHs are shown for the different conformations. n_H is HILL's number for each of the transitions
(Reprinted from *J. Mol. Biol.*, **125**, THIRY L. & HERVÉ G., The stimulation of *Escherichia coli* aspartate transcarbamylase activity by adenosine triphosphate: relation with the other regulatory conformational changes; a model, 515. © (1978) with permission from Elsevier)

Synergistic action of CTP and UTP

◤ It has been shown that the two final metabolites, CTP and UTP, decrease the aspartate binding to the catalytic sites in a synergistic manner. UTP alone does not have an effect, even though it binds to the regulatory sites. CTP has a limited effect. In the presence of the two nucleotides, ATCase is completely inhibited (WILD *et al.*, 1989). The synergy phenomenon manifests itself entirely in the nucleotide binding to the two regulatory sites of the regulatory dimers (Fig. 13.31 below). UTP and CTP bind in a competitive manner to the same regulatory sites. CTP binding to the first

site decreases the affinity of the second site for CTP, but increases its affinity for UTP and conversely (ENGLAND & HERVÉ, 1992). The phenomenon is identical if the regulatory subunits (dimers of regulatory chains) are associated or not to the catalytic subunits, showing that in this process there is no interaction between the regulatory subunits in the holoenzyme.



$K_{d\,(UTP)}$ = 810 μm → real $K_{d\,(UTP/CTP)}$ = 10 μM

$K_d1_{\,(CTP)}$ = 9 μm → real $K_d1_{\,(CTP/UTP)}$ = 1.7 μM

$K_d2_{\,(CTP)}$ = 260 μm → real $K_d2_{\,(CTP/UTP)}$ = 70 μM

Fig. 13.31 Model of the interactions between regulatory sites within a regulatory dimer in ATCase
The + and – signs correspond to an increase (cooperativity) and to a decrease (anticooperativity) of affinity, respectively
(Reprinted with permission from *Biochemistry*, **31**, ENGLAND P. & HERVÉ G., Synergistic inhibition of *Escherichia coli* aspartate transcarbamylase by CTP and UTP: binding studies using continuous-flow dialysis, 9725. © (1992) American Chemical Society)

The regulatory signals of ATP, UTP and CTP are transmitted by different paths

Extensive use of site-directed mutagenesis has shown that the transmission of different regulatory signals of the three effectors imply different interface zones between the catalytic chains and the regulatory chains, but also between different domains of the same chain (XI et al., 1991; VAN VLIET et al., 1991; DE STAERCKE et al., 1995). At the interface between the C-terminal region of one regulatory chain and the loop containing the residue 240 of a catalytic chain (R1C4 type interface in Fig. 13.23), two contact zones, close to each other, are specifically involved, one in the transmission of the regulatory CTP signal, the other in the transmission of the regulatory ATP signal.

The regulatory chain is constituted of two domains, the allosteric domain where the regulatory site is located, and the zinc domain, which is in contact with the catalytic chains. The interface between the two domains is constituted of a hydrophobic pocket in which a tyrosine residue is inserted (Fig. 13.32 opposite). All of the interface is essential for the transmission of the ATP regulatory signal (XI et al., 1994).

In this manner, ATCase can be considered as a transfer and integration network of regulatory signals (Fig. 13.33). The integration is performed at the catalytic site.

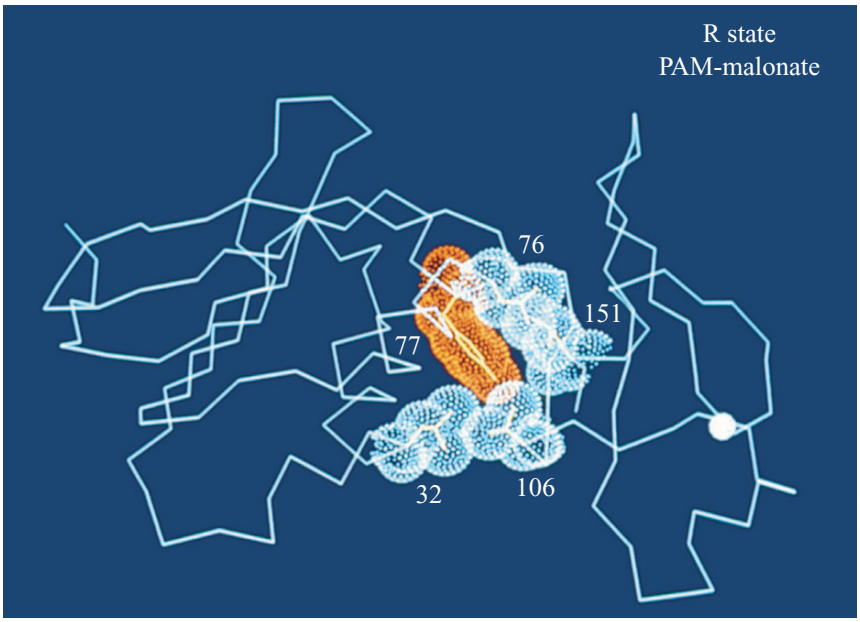


Fig. 13.32 Hydrophobic interface between the two domains of the regulatory chain, in which a tyrosine residue is inserted (orange: *tyrosine 77;* blue: *residues of the hydrophobic pocket)* (Reprinted from *J. Mol. Biol.*, **242**, XI X.G. et al., The Activation of *Escherichia coli* Aspartate Transcarbamylase by ATP: Specific Involvement of Helix H2' at the Hydrophobic Interface Between the Two Domains of the Regulatory Chains, 139. © (1994) with permission from Elsevier)

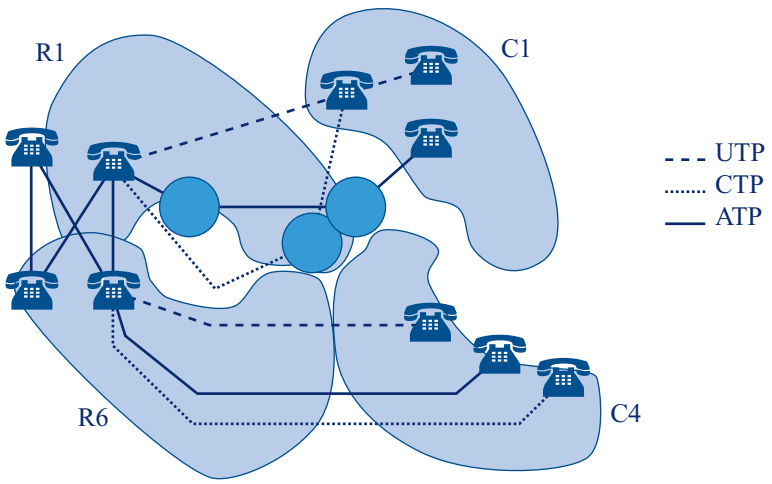


Fig. 13.33 Schematic representation of the transmission of regulatory signals in aspartate transcarbamylase

13.10.4. RIBONUCLEOTIDE REDUCTASE

Ribonucleotide reductase is an enzyme that plays an extremely important role, since it is responsible for the transformation of ribonucleotides into the corresponding desoxyribonucleotides. In the majority of species, the modification occurs on nucleotide diphosphates. The enzyme constitutes a particularly interesting example in the context of allosteric regulation mechanisms. Although in the majority of species the enzyme is oligomeric, it doesn't show any cooperativity, although, it is the subject of a very elaborate allosteric regulation (ERIKSSON & SJÖBERG, 1989).

13.10.4.1. REACTION MECHANISM

The reaction mechanism of the ribonucleotide reductases ensures the direct replacement of the OH-group in the ribose 2' position by a hydrogen atom, which is supplied by NADPH. This process involves, however, the participation of several proteins acting as hydrogen transporters (Fig. 13.34a).

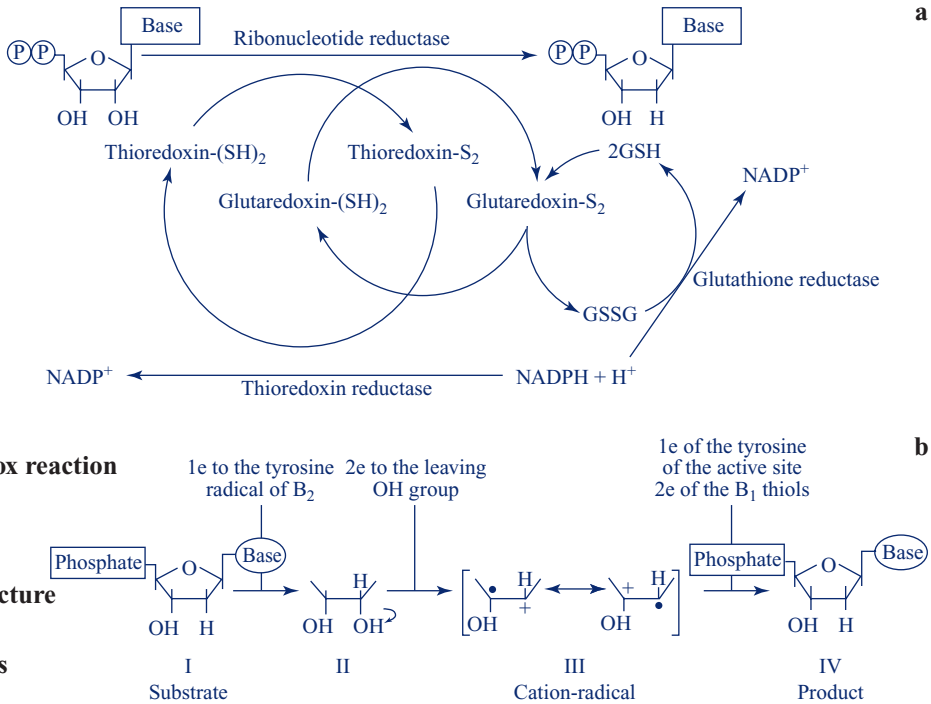


Fig. 13.34 Enzymatic reduction of nucleotides

*(**a**) components involved in the transfer of NADPH electrons – (**b**) reactional mechanism proposed for the ribonucleotide reductase of E. coli* (Reprinted from *Allosteric Enzymes*, ERIKSON S. & SJÖBERG B.M., Ribonucleotide reductase, G. HERVÉ ed., 189. © (1989) CRC Press, Taylor and Francis Group)

In the specific case of the *E. coli* ribonucleotide reductase, the reaction mechanism involves the combined action of a dithiol (two cysteine residues), which is the ob-

ject of a redox process, two iron atoms directly linked to the protein without a heme intermediate, and an organic free radical found on a tyrosine residue buried in the protein structure (FONTECAVE et al., 1992).

The first step of the reaction consists of the detachment of the hydrogen atom in position 3′ of the ribose by the tyrosine's free radical (Fig. 13.34b). In the following step, the hydroxyl group in position 2′ is torn away and gives rise to an intermediate cation radical. Finally, the product formation involves the regeneration of the tyrosine's free radical, by transfer of an electron to the intermediate cation rad-ical and a hydrogen atom binding in 2′ of the substrate, at the expense of the dithiol present in the enzyme. Excepting a few details, the mechanism is common to the different types of ribonucleotide reductases. In the majority of cases, the redox system associated with the reaction involves thioredoxin, glutathione and glutaredoxin (Fig. 13.34).

The FONTECAVE and REICHARD groups have shown that the formation of the stable free radical on tyrosine 122 of the *E. coli* enzyme results from the action of a NADH-flavin reductase. It acts *via* the intermediate of the redox centre that constitutes the dithiol (FONTECAVE et al., 1987; COVÈS et al., 1993).

13.10.4.2. RIBONUCLEOTIDE REDUCTASE STRUCTURE

To date, two types of enzymes have been found depending on the organisms studied. The enzymes of the first type are monomeric and their activity depends on the presence of a cofactor: adenosylcobalamine. It is the type of enzyme that is found in *L. leichmanii* (Fig. 13.35a below). In the majority of cases, these enzymes use the nucleotide triphosphates as substrates.

The second type of ribonucleotide reductases is found for example in *E. coli* and higher organisms. These enzymes have a tetrameric structure of the $\alpha_2\beta_2$ type (Fig. 13.35b). α corresponds to the subunits called B_1, of about 85 kDa molecular mass and β corresponds to the subunits called B_2, also of about 85 kDa molecular mass. The proteins easily dissociate in pure solution but the presence of Mg^{++} ions favours the associated form. The iron atoms and the free radical present on tyrosine 122 are located in the B_2 subunit of the *E. coli* protein. The dithiol is localised in the B_1 subunit however, which suggests that the catalytic site is situated at the interface between the two types of subunits. The allosteric effector binding sites are also found in the B_1 subunits. All these enzymes show a strong sequence homology.

The crystallographic structure of the B_2 subunit of ribonucleotide reductase from *E. coli* was determined by the EKLUND group with a resolution of 2.2 Å (NORDLUND & EKLUND, 1993). The structure is remarkable in that it is constituted of α helices to 70%, some of which contain more than 30 residues. Only one loop contains a short antiparallel β sheet (Fig. 13.36a below). Tyrosine 122, which bears the free radical, is buried at 10 Å from the surface of the protein and at 5 Å from the closest iron atom (Fig. 13.37 below). The two iron atoms are separated by 25 Å.

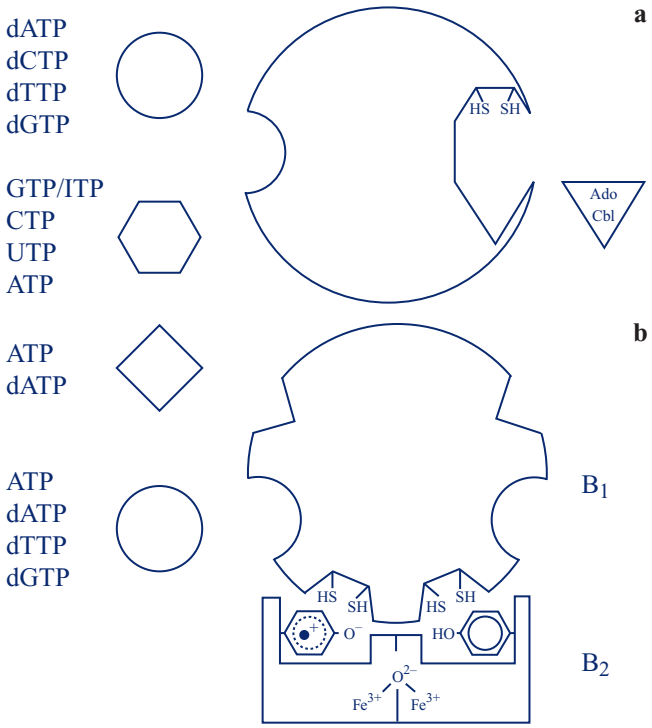


Fig. 13.35 Schematic representation of the ribonucleotide triphosphate reductase (a) from L. leichmanii – (b) from E. coli. (Reprinted from *Allosteric Enzymes*, ERIKSON S. & SJÖBERG B.M., Ribonucleotide reductase, G. HERVÉ ed., 189. © (1989) CRC Press, Taylor and Francis Group)

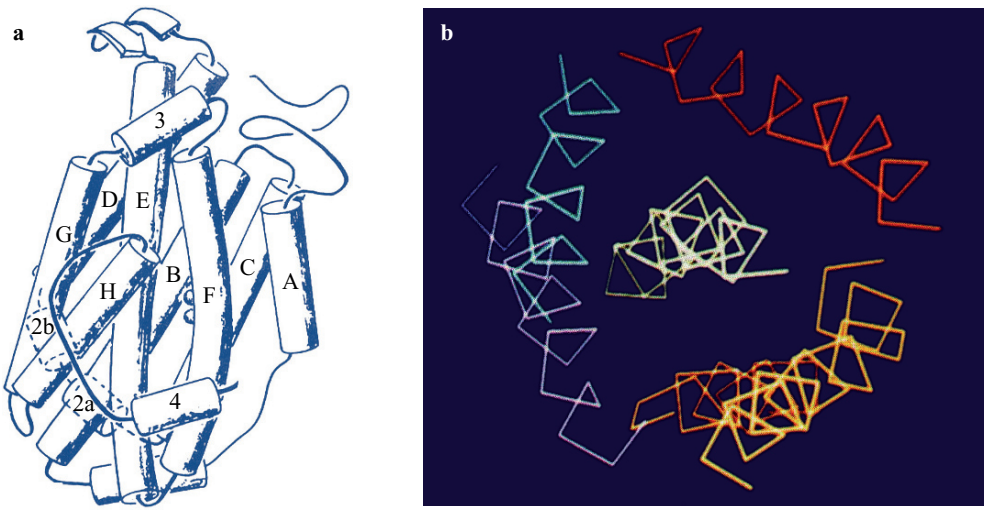


Fig. 13.36 (a) schematic representation of the E. coli ribonucleotide reductase B_2 subunit secondary structure. The two iron atoms are represented by small spheres (b) α helices of the E. coli ribonucleotide reductase B_2 subunit, showing the β barrel
(Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **345**, NORDLUND P. et al., 594. © (1990))

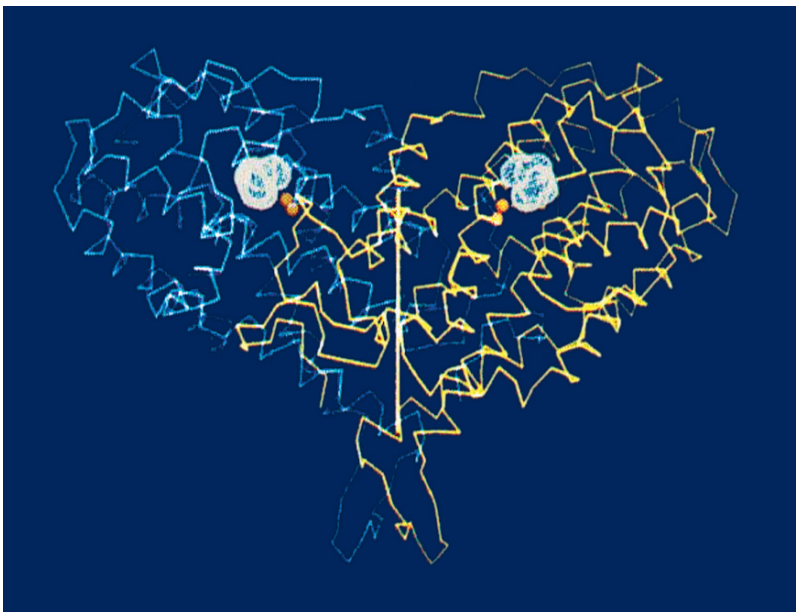


Fig. 13.37 The B_2 dimer of ribonucleotide reductase
The C^{α} of one subunit chain are in blue*, and those of the other one in* yellow. *The iron atoms in the centre are represented by orange spheres. The VAN DER WAALS surface of Tyr122 is in* light blue. *The 2-fold symmetry axis of the molecule is represented by a vertical* yellow line (Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **345**, NORDLUND P. et al., 594. © (1990))

13.10.4.3. ALLOSTERIC REGULATION

The activity of the different enzymes is regulated by several nucleotide effectors. Substrates and effectors bind to distinct and specific sites. Nevertheless, it has been possible to show the existence of "squatting" (see Sect. 13.11), which involves difference in affinities by a factor 100.

Monomeric enzymes (Lactobacillus leichmanii type)

The only catalytic site of the enzymes uses as a substrate the four nucleotide triphosphates. The reaction is regulated by different end-products. All the allosteric effectors bind to the same regulatory site but lead to specific effects:

- dATP increases the rate of CTP reduction,
- dCTP increases the rate of UTP reduction,
- dTTP increases the rate of GTP reduction,
- dGTP increases the rate of ATP reduction.

These allosteric effects tend to equilibrate the production of the four desoxyribonucleotide triphosphates.

Tetrameric enzymes

The four substrates, nucleotide diphosphates, bind to the same catalytic site. The site is located at the interface between the B_1 and B_2 subunits. The B_1 subunit possesses two types of regulatory sites, which differ, among other things, by their dATP affinity constant; the dissociation constant is 0.5 μM for the l site and 0.03 μM for the h site.

L sites or "activity sites"

These regulatory sites bind ATP and dATP and the effectors act on the catalytic activity and not on the substrate affinity. ATP activates the enzyme and dATP inhibits it, creating an antagonism that tends therefore, in a synergistic manner, to regulate dATP production. The inactivation by dATP is accompanied by an aggregation of the enzyme.

H sites or "specificity sites"

This regulatory site behaves in the same way as the regulatory site of the monomeric ribonucleotide reductases. The binding of a given effector specifically provokes an increase of the catalytic site's affinity for given substrates.

In the absence of an effector, the reaction rate is weak. In the presence of an effector, it increases in a specific manner:

- dATP increases the reduction rate of CDP and of UDP,
- dGTP increases the reduction rate of ADP and of GDP,
- dTTP increases the reduction rate of the four substrates.

In addition, GDP binding to the catalytic site increases the affinity of the h sites for dTTP, which, in return, stimulates the use of the four substrates.

These enzymes therefore show very complex and subtle allosteric regulation properties. However, they display no cooperativity between catalytic sites, even in the presence of effectors.

Ribonucleotide reduction in anaerobic conditions

◤ It has been shown by the FONTECAVE and REICHARD groups, that in anaerobiosis, *E. coli* uses a different type of ribonucleotide reductase for the synthesis of desoxyribonucleotides (FONTECAVE et al., 1989; MULLIEZ et al., 1993). The activity of the enzyme is very sensitive to oxidation. It uses the nucleotide triphosphates as substrates. Its activity involves an iron-sulphur centre, and the free radical is not carried by a tyrosine residue, but probably by a glycine residue located in position 681 of the polypeptide chain. Antibodies prepared against the aerobic ribonucleotide reductase described previously do not react with the anaerobic ribonucleotide reductase, showing that these two enzymes have different structures. This enzyme is activated by ATP and inhibited by desoxyribonucleotide triphosphates, which indicates that it is also subjected to allosteric regulations.

The existence of this enzyme may have an interesting significance from the point of view of evolution. It may come from an enzymatic system which allowed, in anaerobic conditions, the transition from the "RNA world" to the "DNA world" (FONTECAVE et al., 1989).

13.11. "SQUATTING"

The concerted model of MONOD, WYMAN and CHANGEUX postulated that substrates and effectors bind to distinct sites. To date, the prediction has been verified in the case of allosteric enzymes whose structure is known with sufficient precision. The specific binding to distinct sites is easily conceivable in the case where substrates and effectors have very different structures. The situation is more complex when substrates and effectors have a structural similarity. It is the case, amongst others, for phosphofructokinase and ribonucleotide reductase (see Sects. 13.10.2 and 13.10.4), enzymes for which it has been shown that each of the two ligand types may bind to the other type of site, although with a weaker affinity than for the specific site. The phenomenon was called "squatting" by MAZAT (1977), who modelled the effects and showed it could confer very elaborate and advantageous regulation properties for the cell.

The modelling performed by MAZAT was based on the simple two state model, T and R states, of weak and strong affinities for the substrate and the effector(s) and consists in varying the relative ligand concentrations as well as the model parameters (allosteric constant L_0, K_T, K_R etc.). The models explore the regulation possibilities that may result from the existence of squatting, either between different types of regulatory sites, or between catalytic and regulatory sites (Fig. 13.38).

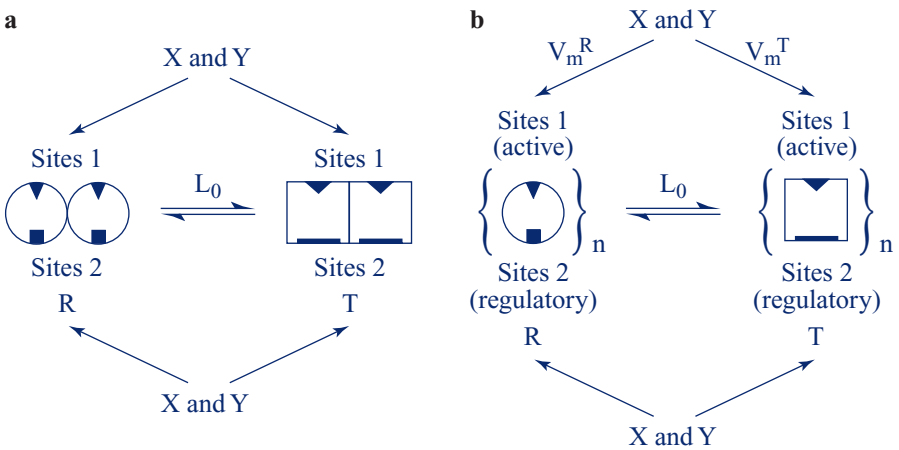


Fig. 13.38 Squatting in the simple two state model

*(**a**) squatting between two effectors – (**b**) squatting between substrate and effector. X and Y are ligands, L_0 the allosteric constant in the absence of any ligand, V_m^R and V_m^T the maximal reaction rates corresponding to the R and T states, respectively*

(Reprinted from *J. Theor. Biol.*, **68**, MAZAT J.P. et al., 365. © (1977) with permission from Elsevier)

When the two ligands considered are effectors binding to different sites, squatting may provoke the following behaviour:

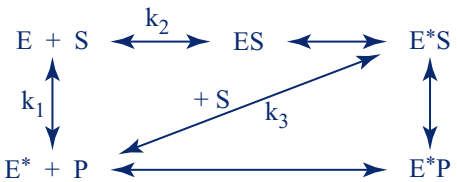
- **Concerted or multivalent inhibition –** In this case the two effectors, which are both weak inhibitors, can provoke the enzyme's total inhibition in a synergistic manner (Fig. 13.39 opposite).
- **Effect reversal –** For certain values of the two state model constants, the simultaneous presence of two effectors can provoke the reversal of their effects. In the case presented in Fig. 13.40a below, the X and Y effectors considered separately are activators, but when they are both present, they provoke a strong inhibition. In the case presented in Fig. 13.40b, X and Y considered separately are good inhibitors, but when they are both present they provoke a strong activation.

The modelling results are also spectacular in the case of squatting between catalytic and regulatory sites and very complex regulatory effects can be observed. The model's main point of interest is the fact that certain allosteric enzymes display effects similar to the model's predictions. It is the case for glycogen phosphorylase, aspartokinase and phosphofructokinase, amongst others. For example, in the case of the aspartokinase from *Pseudomonas testeroni*, the reversal of the effect of lysine, which alone plays the role of an activator, is observed. In the presence of threonine, which has no effect by itself even at high concentration, lysine completely inhibits the enzyme activity.

13.12. "MNEMONIC" ENZYMES

Some monomeric enzymes display kinetic activity that differs from the MICHAELIS law. The variation in reaction rate as a function of substrate concentration is sigmoidal, as in the case of enzymes presenting cooperativity between catalytic sites. For metabolic regulation, the behaviour presents the same advantage as cooperativity, but obviously does not result from interactions between catalytic sites. It depends on an isomerisation process of the enzyme. Amongst the enzymes that manifest the phenomenon, we can cite wheat germ hexokinase L1, β-glycosyl transferase from plant cell walls, glucokinase from rat liver, octopine deshydrogenase from PECTEN.

The principle of the mode of function of these enzymes was presented for the first time by RABIN in 1967: "the conformation of the free enzyme at the end of the catalytic cycle is different from its initial conformation". The term of an enzyme "with memory" was introduced for the first time by WHITEHEAD (1970) with the following meaning: "the enzyme remembers for a certain time the conformation it had when it was still associated to the reaction product(s)". To explain the phenomenon, a more simplified picture was proposed by RABIN in 1967, in which the catalysis involves the enzyme's isomerisation in an activated form E*:



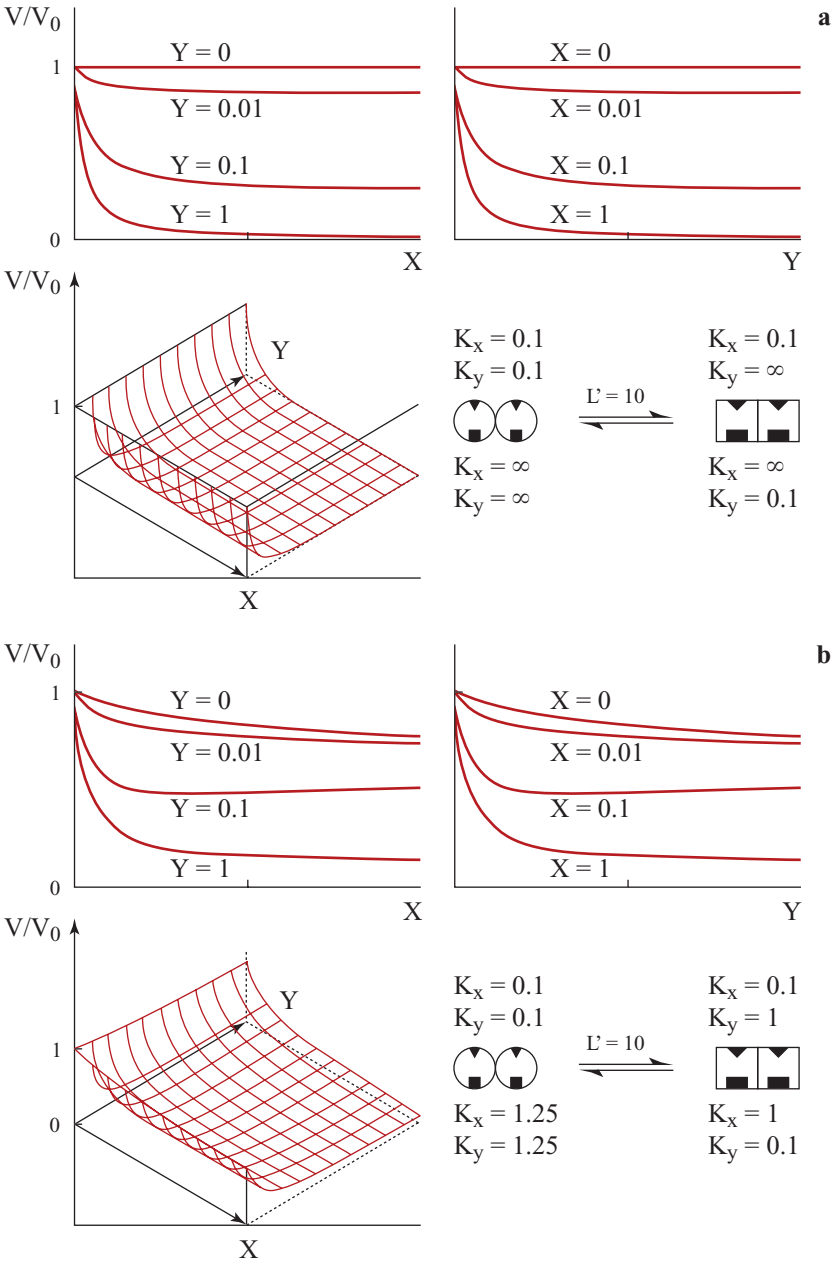
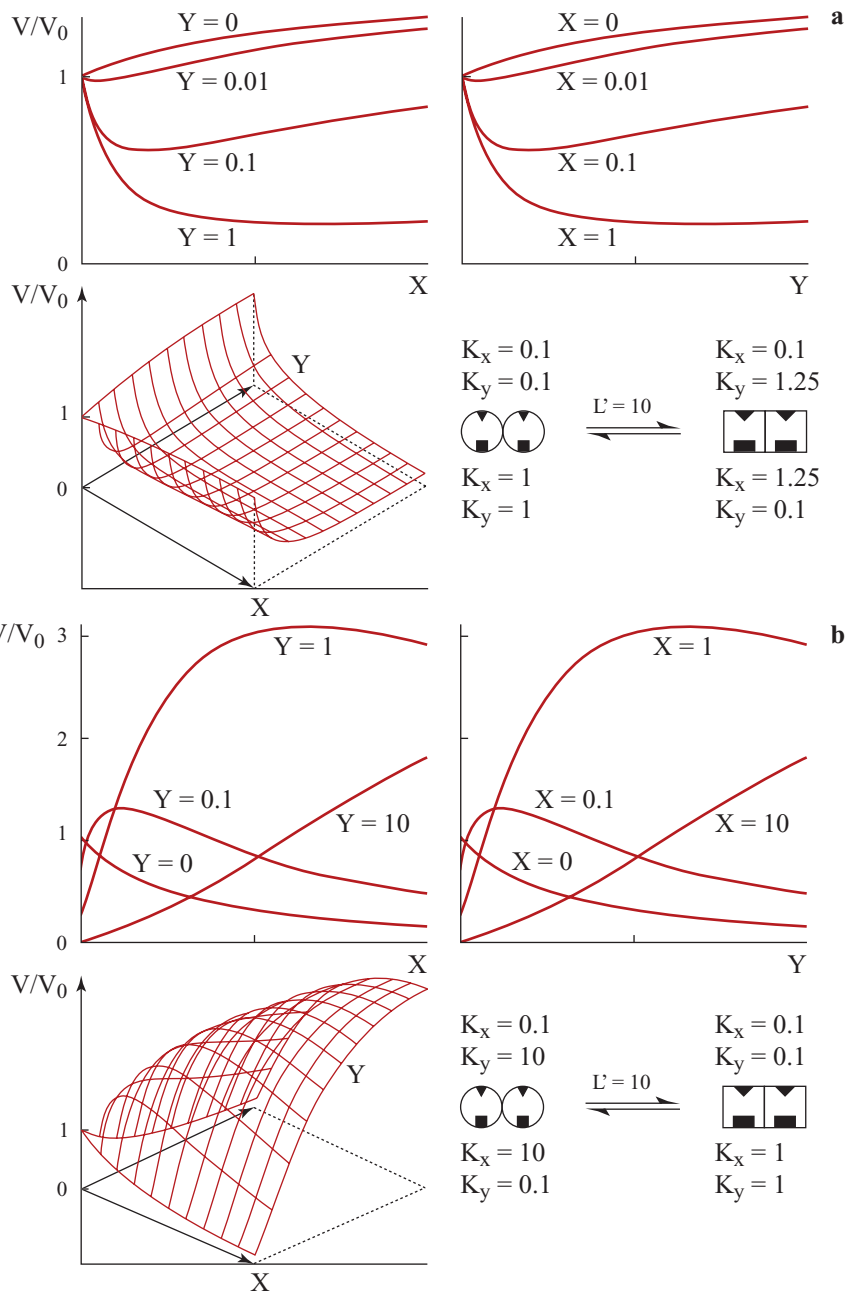


Fig. 13.39 Squatting: concerted inhibition

The dissociation constants are indicated on the figures in arbitrary units. Note that the constant values of figures ***(a)*** *and* ***(b)*** *are very close. The T form is completely inactive.* ***(a)*** *the ligands X and Y alone imply no change in activity –* ***(b)*** *the ligands alone bring about a slight inhibition (34% at saturation)* (Reprinted from *J. Theor. Biol.*, **68**, MAZAT J.P. et al., Double-site enzymes and squatting. A study of the regulation by one or several ligands binding at two different classes of site, 365. © (1977) with permission from Elsevier)



(Reprinted from *J. Theor. Biol.*, **68**, MAZAT J.P. *et al.*, 365. © (1977) with permission from Elsevier)

Fig. 13.40 Squatting: reversal of effects

*The dissociation constants are indicated on the figures in arbitrary units. The T form is completely inactive. (**a**) the ligands X and Y are slight activators. When they are together they induce an important inhibition – (**b**) the ligands X and Y are strong inhibitors (99% inhibition at saturation). In the presence of both however, the activity of the enzyme can be strongly increased*

If the isomerisation rate constant k_1 is small compared to the substrate binding rate k_3 to the form E*, the latter will have a certain chance of binding a substrate molecule before returning to the initial conformation E. The probability of such an event increases with the substrate concentration, and therefore results in an apparent cooperativity.

The above mecanisms were described by WHITEHEAD (1970), FRIEDEN (1970) and RICARD (1974). RICARD's model rests on three postulates:

- the enzyme exists in two conformations at equilibrium,
- substrate binding induces the appearance of a third conformation,
- only one of the two initial conformations is stabilised by the reaction product(s).

An important aspect of the behaviour of "mnemonic" monomeric enzymes is that *the cooperation* **between the two enzyme states is only catalytic and cannot appear in substrate or substrate analog binding experiments.** The point is important as it allows the distinction between **memory** and **cooperativity**.

13.12.1. CASE OF A MNEMONIC ENZYME WITH ONE SUBSTRATE AND ONE PRODUCT

13.12.1.1. KINETIC BEHAVIOUR

The case is illustrated by Fig. 13.41. In the figure, the step corresponding to the liberation of the product P is presented as irreversible. This means that the reaction is analysed in the initial stationary phase conditions, in which the concentration of the product P is negligible. The rate equation is:

$$v = \frac{\delta \dfrac{1}{[A]} + \varepsilon}{\alpha \dfrac{1}{[A]^2} + \beta \dfrac{1}{[A]} + \gamma}$$

with:

$$\alpha = (k_4 + k_{-4})[(k_{-1} + k_{-3})(k' + k_2) + kk_2]$$

$$\beta = (k_1k_4 + k_3k_{-4})(k + k' + k_2)$$

$$\gamma = k_1k_3(k + k' + k_2)$$

$$\delta = kk_2(k_1k_4 + k_3k_{-4})$$

$$\varepsilon = kk_1k_2k_3$$

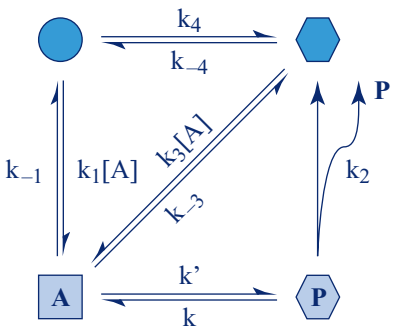


Fig. 13.41 Mnemonic transition for a monomeric enzyme with one substrate and one product
(Reproduced from *Eur. J. Biochem.*, **49**, RICARD J. et al., Regulatory behaviour of monomeric enzymes 1 the mnemonic enzyme, 195. © (1974) with permission of Blackwell Publishing Ltd)

Figure 13.42 illustrates the form of the LINEWEAVER-BURK plot for different rate constant values, k_1 and k_3 being the most determinant:

- if $k_3 > k_1$: the enzyme displays apparent cooperativity,
- if $k_3 = k_1$: the enzyme displays Michaelian behaviour,
- if $k_3 < k_1$: the enzyme displays apparent anti-cooperativity.

The simulations also show that the rate constant of substrate binding to the initial enzyme conformation, k_4, influences the apparent degree of cooperativity measured by HILL's number for instance, but does not affect the equilibrium constant between the two enzyme conformations.

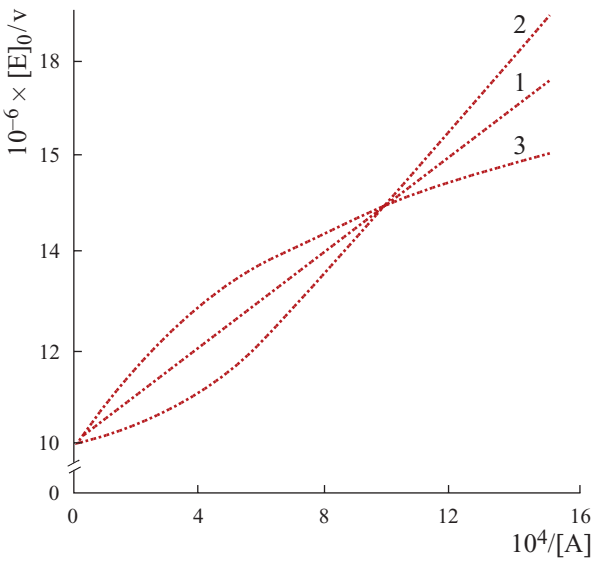


Fig. 13.42 Simulation of mnemonic enzyme behaviour for different rate constant values

LINEWEAVER-BURK plot. 1: Michaelian behaviour; 2: cooperative behaviour; 3: anti-cooperative behaviour (Reproduced from *Eur. J. Biochem.*, **49**, RICARD J. et al., Regulatory behaviour of monomeric enzymes 1 the mnemonic enzyme, 195. © (1974) with permission of Blackwell Publishing Ltd)

13.12.1.2. THERMODYNAMIC ASPECTS

The fact that apparent cooperativity or anti-cooperativity is observed is due to the thermodynamic parameters of the reaction. The activation energy associated to the constants k_1 and k_3 can be decomposed into two contributions. The constant k_1 is the sum of a contribution corresponding to the activation energy associated to simple substrate binding, $\Delta G_s^{\neq}$, which does not take into account the conformational change that accompanies it, and the transconformation contribution $\Delta G_T^{\neq}$ which is the conformational change activation energy:

$$\Delta G_1^{\neq} = \Delta G_s^{\neq} + \Delta G_T^{\neq}$$

In the same way, the constant k_3 is the sum of two corresponding terms:

$$\Delta G_3^{\neq} = \Delta G_s^{\neq} + \Delta G_{T'}^{\neq}$$

The comparison of the two above relations allows the following predictions to be made:

- if $\Delta G_{T'}^{\neq} < \Delta G_T^{\neq}$, k_3 will be larger than k_1, and there will be apparent cooperativity;
- if $\Delta G_{T'}^{\neq} > \Delta G_T^{\neq}$, k_3 will be smaller than k_1, and there will be apparent anti-cooperativity;
- if $\Delta G_{T'}^{\neq} = \Delta G_T^{\neq}$, the enzyme's behaviour will be Michaelian. The degree of co-operativity can be expressed by the relation:

$$\tau = \frac{2k_4}{k_1 - k_3^2}(k_3 - k_1)$$

Figure 13.43 shows the sign and the variation of τ as a function of $\Delta G_T^{\neq} - \Delta G_{T'}^{\neq}$.

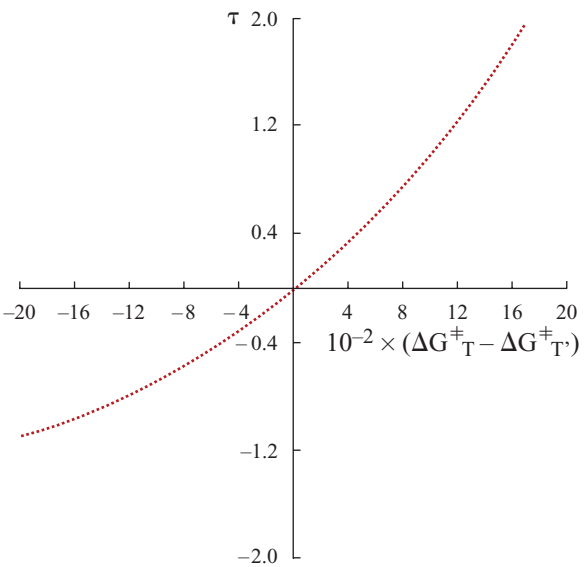


Fig. 13.43 Effects of the activation energies associated with the conformational changes on the cooperativity or anti-cooperativity of a mnemonic enzyme with one substrate and one product

(Reproduced from *Eur. J. Biochem.*, **49**, RICARD J. et al., Regulatory behaviour of monomeric enzymes 1 the mnemonic enzyme, 195. © (1974) with permission of Blackwell Publishing Ltd)

13.12.2. CASE OF A MNEMONIC ENZYME WITH TWO SUBSTRATES AND TWO PRODUCTS

The situation is obviously more complex. It was described by RICARD in the case of an ordered enzymatic mecanism in which the substrate A binds before the substrate B, the liberation of the products happens in the order P, then Q. Figure 13.44 illustrates the reactional diagram.

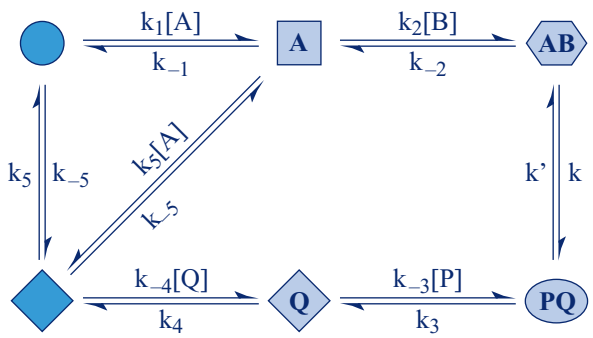


Fig. 13.44 Mnemonic transition for a monomeric enzyme with two substrates and two products in the case of a sequential mecanism

(Reproduced from *Eur. J. Biochem.,* **49**, RICARD J. et al., Regulatory behaviour of monomeric enzymes 1 the mnemonic enzyme, 195. © (1974) with permission of Blackwell Publishing Ltd)

The rate equation is:

$$v = \frac{\delta[B]\frac{1}{[A]} + \varepsilon[B]}{(\alpha_1 + \alpha_2[B])\frac{1}{[A]^2} + (\beta_1 + \beta_2[B])\frac{1}{[A]} + (\gamma_1 + \gamma_2[B])}$$

with:

$$\alpha_1 = k_4(kk_3 + k_{-2}k_3 + k'k_{-2})(k_{-1} + k_{-6})$$

$$\alpha_2 = kk_2k_3k_4(k_5 + k_{-5})$$

$$\beta_1 = k_4(kk_3 + k_{-2}k_3 + k'k_{-2})(k_1k_5 + k_1k_{-6} + k_{-1}k_6 + k_{-5}k_6)$$

$$\beta_2 = k_2(k_1k_5 + k_{-5}k_6)(kk_4 + k'k_4 + k_3k_4 + kk_3) + kk_1k_2k_3k_4$$

$$\gamma_1 = k_1k_4k_6(k'k_{-2} + k_{-2}k_3 + kk_3)$$

$$\gamma_2 = k_1k_2k_6(k_3k_4 + k'k_4 + kk_4 + kk_3)$$

$$\delta = kk_2k_3k_4(k_1k_5 + k_{-5}k_6)$$

$$\varepsilon = kk_1k_2k_3k_4k_6$$

In this case, only the substrate A is important to determine the kinetics, since the substrate B binds to only one enzyme conformation. In the LINEWEAVER-BURK plot, a deviation from linear behaviour will only be observed in the representation of 1/v as a function of 1/[A]. The rate variation as a function of [B] remains Michaelian. In this case, only the values of k_1 and k_6 determine the apparent cooperativity or anti-cooperativity (Fig. 13.45 opposite).

As previously, the cooperativity index τ depends on the rate constants following the relation:

$$\tau = \frac{2k_5}{k_1 - k_6^2}(k_6 - k_1)$$

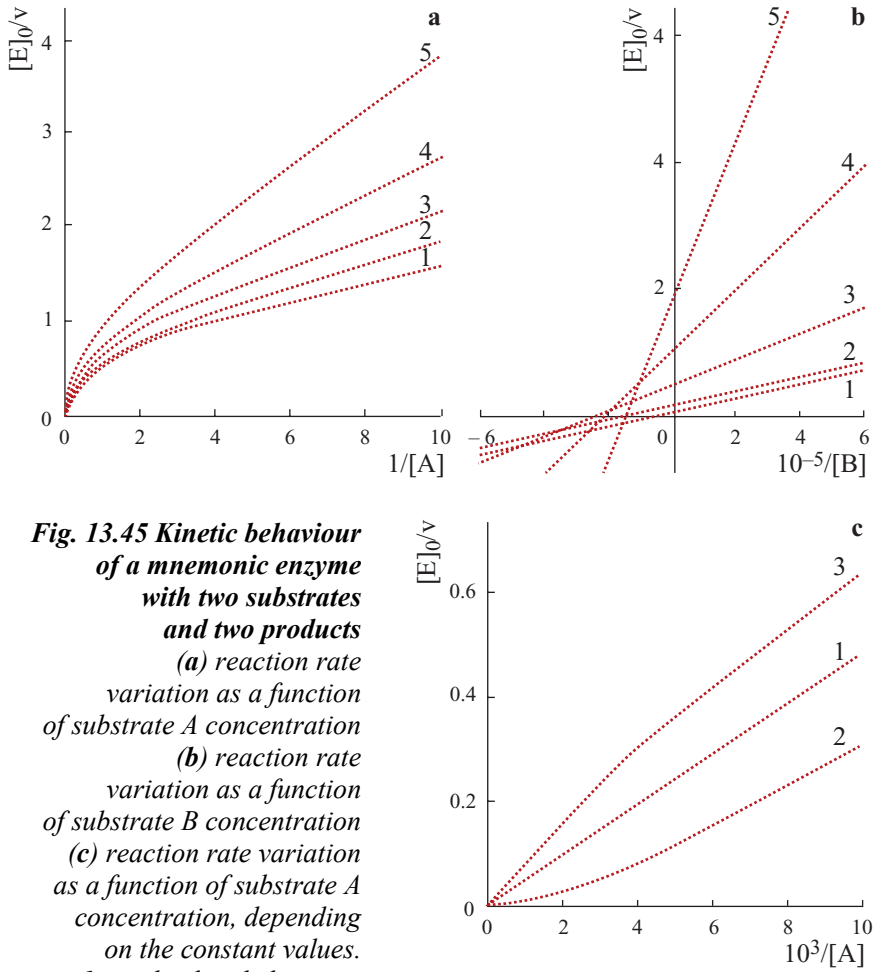


Fig. 13.45 Kinetic behaviour of a mnemonic enzyme with two substrates and two products *(**a**) reaction rate variation as a function of substrate A concentration (**b**) reaction rate variation as a function of substrate B concentration (**c**) reaction rate variation as a function of substrate A concentration, depending on the constant values. 1: michaelien behaviour; 2: cooperative behaviour; 3: anti-cooperative behaviour. LINEWEAVER-BURK plot.*
(Reproduced from *Eur. J. Biochem.*, **49**, RICARD J. et al., Regulatory behaviour of monomeric enzymes l the mnemonic enzyme, 195. © (1974) with permission of Blackwell Publishing Ltd)

The deviation from MICHAELIS law is linked to the activation thermodynamic parameters and:

- if $\Delta G_{T'}^{\neq} > \Delta G_{T}^{\neq}$, k_6 will be smaller than k_1, and there will be apparent anti-cooperativity;
- if $\Delta G_{T'}^{\neq} < \Delta G_{T}^{\neq}$, k_6 will be larger than k_1, and there will be apparent cooperativity;
- if $\Delta G_{T'}^{\neq} = \Delta G_{T}^{\neq}$, the enzyme's behaviour will be Michaelian.

Figure 13.46 below shows the sign and the variation of τ as a function of the value of $\Delta G_{T'}^{\neq} - \Delta G_{'}^{\neq}$.

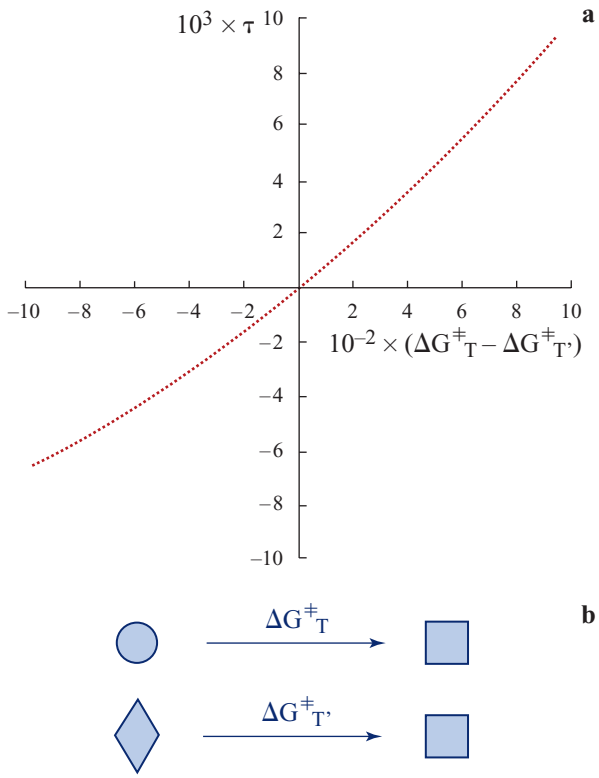


Fig. 13.46 (a) effects of the activation energies corresponding to the conformational changes in the case of mnemonic enzymes with two substrates and two products (b) definition of the activation energies associated to the conformational changes

(Reproduced from *Eur. J. Biochem.,* **49**, RICARD J. et al., Regulatory behaviour of monomeric enzymes 1 the mnemonic enzyme, 195. © (1974) with permission of Blackwell Publishing Ltd)

13.12.3. THE REACTION PRODUCT ACTS AS AN EFFECTOR

It has been previously remarked that the substrate B has no influence on the apparent cooperativity or anti-cooperativity. However, it is interesting to note that the second product Q acts as a reaction effector. Simulations show that there will indeed be cooperativity if:

$$k_6 > k_1(1 + k_4[Q])$$

On the contrary, there will be anti-cooperativity if:

$$k_6 < k_1(1 + k_4[Q])$$

As a result:

- when there is cooperativity in the absence of Q ($k_6 > k_1$), the presence of Q increases the cooperativity or reverses it;
- if there is anti-cooperativity in the absence of Q ($k_6 < k_1$), the presence of this product will increase the anti-cooperativity, an effect which is of interest from a physiological point of view;

- if $k_6 = k_1$, the system remains Michaelian as long as the concentration of the product Q remains negligeable. When the concentration increases however, it induces the appearance of anti-cooperativity, which is also of interest for metabolic regulation.

The simulation of these regulatory effects is presented in Fig. 13.47.

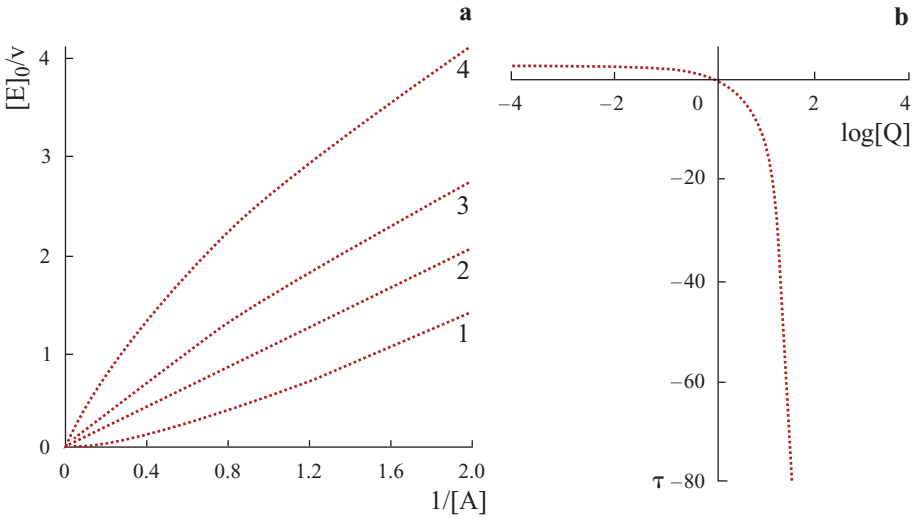


Fig. 13.47 (a) simulation of the inversion of cooperativity by the reaction product Q. The concentrations of product Q are respectively 10^{-5} (1), 0.99 (2), 2 (3), 4 (4). The numerical values are given in arbitrary units (b) effect of the product Q on the degree of cooperativity

(Reproduced from *Eur. J. Biochem.,* **49**, RICARD J. et al., Regulatory behaviour of monomeric enzymes 1 the mnemonic enzyme, 195. © (1974) with permission of Blackwell Publishing Ltd)

An example of mnemonic enzymes, which has been particularly well-documented is that of the hexokinase of wheat germ studied by the RICARD group (MEUNIER et al., 1974). This monomeric enzyme catalyses the phosphorylation of glucose by ATPMg to give glucose-6-phosphate. The reaction proceeds according to an ordered mechanism, in which the hexose binds before the ATPMg. The kinetic studies reveal an apparent cooperativity for glucose, while the variation of reaction rate as a function of ATPMg obeys MICHAELIS law. The profiles observed in the LINEWEAVER-BURK plot (Fig. 13.48 below) are analogous to those represented on Fig. 13.45 (a and b). In addition, glucose-6-phosphate, a reaction product, accentuates the cooperativity.

However, it should be noted that the studies by dialysis at the equilibrium of glucose substrate binding show no cooperativity, which is conform with the model predictions. Also, the cooperativity observed during kinetic studies can be abolished by weak concentrations of denaturing agents, confirming that the observed effects are the consequence of conformational equilibria; the slowness of the conformational relaxation would be due to a certain rigidity of the enzyme molecule.

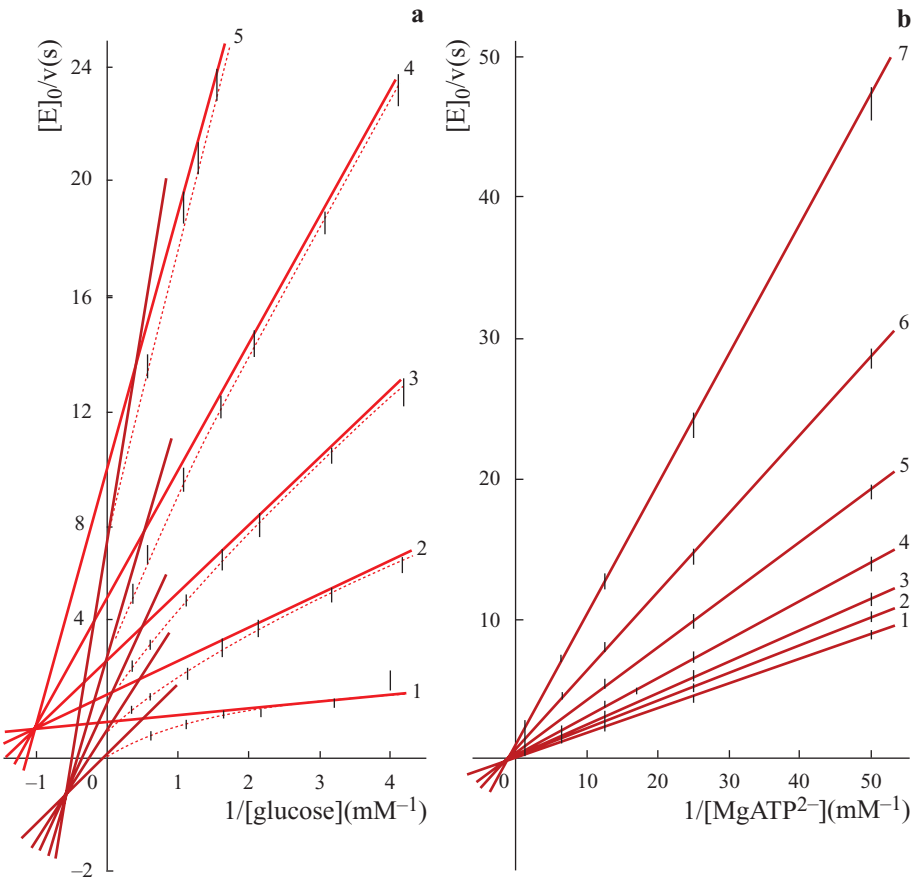


Fig. 13.48 Phosphorylation of glucose by the hexokinase of wheat germ in the presence of ATPMg

*(**a**) reaction rate variation as a function of glucose concentration (LINEWEAVER-BURK plot). The concentrations of ATPMg are respectively 1 mM (1), 0.16 mM (2), 0.08 mM (3), 0.04 mM (4), 0.02 mM (5) – (**b**) variation of the reaction rate as a function of ATPMg concentration. The concentrations of glucose are respectively 32 mM (1), 8 mM (2), 4 mM (3), 2 mM (4), 1 mM (5), 0.5 mM (6), 0.2 mM (7)* (Reproduced from *Eur. J. Biochem.*, **49**, MEUNIER J.C. et al., Regulatory Behavior of Monomeric Enzymes: 2. A Wheat-Germ Hexokinase as a Mnemonical Enzyme, 209. © (1974) with permission of Blackwell Publishing Ltd)

13.13. REGULATION THROUGH PROTEIN-PROTEIN INTERACTION

Different cases of enzymatic activity regulation by small molecules binding to specific regulatory sites have been examined previously. Other mechanisms of activity regulation exist however for several enzymes, which bring into play a specific and transitory association with other proteins, which themselves possess or not an en-

zymatic activity. Amongst the best-known systems are the lipase-colipase system, the regulation of ornithine transcarbamylase by arginase in some organisms, the cAMP dependent protein kinase and the pleiotropic regulation by the calmodulin-calcium complex. It is also the case of protein inhibitors of proteases, which are treated in the following chapter, some of which form covalent associations, others non-covalent ones with the target protease.

13.13.1. THE LIPASE-COLIPASE SYSTEM

The lipase-colipase system offers a particularly interesting example of regulation *via* the non-covalent association of two proteins. The pancreatic lipase or triacylglycerol hydrolase plays a key role in the digestion of lipids, by hydrolysing triglycerides into diglycerides, and then monoglycerides and free fatty acids. DESNUELLE and his collaborators from Marseille had a pioneering role in enzymatic studies of the pancreatic lipase, which they began as early as 1951 by in vitro experiments. The studies, continued by the groups of VERGER, SARDA and CHAPUS, now benefit from the structural knowledge that allows the enzyme's mechanism of action to be better understood.

Lipases exist in different plant and animal organisms and in microorganisms. They belong to the family of serine hydrolases. Their active site comprises a catalytic triad, Ser, His, Asp or Glu, the existence of which was confirmed by recent crystallographic data. The essential role of these amino acids was corroborated by site-directed mutagenesis in the pancreatic lipase, the hepatic lipase and the lipoprotein lipase. The pancreatic lipase's mode of action differs slightly however from that of the classic serine esterases by several characteristics. In contrast to classic serine esterases, the pancreatic lipase is weakly active on monomeric substrates, but displays its full activity on emulsified substrates. It is remarkable that the enzyme acts at the lipid-water interface and presents a high catalytic efficiency. In addition, in vivo, to avoid the inhibitor action of biliary salts, the presence of a cofactor, the colipase, is required. Colipase is a small pancreatic protein that forms with the lipase a non-covalent complex of stoechiometry 1:1 and whose function is to anchor the lipase at the lipidic interface covered with biliary salts.

The three-dimensional structures of several lipases have been resolved, including a mammalian one, the human pancreatic lipase and two from microorganisms. Although these lipases do not display sequence homology, they have in common a core of β sheets, which is also present in the structure of esterases, the α/β fold of hydrolases. In three lipases of known structure, the active site is inaccessible to the solvent because of the presence of a lid, which, in the pancreatic lipase, is formed of an amphiphilic helix. The human pancreatic lipase, which contains 449 amino acids is folded into two domains (WINCKLER et al., 1990). The large N-terminal domain comprises residues 1–335: it is typical of a α/β structure, dominated by the large central β sheet. The C-terminal domain is of the β sandwich type formed by

four anti-parallel segments (Fig. 13.49 below). The limited proteolysis by chymotrypsin of the Phe335-Ala336 bond allows the separation of the two domains. The N-terminal domain contains the catalytic groups Ser152, His263 and Asp176, completely inaccessible in the enzyme's structure. The C-terminal domain isolated by proteolysis is capable of hydrolysing small substrates, but the presence of an interface induces no activation.

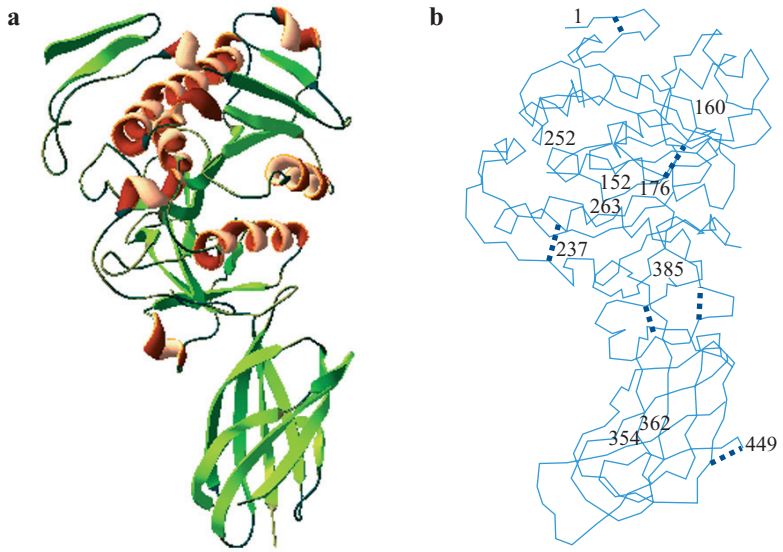


Fig. 13.49 Structure of the pancreatic lipase
*(**a**) rat – (**b**) human lipase. The six disulphide bridges are indicated by dots* (Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **343**, WINKLER F.K. et al., 771. © (1990))

The three-dimensional structures of procolipase, of the lipase-procolipase complex, and of the lipase-procolipase complex in the presence of mixed micelles (1,2-didodecanoyl-*sn*-3-glycerophosphoryl choline and taurodeoxycholate) were obtained at 3 Å resolution by the CAMBILLAU group from Marseille. The procolipase is a small protein of 95 amino acids; the N-terminal pentapeptide, which is not necessary for activity in vitro is cleaved during the activation process. The protein has a structure in three fingers stabilised by five disulphide bridges (Fig. 13.50a opposite). A fourth shorter finger stands out from the core of the molecule. Although many residues adopt the β conformation, they do not form regular β sheets. The global structure of the molecule resembles that of erabutoxin.

Procolipase binds exclusively to the C-terminal domain of lipase *via* the part of the molecule opposite to the fingers. The plane of the colipase is practically perpendicular to the plane of the C-terminal domain β sheet (Fig. 13.50b). Despite the numerous VAN DER WAALS contacts that the two molecules form, there are few interactions between the lipase and the colipase. This is in agreement with the association constant ($K_a = 2 \times 10^6\ M^{-1}$). The lipase-colipase complex corresponds more

to a complex of the type antigen-antibody where a surface complementarity exists, than to a proteinase-protein inhibitor complex.

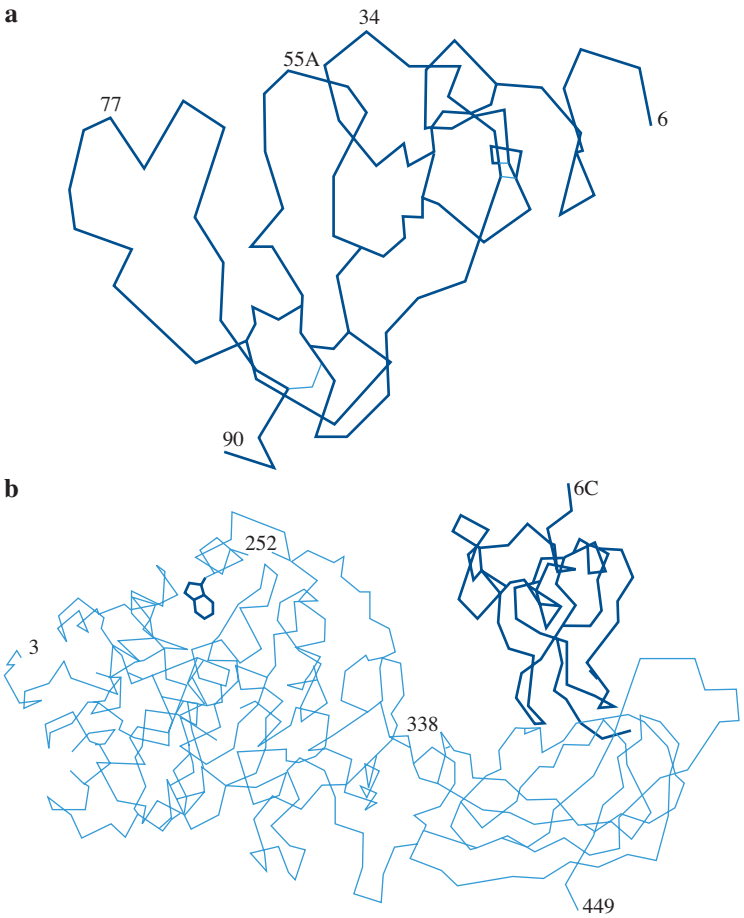


Fig. 13.50 (a) the procolipase structure
(b) the lipase-procolipase complex structure in the closed form
(Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **359**, VAN TILBEURGH H. et al., 159. © (1992))

The important conformational change that the lipase undergoes upon its association to a lipid-water interface was known from solution studies before the resolution of the three-dimensional structure. It also seemed obvious that the active site had to be uncovered to allow the substrate access. The structural data of the lipase-colipase complex in the absence and in the presence of mixed micelles allowed the comparison of the open and closed structures (Fig. 13.51 below). In the open structure, the helix forming the "lid" is partially unwound into two new helices and is turned around on the "body" of the molecule, considerably increasing the hydrophobicity of the active centre. The maximum displacement of the main chain, on the opening of the "lid" is 29 Å for Ile248. The combined movement of the "lid" and the turn β5

changes the environment of the catalytic triad. The active serine becomes completely solvent accessible and the hole of the oxyanion is formed at the very bottom of a hydrophobic crevice, perfectly adapted to the binding of a lipidic substrate.

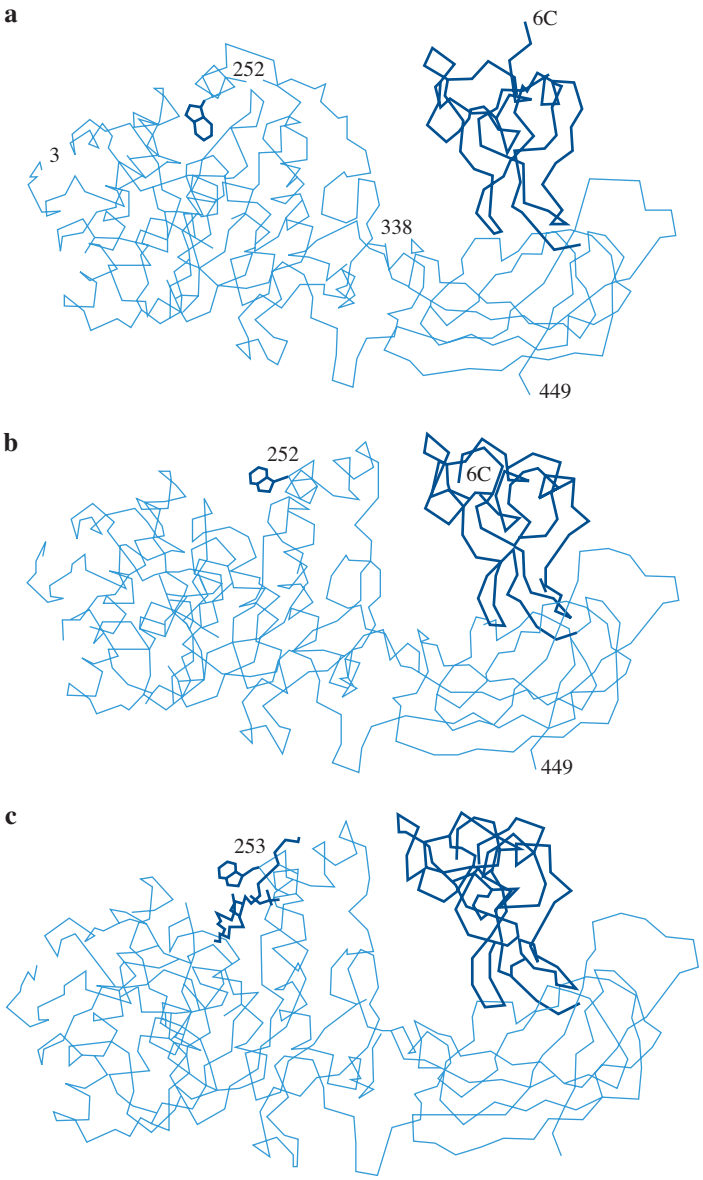


Fig. 13.51 Closed (a) and open (b) structures of the lipase-procolipase complex (c) open structure in the presence of a substrate

(Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **359**, VAN TILBEURGH H. et al., 159. © (1992))

In addition, the catalytic domain interacts with the procolipase by the intermediate of the lid, creating a long continuous hydrophobic plateau of more than 50 Å in length. Such a surface is capable of strongly interacting with the lipid-water interfaces covered by biliary salts, preventing their inhibitor effect. Figure 13.52, obtained by modelling, represents the interaction of lipids with the complex.

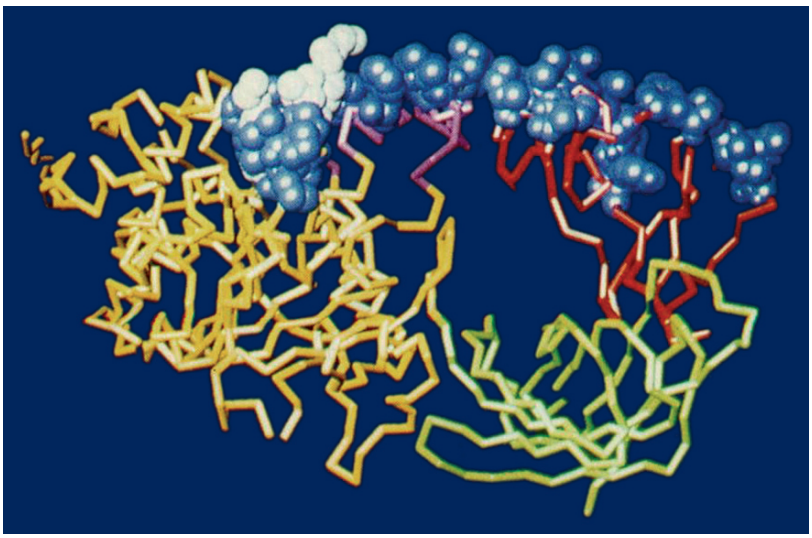


Fig. 13.52 Model of the hypothetical surface of lipid binding (*in white*) to the lipase-procolipase complex

The catalytic domain is in yellow*, the C-terminal domain in green and the colipase* in red*. The hydrophobic residues are represented as blue spheres* (Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **359**, VAN TILBEURGH H. et al., 159. © (1992))

On the described structural basis, a simplified model of the interfacial activation of the lipase is presented in Fig. 13.53, according to CARRIÈRE et al. (1994).

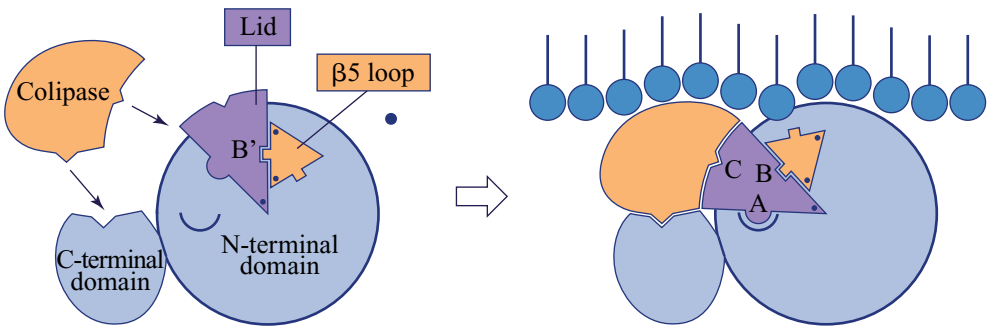


Fig. 13.53 Activation mechanism of the lipase

(Reprinted from *Protein Engineering, Design and Selection*, **7**, CARRIÈRE F. et al., Structure-function relationships in naturally occurring mutants of pancreatic lipase, 563. © (1994) by permission of Oxford University Press)

The activation of the lipase-colipase complex results from the conformational change of the two surface loops (the lid and the β5 loop) and the association of the colipase to the lipase. The lid domain interacts with the core of the protein (A), the β5 loop (B) and the N-terminal domain of the colipase (C). The interactions of the lid domain in the open (B) and closed (B') forms differ, although the tryptophan 252 of the lid domain is always involved in the interactions.

The presented example, well documented today, shows how an interaction between two proteins ensures the regulation of an enzymatic activity in vivo.

13.13.2. REGULATION OF ORNITHINE TRANSCARBAMYLASE FROM SACCHAROMYCES CEREVISIAE BY ARGINASE

Ornithine transcarbamylase catalyses the first reaction of the biosynthesis pathway of arginine. Arginase catalyses the cleavage of arginine into ornithine and urea. The activities are therefore at the origin of two different metabolic pathways (Fig. 13.54 opposite). When the intracellular concentrations of arginine and ornithine are high, the system constitutes a "futile cycle", leading to un-controlled ATP degradation. Such a danger is avoided in certain species such as *Saccharomyces cerevisiae* by the strong but reversible association of ornithine transcarbamylase and arginase.

The phenomenon, discovered by MESSENGUY and WYAME (1969), was studied at the molecular level by HENSLEY (1988). The association of the two enzymes completely abolishes the activity of ornithine transcarbamylase, but does not affect the activity of arginase. These two enzymes are trimeric proteins of molecular weights of 102 000 and 99 000, respectively.

Their association depends on effectors, which are in fact certain of the two enzymes' substrates (Fig. 13.55 below). In the presence of ornithine, these two enzymes associate strongly; the dissociation constant is 2.3×10^{-8} M. The association is also favoured by arginine binding to arginase. Carbamyl phosphate, substrate of ornithine transcarbamylase, reverses the effect of ornithine. Analytical ultra-centrifugation and electron microscopy experiments have established that the association occurs with a 1:1 stoechiometry. Several localised mutations in the arginase abolish its enzymatic activity without altering its capacity to associate with ornithine transcarbamylase.

In *Saccharomyces cerevisiae*, the two enzymes are found in the cytoplasm. In other yeast species, they are found in different cellular compartments. In this case the two enzymes have not acquired the capacity to associate, which constitutes an interesting example of evolutionary relation between cellular localisation and regulation properties.

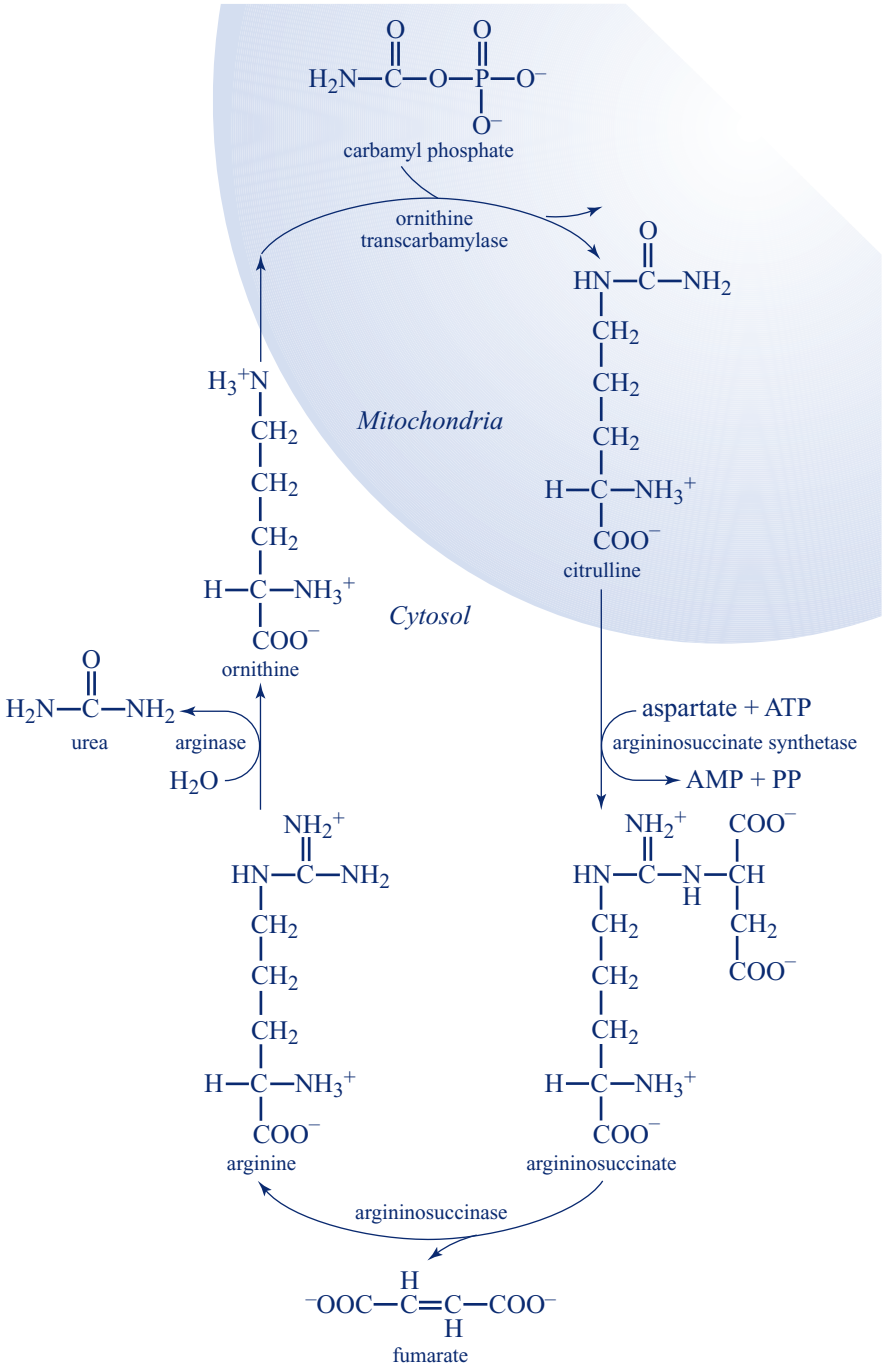


Fig. 13.54 Metabolic pathways of ornithine transcarbamylase and arginase

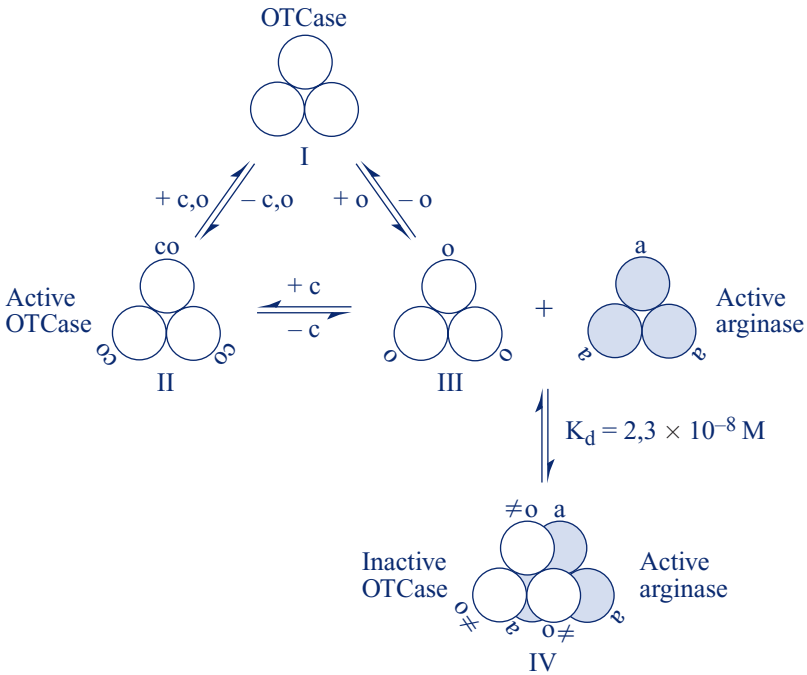
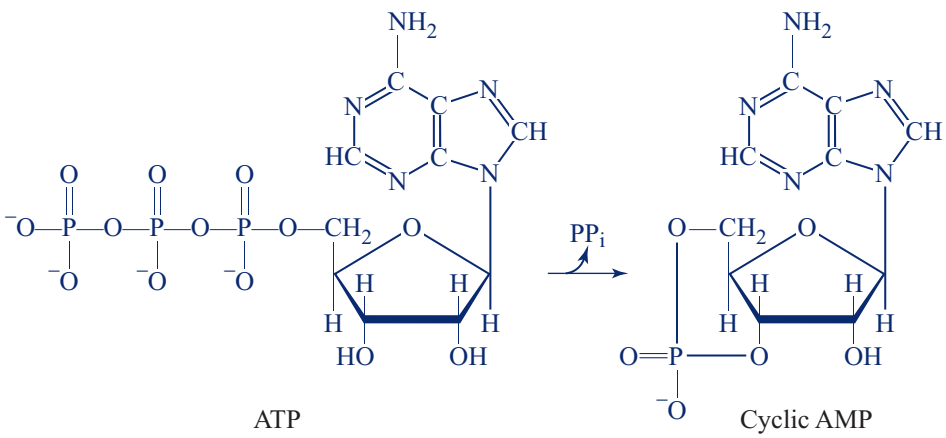


Fig. 13.55 Regulation of OTCase by arginine association

13.13.3. cAMP DEPENDENT PROTEIN KINASES

Cyclic AMP, cAMP, synthesised from ATP by adenylate cyclase plays an important and ubiquitary role in cellular regulation processes:



cAMP is involved, in particular, in the regulation of genetic expression and in the mechanism of action of numerous hormones. The nature of its action in the last case has resulted in it being named the "second messenger". When certain hormones such

as adrenaline or glucagon bind to their receptors, at the surface of their target cell membranes, the adenylate cyclase associated to these receptors synthesise cyclic AMP, which activates a protein kinase. The protein kinase, in turn, specifically phosphorylates proteins playing a regulatory role, in particular at the level of gene expression. The phosphorylation occurs on the hydroxyl group of serine, threonine or tyrosine residues. A consensus sequence around the phosphorylation site of the substrates has been identified: Arg-Arg-X-Ser(Thr)-Y, in which X is a small residue and Y is a bulky hydrophobic group.

The protein kinases are tetramers constituted of two catalytic subunits and two regulatory subunits. In this associated form, the protein kinase is inactive. Its activation by cAMP is a result of nucleotide binding to the two regulatory subunits, which brings about the dissociation of the ensemble:

$$R_2C_2 \text{ (inactive)} + 2 \text{ cAMP} \longleftrightarrow R_2(\text{cAMP})_2 + 2 \text{ C (active)}$$

The catalytic subunit has a molecular mass of 40 kDa. The molecular mass of the regulatory subunit is different depending on if it is a type I (49 kDa) or type II (56 kDa) protein kinase. These two types of enzymes also differ in their mechanism.

Type I protein kinases in the non-dissociated form bind two ATP molecules with a strong affinity (K_d = 10 μM). The binding decreases the regulatory subunits' affinity for cAMP. The activation of the enzymes by dissociation therefore requires fairly high concentrations of cAMP.

Type II protein kinases are capable of auto-phosphorylation. The reaction increases the regulatory subunits' affinity for cAMP. The activation by dissociation of the catalytic and regulatory subunits is therefore effective in the case of low concentrations of cAMP.

In the case of type I protein kinases, the reaction mechanism is therefore:

$$R_2C_2(ATP)_2 \text{ (inactive)} + 2 \text{ cAMP} \longleftrightarrow R_2(\text{cAMP})_2 + 2 \text{ C(ATP)}$$

in which the ATP is bound either to the associated form, or to the dissociated form. Bound to the dissociated form, it allows the phosphorylation of the target proteins. Bound to the associated form, it cannot react with the target protein. In addition, its binding shifts the association equilibrium (K_d = 50 nM) in favour of the associated form. **ATP behaves therefore both as a substrate and as an effector of the protein kinase.**

It appears that in the type I protein kinases, ATP binding involves both a part of the catalytic site and a particular region of the regulatory subunit. cAMP binding to the latter brings about the dissociation of subunits, allowing the correct AMP positioning in the catalytic site for phosphorylation. The activity of the catalytic subunit is also inhibited by its interaction with a natural protein thermo-stable inhibitor, which binds to a sequence of twenty residues close to the N-terminal extremity.

The crystallographic structure of the catalytic subunit of a mammal protein kinase is known with a resolution of 2.2 Å (MADHUSUDAN et al., 1994). The subunit is made up of two domains separated by a large crevice in which the ATP binding sites and the sites that recognise the consensus sequence of the peptide to be phosphorylated are found. The smaller N-terminal domain, is involved in ATP binding; it is predominantly constituted of large anti-parallel β sheets (Fig. 13.56), an unusual structure for a nucleotide binding motif.

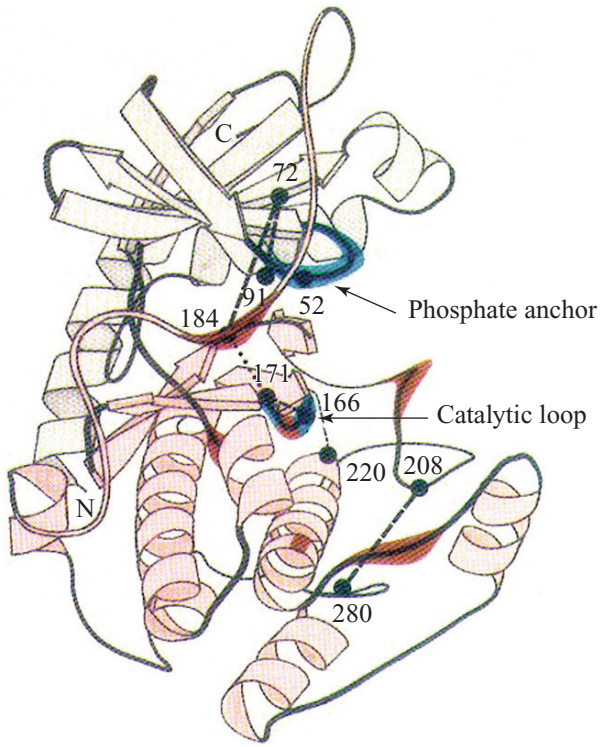


Fig. 13.56 Structure of the cyclic AMP dependent protein kinase subunit
(From *Science*, **253**, KNIGHTON D.R. et al., Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase, 414.
© (1991) reprinted with permission from American Association for the Advancement of Sciences)

The C-terminal domain is essentially constituted of α helices with a single β sheet, placed at the interface between the two domains. The C-terminal domain is involved in the binding of the peptide to be phosphorylated and in catalysis. It is remarkable that the sequence between residues 40 and 280 constitutes a catalytic core common to more than 100 protein kinases! This sequence includes residues Lys72, Glu91 and Asp184 that form the catalytic triad (Fig. 13.57 opposite). The reaction catalysed by the protein kinases proceeds through an ordered mechanism, in which the ATPMg binds before the peptide to be phosphorylated. ATPMg binding induces a movement of the two domains, bringing them closer together, which was confirmed by crystallographic studies of different complexes (MADHUSUDAN et al., 1994). Once the substrates are bound, the catalysis is rapid and the phosphorylated peptide dissociates instantaneously. The limiting step of the reaction is the dissociation of ADP. In the case of the cAMP dependent protein kinase, kinetic studies

suggested that Asp166, which is present in all kinases, is the most probable candidate to play the role of catalytic base. The catalysis occurs by direct transfer in-line with configuration inversion; no phosphoryl-enzyme intermediate is formed.

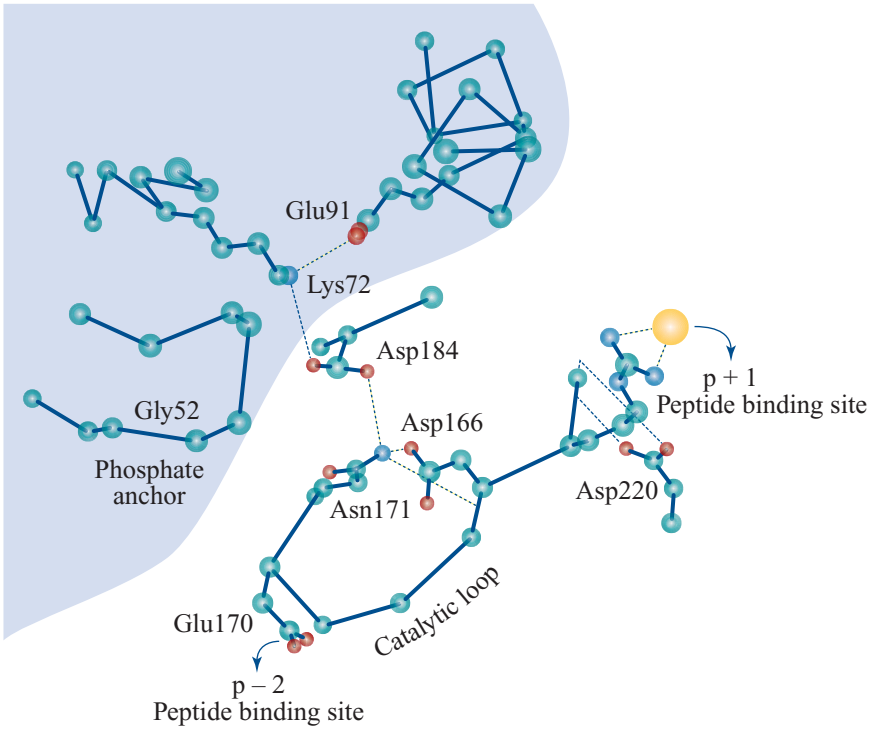


Fig. 13.57 Residues involved in the catalytic centre of the cyclic AMP dependent protein kinase (From *Science*, **253**, KNIGHTON D.R. et al., Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase, 414. © (1991) reprinted with permission from American Association for the Advancement of Sciences)

13.13.4. REGULATIONS BY INTERACTION WITH THE CALMODULIN-CALCIUM COMPLEX

Calmodulin is a small protein of molecular weight of the order of 15 000 depending on the species, capable of binding four calcium ions. It is found in variable concentrations in all tissues in eukaryotic organisms. The calcium-calmodulin complex is involved in the regulation of about thirty proteins identified to date, and the list is certainly not complete. The physiological processes in which calmodulin intervenes are numerous and varied: intracellular calcium concentration, contraction of smooth muscles, neurotransmitter metabolism, cellular division, secretion and motility phenomena, toxicity of certain microorganisms (for example *Bordetella pertussis*, the agent of whooping cough etc.). The described effects require the activation of certain enzymes such as adenylate and guanylate cyclases, the protein

kinases amongst those that assure the phophorylation of glycogen phosphatase, the phosphodiesterases of cyclic nucleotides, phospholipase A2, NO-synthase etc. The activation requires the specific interaction of the calmodulin-calcium complex with these enzymes or with other proteins such as ion transporters, motility proteins, transcription factors and cytoskeleton components. However in the particular case of adenylate cyclase activation of *Bordetella pertussis*, the presence of calcium is not necessary.

The crystallographic structures of several calmodulins from different species have been determined; that of the recombinant protein from vertebrates has been resolved to 1.7 Å by the group of CHATTOPADHYAYA et al., (1992). The protein has a dumb-bell structure (Fig. 13.58); it is constituted of two globular domains linked by a long α helix. The length of the whole protein is 65 Å. The globular domains have a diameter of 30 Å and are made up of three α helices and two small anti-parallel β sheets. The central α helix spreads over eight turns and is solvent exposed.

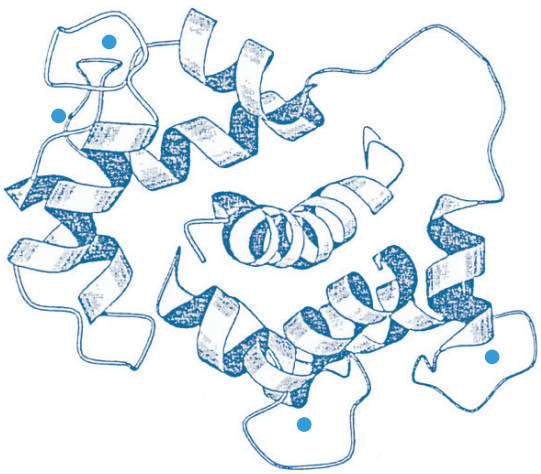
Fig. 13.58 Representation of the calmodulin molecule from the crystallographic structure of BABU* et al. *(1988) (PDB: 1CLL)
The residues of the hydrophobic pockets are shown in black

The four calcium binding sites are located in the globular domains. The average affinity of the sites is high, with a dissociation constant of Kd ~ 10^{-6} M. The affinity varies however with the degree of occupation of these sites. To explain all the phenomena observed on calcium binding, HAIECH *et al*. (1988) proposed a model. According to this model, the binding of the first calcium atom to the C-terminal domain increases the affinity of the second site of the same domain for calcium. The occupation of the two C-terminal domain sites induces a more important conformational change that has repercussions at the level of the N-terminal domain. The third calcium atom can then bind to one of the sites of the N-terminal domain increasing

the affinity of the fourth site for calcium. The properties of the modified forms of calmodulin in which tryptophan residues were introduced by site-directed mutagenesis at each calcium site confirmed the model described by HAIECH (KILHOFFER et al., 1992). The conformational changes that accompany the calcium binding decrease the length of the molecule by a few angströms, through inducing a curvature of the central helix. This is accompanied by an increase of a few percent of the helical percentage of the protein with the formation of a hydrophobic pocket at each globular domain.

Calmodulin interacts with its target proteins *via* the intermediate region of the central helix. Despite the very large diversity of these proteins and their variable accessibility (ATPase-Ca^{++}-Mg^{++} is a membrane protein), the interaction is always very strong with a dissociation constant of the nanomolar order. The peptide corresponding to the interaction zone of several target proteins has been isolated or synthesised. The peptides possess the common characteristic of presenting an organisation in amphiphile and basic helices. The crystallographic structure of several complexes of the peptides with calmodulin has been resolved. This is the case, in particular, of the peptide isolated from the kinase of the light chain of myosine, as well as wasp (mastoparan) and bee (mellitin) toxins. In the complex, the curvature of the central helix is very pronounced and the two globular domains come into contact. The interaction between the peptide and calmodulin is essentially of hydrophobic nature. Figure 13.59 shows how the peptide M13 of the kinase of the light chain of myosin, of helicoidal structure, is inserted in the hydrophobic canal, which passes across the centre of the calmodulin molecule (IKURA et al., 1992).

Fig. 13.59 Three-dimensional structure of calmodulin in the presence of the M13 peptide from the kinase of the myosin light chains
The grey dots represent calcium ions. The peptide is in the centre of the structure. The central helix has lost its α helical structure
(From *Science*, **256**, IKURA M. et al., Solution structure of a calmodulin-target peptide complex by multidimensional NMR, 632. © (1992) reprinted with permission from American Association for the Advancement of Sciences)



Calmodulin-M13 peptide complex

Calmodulin is the object of post-translational covalent modifications such as phosphorylation and trimethylation of the lysine residue in position 115. The post-translational modifications are supposed to be involved in the regulation of the proteolytic degradation of calmodulin.

All the aspects treated in this chapter show the wide variety of mechanisms of non-covalent regulation that living organisms dispose of to modulate enzymatic activity as a function of the requirements of the cell and the environmental constraints.

BIBLIOGRAPHY

SPECIALISED ARTICLES

ADAIR G.S. –1925– *J. Biol. Chem.* **63**, 529.

ALLEWELL N. –1989– *Annu. Rev. Biophys. Chem.* **18**, 71.

AUZAT I., GAWLITA E. & GAREL J.R. –1995– *J. Mol. Biol.* **34**, 13203.

BABU Y.S., BUGG C.E. & COOK W.J. –1988– *J. Mol. Biol.* **204**, 191.

BAILLON J., TAUC P. & HERVE G. –1985– *Biochemistry* **29**, 7182.

BLACK W.J. & WANG J.H. –1968– *J. Biol. Chem.* **243**, 5892.

BLANGY D., BUC H. & MONOD J. –1968– *J. Mol. Biol.* **31**, 13.

BARFORD D. & JOHNSON L.N. –1989– *Nature* **340**, 609.

BARFORD D. & JOHNSON L.N. –1992– *Protein Sci.* **1**, 472.

BETHEL M.R. & JONES M.E. –1969– *Arch. Biochem. Biophys.* **134**, 352.

BUC H. –1967– *Biochem. Biophys. Res. Commun.* **28**, 59.

CARRIERE F., THIRSTRUP K., BOEL E., VERGER R. & THIM L. –1994– *Protein Eng. Des. Sel.* **7**, 563-569.

CHAPUS C., ROVERY M., SARDA L. & VERGER R. –1988– *Biochimie* **70**, 1223.

CHATTOPADHYAYA R., MEADOR W.E., MEANS A.R. & QUIOCHO F.A. –1992– *J. Mol. Biol.* **228**, 1177.

CITRI N. –1973– *Adv. Enzymol. Rel. Areas Mol. Biol.* **37**, 397.

COVÈS J., NIVIÈRE V., ESCHENBRENNER M. & FONTECAVE M. –1993– *J. Biol. Chem.* **268**, 18604.

DE STAERKE C., VAN VLIET F., XI X.G., RANI C.S., LADJIMI M., JACOB A., TRINIOLLE F., HERVÉ G. & CUNIN R. –1995– *J. Mol. Biol.* **266**, 132.

ENGLAND P. & HERVÉ G. –1992– *Biochemistry* **31**, 9725.

ERIKSON S. & SJÖBERG B.M. –1989– Ribonucleotide reductase, in *Allosteric Enzymes*, G. HERVÉ ed., CRC Press, Boca Raton, Florida, 189–215.

EVANS P.R. –1992– Activity and allosteric regulation in bacterial phosphofructokinase, in *The Robert*, A. WELCH Foundation Conference on Chemical Research XXXVI: Regulation of proteins by ligands, Houston, Texas, 39–54.

EVANS P.R. & HUDSON P.J. –1979– *Nature* **279**, 500.

FELTER L., TAUC P., HERVÉ G., MOODY M.F. & VACHETTE P. –1995– *J. Mol. Biol.* **251**, 243.

FONTECAVE M., ELIASSON R. & REICHARD P. –1987– *J. Biol. Chem.* **262**, 12325–12331.

FONTECAVE M., ELIASSON R. & REICHARD P. –1989– *Proc. Natl Acad. Sci. USA* **86**, 2147.

FONTECAVE M., NORDLUND P., EKLUND H. & REICHARD P. –1992– *Adv. Enzymol.* **65**, 147.

FOOTE J., LAURITZEN A.M. & LIPSCOMB W.N. –1985– *J. Biol. Chem.* **260**, 9624.

FOOTE J. & LIPSCOMB W.N. –1981– *J. Biol. Chem.* **256**, 11428.

FRIEDEN C. –1970– *J. Biol. Chem.* **245**, 5788.

GOUAUX J.E., STEVENS R.C. & LIPSCOMB W.N. –1990– *Biochemistry* **29**, 7702.

GRIFFIN J.H., ROSENBUSCH J.P., BLOUT E.R. & WEBER K.K. –1973– *J. Biol. Chem.* **248**, 5057.

HAÏECH J., KILHOFFER M.C., CRAIG T.A., LUKAS T.J., WILSON E., GUERRA-SANTOS L. & WATTERSON D.M. –1989– *Adv. Exp. Med. Biol.* **269**, 43.

HENIS Y., & LEVITZKI A. –1980– *Eur. J. Biochem.* **112**, 59.

HENSLEY P. –1988– *Curr. Top. Cell. Regul.* **29**, 35.

HERVÉ G. –1989– Aspartate transcarbamylase from *Escherichia coli*, in *Allosteric Enzymes*, G. HERVÉ ed., CRC Press, 61-79.

HERVÉ G., MOODY M.F., TAUC P., VACHETTE P. & JONES P.T. –1985– *J. Mol. Biol.* **185**, 189.

HILL A.V. –1910– *J. Biol. Chem.* **63**, 493.

HILL A.V. –1913– *Biochem. J.* **7**, 471.

HONZATKO R.B., CRAWFORD J.L., MONACO H.L., LADNER J.E., EDWARDS B.F.P., EVANS D.R., WARREN S.G., WILEY D.C., LADNER R.C. & LIPSCOMB W.N. –1982– *J. Mol. Biol.* **160**, 219.

HSUANYU Y. & WEDLER F. –1987– *Arch. Biochim. Biophys.* **259**, 316.

HSUANYU Y. & WEDLER F. –1988– *J. Biol. Chem.* **263**, 4172.

IKURA M., CLORE G.M., GRONENBORN A.M., ZHU G., KLEE C.B. & BAX A. –1992– *Science* **256**, 632.

JOHNSON L. –1992– Allosteric regulation of glycogene phosphorylase, in *The Robert*, A. WELCH Foundation Conference on Chemical Research XXXVI: Regulation of proteins by ligands, Houston, Texas, 17–35.

JOHNSON L. & BADFORD D. –1990– *J. Mol. Biol.* **265**, 2409.

KE H., LIPSCOMB W.N., CHO Y. & HONZATKO R.B. –1988– *J. Mol. Biol.* **204**, 725.

KILHOFER M.C., ROBERTS D.M., ADIBI A.O., WATTERSON D.M. & HAIECH J. –1988– *J. Biol. Chem.* **263**, 17023.

KILHOFFER M.C., KUBINA M., TRAVERS F. & HAIECH J. –1992– *Biochemistry* **31**, 8098.

KNIGHTON D.R., ZHENG J., TEN EYCK L.F., XUONG N.H., TAYLOR S. & SOWADSKI J.M. –1991– *Science* **253**, 414.

KOSHLAND D.E., NÉMÉTHY G. & FILMER D. –1966– *Biochemistry* **5**, 365.

KRAUSE K.L., VOLZ K.W. & LIPSCOMB W.N. –1987– *J. Mol. Biol.* **193**, 527.

KUNDROT C.E. & EVANS P.R. –1991– *Biochemistry* **30**, 1478.

LADJIMI M.M. & KANTROWITZ E.R. –1988– *Biochemistry* **27**, 276.

LAURENT M., CHAFFOTTE A.F., ROUCOUS C., TENU J.P. & SEYDOUX F. –1978– *Biochem. Biophys. Res. Commun.* **80**, 646.

LAURENT M. & YON J.M. –1989– Yeast phosphofructokinase, in *Allosteric Enzymes*, Chap. 11, G. HERVÉ ed., CRC Press, Boca Raton, Florida.

LAZDUNSKI M., PETITCLERC C., CHAPPELET D. & LAZDUNSKI C. –1971– *Eur. J. Biochem.* **20**, 124.

LE BRAS G. & GAREL J.R. –1985– *J. Biol. Chem.* **260**, 13450.

LEGER D. & HERVÉ G. –1988– *Biochemistry* **27**, 4293.

LEVITZKI A.–1978– *Mol. Biol. Biochem. Biophys.* **28**, 1.

LIPSCOMB W.N. –1992– Activity and regulation in aspartate transcarbamylase, in *The Robert*, A. WELCH Foundation Conference on Chemical Research XXXVI: Regulation of proteins by ligands, 103–143.

MADUHUSAN, TRAFNY E.A., XUONG N.H., ADAMS J.A., TEN EYCK L.F., TAYLOR S.S. & SOWADKI J.M. –1994– *Protein Sci.* **3**, 176.

MAZAT J.P., LANGLA J. & MAZAT F. –1977– Double-site enzymes and squatting. A study of the regulation by one or several ligands binding at two different classes of site, in *J. Theor. Biol.* **68**, 365.

MAZAT J.P. & MAZAT F. –1986– *J. Theor. Biol.* **121**, 89.

MESSENGUY F. & WYAME J.M. –1969– *FEBS Lett.* **3**, 47.

MEUNIER J.C., BUC J., NAVARRO A. & RICARD J. –1974– *Eur. J. Biochem.* **49**, 209.

MONOD J., JACOB F. & CHANGEUX J.P. –1963– *J. Mol. Biol.* **6**, 306.

MONOD J., WYMAN J. & CHANGEUX J.P. –1965– *J. Mol. Biol.* **12**, 88.

MULLIEZ E., FONTECAVE M., GAILLARD J. & REICHARD P. –1993– *J. Biol. Chem.* **268**, 2296.

NORDLUND P. & EKLUND H. –1993– *J. Mol. Biol.* **232**, 123.

NORDLUND P., SJÖBERG B.M. & EKLUND H. –1990– *Nature* **345**, 593.

PENVERNE B. & HERVÉ G. –1983– *Arch. Biochem. Biophys.* **225**, 562.

PORTER R.W., MODEBE M.O. & STARK G.R. –1969– *J. Biol. Chem.* **244**, 1846.

RABIN B.R. –1967– *Biochem. J.* **102**, 22.

RICARD J., MEUNIER J.C. & BUC J. –1974– *Eur. J. Biochem.* **49**, 195.

RICARD J. –1989– Concepts and models of enzyme cooperativity, in *Allosteric Enzymes*, Chap. 1, G. HERVÉ ed., CRC Press, Boca Raton, Florida.

RUBIN M. & CHANGEUX J.P. –1966– On the nature of allosteric transitions: implications of non-exclusive ligand binding, in *J. Mol. Biol.* **21**, 265.

RYPNIEWSKI W.R. & EVANS P.R. –1989– *J. Mol. Biol.* **207**, 805.

SHIRAKIHARA Y. & EVANS P.R. –1988– *J. Mol. Biol.* **204**, 973.

SHIRMER T. & EVANS P.R. –1990– *Nature* **343**, 140.

STEVENS R., GOUAUX J.E. & LIPSCOMB W.N. –1990– *Biochemistry* **29**, 7691.

STEVENS R. & LIPSCOMB W.N. –1992– *Proc. Natl Acad. Sci. USA* **89**, 5281.

TAUC P., LECONTE C., KERBIRIOU D., THIRY L. & HERVÉ G. –1982– *J. Mol. Biol.* **155**, 155.

THIRY L. & HERVÉ G. –1978– *J. Mol. Biol.* **125**, 515.

VAN TILBEURGH H., SARDA L., VERGER R. & CAMBILLAU C. –1992– *Nature* **359**, 159.

VAN TILBEURGH H., EGLOFF M.P., MARTINEZ C., RUGANI N., VERGER R. & CAMBILLAU C. –1993– *Nature* **362**, 814.

VAN VLIET F., XI X.G., DE STAERCKE C., DE WANNEMAEKER B., JACOBS A., CHERFILS J., LADJIMI M.M., HERVÉ G. & CUNIN R. –1991– *Proc. Natl Acad. Sci. USA* **88**, 9180.

WEBER G. –1965– The binding of small molecules to proteins, in *Molecular Biophysics*, B. PULLMAN & M. WEISSBLUTH, Acad. Press, 369–396.

WEBER G. –1972– *Biochemistry* **11**, 864.

WHITEHEAD E. –1970– *Prog. Biophys*. **21**, 321.

WILD J.R., LOUGHREY-CHEN S.J. & CORDER T.S. –1989– *Proc. Natl Acad. Sci. USA* **86**, 46.

WINCKLER F.K., D'ARAY A. & HUNZIKER W. –1990– *Nature* **343**, 971.

WINKLER F.K., D'ARCY A. & HUNTZINGER W. –1990– *Nature* **343**, 771.

WYMAN J. –1964– *Adv. Prot. Chem.* **19**, 223.

XI X.G., VAN VLIET F., LADJIMI M.M., CUNIN R. & HERVÉ G. –1990a– *Biochemistry* **29**, 8491.

XI X.G., VAN VLIET F., LADJIMI M.M., DE WANNEMAEKER B., DE STAERCKE C., PIERARD A., GLANDSDORFF N., HERVÉ G. & CUNIN R. –1990b– *J. Mol. Biol.* **216**, 375.

XI X.G., VAN VLIET F., LADJIMI M.M., DE WANNEMAEKER B., DE STAERCKE C., GLANDSDORFF N., PIERARD A., CUNIN R. & HERVÉ G. –1991– *J. Mol. Biol.* **220**, 789.

XI X.G., DE STAERCKE C., VAN VLIET F., TRINIOLLES F., JACOBS A., STAS P., LADJIMI M.M., SIMON V., CUNIN R. & HERVÉ G. –1994– *J. Mol. Biol.* **242**, 139.

14 – REGULATION BY COVALENT MODIFICATION

The preceding chapter showed that the regulation of enzymatic activity can be carried out through a variety of mechanisms. Regulation involving covalent processes is even more diverse. These processes include various co- and post-translational events such as limited proteolysis and various chemical modifications (glycosylation, carboxylation, ADP-ribosylation etc.). They also involve mechanisms of enzyme phosphorylation-dephosphorylation, catalysed by specific enzymes, as a function of the cell's metabolic needs or in response to signals.

14.1. MODIFICATIONS BY LIMITED PROTEOLYSIS: ACTIVATION OF PRECURSORS

We refer the reader to Chap. 8, where these aspects were treated, but they are succinctly revisited below.

Modifications *via* specific and limited proteolysis include first of all the removal of the signal peptide, catalysed by a signal peptidase for proteins synthesised at the ribosomes associated with the rough endoplasmic reticulum. Furthermore, some enzymes exist as precursors or zymogens, which must be activated by additional proteolytic cleavage and the release of a propeptide. This is the case, for example, of serine proteases: trypsin, chymotrypsin, elastase, proteases of the blood clotting cascade etc. The propeptide is of very variable length depending on the zymogens. The simple cleavage of Arg15-Ile16 leads to the activation of chymotrypsinogen; the release of a hexapeptide in the N-terminal position transforms trypsinogen into trypsin. The propeptide of the α-lytic protease consists of 174 amino acids however, practically two thirds of the length of the active enzyme. The propeptide of carboxypeptidase Y has a length of 91 amino acids, that of the subtilisin 77.

The existence of zymogens constitutes a protection for organisms, avoiding deleterious proteolysis and allowing enzymatic activity only in very precise tissue or cellular locations.

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_15, © Springer-Verlag Berlin Heidelberg 2010

14.2. PROTEIN INHIBITORS OF PROTEASES

Amongst protein inhibitors of enzymes, the most widespread are protease inhibitors. Inhibitors of other enzymes are also known, including inhibitors of α-amylase, phospholipase A, protein kinases, desoxyribonuclease. They are relatively rare however compared to protease inhibitors.

The protein inhibitors of proteases are ubiquitous; they are present in numerous animal and plant tissues, as well as in micro-organisms. Generally, their physiological function is to intervene in the regulation of proteolytic activity of their target protease in order to prevent harmful proteolysis. It is generally admitted that the large number of protease inhibitors present in the plasma has the effect of regulating blood clotting and other proteolytic cascades such as complement binding, the production and elimination of hormones. The fact that individuals displaying a deficiency in α_1-antiprotease are subject to emphysema suggests a protecting role of the inhibitor against the proteolysis of pulmonary tissues by leucocyte elastase, considered as the target enzyme of the inhibitor. The presence of protein inhibitors in plants may be supposed to have a defensive function against the proteases of parasitic insects. However the precise physiological implications of these protein inhibitors are still very poorly known.

The majority of known and well-characterised protein inhibitors are directed against serine proteases. In the past few years, inhibitors against thiol proteases were discovered: however, we still know very few protein inhibitors directed against metalloproteases and aspartyl-proteases.

14.2.1. SERINE PROTEASE INHIBITORS

LASKOWSKI and KATO (1980) presented a classification regrouping these different inhibitors in distinct families.

14.2.1.1. α_2-MACROGLOBULINS

The α_2-macroglobulin family is a special one, since these inhibitors are not specific and can associate to all proteases and inhibit them. The α_2-macroglobulins are glycoproteins of high molecular weight. The human α_2-macroglobulin has a molecular weight of 720 000 and is constituted of four apparently identical polypeptide chains. α_2-macroglobulins inhibit all proteases *via* a "molecular trap" mechanism. The protease hydrolyses one or several sensitive peptide bonds in the critical region of the α_2-macroglobulin, which brings about a conformational change during which the protease finds itself "trapped" in the α_2-macroglobulin. The critical segment must be large enough to contain peptide bonds covering the specificity of a large number of proteases. One of the characteristics of this mechanism is that, in the complex, the catalytic activity of the protease is inhibited only for protein substrates of large

size. The proteolytic activity towards small substrates is not inhibited. Likewise, small protein inhibitors such as the pancreatic trypsin inhibitor (BPTI), may still bind to the complex and inhibit enzymatic activity, whilst the larger ones such as ovumucoid cannot inhibit it. Thus in the complex, the active site of the protease is accessible to small molecules, but for steric reasons large molecules cannot gain access. It seems that in vivo, the α_2-macroglobulin-protease complexes are very rapidly eliminated; their half-life is 10 min in man. α_2-macroglobulins may therefore play a central role in the regulation of protease levels in blood.

14.2.1.2. PROTEIN PROTEASE INHIBITORS THAT POSSESS A SPECIFIC CLEAVAGE SITE

These inhibitors can be regrouped into at least sixteen families, by sequence and topology similarities, and their mechanism of action. For eleven of them, the three-dimensional structure of several inhibitors is today known.

The majority of these inhibitors react with their target enzyme according to a standard mechanism. They are highly specific. At their surface, there is a peptide bond or "reactive site", which interacts specifically with the target protease and leads to a limited proteolysis between P_1 and P_1' sites (see Chap. 12); the inhibitor behaves like a true protein substrate. The global enzyme-inhibitor interaction mechanism including intermediates, whose existence has been proved, is written as follows:

$$E + I \longleftrightarrow EI \longleftrightarrow C \longleftrightarrow X \longleftrightarrow EI^* \longleftrightarrow E + I^*$$

In the above scheme E is the enzyme, I and I^* are the native and modified inhibitors respectively, EI and EI^* the non-covalent complexes that the enzyme forms with each form of the inhibitor, C a stable complex and X an intermediate with a long lifetime.

The value of k_{cat}/K_m for the hydrolysis of the peptide bond by the enzyme at neutral pH is very high, of the order of 10^4 to 10^6 $M^{-1} . s^{-1}$, whilst for substrates it is generally of the order of 10^3 $M^{-1} . s^{-1}$. For inhibitors however, the k_{cat} value is extremely small, by several orders of magnitude compared to that observed for true substrates. On the other hand, the association constant is very high, of the order of 10^7 to 10^{13} M^{-1}, which explains the value of the k_{cat}/K_m ratio. The stable complex forms very rapidly and dissociates very slowly. The hydrolysis constant $K_{hyd} = (I)/(I^*)$ is close to one at neutral pH.

This group of inhibitors comprises small proteins (or one domain in the inhibitors with several functional domains), of 29 to 190 amino acids. The inhibitors displaying a "canonical" conformational loop all possess a binding loop that has a defined conformation, but their global structure differs. Serpins form a super-family of homologous glycoproteins of large size comprising about 400 amino acids. Like the inhibitors that have a canonical conformational loop, these proteins interact with their target protein *via* an exposed binding loop. The complexes are however transitory and collapse with the liberation of a cleaved form, of different structure and stability.

Inhibitors with a "canonical" conformational loop

Amongst these protein inhibitors of serine proteases, are the KUNITZ type inhibitors, such as the pancreatic trypsin inhibitor (BPTI) and the family of KAZAL type inhibitors including ovomucoids. Some ovomucoids possess different functional domains capable of associating with different proteases; the crocodile ovomucoid, for example, possesses four domains. The turkey inhibitor with three domains possesses two independent reactive sites, one for trypsin, the other for chymotrypsin, subtilisin and elastase. The family of STI-KUNITZ type inhibitors are also in the category of inhibitors with a canonical conformational loop, with the soybean trypsin inhibitor, the SSI type inhibitors such as the subtilisin inhibitor of *Streptomyces*, the inhibitors of potatoes with the PI-1 (to which eglin belongs) and PI-2 families, the chelonianin family, the BOWMAN-BIRK type inhibitors, the family of trypsin inhibitors from squash seeds.

The three-dimensional structure of several of these inhibitors was resolved either by X-ray crystallography or by NMR and, for several of them, that of their complex with the target protein is known (Fig. 14.1).

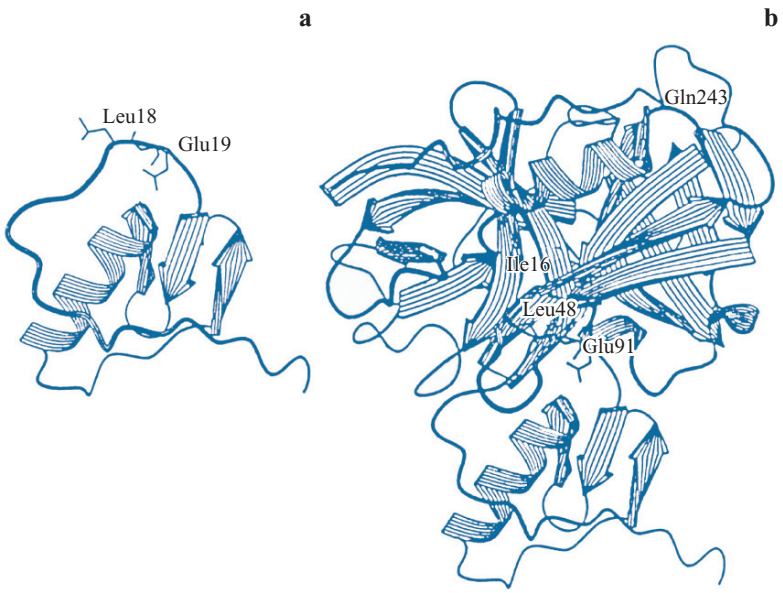


Fig. 14.1 (a) structure of the third domain of turkey ovomucoid, showing the reactive site Leu18-Glu19 in P1–P1′– (b) structure of the complex of leucocyte elastase with the third domain of turkey ovomucoid

(Reproduced from *Eur. J. Biochem.,* **204**, BODE W. & HUBER R., Natural proteinase inhibitors and their interaction with, 433. © (1992) with permission of Blackwell Publishing Ltd)

These inhibitors have in common a canonical conformational binding loop but, between these different families, the rest of the structure is completely different. The loop has a flat form that adjusts to the active centre of the target protease. The

loop possesses a characteristic conformation from P3 to P3′ with a small anti-parallel β segment in P3 (except for BPTI), a polyproline II type structure in P2 and P1′, and a conformation close to a $helix_{3,10}$ in P1, and parallel β segments in P2′ and P3′. The interaction loop is stabilised by residues on either side of its reactive site, which interact with residues of the hydrophobic centre.

In inhibitor complexes with the target trypsin, the reactive site loop of the inhibitor interacts with the residues of the enzyme's catalytic centre, in the same way as protein substrates. The inhibitor reactive site is close to catalytic groups of the enzyme, the carbon of the carbonyl in P1 is in contact with the O^{γ} of the serine, the carbonyl group introduces itself in the oxyanion hole.

In several crystallographic complexes, the reactive peptide bond stays intact but undergoes a deformation, the oxygen of the carbonyl being slightly projected out of the plane. The carbon is not completely trigonal, but distorted, halfway towards a tetrahedral configuration. It had been suggested that the serine O^{γ} of the active centre was the cause of this tetrahedral distortion. This hypothesis was abandoned however, since anhydrotrypsin in which the catalytic serine has been modified into dehydroalanine results in the same geometry. NMR results also lead to excluding the existence of a tetrahedral intermediate, although this remains under discussion. From crystallographic data, these complexes resemble MICHAELIS complexes that have been subjected to a distortion. This was observed in particular in the complex of trypsin with BPTI, in the subtilisin-eglin complex, but not in complexes that the ovomucoid forms. In complexes of specific inhibitors with chymotrypsin, the interaction of P1 with the S1 specificity site is the most important from an energetic point of view. In fact, as for the substrates, the P1 site confers its specificity to the inhibitor, Lys and Arg for trypsin, Trp, Tyr or Phe for chymotrypsin. Consequently, the substitution of a residue in P1 brings about a change in the specificity of the inhibitor. The interactions with the S1 site are less important in subtilisins, while the adjustment of the P4 residue becomes particularly important. Other than these primary interactions, the inhibitors contract additional interactions with their target protease. The fit between the active site of the enzyme and the reactive loop of the inhibitor is optimised.

Figure 14.2 below illustrates in a schematic way the mode of inhibition of a protein inhibitor with a canonical conformational loop in the complex between trypsin and the third domain of ovomucoid.

In the majority of inhibitors that obey the standard mechanism, the reactive loop is maintained by at least one disulphide bridge, which could contribute to stabilising the molecule upon conversion of the intact inhibitor to the modified inhibitor. Figure 14.3 below shows the localisation of the disulphide bridges in relation to the cleavage site in a few of these inhibitors. However a strong non-covalent association can make up for the absence of a disulphide bridge. Eglin, which possesses no disulphide bridges is a good inhibitor.

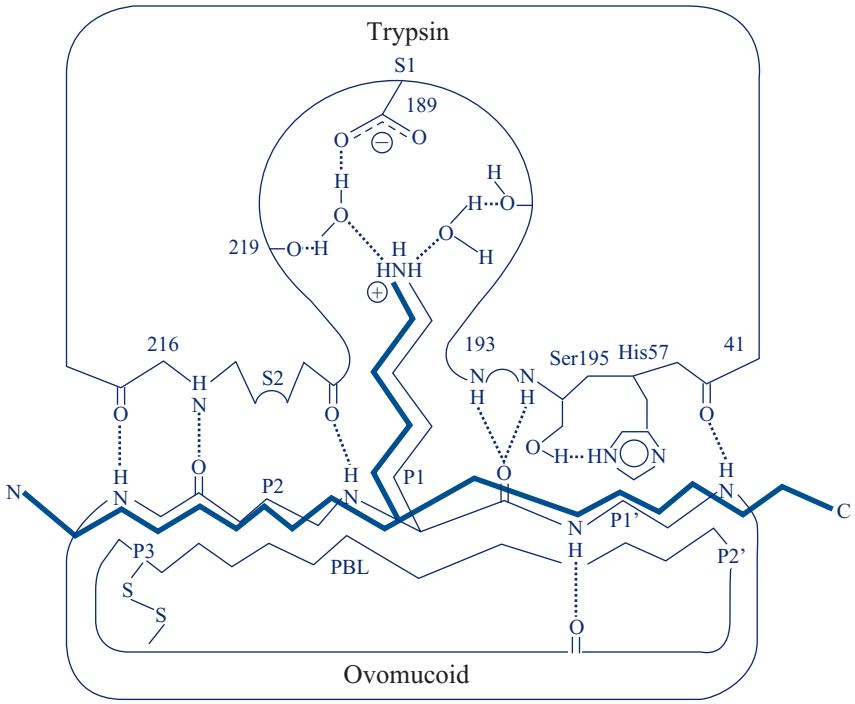
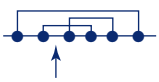


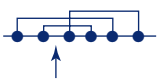
Fig. 14.2 Schematic representation of the mode of action of an inhibitor with the canonical conformational loop in the complex between trypsin and the third domain of ovomucoid (Reproduced from *Eur. J. Biochem.*, **204**, BODE W. & HUBER R., Natural proteinase inhibitors and their interaction with, 433. © (1992) with permission of Blackwell Publishing Ltd)

Bovine pancreatic trypsin inhibitor (BPTI)

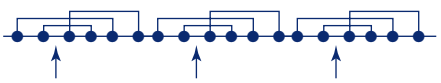


Turtle egg chelonianin

Pancreatic trypsin inhibitor (KAZAL type)



Ovomucoid



Bean inhibitor (BOWMAN-BIRK)

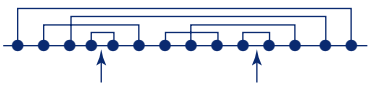


Fig. 14.3 Localisation of disulphide bridges in a few inhibitors that have the canonical conformational loop *The* dots *correspond to the half-cystines, the* arrow *indicates the cleavage site*
(Reprinted, with permission, from the *Annu. Rev. Biochem.*, **49**, LASKOWSKI M. & KATO I., 593. © (1980) by Annual Reviews www.annualreviews.org)

Serpins

More than forty proteins that are members of the serpin family have been identified in higher organisms, but also in plants and viruses. These proteins result from diverging evolution that has been spread over about five hundred million years. The majority of these proteins have conserved the presumed function of protease inhibitors of their ancestral protein. A few however have also developed other biological functions, either as transporters of lipophile molecules (globulins that bind thyroxine or cortisol), or as hormone precursors (angiotensinogen); others such as ovalbumin possess no known function.

The better studied members of this family are the plasma serpins, which display different specialisations, illustrating the fact that serpins have evolved in parallel with their target protease. Antithrombin would therefore have evolved with thrombin, the C1 inhibitor with the C1 protease, antiplasmin with plasmin. Some serpins play a key role in the process of blood clotting. Antithrombin inhibits clotting and the inhibitor of the system, which ensures the activation of plasminogen, inhibits fibrinolysis. The archetype of serpins, α_1-antitrypsin or α_1-antiprotease, is an effective trypsin inhibitor, but its physiological role is to inhibit leucocyte elastase.

The three-dimensional structures of four inhibitors of the serpin family are known at high resolution in their modified state; there are two types of α_1-antitrypsin, the inhibitor of horse leucocyte elastase and α_1-antichymotrypsin. Structures of ovalbumin and its proteolysed form plakalbumin are also resolved. Figure 14.4a below shows the structure of cleaved α_1-antitrypsin; the specific cleavage site is between Met358 and Ser359. The terminal segments resulting from the cleavage S4A and S1C insert themselves in two different β sheet regions, A and C. P1 and P1′ are found at a distance of 70 Å. The result is a very compact structure of the molecule, in agreement with its high thermodynamic stability compared to that of the native inhibitor. The reactive loop stretches from P1 to P14, whilst it goes from P1 to P5 in canonical inhibitors. The loop interacts with the binding site in the "canonical" conformation, conferring a high stability to the complex. These inhibitors undergo therefore a fundamental conformational rearrangement on proteolysis. However, neither the loop structure of native isolated serpins, nor the geometry of their complex is known. Data suggests that the reactive loop of serpins has a high flexibility. In native ovalbumin, the corresponding loop is partially structured in a helix, and in the proteolysed form (plakalbumin) the resulting segments are projected outside the molecule, explaining the absence of inhibitor capability (Fig. 14.4b).

The CARELL group showed (1991) that serpins can exist in a latent inactive form, which is more stable than the active form. It seems that, in this form, as in the proteolysed form, the A4 segment is incorporated in the β sheet A. The GHÉLIS group showed that the formation of the α_1-antiprotease complex with elastase brings about different local effects in the serpin molecule; the region close to the tryptophan residues is destabilised, while another region is stabilised due to the insertion of the A4 segment in β sheet A. This high conformational flexibility of serpins may allow

a modulation of their inhibitory activity and would ensure the protection of the circulating inhibitor against proteolytic attacks.

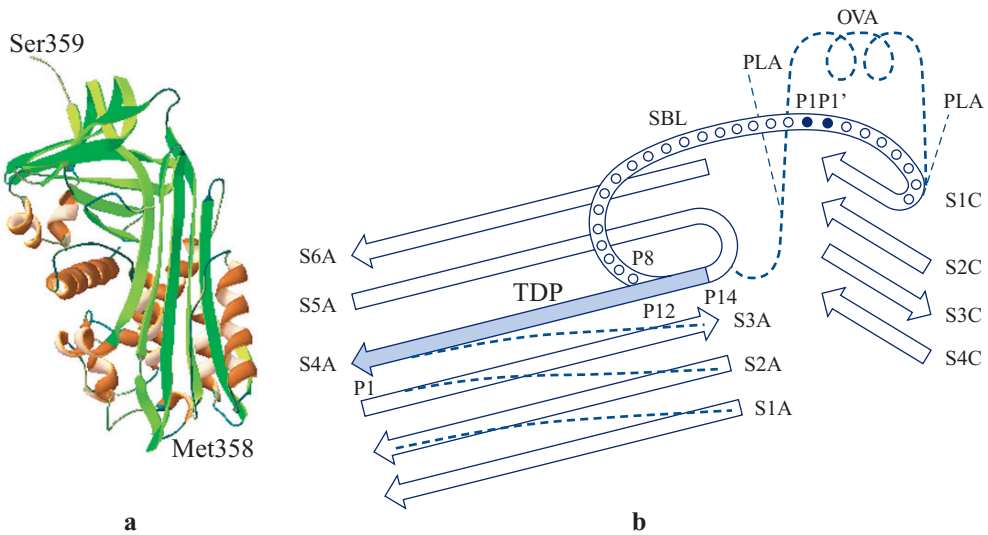


Fig. 14.4 (a) ribbon structure of cleaved α_1-antiproteinase (PDB: 2ACH) (b) schematic representation of β sheets, A and C of serpins

*The serpin binding loop (SBL) of native serpins is represented in its presumed position; it is supposed to be partially inserted in the A sheet (P12–P14), P1–P1′ represents the cleavage site, all of the part from P1 to P14 (TDP) inserts itself in the A sheet (*grey arrow *S4A). The native ovalbumin loop (OVA) is also represented, as well as the cleavage sites that lead to plakalbumin (PLA).* (14.4b – Reproduced from *Eur. J. Biochem.*, **204**, BODE W. & HUBER R., Natural proteinase inhibitors and their interaction with, 433. © (1992) with permission of Blackwell Publishing Ltd)

The formation of these protease-inhibitor complexes also has energetic consequences on the stability of the protease. One of the consequences of the complex formation is protease destabilisation; this destabilisation is of 5 kcal . mol^{-1} for elastase in its covalent complex with α_1-antiprotease (HERVÉ & GHÉLIS, 1991).

Hirudin interaction with thrombin

Hirudin is a small protein of 65–66 amino acids, which selectively binds to thrombin. The structural analysis of two recombinant hirudins performed by bidimensional NMR showed that the segment 3–47 forms a compact globular domain (with the exception of amino acids 31–36), while the two first residues and the 18 residues at the C-terminal extremity are flexible in solution. The three-dimensional structures of the two complexes formed by thrombin and two hirudin variants obtained by genetic recombination were resolved by X-ray crystallography at 2.3 Å and 2.95 Å, respectively. In these complexes, hirudin associates in an extended conformation in the crevice of the enzyme's active site (Fig. 14.5 opposite). The amino-terminal 1-3 segment forms a parallel β sheet with the segment Ser214-Gly219 of thrombin, contrary to

the antiparallel structure observed in the protein inhibitors that have the canonical conformational loop. In this complex, the catalytic serine of thrombin is not blocked, no more than the specificity pocket, which is occupied by several water molecules. The C-terminal extremity of hirudin in extended conformation stretches along a long groove, which starts at the active site of thrombin, forming many hydrophobic contacts with the enzyme and a few salt bridges with charged groups of the secondary fibrinolysis binding site. The contact surface thrombin-hirudin is of 180 Å^2.

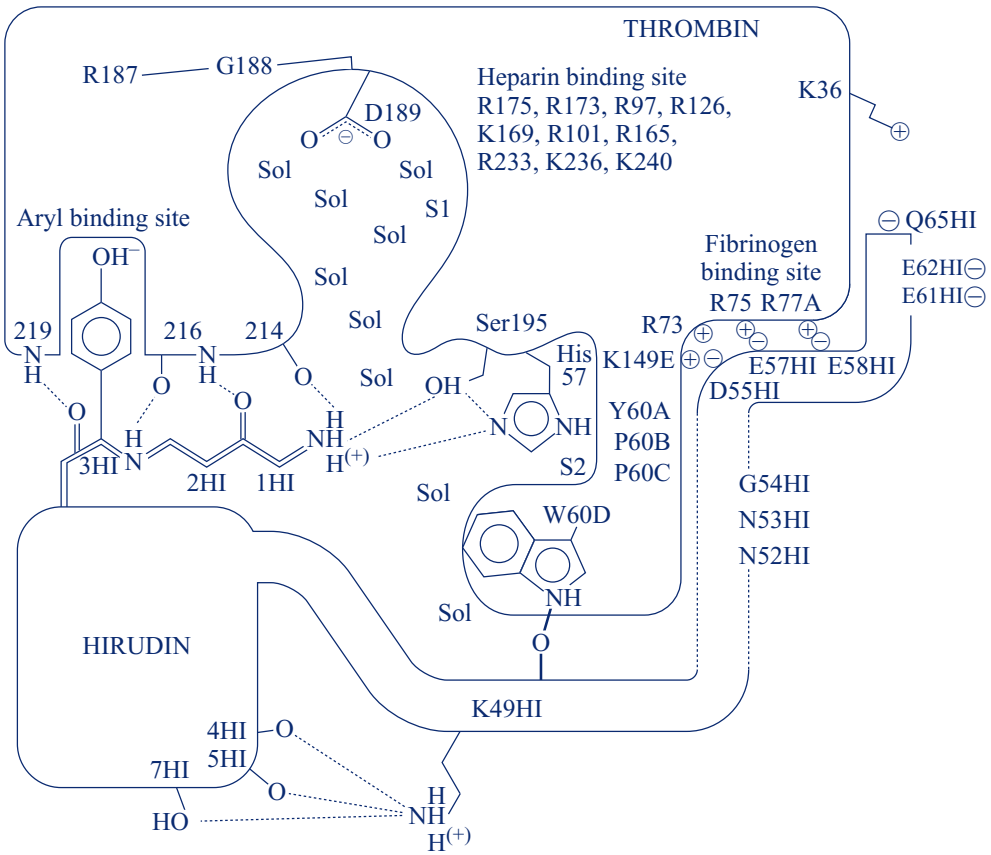


Fig. 14.5 Schematic representation of the interaction of hirudin with thrombin
(Reproduced from *Eur. J. Biochem.,* **204**, BODE W. & HUBER R., Natural proteinase inhibitors and their interaction with, 433. © (1992) with permission of Blackwell Publishing Ltd)

14.2.2. THIOL PROTEASE INHIBITORS

Cystatins and stefins are thiol protease inhibitors, which bind strongly to the enzyme in a reversible manner. The X-ray structures of two of them, hen egg cystatin and B stefin in complex with S-carboxymethylated papain were resolved. These inhibitors possess a long helix enveloped in a β sheet with five antiparallel segments,

with an additional helix in cystatin and a β strand in stefin. A first exposed loop, constituted of highly conserved amino acids, is flanked on one side by the amino-terminal segment, and on the other by a second loop. The hydrophobic side of the inhibitor has a conformation that is complementary to that of the papain active site. In the complex, the two loops interact with the adjacent subsites of the enzyme's catalytic residues. The N-terminal part covers the catalytic Cys25 and interacts with the two residues of S2 and S3 subsites. Contrary to the substrates, the "trunk" of the inhibitor is far away from the catalytic residues and, as a consequence, is not proteolysed. The primary hydrophobic interactions with the first loop confer the major part of its stability to the complex (Fig. 14.6).

A few other thiol protease inhibitors (inhibitors of papain, ficin, B1 cathepsin and bromelain) from different animal or plant origins have also been identified. Very little data exists about these proteins however, and their mechanism of action is not yet known.

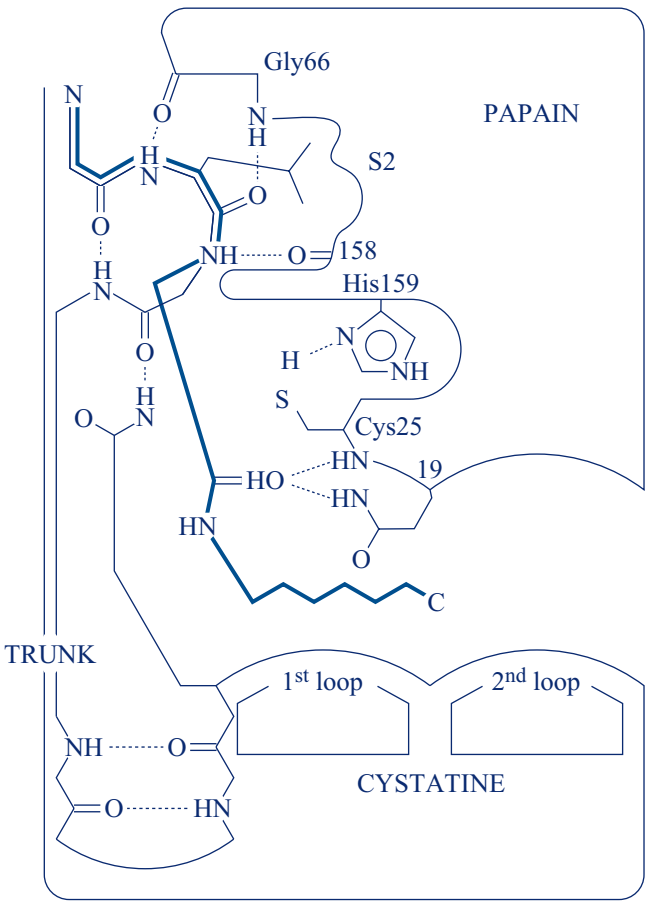


Fig. 14.6 Schematic representation of the interaction of papain with cystatin

(Reproduced from *Eur. J. Biochem.,* **204**, BODE W. & HUBER R., Natural proteinase inhibitors and their interaction with, 433. © (1992) with permission of Blackwell Publishing Ltd)

14.2.3. PROTEIN INHIBITORS OF METALLOPROTEASES

The protein inhibitors of metalloproteases are even more poorly known. A few have been identified, collagenase inhibitors, and some small inhibitors of carboxypeptidases. One of 4 300 molecular weight, from the potato, inhibits carboxypeptidases A and B, the other of 7 530 molecular weight, which comes from *Ascaris lumbriocoides* seems to specifically inhibit carboxypeptidase A.

14.2.4. ASPARTYL PROTEASE PROTEIN INHIBITORS

Very little data exists about these inhibitors. Several have been identified, such as pepsin inhibitor, cathepsins D and E inhibitors, but again the mechanism of action remains unknown.

14.3. REGULATION VIA CHEMICAL MODIFICATION

The activity of many enzymes is regulated by discrete chemical modifications catalysed by specific enzymes. These reactions consist of the binding on well defined amino acids, of groups whose presence alters, in a discrete manner, the conformation of the targeted enzyme, and therefore its activity. Depending on the case, this may have as a consequence either the activation or the inhibition of the enzyme. These chemical modifications, many of which are reversible, are of various types: phosphorylation, adenylation, uridylation, methylation, ADP-ribosylation, succinylation, guanidylation, glycosylation etc. (see Sect. 8.1.2).

14.3.1. PHOSPHORYLATION

Mechanisms of phosphorylation-dephosphorylation are implicated in the regulation of many enzymes and other proteins. The process is dynamic and intervenes in the regulation of many cellular functions: signal transmission, the cell cycle, oocyte maturation, cellular adhesion, muscular contraction, transcription regulation, lymphocyte activation, secretion, ionic channel function etc. Phosphorylations, that is, the transfer of a phosphate group from a donor to an acceptor amino acid of a protein, are catalysed by kinase proteins. The γ-phosphate group of ATP or of another nucleoside triphosphate is generally the transferred group. The reversibility of these modifications is ensured by phosphatases, whose activity is also regulated by phosphorylation.

Kinase proteins have multiple substrates; their classification was based on the acceptor amino acid rather than on the specific protein. The official terminology distinguishes five classes of phosphotransferases (HUNTER, 1991):

- class 1 comprises phosphotransferases for which the acceptor group of the protein is the serine or threonine alcohol function, which is transformed into a phosphate ester: these are serine/threonine kinases;
- class 2 includes those for which the acceptor group is a phenol, which generates a phosphate ester; these are the tyrosine kinases;
- class 3 includes enzymes for which the acceptor group is a histidine (nitrogen in position 1 or 3), the guanido group of arginine or the ε-amine of lysine, which leads to the formation of a phosphoramidate; these are histidine kinases;
- class 4 includes phosphotransferases for which the acceptor group is the SH of a cysteine that forms a phosphate thioester, these are the cysteine kinases;
- class 5 includes phosphotransferases for which the acceptor group is an acyl group which leads to the formation of a mixed phosphate-carboxylate anhydride; these are the aspartyl or glutamyl kinases.

The enzymes from the first two categories are well known, whilst those from the last three are less well characterised.

Protein tyrosine kinases are regulated enzymes, some of them by cAMP, others by calmodulin, others by diacylglycerol (see Chap. 13). Some protein tyrosine kinases are regulators of growth factors and are regulated by the binding of these. The regulator subunits can inhibit kinases by hiding the active centre. This property is called "intrasteric" regulation, since it is directed to the active centre, by opposition to allosteric regulation, in which the ligand binds to another site. About a hundred protein kinases are now known, in eukaryotic and prokaryotic organisms. They differ by their size, their specificity and their cellular location. Protein Ser/Thr kinases are more varied than the protein Tyr kinases. Despite their diversity, all the protein kinases have a highly conserved "catalytic core" (HANKS & QUINN, 1991). The three-dimensional structure of a few protein kinases was resolved (see Sect. 13.3). From the data available, a few characteristics common to all protein kinases emerge. The catalytic core is constituted of two lobes, separated by a groove that binds protein substrates; the binding site of nucleotides is dominated by β structures. The large essentially helicoidal domain is associated to the binding of protein substrates; it contains two β segments, which converge towards the active site at the interface of the two domains. Two highly conserved loops, one in each domain, also converge towards the active site. Contrary to these characteristics that are common to all kinases, the recognition of the protein by the catalytic subunit implies non-conserved amino acids and the peptide binding site stretches over different and separated regions at the surface of the enzyme.

The existence of regulations *via* phosphorylation was discovered in the case of enzymes implicated in glycogen metabolism (glycogen phosphorylase, glycogen synthetase) (FISHER & KREBS, 1955). One of the enzymes whose regulation by phosphorylation has been the most studied is glycogen phosphorylase, responsible for the break-down of glycogen. This dimeric enzyme with pyridoxal phosphate is also the subject of an allosteric regulation, which was described in the preceding chapter (Sect. 13.10.1). The regulation of glycogen phosphorylase by phosphorylation is

part of a complex phosphorylation-dephosphorylation cascade system, which is presented further on (Sect. 14.4.2).

◤ The structural modifications that result from the phosphorylation of glycogen phosphorylase were determined in detail by the groups of MADSEN, JOHNSON and FLETTERICK, who determined and compared crystallographic structures of phosphorylated (glycogen phosphorylase a) and dephosphorylated (glycogen phosphorylase b) forms of the enzyme (SPRANG et al., 1988; BARFORD et al., 1991). The phosphorylation occurs on serine 18 of the N-terminal region of the protein, and the structural modification that it induces is essentially localised in this region. The N-terminal region possesses a disordered structure in non-phosphorylated glycogen phosphorylase b. The phosphorylation leads to the organisation of this region in an α helix, and its interaction with the complementary subunit. The phosphorylated helix inserts itself in the large groove, found at the interface between the two subunits and forms many VAN DER WAALS interactions and hydrogen bonds (Fig. 14.7; see also Fig. 12.46).

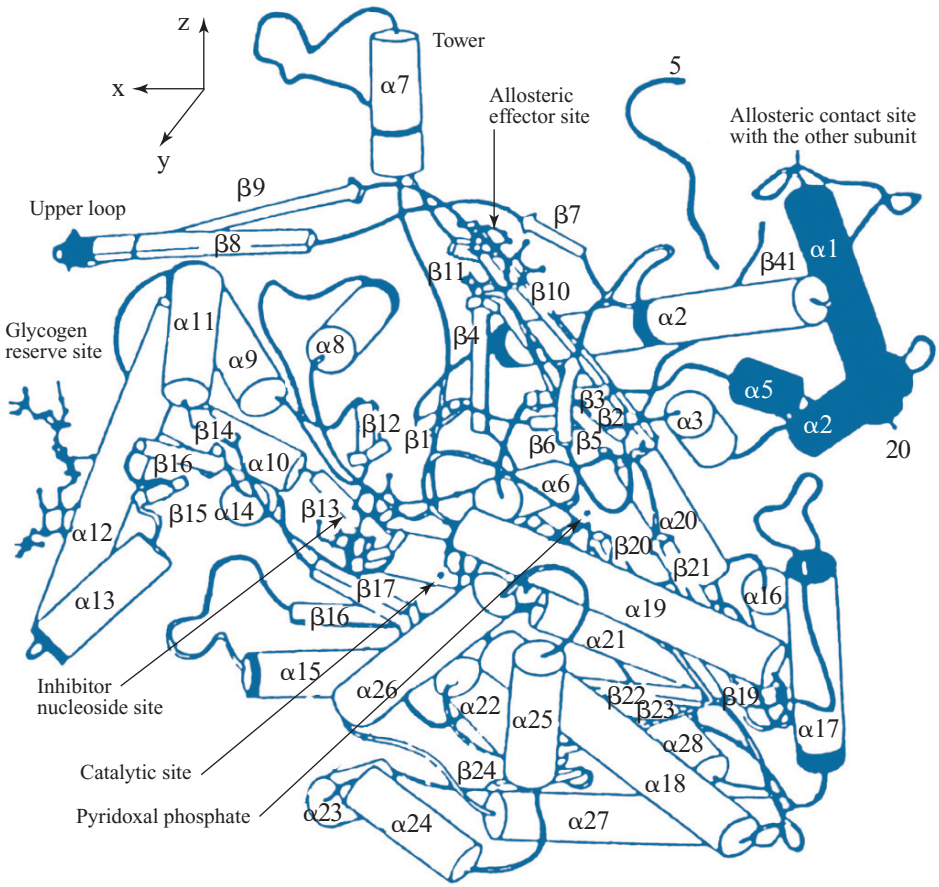


Fig. 14.7 Schematic representation of one glycogen phosphorylase subunit, showing different regions (*in black*) where the majority of C^{α} differ by more than 0.6 Å between phosphorylase a and phosphorylase b

(Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **336**, SPRANG S.R. et al., 215. © (1988))

Eight hydrogen bonds form within the same subunit and eight others between the two subunits. Phosphoserine 18 becomes involved in two ionic bonds, one with Arg69 found in the neighbouring helix h2 of the same subunit, the other with Arg43' found in the "cap" region of the neighbouring subunit. The phosphoserine also interacts with the amide group of valine 15 and, *via* a water molecule, with the side-chain of Gln72, these two residues belonging to the same chain. Some hydrophobic interactions also contribute to the stabilisation of the helix at the interface. Isoleucine 13 and valine 15 are buried in a hydrophobic pocket belonging to the complementary subunit. This complex ensemble of interactions therefore strongly increases the association between the two subunits and the cohesion of the dimer, in particular when the two chains are phosphorylated. It is remarkable that the phosphoserine 18 stays accessible to the action of glycogen phosphorylase phosphatase for the reconversion of glycogen phosphorylase a in glycogen phosphorylase b. ◢

Concerning enzymatic activity, phosphorylase a is active even in the absence of AMP. The presence of this effector leads only to an increase of 10% in activity. In addition, glycogen phosphorylase a is insensitive to inhibition by ATP, ADP and glucose-6-P. It is sensitive to inhibition by adenosine however, the physiological significance of which is not totally understood.

Phosphorylase a is nevertheless subject to the T ⟷ R transition, described in the preceding chapter (Sect. 13.10.1) in the case of phosphorylase b.

14.3.2. ADP-RIBOSYLATION

Nicotinamide mononucleotide (NMN) had been known for several years as a coenzyme (see Sect. 11.3), when in 1966 CHAMBON and collaborators discovered that the compound was also at the base of polyadenylic acid synthesis by a nuclear enzyme (CHAMBON et al., 1966). This observation, as well as the study of the mode of action of certain bacterial toxins, led to the observation that the reaction constitutes a regulation mechanism, which is involved in various and important ways in the physiology of eukaryotic organisms, of bacteria and in certain viruses. It became apparent that this mode of regulation is involved in many processes such as the DNA repair, the cell cycle regulation, transformation, rearrangement of genes and transposition, but also that it intervenes in the function of membrane proteins, the binding of nitrogen by nitrifying bacteria, the mechanism of action of certain toxins and the development of bacteriophages. All these effects are the result of the covalent modification of enzymes or proteins by ADP-ribosylation.

14.3.2.1. ENZYMATIC ADP RIBOSYLATIONS

The reaction of ADP-ribosylation is catalysed by specific enzymes, ADP-ribosyl transferases, which use NAD^+ as a substrate. During this transfer, the β-N glucosidic bond of NAD^+ is broken. The action of this enzyme consists of three reactions:

- the covalent binding of an ADP-ribose residue on a specific amino acid of the protein to be regulated (Fig. 14.8 opposite);

- the successive addition of new ADP-ribose residues in position 2′-hydroxyl of adenine-ribose already linked to the target protein (Fig. 14.9 below);
- the addition of an ADP-ribose residue to the 2′-hydroxyl of what was the ribose of the nicotinamide, reaction which generates a point of branching for the rest of the synthesis (Fig. 14.9).

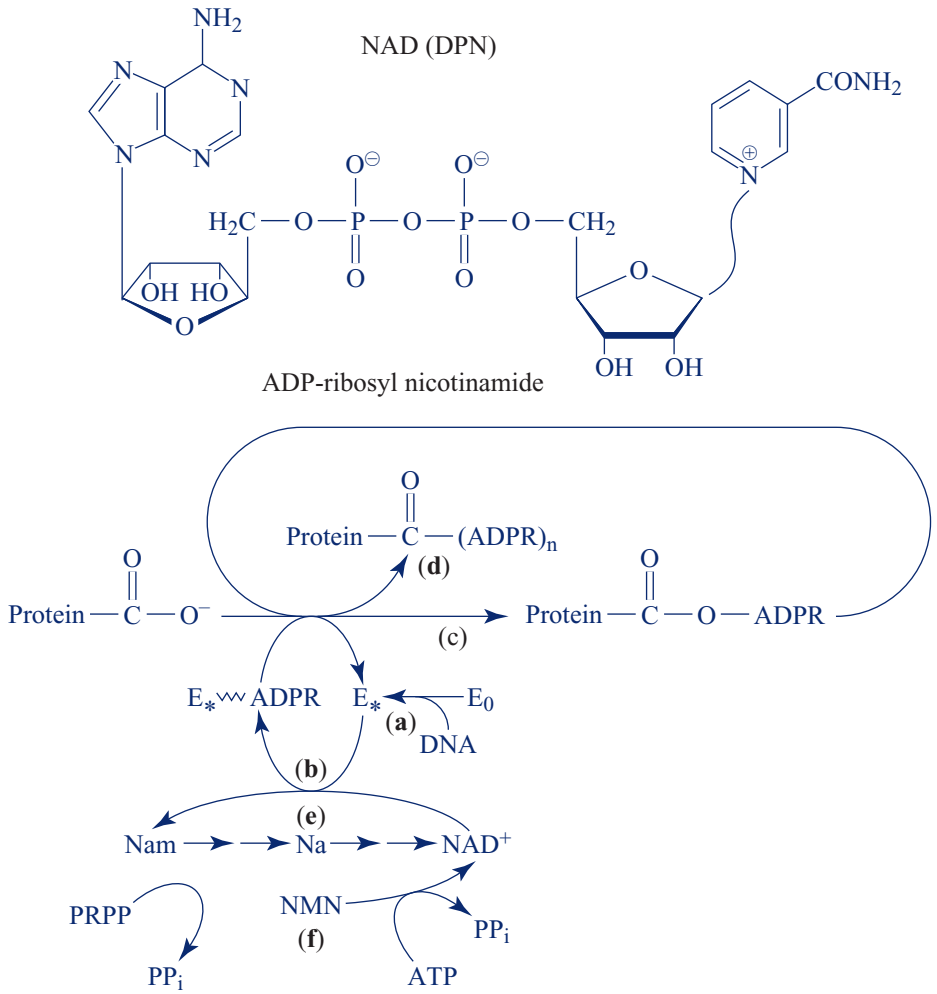


Fig. 14.8 Reaction of ADP-ribosylation

The ADP-ribosyl transferase is inactive (E_0) as long as it is not associated to DNA (E). The active enzyme catalyses the cleavage of NAD^+ into Nam and ADP-ribose, which is attached to the enzyme's catalytic site (E_*–ADPR). The ADP ribose is then transferred to the protein to give the conjuguated derivative containing either only one (c), or multiple (d) ADP-ribose groups. The NAD^+ is regenerated either via nicotinamide (e), or via NMN (f). PRPP: phosphoribosyl pyrophosphate; NMN: nicotamide mononucleotide; Nam: nicotamide; Na: nicotinate* (Reprinted from *Trends Biochem. Sci.*, **11**, GAAL J.C. & PEARSON C.K., Covalent modification of proteins by ADP-ribosylation, 171. © (1986) with permission from Elsevier)

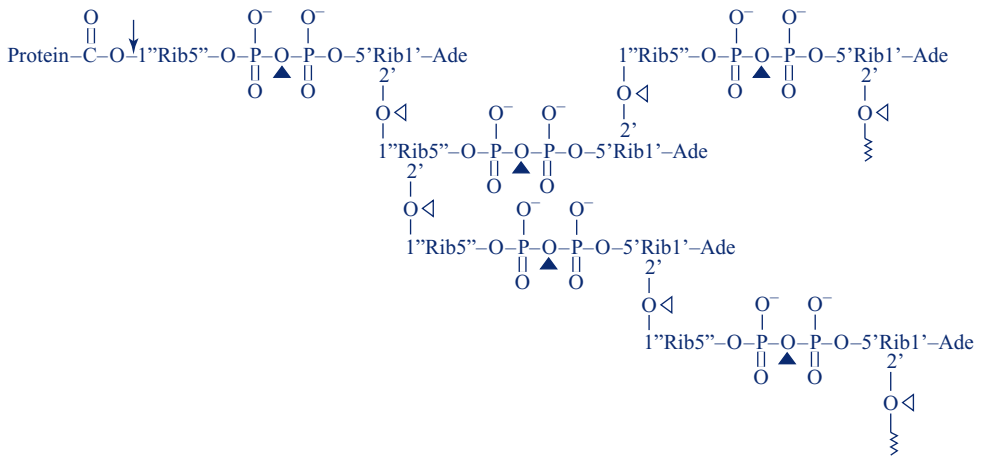


Fig. 14.9 Structure of poly(ADP-ribose) showing the branching and enzymatic degradation

◁: cleavage sites recognised by poly(ADP-ribose) glycohydrolase; ▲: and by phosphodiesterase; ↓: ADP-ribosyl protein lyase attack site; Rib: ribose; P: phosphate; Ade: adenine (Reprinted from *Trends Biochem. Sci.*, **11**, GAAL J.C. & PEARSON C.K., Covalent modification of proteins by ADP-ribosylation, 171. © (1986) with permission from Elsevier)

Together, the two described reactions can lead to the formation of a polymer containing more than 50 ADP-ribose residues. The polymer can be degraded by poly(ADP-ribose) glycohydrolase in vivo and by phosphodiesterase in vitro (Fig. 14.9).

It is convenient to distinguish mono-ADP-ribosyl transferases, which catalyse only the first of these reactions, from poly-ADP-ribosyl transferases, which catalyse the three stages. Poly-ADP-ribosyl transferases are only found in the nucleus and intervene in different aspects of DNA metabolism. Mono-ADP ribosyl transferases are found in other cellular compartments (cytoplasm, mitochondria, membranes) and microorganisms.

The ADP-ribosylation reaction affects carboxyl groups of particular residues of glutamic acid or lysine and arginine located at the C-terminal of the target protein. The reaction implies transitory auto-ADP-ribosylation of the transferase.

The structural analysis of poly-ADP-ribosyl transferases showed that it is the C-terminal region of these enzymes that is implicated in the three partial reactions described above. The N-terminal region contains a nuclear localisation signal and two zinc finger structures, which are necessary for the binding of the enzyme to DNA. It is by the N-terminal region that the ADP-ribosyl nuclear transferases recognise DNA cleavages. The enzyme binding on these cleavages increases its efficiency by a factor 500. This led SHALL (1995) to propose that the cleavage of DNA behaves as an allosteric effector of transferase.

14.3.2.2. PHYSIOLOGICAL EFFECTS

In the cellular nucleus, poly-ADP-ribosyl transferase intervenes at three levels:

- ADP-ribosyl transferase is activated by DNA cleavages, which is at the basis of its intervention in the processes of DNA repair, of transformation and cellular differentiation;
- ADP-ribosyl transferase modifies histones, which produces changes in chromatin structure;
- ADP-ribosylation modulates the activity of several enzymes implicated in DNA metabolism: DNA ligase III, terminal deoxynucleotidyl transferase, α and β DNA polymerases, topoisomerases, Ca^{++} and Mg^{++} dependant endonucleases. DNA ligase III activity is either inhibited, or activated depending on the ADP-ribosylation site. In relation to these effects, the activity of ADP-ribosyl transferase is significantly higher in tumoural cells than in normal cells.

Outside the nucleus, ADP-ribosylation intervenes also in mitochondria. For example, the modification of a protein found in the internal membrane of a mitochondrion modifies calcium permeation. The ADP-ribosylation reactions also play a role in the toxicity of certain bacteria. Diphtheria toxin inhibits the EF2 elongation factor by mono-ADP-ribosylation, which blocks protein synthesis in the infected cell. Choleric and pertussis toxins provoke ADP-ribosylation of a G protein implicated in the regulation of adenylate cyclase activity, provoking in this way an important increase in the cellular ratio of cyclic AMP.

These types of regulation are also involved in the process of virus infections. In this manner, the bacteriophage T4 inhibits DNA transcription in the host cell by ADP-ribosylation of the RNA-polymerase α subunit. ◢

14.3.3. GLYCOSYLATION

Protein glycosylation is a post-translational modification as widespread as phosphorylation. The majority of membrane proteins, but also cytosolic proteins of animal cells, contain saccharide chains. Protein glycosylation plays an important role in many biological processes, such as cellular adhesion, immune response, intracellular transport, the binding of certain hormones, but also in protein stabilisation.

There are two major types of glycosylation, N- and O-glycosylations, which involve the binding of the saccharide chain to an asparagine and a serine or a threonine, respectively (see Sect. 11.3).

◤ Practically all the N-glycosylations occur on an asparagine belonging to a sequence Asn-X-Ser/Thr, where X can be any amino acid except proline or aspartic acid. All the saccharide chains contain the pentasaccharide Manα1 ⟶ 6Manα1 ⟶ 3Manβ1 ⟶ 4GlcNacβ1 ⟶ 4GlcNac as a common motif (Fig. 14.10 below). Other residues add on to this motif and the different branched complexes formed in this way can be subdivided into three sub-groups as shown in Fig.14.10. The first sub-group is the complex type; it does not contain any other mannose residues apart from the trimannosyl motif. The second sub-group contains a high mannose content. The third sub-group is of the hybrid type, the saccharide chains show characteristics of the two preceding sub-groups. The presence of a α-fucosyl residue bound in position C6 of

the proximal N-acetyl glucosamine and of the β-N-acetyl glucosamine bound in position C4 of the β-mannosyl of the trimannosyl motif can contribute to structural variations in the complex and hybrid type sub-groups. The complex type sub-group displays the largest structural variations following the addition of 1 to 5 chains to the trimannosyl motif, forming mono-, tri-, tetra- and penta-antenna structures. Other saccharide chains such as sialic acid can also be added.

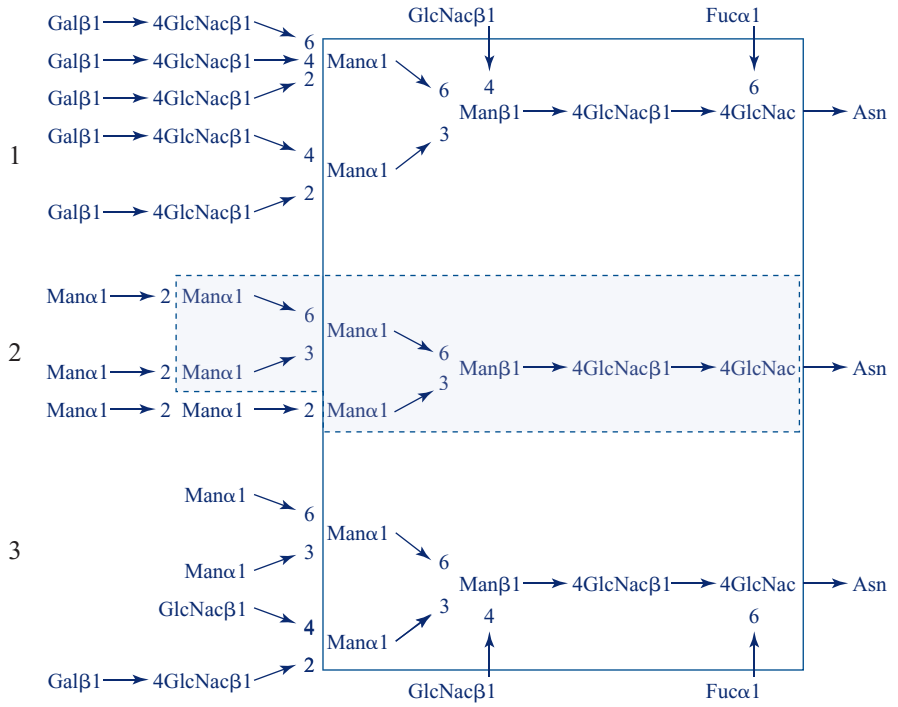


Fig. 14.10 Diagram of different N-glycosylation types

O-glycosylations occur on hydroxyl groups of a serine or threonine side-chain, found in a β-turn. There are different types of O-glycans. All these post-translational modifications require the intervention of numerous glycosyl transferases, some of which are not only very specific towards the substrate, but also towards the branched structure.

In addition to these two categories, several types of glycosylation were described in eukaryotes (HAYES & HART, 1994). The O-N-acetyl glucosamine motif, which is the result of the addition of a simple N-acetyl glucosamine residue to a serine or threonine residue was found in numerous structure proteins, but also in RNA polymerase II and in transcription factors. It seems that O-GlcNac is a regulatory modification intervening in an analogous manner to phosphorylation; for some proteins this modification is in fact transitory. In addition, this type of modification appears in proteins forming transitory multimeric complexes. All of the O-GlcNac proteins identified to date are also phosphoproteins, suggesting a possible relation between phosphorylation and glycosylation. GUPTA and collaborators (RAY et al., 1992) suggested that the residues O-GlcNac contribute to the regulation of protein biosynthesis. The addition of α-glucose-1-phosphate to a mannose bound to a serine or a threonine to form

a Glc-PO_4-Man_n motif on cytoplasmic proteins was described in the case of phosphoglucomutase (MARCHASE et al., 1993). It was suggested that this modification could either have a regulatory role of enzymatic activity, or could intervene in the enzyme's intracellular localisation.

Other types of modifications were described and it is possible that new forms of glycosylation will be shown in the future. Generally however, these modifications do not seem to be involved in the regulation of enzymatic activity. In the case of glycosylated enzymes, glycosylation would mostly have a stabilising role on the protein, protecting it against proteolytic degradation. ◢

14.3.4. ADENYLATION, URIDYLATION

This type of covalent modification consists of the binding of an adenyl or an uridyl group to a well-defined tyrosine residue of a protein. Adenylation is more common than uridylation.

The most well known and spectacular example is that of glutamine synthetase, whose regulation involves these two processes. Glutamine synthetase catalyses the following reaction:

$$\text{glutamate} + \text{ammonium} + \text{ATP-Mg} \longleftrightarrow \text{glutamine} + \text{ADP} + P_i$$

This enzyme occupies an important position in cellular metabolism because glutamine is the principal source of amino groups for the biosynthesis of many amino acids, purine and pyrimidine nucleotides, glucosamine-6-phosphate and NMN. The amino group transfer involves the intervention of many reactions catalysed by subunits or domains of glutaminase, which extract the amino group from the glutamine and transfer it to the catalytic site of the targeted enzyme. It is therefore not surprising that the activity of glutamine synthetase is the subject of many regulatory processes. These complex mechanisms were studied in detail by the STADTMAN group in the case of the *E. coli* enzyme. Table 14.1 indicates the group of effectors that have direct or indirect effects on the activity of glutamine synthetase.

Table 14.1 Direct and indirect effectors of glutamine synthetase activity

Direct effectors	*Indirect effectors*	*Indirect effectors*
Alanine	Glutamine	P-enolpyruvate
Glycine	α-cetoglutarate	CoA
Serine	Orthophosphate	Oxaloacetate
Histidine	Glutamate	Citrate
Tryptophan	Tryptophan	Pyruvate
Carmabyl-P	Methionine	Succinate
Glucosamine-6-P	ATP	Fumarate
ATP	UTP	Malate

Direct effectors	*Indirect effectors*	*Indirect effectors*
CTP	NH_4^+	PP_i
AMP	3-phosphoglycerate	Mg^{++}
GMP	Fructose-6-P	Mn^{++}
IMP	Fructose-1-6-P2	RNA
Glutamate	CMP	
Mg^{++}	UMP	
Mn^{++}	GMP	
NH_4^+	TMP	

The activity of glutamine synthetase is regulated in an integrated manner by a series of processes including:

- an integrated allosteric cumulative retro-inhibition by several tens of metabolites,
- a rapid renewal by biosynthesis and degradation after oxidative inactivation,
- a cascade regulation by adenylation and uridylation of a regulatory protein (see Sect. 14.4).

Glutamine synthetase from *E. coli* is a dodecamer formed by the superposition of two hexamers, in which the protomers are associated in a circle with a C6 symmetry (Fig. 14.11).

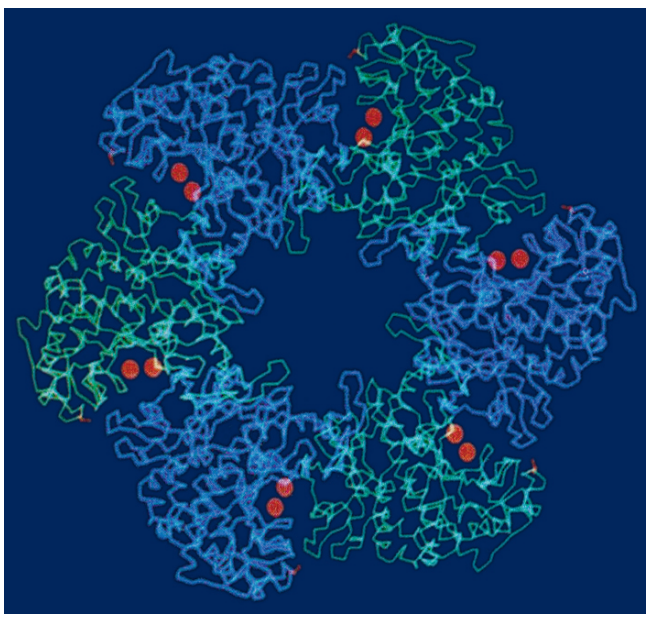


Fig. 14.11 Structure of glutamine synthetase

The structure of the upper half is projected perpendicular to the 6-fold symmetry axis. The external diameter is 143 Å. The active sites are marked by the pairs of Mn^{++} *(*in red*). The Tyr397 side-chain (*in red*) indicates the adenylation site* (Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **323**, ALMASSY R.J. et al., 306. © (1986))

It appears however that the interactions are stronger between monomers that do not belong to the same hexamer. The catalytic site is at the interface between two monomers that belong to the same hexamer. The catalytic site includes the loop containing tryptophan 57, the elements of β structure 44–52 and 59–70, as well as the tyrosine 397 region (Fig. 14.12). In the oligomer, tyrosine 397 of one of the protomers is at a distance of 6 Å from tryptophan 57 of the neighbouring protomer, which contributes to forming the catalytic site. The substrate ATP binding site is positioned near the N-terminal extremity. Enzymatic activity requires the presence of Mg^{2+} and Mn^{2+}, whose the binding sites are distant by 5.8 Å. Catalysis involves the movement of the loop containing tryptophan 57 of the N-terminal domain from one protomer to the C-terminal domain of the neighbouring subunit.

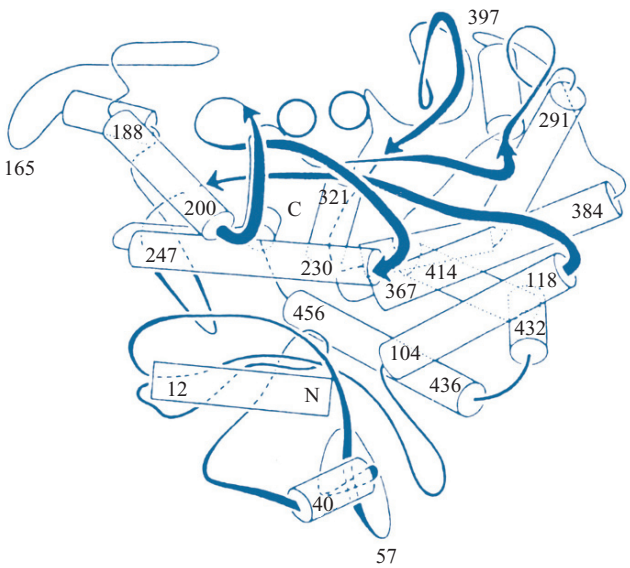


Fig. 14.12 Schematic representation of a glutamine synthetase subunit
The cations are represented by cercles. Arrows *indicate β segments. The junction between C- and N-terminal domains is near residue 104, which is the start of the third helix. The central loop is at the* top left (Reprinted by permission from Macmillan Publishers Ltd: *Nature*, **323**, ALMASSY R.J. et al., 307. © (1986))

Figure 14.13 below schematically presents the cascade regulation of the activity of glutamine synthetase from *E. coli*, by adenylation of the enzyme and uridylation of a regulatory protein (see also Sect. 14.4.3). The enzyme's activity is inhibited by adenylation of tyrosine 397. This effect seems to be the result of the interaction between the adenyl group and the region of tryptophan 57, interaction that prevents the relative movement of the two domains associated with catalysis. The inactivation is accompanied by the change in specificity of the enzyme for the divalent cations. The enzyme activity is inversely proportional to the number of adenylated subunits. There is therefore no discernable cooperativity in the adenylation effects, although from the statistical distribution of possible adenylated sites, 382 forms of

the enzyme are possible. The form in which the twelve protomers are adenylated is totally inactive and contains only Mg^{2+} ions. Their replacement by Mn^{2+} restores 25% of maximal activity. Adenylation also provokes an increase in the enzyme's sensitivity to some of its allosteric effectors.

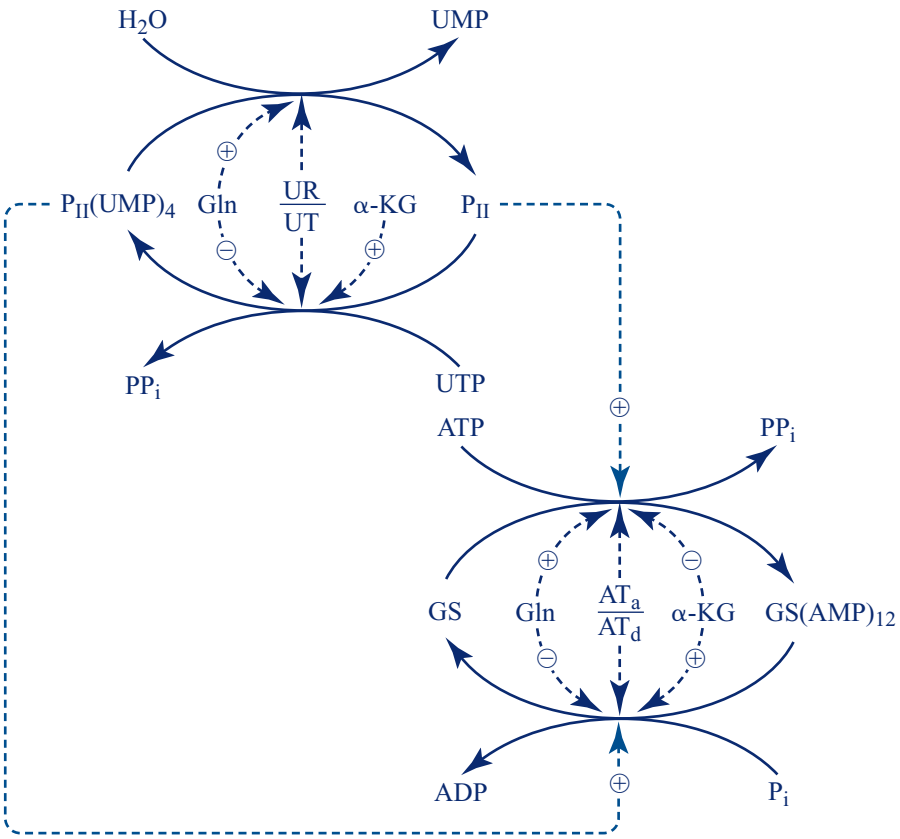
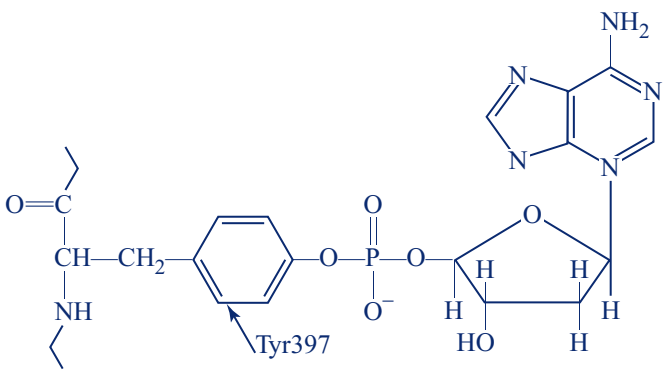


Fig. 14.13 Regulation cascade of glutamine synthetase

Adenylation is catalysed by an adenylate transferase and occurs on tyrosine 397:

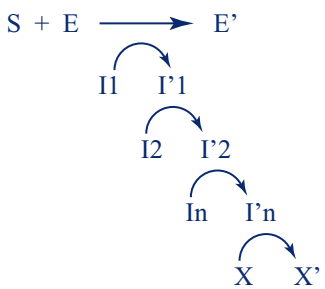


Glutamine synthetase can be reactivated by phosphorolysis that induces adenylate release. The reaction is also catalysed by adenylate transferase. The direction in which the reaction occurs is determined by the binding of a regulatory protein, PII, which itself exists in two forms, one uridylated and one non-uridylated. The first form favours phosphorolysis, the second the adenylation of glutamine synthetase. The uridylation of the PII protein is catalysed by a tetrameric enzyme, uridyl transferase. The deuridylation however, is catalysed by a different monomeric enzyme. These two reactions are under the control of allosteric effectors. Glutamine inhibits uridylation and activates deuridylation. α-cetoglutarate on the other hand favours uridylation. This regulation is more complex and integrated because of the fact that these same allosteric effectors have inverse effects on adenyl transferase itself. Glutamine activates adenylation and inhibits phosphorolysis while the α-cetoglutarate acts in the inverse way (see Fig. 14.13).

14.4. CASCADE SYSTEMS – MECHANISM OF ACTION

14.4.1. DEFINITION

The enzymatic regulation systems known as "cascade systems" have in common the fact that they involve covalent modification of at least one enzyme. This covalent modification is accompanied by an important change in the activity of the enzyme; in general the enzyme is initially in an inactive form and becomes active as a result of the covalent modification, leading to the enzymatic cascade. Table 14.2 below gives a few examples of chemical modifications involved in cascade systems. The interconversion enzyme is itself activated either by an effector or by a covalent modification. These enzymes are therefore part of more or less complex enzymatic pathways that can be represented in the following way:



hence the name cascade system.

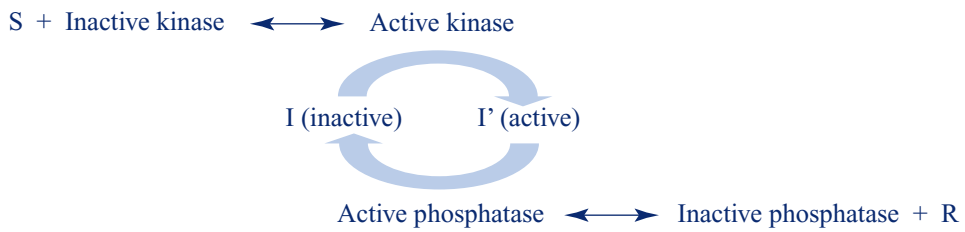
One can distinguish two types of cascade systems. The first corresponds to irreversible and uni-directional systems in which, in response to a primary stimulus, a very strong amplification is produced, which ends in an explosive type of response. It is the case of two complex systems of vital importance for higher organisms: the cascade system of blood clotting and that of complement system. These systems must

respond to urgent situations. In these two systems, the covalent modifications are limited proteolysis.

Table 14.2 Cascade systems

Modification	*Enzyme*	*Residue*
Proteolysis	Blood clotting Complement system	
Phosphorylation	Glycogen phosphorylase Glycogen synthase Pyruvate dehydrogenase Pyruvate kinase Phosphorylase kinase c-AMP dependent protein kinase	Ser
Nucleotidylation	Protein PII (uridylation) Glutamine synthetase (adenylation)	Tyr
ADP-ribosylation	RNA-polymerase	Arg
Carboxymethylation	Membrane proteins	Asp, Glu

The second type of cascade is that of reversible systems in which enzymes are interconvertible. It is the case for example where the chemical modification is phosphorylation. The phosphorylated enzyme can be de-phosphorylated by a phosphatase. Thus, the interconvertible enzyme describes a cycle, where the name mono-, bi- or multi-cyclic comes from. Suppose for example the following monocyclic cascade:



R is the reversion enzyme. Depending on the physiological state of the cell, therefore depending on metabolite concentrations, the ratio I/I′ can vary, as well as the activity of the key enzyme: it is a system of enzymatic regulation. Very often, the system is also submitted to allosteric regulations.

14.4.2. IRREVERSIBLE CASCADE SYSTEMS

14.4.2.1. BLOOD CLOTTING CASCADE

When a vascular lesion occurs, platelets are activated and adhere to the lesion area, then aggregate. The platelet activation releases numerous proteins that accelerate

the platelet plug formation, which reduces the blood loss and starts the tissular reparation process. The VON WILLEBRAND factor plays an important role in platelet adhesion. These reactions allow the set up of the clotting cascade, by making accessible negatively charged phospholipids such as phophatidylserine to the surfaces of the activated platelets and the damaged cellular membranes. A series of reactions is then triggered, which lead to the formation of an insoluble fibrin clot that reinforces the platelet plug.

The blood clotting cascade is divided into two pathways: the intrinsic and extrinsic pathways. The intrinsic pathway relates to reactions which only use factors from plasma, while the extrinsic pathway has both a tissue factor, which comes in contact with the blood upon vascular lesion, and plasma factors. The two mechanisms are represented in Fig. 14.14.

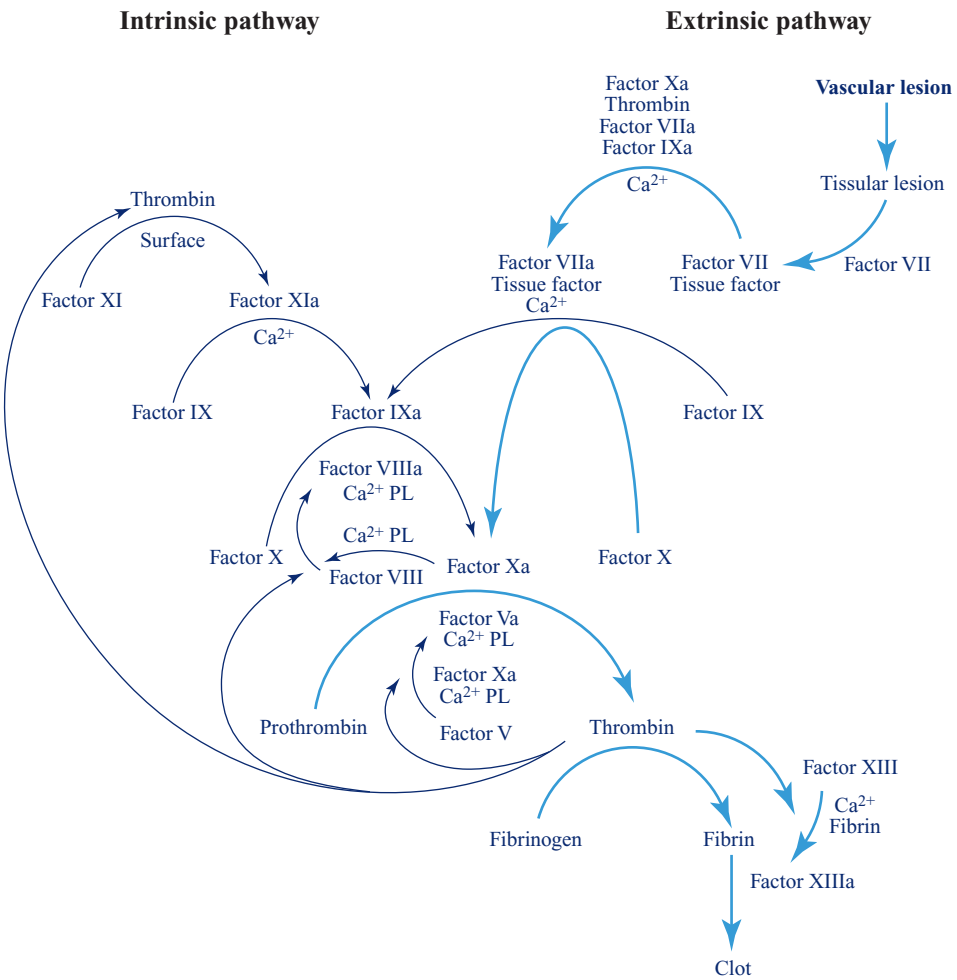


Fig. 14.14 Blood clotting cascade, by intrinsic and extrinsic pathways

(Reprinted with permission from *Biochemistry*, **30**, DAVIE E.W. et al., 10364. © (1991) American Chemical Society)

The majority of clotting enzymes are serine proteases, which come from zymogen precursors by sequential activation; the activity of the majority of them requires the presence of Ca^{++} and phospholipids. The calcium binding site is formed by the addition of a second carboxyl group to glutamic acid residues to form γ-carboxyglutamate groups that display a high affinity for calcium. Most of these factors are glycoproteins.

In the extrinsic pathway, when a vascular lesion occurs, the tissue factor, which is an integral membrane glycoprotein, forms a 1:1 complex of high affinity with the factor VII. The factor VII is then transformed into factor VIIa by the cleavage of a Arg_{152}-Ile peptide bond, which is catalysed by traces of proteases circulating in the blood (thrombin, factor Xa, factor VIIa, factor IXa). The factor VIIa-tissue factor complex converts factor X into factor Xa, by cleavage of the Arg_{52}-Ile bond. The tissue factor functions as a co-factor in the factor X activation. The factor Xa forms a 1:1 complex with the factor Va in the presence of Ca^{++} ions and phospholipids. The complex converts prothrombin into thrombin *via* the hydrolysis of two peptide bonds, Arg_{271}-Thr and Arg_{320}-Ile, which reduces the molecular weight of the precursor from 71 600 to 39 000. The factor Va plays the role of cofactor in the activation of prothrombin by increasing the V_{max} value about 1 000 times; the phospholipids, which come from activated platelets, decrease the K_m value of prothrombin 1 400 times. The factor V activation is probably the result of action by the factor Xa in the presence of Ca^{++} and phospholipids; the newly produced thrombin can also activate factor V by the cleavage of two peptide bonds (Arg_{710}-Ser and Arg_{1545}-Ser), which liberates the connecting peptide between the two chains of factor V (Fig. 14.15 opposite). Thrombin, once formed, converts fibrinogen into fibrin by the cleavage of a peptide bond in each of the two α chains (Arg_{18}-Gly), and in each of the β chains (Arg_{16}-Gly). A polymerisation follows which leads to the formation of the insoluble fibrin clot. Fibrin accelerates the transformation of factor XIII into factor XIIIa, by thrombin in the presence of Ca^{++}. Factor XIIIa catalyses the formation of ε-(γ-glutamyl) lysine bonds between the two fibrin molecules leading to the formation of a very solid clot.

When the quantity of tissue factor is limiting after the vascular lesion, it is the activation of factor IX rather than factor X, by the factor VIIa-tissue factor complex, which triggers the clotting cascade. The activation of factor IX involves the cleavage of two peptide bonds (Arg_{145}-Ala and Arg_{180}-Val). Factor IXa forms a complex with factor VIIIa, which activates factor X in the presence of Ca^{++} and phospholipids. The presence of factor VIIIa increases the maximal speed about 200 000 times. A deficiency in factor VIII is the cause of haemophilia A, and a deficiency in factor IX causes haemophilia B. The factor VIII activation is the result of a limited proteolysis of Arg_{372}-Ser, Arg_{740}-Ser and Arg_{1689}-Ser bonds, bringing about the loss of the connecting region (Fig. 14.15). The extrinsic clotting pathway is inhibited by a clotting inhibitor associated with a lipoprotein or LACI (for lipoprotein-associated coagulation inhibitor).

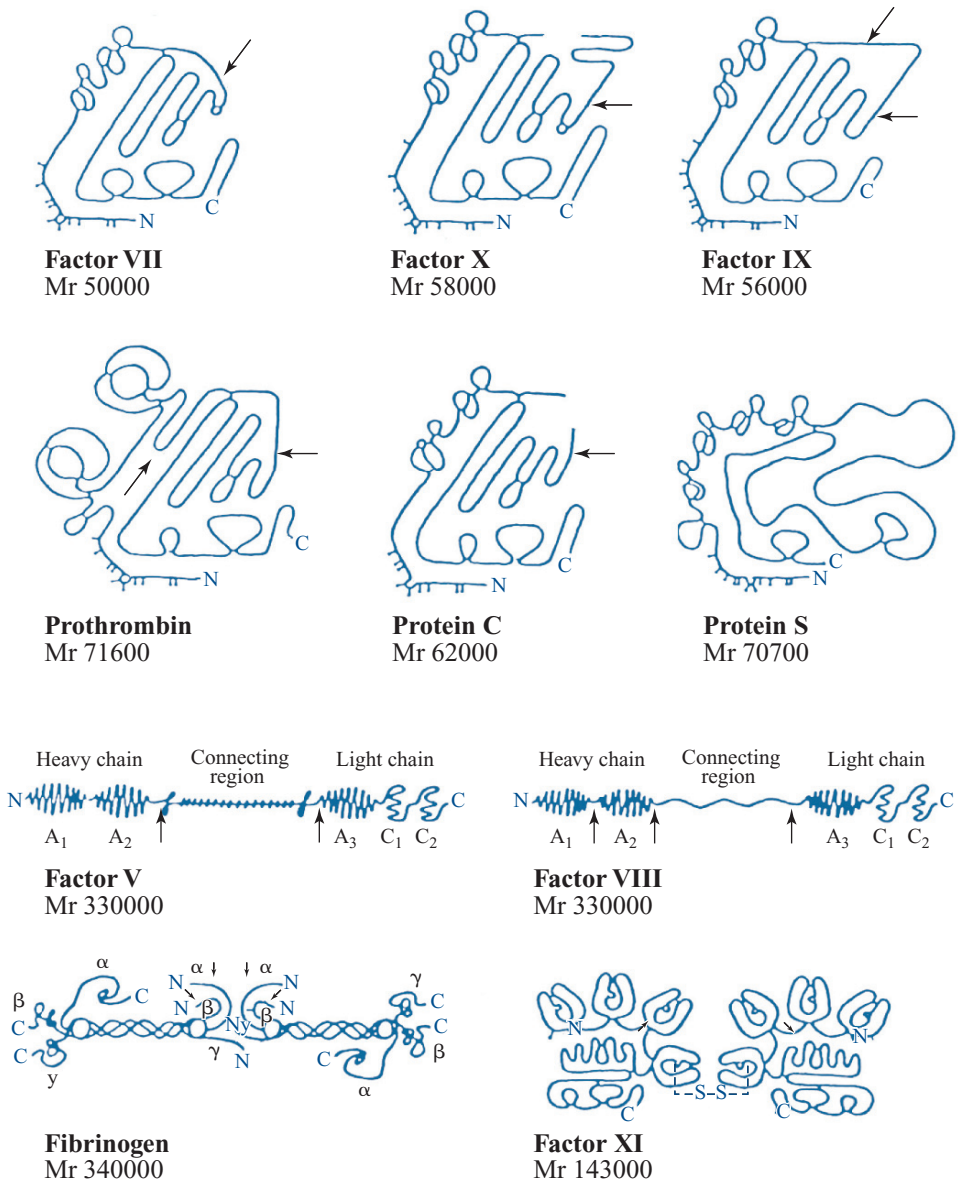


Fig. 14.15 Schematic representation of glycoproteins that participate in the clotting cascade and its regulation

Arrows *indicate proteolysis sites, and* lines *perpendicular to the* chain *indicate γ-carboxylic acid site* (Reprinted with permission from *Biochemistry*, **30**, DAVIE E.W. et al., 10365. © (1991) American Chemical Society)

The intrinsic pathway involves a plasmic glycoprotein, factor IX, which circulates in blood. It is activated by thrombin in the presence of a negatively charged surface, such as heparin or dextran sulphate. The activation is a result of the cleavage of the

Arg_{369}-Ile peptide bond in each of the two subunits of factor XI. Factor XIa activates factor IX into IXa. The newly formed thrombin stimulates the intrinsic pathway by activating factor V and factor VIII.

One of the most striking aspects of the blood clotting cascade is the almost explosive formation of thrombin. In addition, the clotting system is subject to a very precise regulation, which is the result of the fact that the involved enzymes are linked sequentially and that most of the activation reactions are produced on phospholipid surfaces. Phospholipids resulting from the activation of platelets limit the process to the vascular lesion site. There is also a regulation by feedback at the different stages of the coagulation mechanism. The regulation also involves various inhibitors. The LACI factor, which is a KUNITZ type multivalent protease inhibitor, is capable of displacing the extrinsic pathway towards the intrinsic pathway. The plasma also contains other protease inhibitors that can slow down fibrin formation; the most important is antithrombin III, a plasma protein that inhibits thrombin by forming an irreversible complex with it. Antithrombin III belongs to the serpin family; it resembles α_1-antitrypsin, but inhibits thrombin much more strongly than elastase. Antithrombin III also blocks factors XIa, IXa and Xa; its inhibitory action is activated by heparin. Another glycoprotein, activated protein C, plays an important role in the regulation of blood coagulation. It inactivates both factor Va and factor VIIIa by a limited proteolysis in the presence of Ca^{++} and phospholipids. Protein C is activated by protein S. Protein C circulates in blood as a precursor; its activation is the result of the cleavage of the Arg_{12}-Leu bond by thrombin in the presence of a membrane cofactor, thrombomodulin. In this way, thrombin, a clotting factor, is transformed into an anticoagulant factor by thrombomodulin. Most of the coagulation factors, including thrombin, are absorbed into the fibrin clot; in addition, the large quantities of thrombin produced during the clotting cascade stop the activity of factors V and VIII, slowing down in this way an increased thrombin production.

Structural studies of clotting factors have progressed significantly over the past years. The structure of thrombin was determined in 1992 at 1.9 Å resolution (BODE et al., 1992). The B chain of thrombin is very similar to the catalytic domains of factors Xa, IXa and XIa. These factors have, amongst others, a domain that possesses γ-carboxylic residues and two domains that resemble the N-terminal domain of the epidermal growth factor (EGF). The specificity of thrombin is very similar to that of trypsin, but with a clear preference for Arg-X bonds. The three-dimensional structure of factor VIIa, inactivated by a chloromethyl ketone was resolved in 1999 by the HARLOS group. The enzyme consists of four distinct domains, including the catalytic domain.

Factors V and VIII are derived from precursors displaying a similar architecture, A_1-A_2-B-A_3-C_1-C_2, as schematically represented in Fig. 14.15. The association of factor V to the membrane involves the C_2 domain, whose three-dimensional structure was determined in 1999. The domain is folded in a β structure and displays a surface for the stereospecific recognition of the phosphatidyl serine groups of the membrane. The structure of the C2 domain of factor VIII is also a "β sandwich":

one of its motifs, comprising hydrophobic parts and charged residues, suggested a membrane binding mechanism.

The structures of the tissue factor, of fibrinogen and of factor XIII were resolved in 2001. All of these structural data allow a better comprehension of the mechanisms involved in the different steps of blood clotting.

14.4.2.2. THE COMPLEMENT SYSTEM

The complement system can be compared to the blood clotting cascade. It constitutes one of the main ways by which antibodies defend vertebrates against bacterial infections, leading to cellular lysis of the bacteria. The complement is a complex system of proteins, comprising about twenty components in interaction, called C1 (which is a complex of three proteins C1q, C1r, and C1s), C2, C3… to C9, factor B, factor D and a variety of regulatory proteins. These proteins are soluble and circulate in blood and in extra-cellular fluids. Their molecular weights vary between 24 000 and 790 000. The activation of the complement implies a sequential proteolytic cascade. One of the consequences of this activation is the sequential assembly of late components (C5, C6, C7, C8 and C9) in a complex that leads to cellular lysis. The association of these components is triggered by a sequence of proteolyses, involving the early complement factors (C1, C2, C3, C4, factor B and factor D), the majority of which are zymogens as in the clotting cascade. Most of these active components bind strongly to membranes.

The central component of the proteolytic cascade is the C3 component. It can be activated by two different pathways, the classical pathway and the alternative pathway. In both cases, C3 is proteolysed by a complex enzyme, the C3-convertase. A different C3-convertase is produced in each of these pathways, but in both cases, it is formed by the association of two of the components already activated in the cascade. The C3-convertase cleaves the C3 component into two fragments; the largest, C3b, binds to the membrane of the target cell to form an enzymatic complex with a modified specificity, C5-convertase. C5-convertase cleaves the C5 compound, which leads to the spontaneous assembly of the last components C5, C6, C7, C8 and C9, creating in this way the lytic complex, which forms the transmembrane channel of the target cell.

The **classical pathway** is activated when IgG or IgM antibodies bind to the antigens at the surface of the microorganism. It involves the C1, C2 and C4 components; C1 is itself formed of three proteins C1q, C1r and C1s. C1q, the first component of the cascade, is a large oligomeric protein of 790 000 molecular weight. It is a hexamer made of two heterotrimers. Each subunit has a globular head and a stem that has a collagen type structure. Each globular head binds a constant region of the antibody, on condition that it has already bound the antigen. The binding of the antigen-antibody complex activates C1q, which triggers the cascade. C1q activates C1r, which in turn activates C1s by proteolytic cleavage of the Arg426-Ile427 bond. C1s proteolyses sequentially C4, then C2 at a unique Arg-X peptide bond.

These proteases have a modular structure. From the N-terminal extremity, C1s comprises a module called CUB, then a module analogous to the epidermal growth factor, then a second CUB module and two modules of complement control or CPP (for Complement Control Protein). The C-terminal protease domain displays a structure analogous to that of chymotrypsin. A fragment of the human C1s protein comprising the catalytic domain and the CCP2 module (residues 342–406 and 410–668) was obtained by genetic recombination (GABORIAUD et al., 2000). Its three-dimensional structure was resolved at 1.7 Å. The CCP2 module is folded into a very compact hydrophobic core surrounded by six β strands and stabilised by two disulphide bridges. Very rigid, it is oriented perpendicular to the protease forming a sort of handle to it. This relative orientation is in agreement with the fact that the CCP2 module furnishes additional recognition sites for the protein C4. The structure shows that access to the sub-sites of C1s is very restricted compared to what is seen in proteases such as trypsin. This situation is reminiscent of that of the very specific proteases of the blood clotting cascade.

The protein C4 binds to the membrane then associates with C2. The complex C4,2 is the C3-convertase in the classic pathway. C4,2 cleaves C3 in two fragments C3a and C3b; C3b rapidly binds to the membrane close to C4,2 to form the C4,2,3b complex, which is the C5-convertase of this activation pathway. The C5-convertase cleaves C5 in C5a and C5b, and C5b associates to C6, beginning in this way the association of the last components to form the lytic complex (Fig. 14.16).

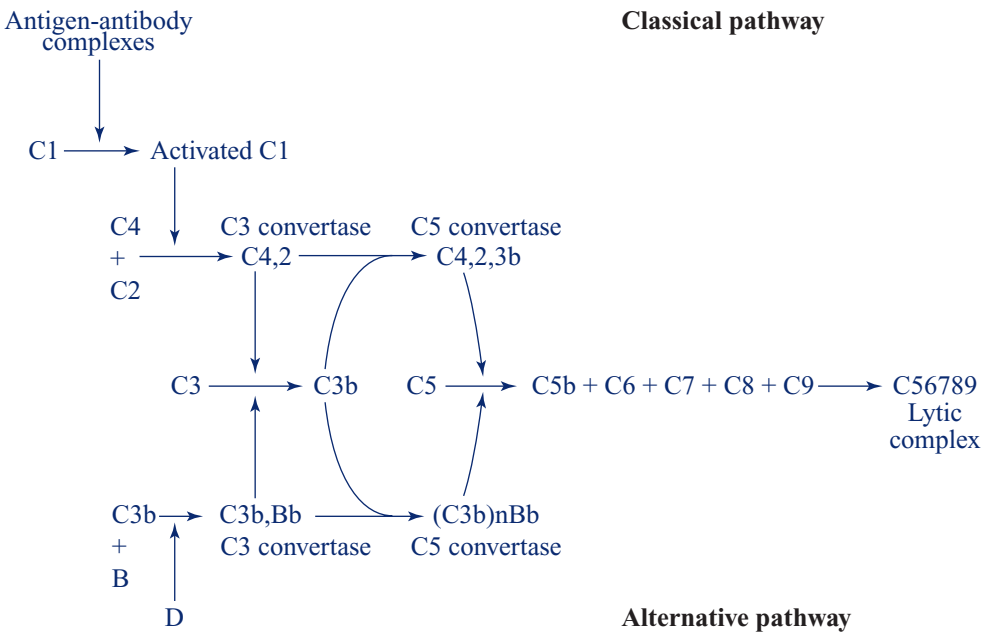


Fig. 14.16 Complement activation

(Reprinted from *Molecular Biology of the Cell*, ALBERTS B. et al. © (1983) Garland pub., Taylor and Francis Group)

The **alternative** or **lectin pathway** acts as positive feedback to amplify the production of C3b (Fig. 14.16). It can also be activated in the absence of antibodies by polysaccharides present in the cellular envelopes of bacteria, yeast and protozoa, ensuring the first line of defence in the organism against infection before the immune response can be initiated. In the lectin pathway, mannose type oligosaccharides associate to a protein, MBL (for Mannose Binding Lectin), which itself is capable of associating with different proteases called MASPs (for MBL Associated Serine Protease), one of which has the same specificity as C1 towards proteins C2 and C4. There is another lectin pathway that recognises oligosaccharides comprising the N-acetyl glucosamine and where picoline intervenes, leading to the activation of MASP2, which produces the same effects as those which are the result of protein C1.

The last act of the complement system is the formation of the lytic complex that starts at the proteolytic cleavage of C5. This complex has a molecular weight of 2 million. It destabilises the lipid bilayer of the bacterial membrane by forming a transmembrane channel as shown in Fig. 14.17. This system is so efficient that the binding of a small number of lytic complexes, or even only one, can be enough to kill a cell.

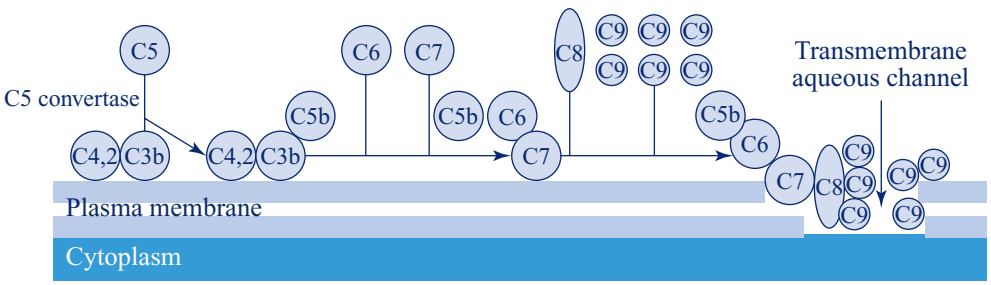


Fig. 14.17 Formation of the lytic complex

(Reprinted from *Molecular Biology of the Cell*, ALBERTS B. et al. © (1983) Garland pub., Taylor and Francis Group)

The complement cascade is submitted to a very strict regulation. The inflammatory and destructive properties that it brings about require that the key components are rapidly inactivated after their formation. The de-activation is ensured in part by the presence of protein inhibitors present in blood, in part by the instability of some of the active components, in particular C4 and C3b, which have lifetimes of less than 0.1 ms. They must therefore be associated to membrane sites very close to the other complement components that activate them.

14.4.3. CYCLIC CASCADES

14.4.3.1. MONOCYCLIC CASCADES

Analysis of a monocyclic cascade system in the steady state

The theoretical analysis of cyclic enzymatic cascades was performed by STADTMAN and CHOCK (1978). These authors showed that such systems are particularly adapted

for the allosteric control of metabolism key enzymes. The essential elements of a monocyclic cascade are represented on Fig. 14.18. The interconvertible enzyme is transformed by the coupling of two irreversible opposite reactions in the cascade system. The reaction in the forward direction is induced by the binding of an allosteric effector e_1 to the inactive form E_i of the enzyme E, which is transformed in the active enzyme E_a; E_a catalyses the chemical modification (phosphorylation) of the interconversion enzyme I, which goes from an original o-I form to a modified m-I form. The regeneration cascade opposes this first cascade; it is induced by the binding of an allosteric effector e_2 to the inactive form, R_i, of another conversion enzyme R, which is transformed into an active enzyme R_a. R_a catalyses the regeneration of m-I into o-I. The dynamical coupling of these cascades in the forward and backward directions leads to the interconversion of o-I into m-I and the transformation of ATP into ADP and P_i in the case of a phosphorylation.

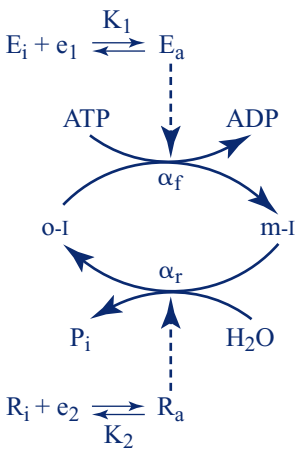


Fig. 14.18 Schematic representation of a monocyclic cascade
The modification is a phosphorylation that results in the transformation of ATP into ADP
(Reprinted from *Curr. Top. Cell. Regul.*, **13**, STADTMAN E.R. & CHOCK P.B., Interconvertible enzyme cascades in metabolic regulation, 53. © (1978) with permission from Elsevier)

- Making the following hypotheses:
  - the formation of the enzyme-enzyme and enzyme-effector complexes is in rapid equilibrium;
  - the concentrations of the enzyme-enzyme complexes are negligible compared to those of the active and inactive enzymes ($[E] = [E_i] + [E_a]$, $[I] = [o\text{-}I] + [m\text{-}I]$, $[R] = [R_i] + [R_a]$ where [E], [I] and [R] are the total concentrations of the enzymes E, I, and R, respectively);
  - the concentrations of the allosteric effectors are maintained constant, the interconversion reactions of o-I and m-I can be written:

$$o\text{-}I + E_a \underset{}{\overset{K_f}{\rightleftharpoons}} o\text{-}I \,.\, E_a \xrightarrow{k_f} E_a + m\text{-}I$$

$$m\text{-}I + R_a \underset{}{\overset{K_r}{\rightleftharpoons}} m\text{-}I \,.\, R_a \xrightarrow{k_r} R_a + o\text{-}I$$

K_f and K_r are the dissociation constants of the complexes and k_f and k_r are the catalytic constants. By writing $\alpha_f = k_f/K_f$ and $\alpha_r = k_r/K_r$, the equation of the steady state is written:

$$[o\text{-}I][E_a]\alpha_f = [m\text{-}I][R_a]\alpha_r$$

The concentrations of E_a and R_a depend on the equilibria dissociation constants:

$$E_a \overset{K_1}{\longleftrightarrow} E_i \,.\, e_1 \quad \text{and} \quad R_a \overset{K_2}{\longleftrightarrow} R_i \,.\, e_2$$

The fraction of interconvertible enzyme in the steady state is given by the following relation:

$$\frac{[\text{m-I}]}{[\text{I}]} = \left[\frac{\alpha_r [R][e_2](K_1 + [e_1])}{\alpha_f [E][e_1](K_2 + [e_2])} + 1 \right]^{-1}$$

In this expression, the role of ATP in the cyclic cascade does not appear because its intracellular concentration is maintained at a constant level, superior by several orders of magnitude to the concentration of the interconvertible enzyme. The expression of [m-I]/[I] is a function of 10 different parameters since α_r and α_f are the ratios of two constants. These parameters include the concentration of allosteric effectors that trigger the cascade. The result is that the interconvertible enzymes are very sensitive to variations in the concentrations of multiple metabolites, and display a higher flexibility in their control system than other types of regulatory enzymes.

Figure 14.19 illustrates three important characteristics of cyclic cascade systems:

- the maximal quantity of interconvertible enzyme that can be modified in the presence of a saturating concentration of the first effector e_1, can vary (curves d and e);
- the effector concentration necessary to produce a significant modification of the interconvertible enzyme can be much weaker than that required to bring about a comparable activation of the conversion enzyme (curves a and c);
- the increment of effector concentration necessary to produce a certain change in the ratio [m-I]/[I] can vary (curves b and c).

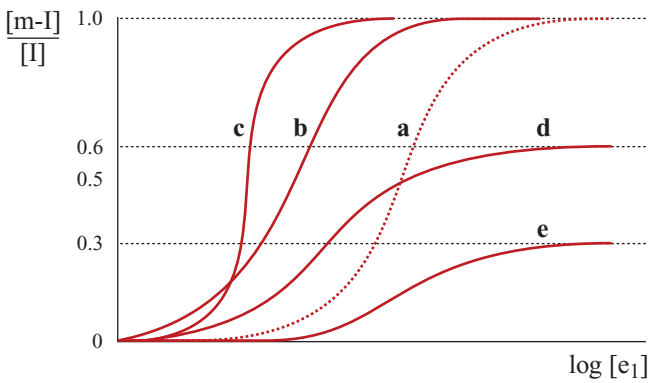


Fig. 14.19 Variation in the production of the interconvertible enzyme as a function of effector e_1 concentration (curves b, c, d, e)

The curve (a) represents the fractional activation of the conversion enzyme. The curves were simulated for different values of the parameters (Reprinted from *Curr. Top. Cell. Regul.*, **13**, STADTMAN E.R. & CHOCK P.B., Interconvertible enzyme cascades in metabolic regulation, 53. © (1978) with permission from Elsevier)

Taking into account these three properties, several parameters are defined to characterise cascade systems, the amplitude (M), the signal amplification factor (SA) and the sensitivity (S). **Amplitude** is the maximal value of the ratio [m-I]/[I] that can be obtained for saturating effector concentrations. In Fig. 14.19 the M values for the curves b, d and e are respectively 1.0, 0.6 and 0.3. The **signal amplification factor** SA is the ratio of effector concentration necessary for an activation of 50% of the conversion enzyme E, to the effector concentration for which the ratio [m-I]/[I] is 50%: $SA = [e]_{0.5E}/[e]_{0.5I}$. It is clear that in the absence of a significant concentration of the conversion enzyme R, or in the absence of effector e_2, the amplification factor of the signal becomes infinite. The **sensitivity** is defined as the ratio: $8.89\ [e_{0.5M}]/([e_{0.9M}] - [e_{0.1M}])$, expression in which $[e_{0.5M}]$, $[e_{0.9M}]$ and $[e_{0.1M}]$ represent the effector concentration necessary to attain 50, 90 and 10% of the max-imal amplitude, respectively; the factor 8.89 was introduced in reference to the value of 1 for the hyperbolic functions as in the MICHAELIS law. The index of sensitivity S is comparable to HILL's number for allosteric enzymes. However its meaning differs slightly. The sigmoidal responses observed in cascade systems do not reflect the cooperativity in the effector binding to the enzyme's sites, but are due to the fact that the effector influences more than one step in the cascade. In the same way, values of sensitivity less than 1 do not reflect anticooperativity of effector binding to an enzyme, but are the consequence of antagonistic effects on more than one step of the cascade.

The cascade system presented in Fig. 14.18 represents a case where the effectors of the conversion enzymes E and R are different. Supposing that the effector e_1 is always required for the activation of E, four different diagrams of monocyclic regulation cascades are presented in Fig. 14.20 opposite, illustrating the wide variety of regulation mechanisms. Case 1 is identical to that of Fig. 14.18. In case 2, the effector e_2 inactivates the enzyme R. In case 3, the same effector e_1 activates the enzyme E, but inhibits the conversion enzyme R. In case 4, the effector e_1 activates both conversion enzymes. The equations of the steady state are different.

In case 2, the equation becomes:

$$\frac{[\text{m-I}]}{[\text{I}]} = \left[\left(\frac{K_1 + [e_1]}{K_2 + [e_2]}\right)\left(\frac{\alpha_r [R][e_2]}{\alpha_f [E][e_1]}\right) + 1\right]^{-1}$$

In case 3:

$$\frac{[\text{m-I}]}{[\text{I}]} = \left[\left(\frac{K_1 + [e_1]}{K_2 + [e_2]}\right)\left(\frac{\alpha_r [R]K_2}{\alpha_f [E][e_1]}\right) + 1\right]^{-1}$$

And in the fourth case:

$$\frac{[\text{m-I}]}{[\text{I}]} = \left[\left(\frac{K_1 + [e_1]}{K_2 + [e_2]}\right)\left(\frac{\alpha_r [R]}{\alpha_f [E]}\right) + 1\right]^{-1}$$

Figure 14.21 opposite illustrates the behaviour of these different cases of monocyclic cascade systems; the curves are simulated with the following parameter values: $K_1 = 10^{-5}$, $K_2 = 10^{-4}$, $[e_2] = 10^{-4}$, $[E] = [R]$ and $\alpha_f = \alpha_r$.

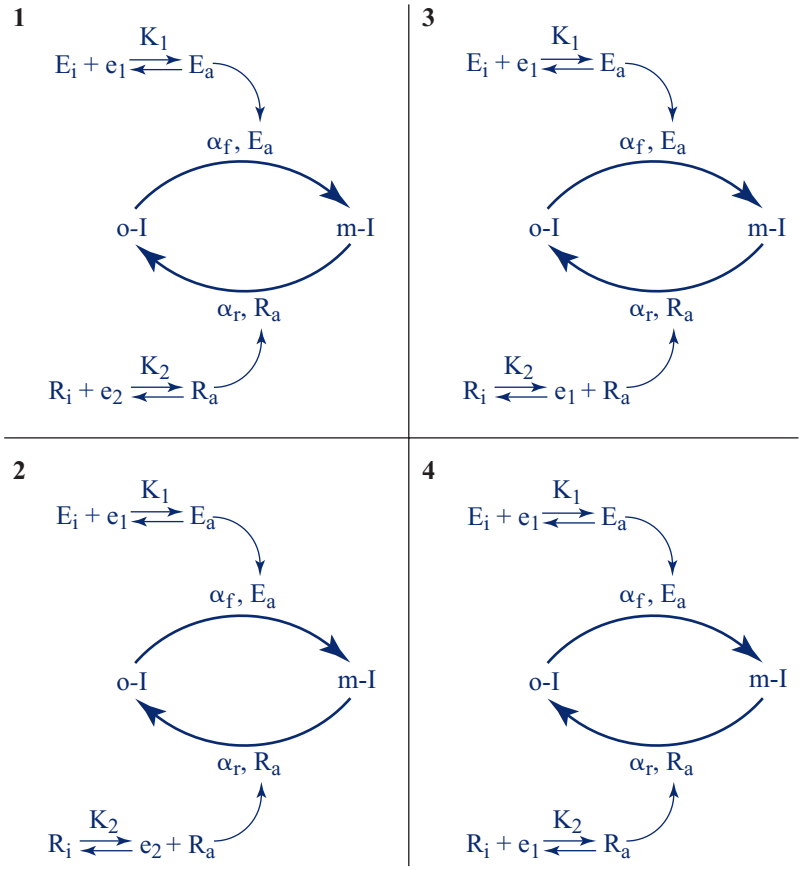


Fig. 14.20 Four different cases of monocyclic cascades

For simplicity, ATP is not represented on the diagrams. Case 1 is identical to that in Fig. 14.18 (Reprinted from *Curr. Top. Cell. Regul.*, **13**, STADTMAN E.R. & CHOCK P.B., Interconvertible enzyme cascades in metabolic regulation, 53. © (1978) with permission from Elsevier)

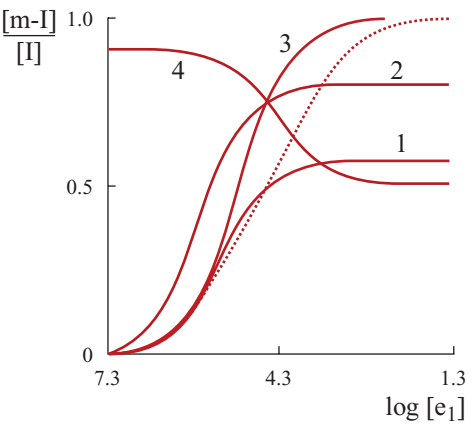


Fig. 14.21 Variation of the proportion of interconvertible enzyme as a function of effector e_1 concentration in the different cases

The curves are simulated with the given values for the parameters (Reprinted from *Curr. Top. Cell. Regul.*, **13**, STADTMAN E.R. & CHOCK P.B., Interconvertible enzyme cascades in metabolic regulation, 53. © (1978) with permission from Elsevier)

In the theoretical studies, it was supposed that the cyclic cascade systems are devoid of chemical flux. However, each complete cycle of phosphorylation-dephosphorylation is associated with the decomposition of ATP in ADP and P_i. In the steady state, it is legitimate to not take this into account since the concentrations of nucleotides are maintained at practically constant levels, several orders of magnitude higher than the concentrations of interconvertible enzymes. ATP hydrolysis is however necessary to maintain a defined steady state of the interconversion enzyme. The ATP flux, J_{ATP}, corresponding to the ATP concentration per unit of time, depends on the concentration of all the enzymes and effectors of the cyclic cascade, following the relation:

$$J_{ATP} = \frac{\alpha_f \alpha_r [e_1][e_2][E][R][I]}{\alpha_r [R][e_2](K_1 + [e_1]) + \alpha_f [E][e_1](K_1 + [e_2])}$$

It is clear that for a given steady state, the rate of ATP decomposition can vary *via* the interaction between allosteric effectors and conversion enzymes. ◢

A monocyclic cascade: pyruvate dehydrogenase

Pyruvate dehydrogenase represents an example of a monocyclic cascade. This enzyme catalyses the transformation of pyruvate into acetyl-CoA following the reaction:

$$\text{pyruvate} + \text{CoA} + \text{NAD}^+ \longleftrightarrow \text{acetyl-CoA} + \text{NADH} + \text{CO}_2$$

Its activity is modulated by phosphorylation-dephosphorylation reactions. The phosphorylation by a specific kinase leads to the inactivation of the enzyme, while the dephosphorylation by a highly specific phosphatase reactivates the enzyme. The activities of the kinase and phosphatase are under the control of allosteric effectors comprising mono- and divalent cations, pyruvate, coenzyme A, acetyl-CoA, ATP, ADP, NAD and NADH. Three of the four diagrams of the regulation cascade represented on Fig. 14.22 opposite are used in the regulation cascade of pyruvate dehydrogenase of mammals. The acetyl-CoA activates the kinase and Ca^{++} ions activate the phosphatase; these effectors have identical roles to e_1 and e_2 in case 1. NADH activates the kinase, but inhibits the phosphatase and plays the role of e_1 in case 3. Mg^{++} however activates at the same time the two conversion enzymes and plays the same role as e_1 in case 4. This example shows the complexity of regulation mechanisms that can be generated by cascade systems.

14.4.3.2. BICYCLIC CASCADE SYSTEMS

When the modified form of the interconvertible enzyme from one cycle catalyses the covalent modification of the interconvertible enzyme from another cycle, the two cycles become coupled. The fractional modification of the interconvertible enzyme from the second cycle is a function of all the parameters of the two cycles. Compared to monocyclic cascades, these systems display an increased flexibility in allosteric control and a higher capacity of response amplification to the initial stimuli.

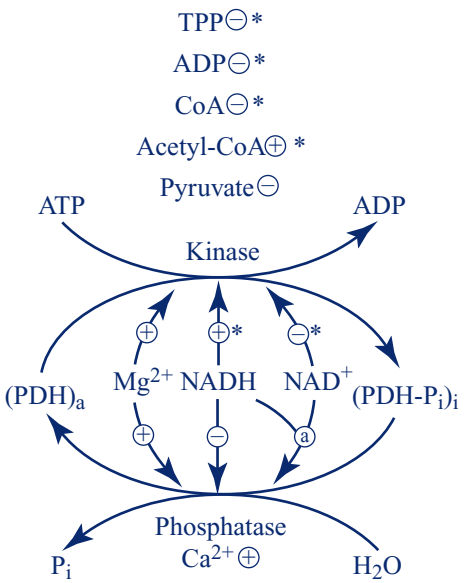


Fig. 14.22 Role of effectors in the covalent interconversion of pyruvate dehydrogenase from mammals
Asterix indicate that the effect is observed only in the presence of K^+ or NH_4^+ ions. (a) indicates the antagonism between NAD^+ and NADH. The + and – signs indicate activation and inhibition, respectively
(Reprinted from *Curr. Top. Cell. Regul.*, **13**, STADTMAN E.R. & CHOCK P.B., Interconvertible enzyme cascades in metabolic regulation, 53. © (1978) with permission from Elsevier)

Two types of bicyclic cascades are involved in the regulation of metabolism key enzymes: open bicyclic cascades and closed bicyclic cascades (Fig. 14.23 below).

Open bicyclic cascades

◤ Such a system is illustrated in Fig. 14.23a. The first cycle of this model is identical to the case 1 of the monocyclic system of Fig. 14.20; to simplify, the role of ATP is not shown. The second cycle differs from the first in that the modified chemical form of the interconvertible enzyme m-I_1, becomes the conversion enzyme that catalyses the modification of the o-I_2 form of the second interconvertible enzyme I_2 into its m-I_2 form. The authors make the same hypotheses as for monocyclic cascades and assume that the binding of the allosteric effectors e_1, e_2 and e_3 to the inactive forms E_i, R_{1i} and R_{2i} of the conversion enzymes is necessary to generate their respective active forms. In these conditions, the fractional modification of the enzyme I_2 in the steady state is written:

$$\frac{[\text{m-I}_2]}{[\text{I}_2]} = \left[\frac{\alpha_{1r}\alpha_{2r}(K_1+[e_1])[R_1][R_2][e_2][e_3]}{\alpha_{1r}\alpha_{2f}(K_2+[e_2])(K_3+[e_3])[E][I_1][e_1]} + \frac{\alpha_{2r}[R_2][e_3]}{\alpha_{2f}[I_1](K_3+[e_3])} + 1 \right]^{-1}$$

This equation shows that the $[\text{m-I}_2]/[\text{I}_2]$ value is a function of 18 parameters. From simulated curves for given parameter values, STADTMAN and CHOCK (1978) showed that an amplification of 320 times obtained in the first cycle in response to a stimulation by e_1, leads to an amplification of 102 400 times in the second cycle.

The cascade involved in glycogen phosphorylase regulation is an example of an open bicyclic cascade (Fig. 14.24 below). This cascade is triggered by the binding of c-AMP to the inactive form of a protein kinase. The protein kinase, once activated, catalyses

the phosphorylation of the phosphorylase kinase, which becomes active and, in turn, catalyses the conversion of phosphorylase b into phosphorylase a. The cascade comprises two phosphorylation cycles; each phosphorylation reaction is reversed by a specific phosphatase. The structural aspects of this system were treated in detail previously (see Sect. 14.3.1).

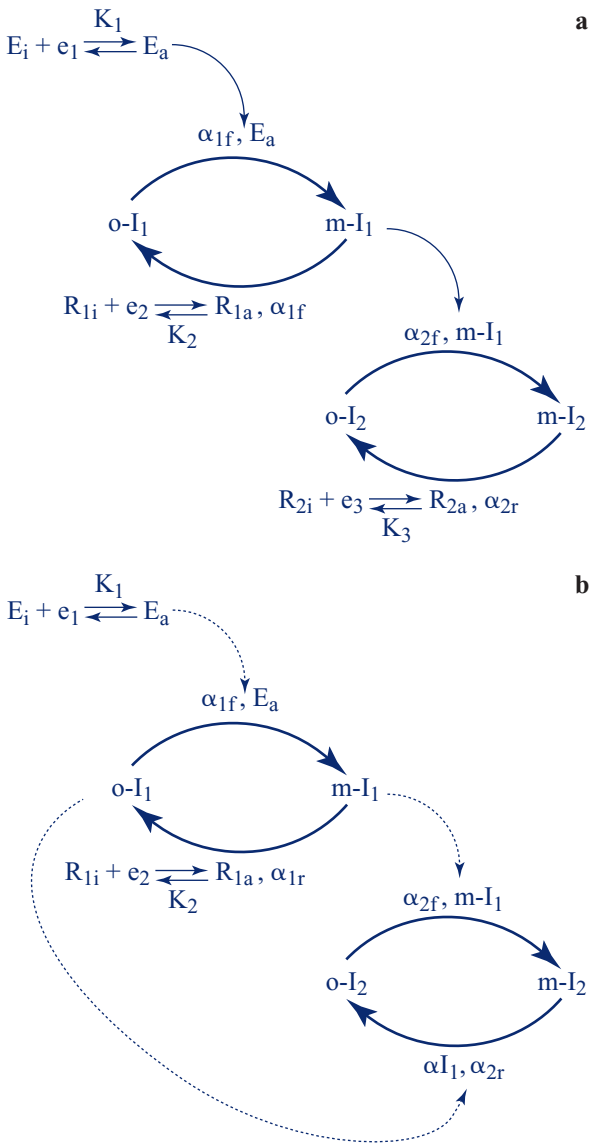


Fig. 14.23 Bicyclic cascades

(a) open – (b) closed (Reprinted from *Curr. Top. Cell. Regul.*, **13**, STADTMAN E.R. & CHOCK P.B., Interconvertible enzyme cascades in metabolic regulation, 53. © (1978) with permission from Elsevier)

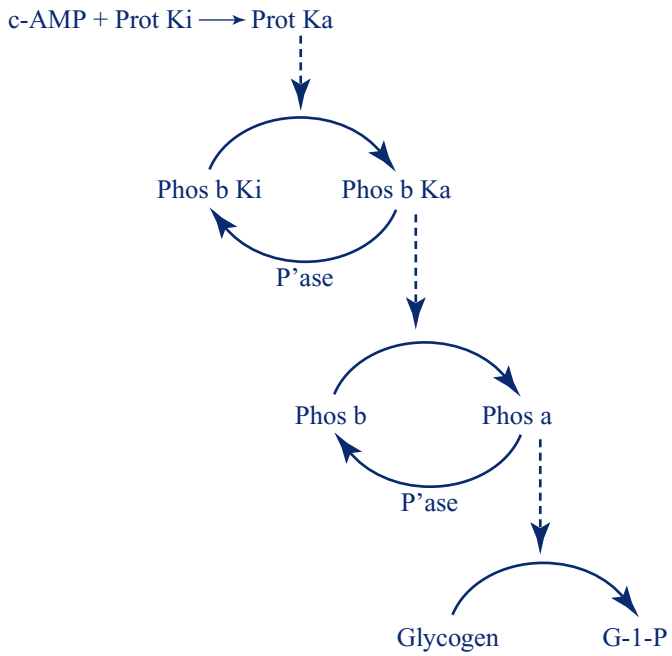


Fig. 14.24 The glycogen phosphorylase cascade

(Reprinted from *Curr. Top. Cell. Regul.*, **13**, STADTMAN E.R. & CHOCK P.B., Interconvertible enzyme cascades in metabolic regulation, 53. © (1978) with permission from Elsevier)

Closed bicyclic cascades

Figure 14.23b illustrates a closed bicyclic cascade. In this system the interconvertible enzyme o-I_1 intervenes in its original form for the conversion of m-I_2 into o-I_2. The fractional modification of I_2 in the steady state is determined by the ratio [m-I_1]/[o-I_1] rather than the total concentration of I_1. By making the same hypotheses as previously, the fractional modification of I_2 in the steady state is given by the following relationship:

$$\frac{[\text{m-I}_2]}{[\text{I}_2]} = \left[\frac{\alpha_{1r}\alpha_{2r}[\text{R}_1][\text{e}_2](\text{K}_1+[\text{e}_1])}{\alpha_{1f}\alpha_{2f}[\text{E}][\text{e}_1](\text{K}_2+[\text{e}_2])}+1\right]^{-1}$$

This equation depends on 14 parameters instead of 18 in an open bicyclic system, showing that the closed bicyclic cascades are less flexible than the open bicyclic cascades.

The regulation system of glutamine synthetase from *E. coli* offers an example of a closed bicyclic cascade. It was described in detail in Section 14.3.4.

14.4.3.3. MULTICYCLIC CASCADE SYSTEMS

Bicyclic cascades display an increased flexibility in allosteric control and a larger amplification capacity than monocyclic cascades. These characteristics increase with the number of cycles. To generalise, STADTMAN and CHOCK (1978) performed the

theoretical treatment of the steady state of a cascade with n cycles. We will not give the obtained complex equation. To give some idea, **it is interesting to mention that the authors, using the obtained curves simulated for given parameter values, showed that the amplification factors for one, two and four cycles are respectively 9, 83, 833 and 8333**. As a consequence of the amplification effects obtained when several parameters of the cascade are simultaneously altered, the interconvertible enzymes can respond to effector concentrations well below the dissociation constant of the effector-allosteric enzyme complexes. ◢

Cascade systems constitute regulation mechanisms which are rigorous and extremely flexible, capable of integrating a large number of signals and of amplifying them considerably. The few examples of cyclic cascades presented above are involved in the regulation of metabolism key enzymes. Such mechanisms can also be involved in other biological processes for which signal amplification is important, for instance sensory perception, nerve transmission, and responses to hormonal action.

14.5. IRREVERSIBLE INACTIVATIONS

Cellular physiology and some differentiation processes require that at well defined stages certain enzymatic activities disappear. These phenomena bring into play rapid inactivation processes, followed sooner or later by proteolytic degradation of the inactivated enzymes.

14.5.1. PROTEASOMES

Mechanisms implicated in the intracellular degradation of proteins have been the subject of important progress (COUX et al., 1996). In eukaryotic cells, the proteasome is responsible for the degradation of most of the cell's proteins and is necessary for the cell's viability. Proteasomes are present in the nucleus, the cytosol and some are sometimes found associated with the endoplasmic reticulum and the cytoskeleton. The proteasome is an essential element of ATP-dependent proteolytic degradation; it catalyses the rapid degradation of many enzymes, transcription regulators and crucial regulatory proteins such as cyclins. It is also essential for the rapid elimination of abnormal proteins, which are the result of either mutations or post-translational damage. Proteasomes degrade proteins into short peptides that are then rapidly hydrolysed by cytoplasmic exopeptidases.

There are two types of proteasomes that are large complexes formed of several subunits, the 20s proteasome (700–705 kDa) and the 26s proteasome (2 000 kDa). The latter selectively degrades ubiquitin conjugated proteins following a mechanism that depends on ATP. In eukaryotic cells, most of the cytosol and nucleus proteins are degraded *via* the ubiquitin-proteasome system. The 26s complex contains the 20s complex as its catalytic core.

14.5.1.1. 20S PROTEASOME

20s proteasomes are ubiquitous; they are found in all species, prokaryotes, eukaryotes and archaea. They all contain 28 subunits, but depending on the species their organisation differs in its complexity. Sequences of more than 70 subunits of proteasomes from different species have been determined. They are classed in two groups α and β that have only a small sequence similarity; it seems however that they derived from a common ancestor during evolution. The subunits have molecular masses between 20 and 35 kDa. Proteasomes from prokaryotes are constituted of two copies of 7 different α subunits and 7 different β subunits; they have the topology (α7β7β7α7). Those from eukaryotes have a more complex arrangement (α1–α7)(β1–β7)(β1–β7)(α1–α7).

The first structural data were obtained by electron microscopy; they showed that the 20s proteasomes have a cylindrical shape constituted of four overlaid rings forming an internal cavity. β subunits form the two internal rings and the α subunits the two external rings, and are organised following a 7 fold symmetry. The three-dimensional structure of the proteasome from an archaebacteria, *Thermoplasma acidophilum*, was determined by X-ray crystallography; the structure was resolved by the HUBER group at 3.4 Å (LÖWE et al., 1995). The structure from *Saccharomyces cerevisiae* was resolved to 2.4 Å in 1997 (GROLL et al., 1997). The composition of the 20s proteasome from eukaryotes is a little more complex, the 14 subunits being different. Almost the same type of structural organisation is found however, with a pseudo 7 fold symmetry. The four rings surround three internal compartments, two "antechambers" and a central proteolytic chamber formed by the β subunits. In this proteasome, the entrance is obstructed by the N-terminal extremities of the α subunits, suggesting that the association with the 19s regulatory complex is necessary to render the interior of the complex accessible. The α and β subunits have similar folds, comprising two β sheets consisting of five antiparallel strands, flanked by two helices on one side and three on the other (Fig. 14.25).

The catalytic mechanism of the proteasome seemed mysterious for a long time. Now, several approaches have demonstrated that the nucleophile of the active site is the OH group of the N-terminal threonine of the β subunit. The replacement by mutagenesis of this threonine by an alanine suppresses proteolytic activity; its replacement by a serine conserves it. This threonine is conserved in all species. In addition, the deletion of the propeptide is essential to enzymatic activity. The N-terminal amino group probably intervenes to capture the proton of the threonine. The residues Glu17 and Lys33 are also essential to proteolytic activity. Residues Ser129, Ser169 and Asp166 are required for structural integrity, but are not involved in catalysis. The proteasome possesses trypsin, chymotrypsin and peptidyl glutamyl peptidase activities, each of which is carried by a specific β subunit. The distribution of the peptide size after proteolysis varies between 4 and 25 residues, with an average of 7 to 9.

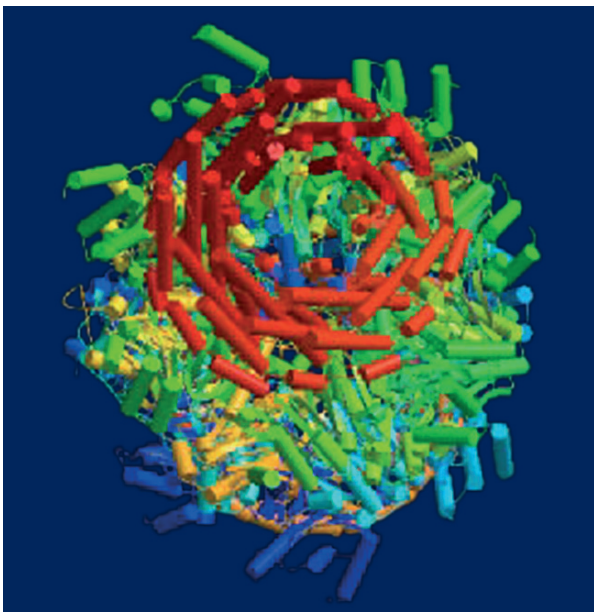


Fig. 14.25 Schematic representation of the three dimensional structure of the proteasome seen from the top of the molecule

In the cell, regulation mechanisms have evolved to allow a selective proteolysis and to avoid a generalised degradation of all proteins. In eukaryotic organisms, the association of the 19s complex allowing the conjugation with ubiquitin of the substrate ensures a selective degradation. The degradation by the proteasome of small peptides, but not denatured or ubiquitinylated proteins, is stimulated by an activator protein, PA28 or regulator 11s of molecular mass of 28 kDa. There are two different but homologuous molecules referred to as PA28α and PA28β found in PA28 complex. This complex is a heterohexamer composed of both α and β subunits with a stoichiometry α3β3 in alternating order (AHN et al., 1996). It binds to the extremities of the 20s proteasome. In addition, some cellular inhibitors maintain the proteasome in an inactive form. Several of them possess a therapeutic potential as anti-inflammatory agents.

14.5.1.2. 26S PROTEASOME

In eukaryotic cells, the degradation of the majority of proteins requires their initial modification by conjugation with ubiquitin. Proteins conjugated with ubiquitin are quickly degraded by the 26s proteasome. The 26s complex is present in the nucleus and in the cytosol. It is the result of the association of the 20s proteasome with the PA700 activator (also called the 19s complex), which is ATP dependent and constituted of 18 proteins whose molecular mass varies between 25 and 110 kDa. The complex binds to the two extremities of the 20s proteasome in opposite orientations, showing a 2 fold rotational symmetry. The 26s proteasome is a fairly labile com-

plex, which explains why it was only discovered later. It is stabilised by ATP. In addition, it comprises several binding sites for chains conjugated by ubiquitin. The basic complex comprises six subunits displaying a strong ATPase activity and two large subunits. It seems that these ATPases form a ring at the interface between the 20s proteasome and the 19s complex. The role of these ATPases is to promote the assembly of the two complexes. In addition, they serve to open the complex and to assist the translocation of denatured proteins into the central cavity of the 20s proteasome. Some of these ATPases are only involved in the degradation of particular substrate types. The 26s proteasome can be easily dissociated and reassembled, a process which happens frequently under physiological conditions. The assembly requires the presence of ATP and CTP. The degradation of proteins conjugated by ubiquitin also requires the presence of ATP, which can be partially replaced by CTP, GTP or UTP. The two processes are catalysed by different ATPases.

Before proteolytic degradation, proteins are conjugated with ubiquitin, a small 76 amino acid protein, which exists in all eukaryotic cells and which has been extremely conserved during evolution. Its three-dimensional structure was resolved: it is of the αβ roll type, which belongs to the superfolds defined by the THORNTON group, like ferredoxin and G protein. The protein has a N-terminal methionine and a C-terminal glycine. The conjugation of proteins with ubiquitin occurs by the formation of an isopeptide bond between the C-terminal of the Gly76 and the ε-amino group of a lysine of the substrate protein. These associations require three types of enzymes, an ubiquitin activating protein E1, an ubiquitin transport protein E2 and an ubiquitin ligase E3. In several subsequent steps, the same reaction occurs on ubiquitin molecules which are bound by Lys48. Several ubiquitin molecules are thereby bound to the protein to be degraded. This reaction requires the ATP hydrolysis. The reaction can be reversed by enzymes that catalyse the release of ubiquitin. The degradation of ubiquitin conjugated proteins involves the release of the ubiquitin molecules catalysed by isopeptidases, which are thiol proteases used to recycle ubiquitin. These activities reside in components of the 19s complex. There are at least fifteen constituents in yeast. Proteolysis requires ATP hydrolysis. Figure 14.26 below sums up the degradation mechanism of proteins by the 26s proteasome. About 90% of proteins are degraded by this system.

The inhibition of proteases by the serpin type inhibitors (see Sect. 14.2.1.2) represents another example of irreversible inhibition.

14.5.2. CASPASES AND APOPTOSIS

Apoptosis, which consists of a regulated destruction of cells, is a very complex process that is not yet completely understood. Apoptosis requires coordinated activation and the execution of multiple sub-programs. The majority of the morphological changes result from the action of cysteinyl proteases, the caspases. Over the past years, a whole family of caspases (for **c**ysteinyl **asp**artate specific prote**ases**), was

identified. These "killing" proteases play a fundamental role in apoptosis or programmed cellular death, which is an essential process for the maintenance of integrity and homeostasis in multicellular organisms.

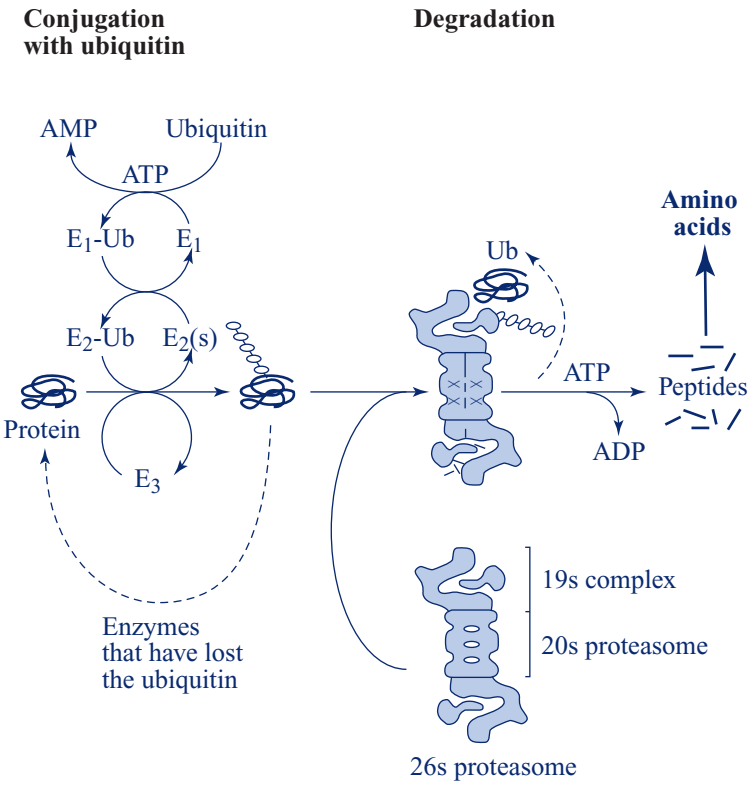


Fig. 14.26 Degradation pathway of proteins, dependent on ubiquitin
The substrate protein is first poly-conjugated with ubiquitin in a reaction that involves three types of proteins: the ubiquitin activating protein E1, the ubiquitin transporter E2 and the ubiquitin ligase E3. Isopeptidases can catalyse the reverse reaction. After ubiquitin binding, the substrate is quickly degraded into small peptides by the 26s proteasome in a reaction that depends on ATP hydrolysis (Reprinted, with permission, from the *Annu. Rev. Biochem.*, **65**, COUX O. et al., 801. © (1996) by Annual Reviews www.annualreviews.org)

Even though the exact details of the pathways of programmed cellular death are not yet completely determined, it has been established by biochemical and genetic approaches that caspases play an essential role in different steps of apoptosis. Figure 14.27 opposite is a schematic representation of the principal apoptotic pathways. These pathways are common to most mammal cells.

In the **extrinsic pathway**, the apoptotic signal is triggered by the trimerisation of receptors, induced by the binding of the "ligands of death". By the intermediate of adaptation proteins, this induces the activation of caspases 8 or 10 and consecutively that of caspases 3, 7 and maybe 6, which are responsible for proteolytic processes

characteristic of cell destruction. The ***intrinsic pathway*** is activated by genomic alterations caused by mutagenic or pharmacologic agents, or even by ionising radiation. In the intrinsic pathway, cytochrome c is liberated from mitochondria and procaspase 9 is activated in its complex with Apaf-1 (for Apoptotic protease activating factor). Caspase 9 activates in turn caspases 3, 7 and maybe caspase 6. Alternatively, caspase 12, localised in the endoplasmic reticulum can be activated by calcium, leading to the activation of destructive caspases and the limited proteolysis of substrates. The ***cytokine pathway*** leading to inflammation is activated by caspase 1.

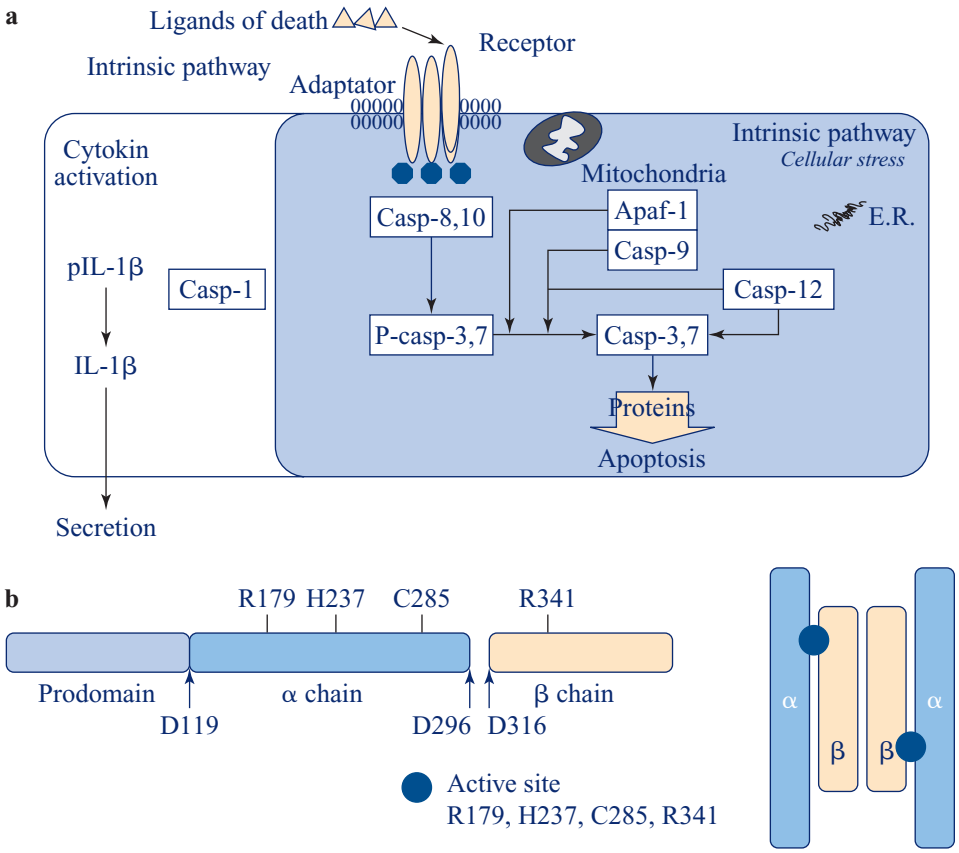


Fig. 14.27 (a) role of caspases in different apoptotic pathways; E.R.: endoplasmic reticulum – (b) schematic representation of caspase activation

(Reprinted from *Curr. Op. Struct. Biol.*, **10**, GRÜTTER M.G., Caspases: key players in programmed cell death, 649. © (2000) with permission from Elsevier)

The two pathways converge at the level of the procaspase 3 activation. Upstream from caspase 3, the apoptotic program is sub-divided in a multitude of sub-programs. The targets of caspases during apoptosis are enzymes implicated in genome function, such as poly(ADP-ribose)polymerase, the catalytic subunit of the DNA

dependent protein kinase, as well as ribonucleoproteins, some cellular cycle regulators, and structural proteins from the nucleus and the cytoskeleton. Caspases are also involved in the incorrect apoptotic phenomena that occur in the etiology of dangerous pathologies such as neurovegetative disorders and cancer (for a review see NICHOLSON & THORNBERRY, 1997). The fragmentation of DNA is the result of this process.

Caspase 1, or mammal interleukin β conversion enzyme (ICE), and caspase 3, product of the gene CED3 necessary for apoptosis in the nematode *Caenorhabditis elegans* were amongst the first members of this family to be discovered. To date, ten caspases that are subdivided into two or three groups were identified in man. The members of the first group (including caspase 1) are for the most part involved in inflammatory processes. The other caspases (including caspase 3) are almost exclusively implicated in apoptosis. All caspases derive from a proenzyme whose organisation is comparable, however the pro-domains are of different lengths. At the time of maturation, the pro-domain is cleaved by a protease, as well as a short internal segment that is eliminated leading to the formation of a large and a small subunit; the resulting active caspase is a heterodimer (see Fig. 14.27b). All the maturation sites are Asp-X bonds. The activation process in vivo seems to be produced by at least two different mechanisms. Some results suggest that caspases are the result of a proteolytic cascade and are activated either by other caspases, or by proteases of similar specificity. In addition, the protease granzyme B of the cytotoxic T-lymphocytes, which is a serine protease of structure analogous to that of chymotrypsin, is capable of activating several caspases in vitro. Lastly, some arguments suggest that several caspases could be activated by autolysis.

The three-dimensional structures of caspases 1, 3, 8 and 7 were resolved by X-ray crystallography. They display high similarities. The large and small subunits are closely associated, each of them bearing residues that form the catalytic centre. The comparison of primary structures of all the caspases show that the residues implicated in catalysis (Cys285 close to His237) are conserved. Each heterodimer comprises six β strands, five of which are parallel and one anti-parallel, which form a β sheet flanked on one side by two α helices and on the other by three α helices. The cavity of the active site is found at the C-terminal extremity of parallel β strands in each heterodimer. The heterodimers are associated following a 2 fold axis. All the caspases show the same type of fold.

The catalytic mechanism proceeds *via* the formation of a tetrahedral intermediate that requires the activation of Cys285, which is polarised by the imidazole group of His237. The oxyanion that forms in the transition intermediate is stabilised by hydrogen bonds with amide protons of the polypeptide chain of Cys285 and Gly238. His237 participates in the formation of the acyl-enzyme by a general acid catalysis. The oxygen of the Arg177 carbonyl (or the equivalent in other caspases) can be used as a third component of the catalytic triad of caspases. Four residues appear to be implicated in the stabilisation of Asp in the P1 position of the substrate. A few

differences are apparent between caspases 1 and 3 in the substrate binding in position P4, which could be responsible for their different specificities.

These different modes of covalent regulation underline once again the multiple mechanisms that cells can use to ensure their survival and to respond to the numerous aggressions to which they may be subjected. The diversity of these covalent and non-covalent regulation modes, developed during evolution with all their complexity, is at the basis of homeostasis. Any dysfunction of these modes of regulation can lead to pathological consequences.

BIBLIOGRAPHY

BOOKS

ALBERTS B. et al. –1983– *Molecular Biology of the Cell*, Garland publisher, New York.

GENERAL REVIEWS

COUX O., TANAKA K. & GOLDBERG A.L. –1996– Structure and function of the 20s and 26s proteasome, in *Annu. Rev. Biochem.* **65**, 801.

GRÜTTER M.G. –2000– Caspases: key players in programmed cell death, in *Curr. Op. Struct. Biol.* **10**, 649–655.

HAYES B.K. & HART G. –1994– Novel forms of protein glycosylation, in *Curr. Op. Struct. Biol.* **4**, 692–696.

HENGARTNER M.O. –2000– The biochemistry of apoptosis, in *Nature* **407**, 770–776.

LASKOWSKI M. JR & KATO I. –1980– Protein inhibitors of proteinases, in *Annu. Rev. Biochem.* **49**, 593–626.

STADTMAN E.R. & CHOCK P.B. –1978– Interconvertible enzyme cascades in metabolic regulation, in *Curr. Top. Cell. Regul.* **13**, 53–95.

SPECIALISED ARTICLES

AHN K., ERLANDER M., LETURCQ D., PETERSON P.A., FRÜH K & YANG Y. –1996– *J. Biol. Chem.* **271**, 18237.

ALMASSY R.J., JANSON C.A., HAMLIN R., XUONG N.H. & EISENBERG D. –1986– *Nature* **323**, 304.

BARFORD D., HU S.H. & JOHNSON L.N. –1991– *J. Mol. Biol.* **218**, 233.

BODE W. & HUBER R. –1992– *Eur. J. Biochem.* **204**, 433.

CARRELL R.W., EVANS D.L. & STEIN P.E. –1991– *Nature* **353**, 576.

CHAMBON P., WEILL J.D., DOLY J., STROSSER M.T. & MANDEL P. –1966– *Biochem. Biophys. Res. Commun.* **25**, 638.

DAVIE E.W., FUJIKAWA K. & KISIEL W. –1991– *Biochemistry* **30**, 10363.

FISHER E.H. & KREBS E.G. –1955– *J. Biol. Chem.* **216**, 121.

GAAL J.C. & PEARSON C.K. –1986– *Trends Biochem. Sci.* **11**, 171.

GABORIAUD G., ROSSI V., BALLY I., ARLAND G.J. & FONTECILLA-CAMPS J.C. –2000– *Embo. J.* **19**, 1755.

GROLL M., DITZEL L., STOCK D., BOCHTLER M., BARTUNIK H. & HIBER R. –1997– *Nature* **386**, 463.

HANKS S.K. & QUINN A.M. –1991– *Methods Enzymol.* **200**, 38.

HERVÉ M. & GHÉLIS C. –1991– *Arch. Biochem. Biophys.* **285**, 142.

HUNTER T. –1991– *Methods Enzymol.* **200**, 3.

KNIGHTON D.R., ZHENG J., TEN EYCK L.F., ASHFORD V.A., XUONG N.H., TAYLOR S.S. & SOWADSKI J.M. –1991– *Science* **253**, 407 and 414.

KOBATA A. –1995– Glycobiology, in *Molecular Biology and Biotechnology*, R.A. Meyers ed., New York, 382–385.

LÖWE J., STOCK D., JAP B., AWICKER P., BAUMEISTER W. & HUBER R. –1995– Science **268**, 533.

MARCHASE R.B., BOUNELIS P., BRUMLEY L.M., DEY N., BROWNE B., AUGER D., FRITZ T.A., KULESZA P. & BEDWELL D.M. –1993– *J. Biol. Chem.* **268**, 8341.

NICHOLSON D.W. & THORNBERRY N.A. –1997– *TIBS* **22**, 299.

RAY A.K., DATTA B., CHAKRABORTY A., CHATTOPADHYAY A., MEZA-KEUTHER S. & GUPTA N.K. –1992– *Proc. Natl. Acad. Sci. USA* **89**, 539.

SHALL S. –1995– *Biochimie* **77**, 313.

SPRANG S.R., ACHARYA K.R., GOLDSMITH E.J., STUART D.I., VARVILL K., FLETTERICK R.J., MADSEN N.B. & JOHNSON L.N. –1988– *Nature* **336**, 215.

15 – MULTIFUNCTIONAL ENZYMES, MULTI-ENZYMATIC COMPLEXES AND METABOLIC CHANNELLING

The existence of multifunctional enzymes and of multi-enzymatic complexes that catalyse consecutive reactions of a metabolic pathway is sometimes accompanied by important functional consequences for enzymatic activity regulation. The organisation of enzymes into multifunctional complexes can represent several advantages *in vivo*, for catalysis and its regulation. One of these consequences is the phenomenon of channelling, in which the product formed in a first reaction is channelled directly to the active site that catalyses the next reaction, thus avoiding the free diffusion of metabolites. This results in an increase in the global reaction rate. At the branching of metabolic pathways, several enzymes enter into competition for the transformation of the same metabolite; the direct transfer of this metabolite from one enzyme to another can constitute an effective means of regulation.

However, the existence of multifunctional enzymes and of multi-enzymatic complexes does not always cause an acceleration of reaction rates. In many metabolic pathways, catalytic efficiency that the enzymatic concentration represents is high enough to ensure the metabolic flow necessary to the cell. The question then arises why such systems have been selected during evolution. The most plausible hypothesis is that these associations within a same molecule or a complex ensure a higher thermodynamical stability than if the enzymes were separated.

A few examples amongst the best-characterised multifunctional enzymes and multi-enzymatic complexes are presented here. Two bifunctional enzymes from the tryptophan synthesis pathway, phosphoribosylanthranilate isomerase-indoleglycerol synthetase, a monomeric enzyme, and tryptophan synthetase, a hetero-oligomeric enzyme, are the subject of detailed studies. Amongst the multi-enzymatic systems, the protein CAD, pyruvate dehydrogenase and the fatty acid synthase enzymatic complex are treated. Lastly, the possibility of channelling in transitory multi-enzymatic complexes is discussed.

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_16, © Springer-Verlag Berlin Heidelberg 2010

15.1. PHOSPHORIBOSYLANTHRANILATE ISOMERASE-INDOLE GLYCEROLPHOSPHATE SYNTHASE

Tryptophan synthesis from chorismate involves several intermediates, as shown in Fig. 15.1.

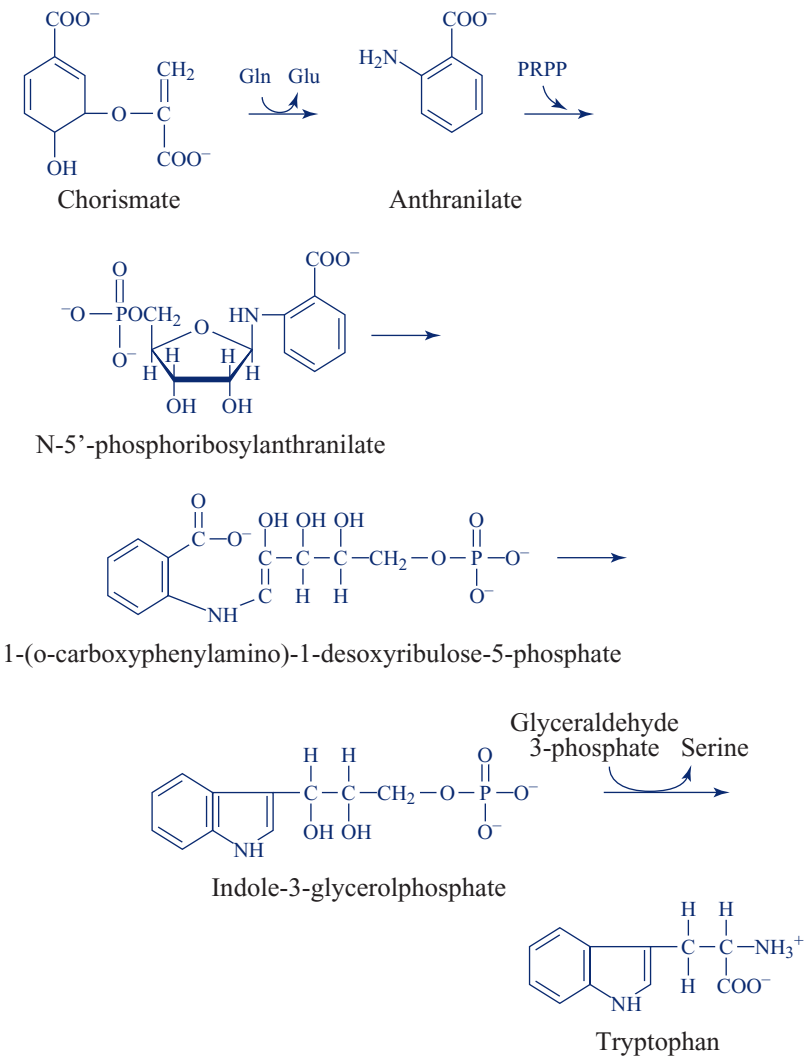


Fig. 15.1 Tryptophan biosynthetic pathway

In the presence of glutamine, chorismate is transformed into anthranilate with the release of glutamate; this reaction is catalysed by anthranilate synthase. The anthranilate is then phosphoribosylated by the 5-phosphoribosyl-1-pyrophosphate; the corresponding enzyme is the anthranilate 5-phosphoribosyl-1-pyrophosphate transferase. Two reactions then follow:

$$\text{PRA} \longrightarrow \text{CdRP}$$

$$\text{CdRP} \longrightarrow \text{IGP} + \text{CO}_2 + \text{H}_2\text{O}$$

PRA is the N-(5′-phosphoribosyl)anthranilate that is transformed into CdRP or 1-(o-carboxyphenylamino)-1-desoxyribulose-5-phosphate by the phosphoribosylanthranilate isomerase (PRA isomerase). In the following reaction, the CdRP is transformed into indole-3-glycerol phosphate (IGP) by indole glycerolphosphate synthase. In *E. coli*, these two enzymatic activities are carried by the same molecule, which is a bifunctional enzyme. The *E. coli* enzyme is coded by the trp CF gene of the tryptophan operon. In other organisms such as *Bacillus subtilis*, *Pseudomonas putida* and *Salmonella typhimurium*, these reactions are catalysed by two different enzymes. In yeast, *Saccharomyces cerevisiae*, PRA isomerase is a monofunctional enzyme while the IPG synthase activity is carried by a bifunctional subunit that is part of a multi-enzymatic complex. The *E. coli* enzyme possesses a very remarkable structure and has been the object of numerous studies.

15.1.1. STRUCTURE OF THE E. COLI ENZYME

The three-dimensional structure of PRA isomerase-IGP synthase from *E. coli* was solved at a resolution of 2 Å by the JANSONIUS group from Basel (1992). The enzyme, of molecular weight 49 500, is constituted of a single polypeptide chain of 552 amino acids folded into two different domains. The N-terminal domain that comprises residues 1–255, bears the IGP synthase activity. The C-terminal domain (256–452) is responsible for the PRA isomerase activity. The three dimensional structure of the bifunctional enzyme consists of two β/α barrels, connected by a short segment going from residues 254 to 256 (Fig. 15.2 below). The two domains display a very similar fold to the motif (αβ)8, with the exception of helix α5 that is absent in the PRA isomerase domain. This type of structure, consisting of eight parallel β strands in the centre of the molecule, surrounded by eight helices that link the consecutive β strands, makes each of the domains of the bifunctional enzyme similar to the triose-phosphate isomerase family (see Chap. 12). The crystallographic structures of 18 type (αβ)8 proteins have been resolved. The structure of PRA isomerase-IGP synthase was the first example of a protein containing two different (αβ)8 barrels. It is interesting to note that the α subunit of the tetrameric tryptophan synthase (α2β2) that catalyses the last but one step of tryptophan biosynthesis displays a similar fold. The superposition of the structures on the basis of their polypeptide skeleton or of their α carbon positions shows that between these three enzymes, 50–75% of the amino acids are structurally equivalent, although the sequence identity is only of 13.6–17.8%. Figure 15.3 below represents a comparison of the secondary structural elements and loops for these three enzymes that contain 8 α/β units.

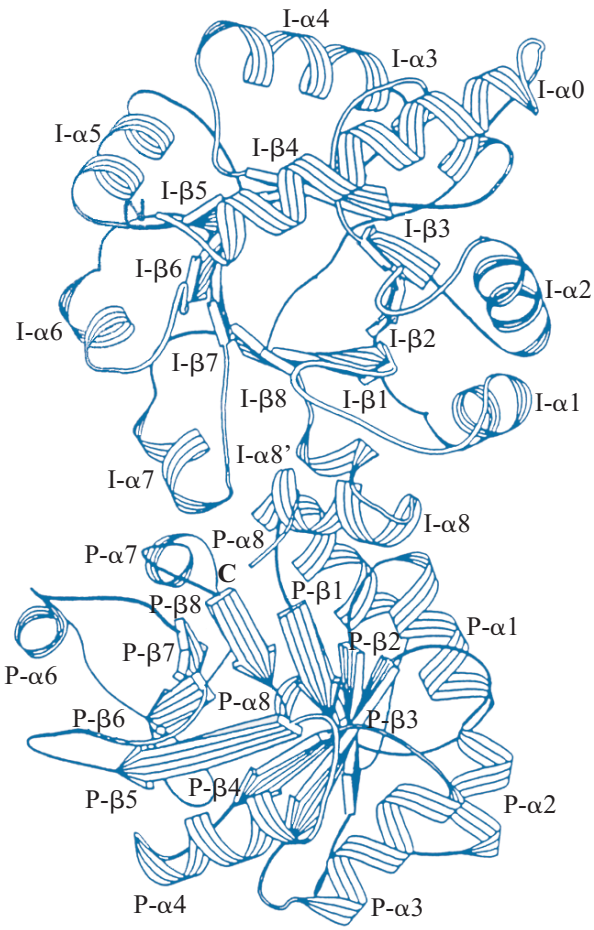


Fig. 15.2 PRA isomerase-IGP synthase structure

The letters I and P correspond to the IGP synthase and PRA isomerase domains, respectively (Reprinted from *J. Mol. Biol.*, **223**, WILMANNS M. et al., Three-dimensional structure of the bifunctional enzyme phosphoribosylanthranilate isomerase: indoleglycerolphosphate synthase from Escherichia coli refined at 2.0 Å resolution, 484. © (1992) with permission from Elsevier)

The IGP synthase domain is surrounded by eight helices, the longest is the helix α4 with 12 residues. The helix 8′ is positioned between the segment β8 and the very short helix α8 (4 residues). As well as the helices of the α/β barrel, the IGP synthase domain contains a helix with 5 turns and 19 residues (4–24), which represents an important fraction of the additional N-terminal part (residues 1–48). The α subunit of the tryptophan synthase also contains an N-terminal extension that does not form part of the (αβ)8 barrel. The central β barrel of the PRA isomerase is surrounded by 7 helices; the residues between the segments β5 and β6 do not form a helix, probably due to the presence of a proline (Pro366) that prevents the formation of one. Four helices of this domain contain at least 10 residues, while two helices α6 (4 residues) and α7 (5 residues) are very short.

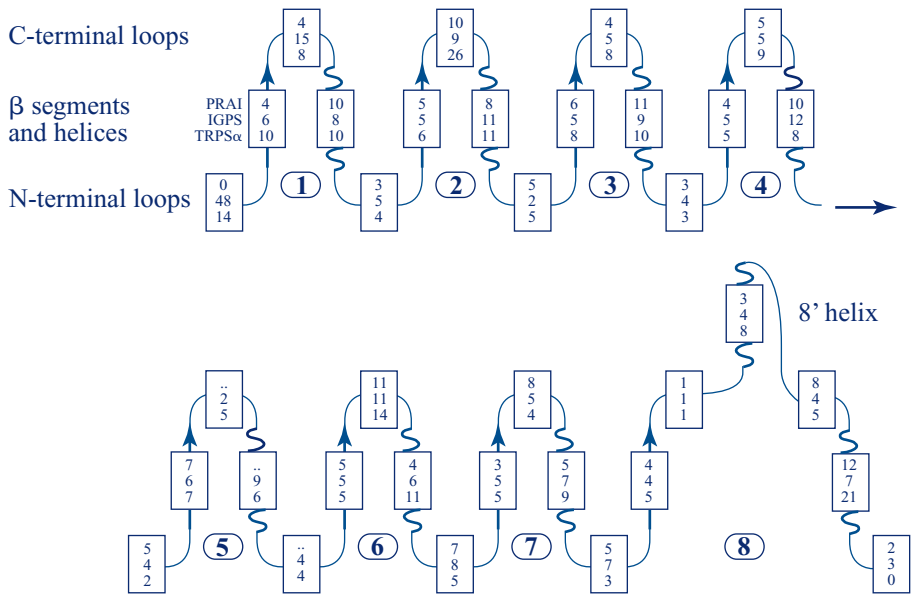


Fig. 15.3 Comparison of the length of secondary structures and loops of PRA isomerase (PRAI), IGP synthase (IGPS) and the α subunit of tryptophan synthase (TRPSα)
The length of these elements, in number of residues, is indicated in the boxes
(Reprinted with permission from *Biochemistry*, **30**, WILMANNS M. et al., 9165. © (1991) American Chemical Society)

The variation observed in the helix lengths is larger in this domain than in the IGP synthase domain. The additional helices 6 and 8 are 3_{10} helices. The interface between the two domains comprises residues found in the loop α0β1, the helix α7, and the helix α8 of the IGP synthase domain, as well as residues of the helix α1, the loop α2β2, the loop α7β8 and of the helix α8 from the PRA isomerase domain. The angle that the axes of the two (αβ)8 barrels form is about 70°.

15.1.2. STRUCTURE OF THE ACTIVE SITE

It is known that the active sites of the α/β barrel type enzymes are found in the C-terminal parts of the central β segments and involve residues of the loops that link the β segments and α helices. It was proposed that the large variations observed in the lengths of a few C-terminal loops were at the origin of the catalytic function diversity of the different α/β barrel type enzymes (WILMANNS et al., 1991).

Although the structure of the enzyme in the presence of substrate analogues has not yet been determined to high resolution, some indications concerning the position of the active centres were obtained. In particular, the IGP synthase domain of the bifunctional enzyme was produced by recombination and a complex between this domain and a substrate analogue, the N-(5′-phosphoribit-1-yl)-anthranilate, was crystallised and the structure determined at a resolution of 4 Å (WILMANNS et al., 1992).

In addition, two phosphate ions were detected in the high resolution structure of the bifunctional enzyme. They are located in the two active centres in equivalent positions on the C-terminal side of the two (αβ) motifs (PRIESTLE et al., 1987). These phosphate ions are bound by hydrogen bonds to the polypeptide chain amido groups of the N-terminal residues of two helices 8′ and of the preceding loop, and also to the main chain amido groups of the first residues linking segment β7 and helix 7. It is interesting to underline that the phosphate part of indolepropanolphosphate, a competitive inhibitor of tryptophan synthase, occupies a similar position in this enzyme. Figure 15.4 shows the position of the phosphate ion in the active sites of these three enzymes.

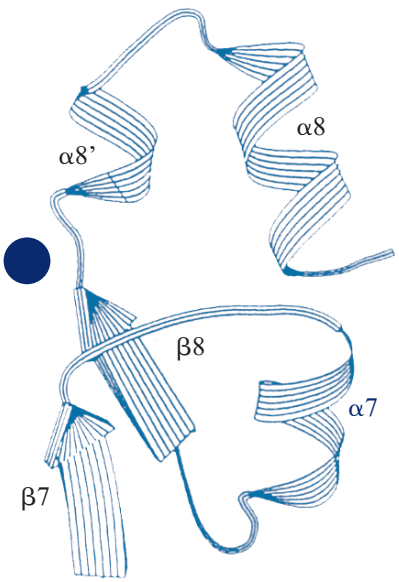


Fig. 15.4 Elements of the structure participating in the phosphate binding site
The phosphate is represented by the black dot
(Reprinted with permission from *Biochemistry*, **30**, WILMANNS M. et al., 9165. © (1991) American Chemical Society)

The active site of the IGP synthase domain is almost entirely covered by the N-terminal part of helix α0 and by the long loops β1α1, β2α2 and β6α6, all found in the C-terminal part of the central β barrel. The cavity is delimited on one side by the phosphate binding site, and on the other by hydrophobic residues. At the phosphate binding site two positive charges, those of lysine residues Lys55 and Lys61, compensate the negative charges of the phosphate ion. Probably the macro-dipole of helix 8′ contributes to its binding. It is remarkable that four residues of the phosphate binding site are invariant in the 16 known sequences, Lys55, Ser215, Gly216 and Leu234. Lys61 is conserved in 15 of the 16 sequences. The hydrophobic pocket contains several invariant residues in all of the known IGP isomerase sequences (Val4, Leu5, Ile8, Pro59, Phe93, Phe116, Leu188 and Leu191). The involvement of the N-terminal extremity of the helix α0 in the active site explains why this N-terminal segment (1–48) is present in all the known IGP synthases. The active site also contains a primarily hydrophobic channel linking the hydrophobic pocket to the solvent. Between the phosphate site and the hydrophobic site, there are 8 conserved residues, whose side chains are oriented towards the interior of the cavity.

The active site of the PRA isomerase domain is covered by four long loops positioned in the central part of the β barrel, those that link β2 and α2, β5 and α6, β6 and α6, as well as β8 and α8. Four of the residues that contribute to the phosphate ion binding site are conserved (Ala405, Gly406, Gly407 and Asp425). A channel positioned close to the hydrophobic pocket links the active site to the solvent.

There are strong similarities between the active sites of PRA isomerase and IGP synthase. In both enzymes, as well as in the 16 other enzymes of (αβ)8 type, the active sites are found in the C-terminal part of each β barrel. The active sites of both enzymes possess a very similar phosphate site and a hydrophobic pocket. The hydrophobic pockets of both domains probably have a similar function, the binding of the anthranilate part of the substrate. The cavity size of PRA isomerase is slightly smaller than that of IGP synthase. Such a difference is probably in relation to the substrate size; the phosphoribosylanthranilate with a closed ribose cycle has a more compact structure than the (o-carboxyphenylamino)-1-desoxyribulose-5-phosphate, whose ribulose part is linear. In both domains, the channel that links the active sites to the solvent is formed by the residues of the same loops (β2α2, β6α6). Three of the glycine residues of the loop β6α6 of PRA isomerase are conserved in all the sequences. According to the temperature factors, this loop is more mobile in PRA isomerase than in IGP synthase. Comparisons with other enzymes of (βα)8 type structure, with and without a substrate analogue (triose phosphate isomerase and ribulose-1,5-bisphosphate carboxylase/oxygenase) suggest that this loop could be involved in the passage of substrates towards the active site.

The fact that the three dimensional structures of PRA isomerase and IGP synthase, as well as their active sites, are particularly well-conserved suggests that the two domains of the *E. coli* bifunctional enzyme are the result of gene duplication. More generally, it was proposed by HOROWITZ (1945) that the biosynthetic pathways evolved by gene duplication, followed by an independent evolution of one of the gene copies. This hypothesis, which seems verified for enzymes catalysing two successive steps in the biosynthetic pathway of methionine in *E. coli*, also seems valid for PRA isomerase and IGP synthase. In the latter case however, in *E. coli*, the evolutive process would have been followed by the fusion of the genes.

15.1.3. FUNCTIONAL PROPERTIES

The two active sites of the bifunctional enzyme seem to catalyse their respective reactions independently. The position of the active sites is such that substrate channelling appears impossible. The competitive inhibitor of IGP synthase, N-(5′-phosphoribit-1-yl)-anthranilate, binds to both enzyme sites. Its binding to IGP synthase is a two-step process that involves a slow conformational change, which is not transmitted to the PRA isomerase site.

The kinetic parameters of the reaction in the steady state were determined with precision for IGP synthase, but only semi-quantitative data exist for the reaction

catalysed by PRA isomerase due to the instability of the substrate. It is not known which of the structural isomers of phosphoribosylanthranilate (α or β anomer) is used as a substrate. However, the catalytic constant of PRA isomerase is about 10 times higher than that of IGP synthase (Table 15.1). Phosphate is a competitive inhibitor of IGP synthase; its binding leads to a conformational change of the protein.

Table 15.1 Catalytic constants of PRA isomerase-IGP synthase

(Reprinted from *Methods in Enzymol.*, **142**, KIRSCHNER K. et al., Phosphoribosylanthranilate isomerase-indoleglycerol-phosphate synthase from Escherichia coli, 386. © (1987) with permission from Elsevier)

Enzyme	*Substrate*	k_{cat} (s^{-1})	K_m(μM)
IGPS	CdRP	3.1	1.12
		0.8*	38*
PRAI	PRA	39	7

* *values obtained in phosphate buffer, the other values are in a tris buffer.*

PRA isomerase and IGP synthase catalyse reactions of phosphorylated substrates that differ in the configuration of the saccharide part. It is therefore likely that the phosphate part of the substrates binds to the phosphate site. Likewise the binding sites of the anthranilate part are very similar. The catalytic mechanism of these enzymes is not known. For the reaction catalysed by PRA isomerase, it was proposed that the PRA conversion into CdRP is performed following an AMADORI rearrangement. The mechanism involves the protonation of the oxygen $O_{4'}$ of the ribose cycle and the abstraction of a proton from the $C_{2'}$ carbon. Studies performed by KIRSCHNER and collaborators on mutants of PRA isomerase suggest that residues Cys260, Glu332 and Asp379 play a catalytic role, while His334 seems to be implicated in substrate binding. In addition, variants of the yeast PRA isomerase obtained by circular permutation of the (βα) motif display a catalytic activity comparable to that of the wild type (LUGER et al., 1989). The fusion of the N- and C-terminal extremities and the cleavage of a surface loop do not affect the structural integrity of the enzyme.

In the reaction catalysed by IGP synthase, the 5-branch heterocycle of tryptophan is formed by the cyclisation of the oxygen of the carbon from the anthranilate aromatic cycle, which is substituted by a carboxylate and the carbon atom of the α-arylamino-ketone of the CdRP substrate. The elimination of a molecule of CO_2 and one of H_2O occurs. The mechanism is not known. It was suggested that the cycle closing is facilitated by a strong polarisation of the α-arylamino-ketone group and by the deprotonation of the O-carboxylate group that is released in the form of CO_2 during the reaction. Kinetic studies as a function of pH suggest that the protonation of the α-amino group accelerates the reaction. Even though the exact position of the substrate in the active site of IGP synthase is not known, the structure suggests that residues Ser60, Lys114, Arg186, Glu163 and Asn184 may play a role in catalysis. Lys55 that participates in phosphate binding likely has another role; its

replacement by site-directed mutagenesis by a serine leads to a decrease in the constant k_{cat}/K_m of 10 000 times compared to the wild type enzyme.

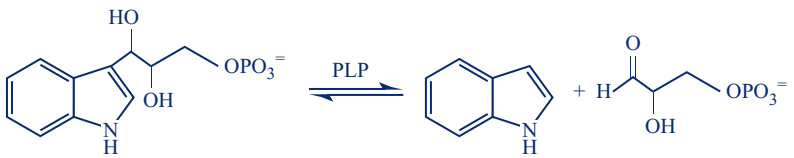
On the basis of the remarkable conservation of their three-dimensional structure, the bifunctional enzyme PRA isomerase-IGP synthase formed of two $(\alpha\beta)8$ domains, which catalyse two successive reactions of the tryptophan synthesis pathway, as well as the α subunit of tryptophan synthase, which catalyses the next reaction, seem to have evolved from a common ancestor. It should be noted that even though it is bifunctional, the PRA isomerase-IPG synthase does not display channelling, while the tetramer of tryptophan synthase allows channelling of the first reaction product to the β subunit active site.

15.2. TRYPTOPHAN SYNTHASE

Although many examples of channelling have been cited in the literature, there are only few cases where data show without ambiguity the existence of such a substrate transfer between two catalytic sites. Tryptophan synthase represents unquestionably an excellent example of a hetero-oligomeric bifunctional enzyme for which the existence of channelling has been rigorously supported by functional studies and structural data.

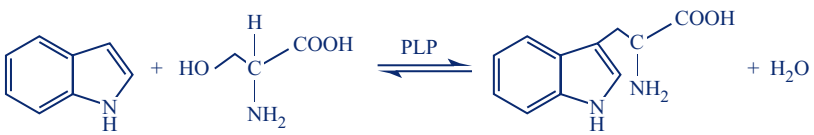
15.2.1. FUNCTIONAL PROPERTIES

Tryptophan synthase catalyses the final reaction of tryptophan biosynthesis. The bacterial enzyme is a bifunctional heterotetramer, $\alpha_2\beta_2$, of molecular mass 143 kDa. The isolated α subunit (29 kDa) catalyses the cleavage of indole 3-glycerolphosphate to give indole and glyceraldehyde 3-phosphate (G3P) following the reaction:



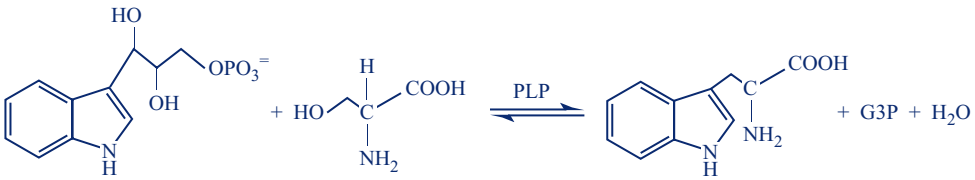
IGP Indole G3P

The isolated β subunit (43 kDa) is a dimer and contains one molecule of pyridoxal phosphate per monomer. The β_2 dimer catalyses the synthesis of L-tryptophan from indole and L-serine, *via* the intermediate of the pyridoxal phosphate:



Indole L-serine L-tryptophan

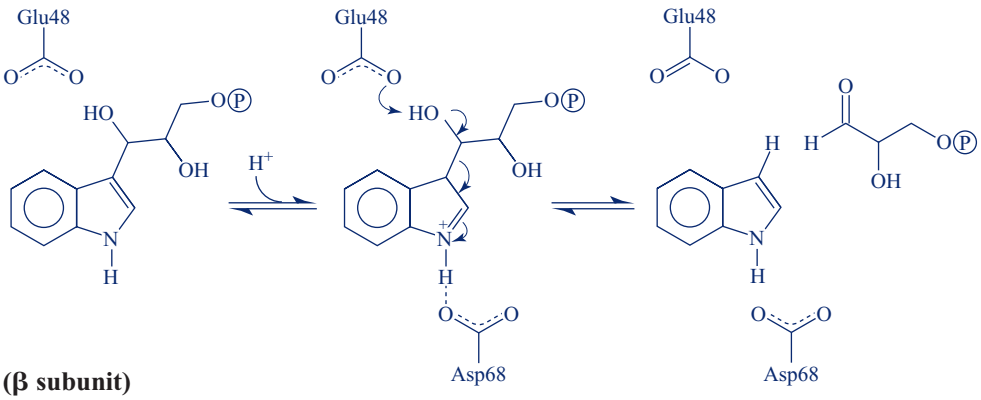
The tetrameric enzyme catalyses the global reaction:



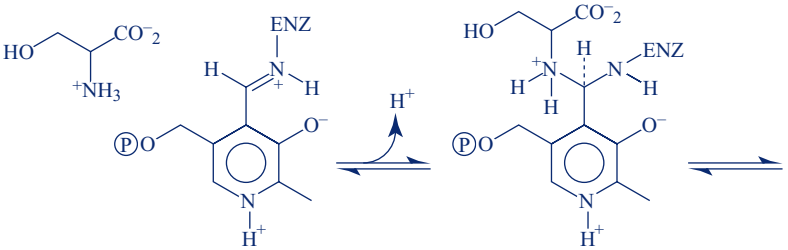
IGP L-serine L-tryptophan

which is the sum of the two preceding reactions. The reactions catalysed by the β subunit proceed *via* intermediates that have distinct spectroscopic properties, allowing their characterisation. They are produced in two steps. The pyridoxal phosphate that intervenes in these reactions is bound to the enzyme *via* an internal SCHIFF base to lysine 87, as have shown studies performed by site-directed mutagenesis. All of these reactions are presented on Fig. 15.5.

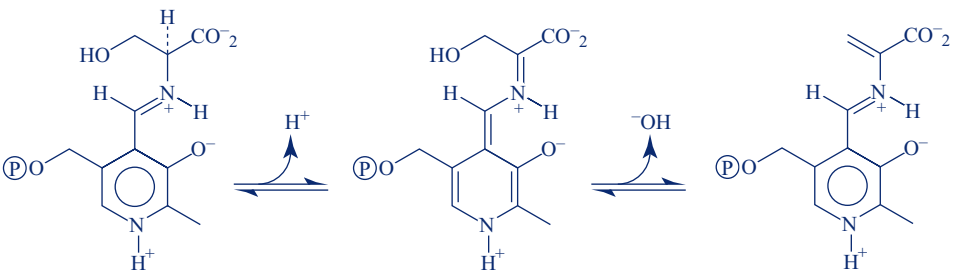
(α subunit)



(β subunit)
Step I



E (Aln) 410 nm E (GD) 310-340 nm



E (Aox1) 422 nm E (Q1) 460 nm ? E (A-A) 360 nm
shoulder 380-630 nm

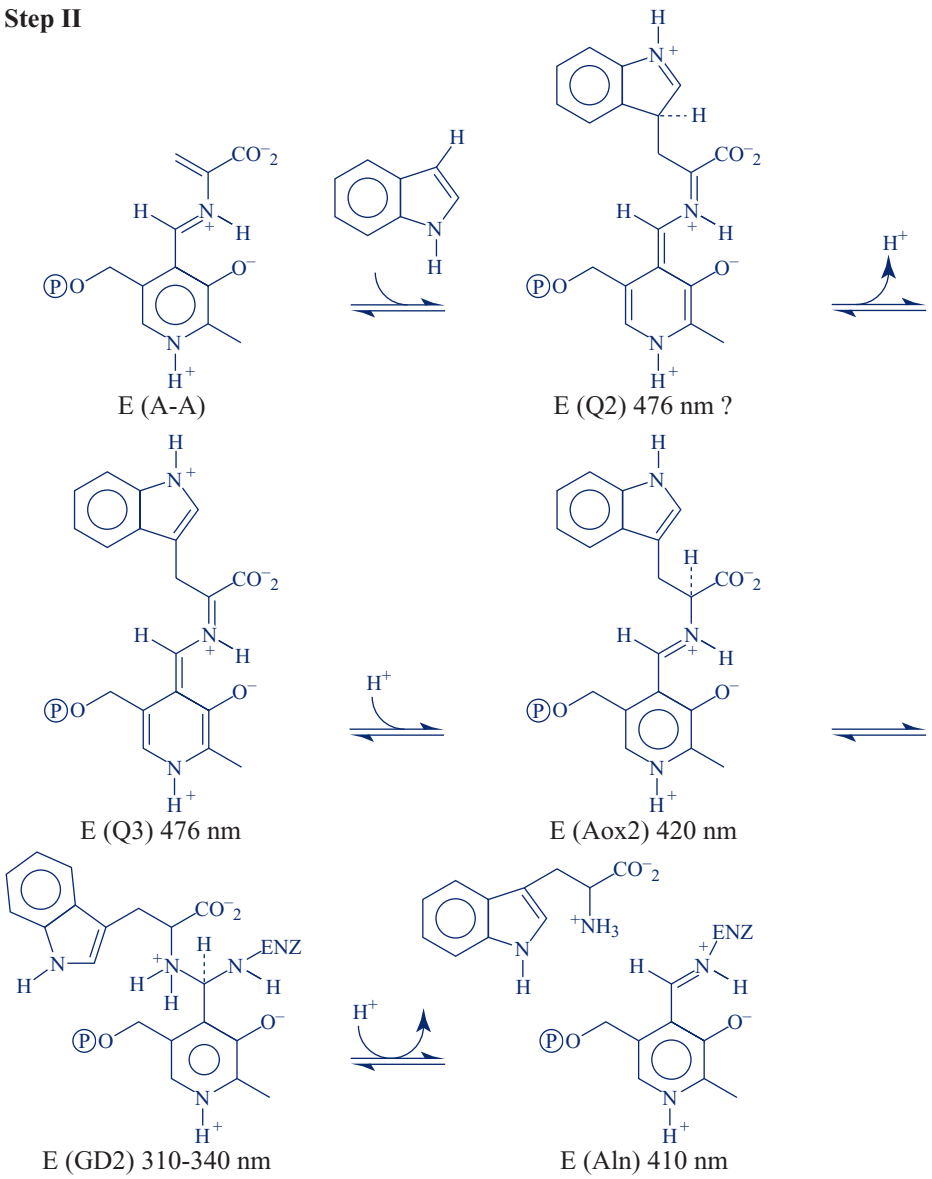


Fig. 15.5 Diagram of the reactions catalysed by each of the tryptophan synthase subunits showing the different reaction intermediates

E (Aln) indicates the enzyme's internal SCHIFF base (aldimine) with the pyridoxal phosphate (PLP), E (GD) the first diamine species geminated, E (Aox1) the SCHIFF base between PLP and serine, E (Q1) the corresponding quinoid form, E (A–A) the α-amino acrylate SCHIFF base, E (Q2) the quinoid form formed by the indole addition, E (Q3) the quinoid form resulting from E (Q2) by abstraction of a proton, E (Aox2) the external aldimine with tryptophan, E (GD2) the diamine geminated with the tryptophan and the enzyme. In the final step E (Aln) is regenerated and the tryptophan liberated

(Reprinted with permission from *Biochemistry,* **35**, HUR O. et al., 7378. © (1996) American Chemical Society)

Kinetic studies in the steady state have shown that the α subunit reaction is twenty times slower in the absence of serine than the global rate of the $\alpha_2\beta_2$ tetramer to form tryptophan. In addition, during a quenched-flow experiment corresponding to a single enzyme turn over, on conversion of radioactive indole glycerolphosphate into tryptophan, no significant trace of indole was detected in the medium (ANDERSON et al., 1991). This implies that the indole reaction to form tryptophan is very rapid ($\geq 1\ 000\ s^{-1}$). The reaction rate of indole from the solution is also too slow to explain the results. As a consequence, the indole produced in the active site of the α subunit must be quickly and directly transferred to the active site of the β subunit where it reacts to form tryptophan.

In the absence of serine, the cleavage of indole glycerolphosphate to give indole is limited by the conformational change of the protein that occurs at a rate of $0.16\ s^{-1}$. When the global reaction of tryptophan synthase is studied by stop-flow, by mixing the enzyme with indole glycerolphosphate and serine, a latent phase is observed both for the indole glycerolphosphate cleavage and for the appearance of tryptophan. The kinetics of this latent phase corresponds to the rate of aminoacrylate formation in the reaction of L-serine with the pyridoxal phosphate at the active centre of the β subunit ($45\ s^{-1}$) (Fig. 15.6).

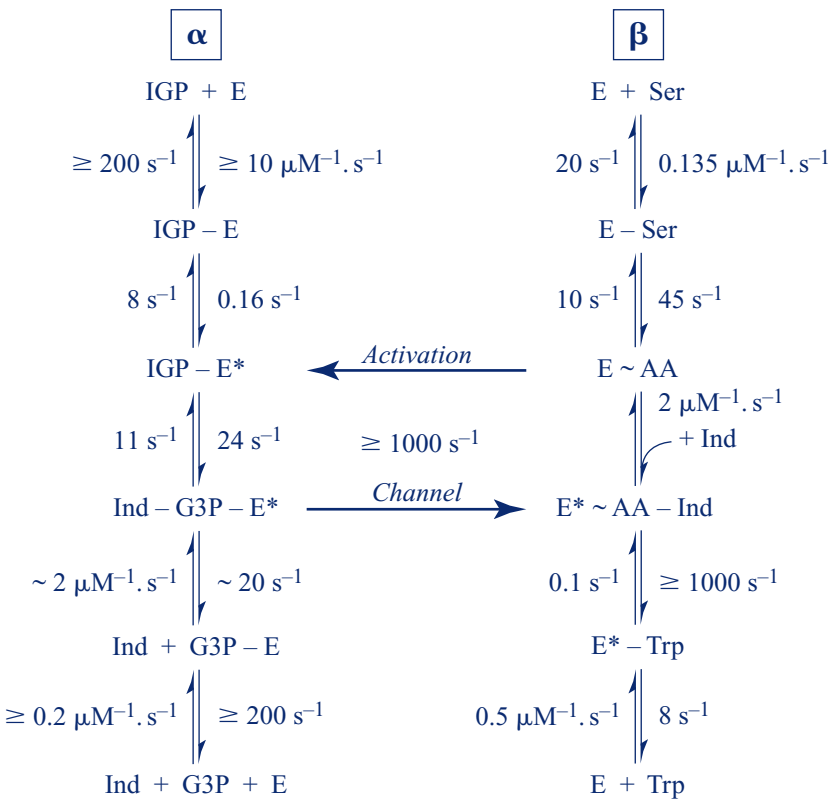


Fig. 15.6 Kinetic scheme of the reactions catalysed by tryptophan synthase, showing the reaction rates

(From *J. Biol. Chem.*, **266**, ANDERSON K.S. *et al.,* 8020. © (1991) with permission from The American Society for Biochemistry and Molecular Biology)

This reaction is accompanied by a change in fluorescence of the protein, indicating a conformational variation, which is produced with the same rate constant. The substitution of the L-serine by an L-cysteine increases the latent time in the kinetics of indole glycerolphosphate transformation and lowers the rate of aminoacrylate formation. These different results show that the reaction of serine at the active site of the β subunit to form the aminoacrylate modulates the reaction of the α subunit by provoking a conformational change that is transmitted to the active centre of the α subunit; this has the effect of increasing the rate of indole glycerolphosphate cleavage 150 times. The role of glutamate 109 in the de-protonation of indole to facilitate its reaction with aminoacrylate was shown by the study of a mutant E109D. All of the data obtained by the group of ANDERSON et al. (1991) led the authors to propose a complete kinetic scheme of the reaction (Fig. 15.6).

The kinetic data show without ambiguity channelling in tryptophan synthase. The rate of indole passage through the channel is very rapid. The indole reaction to form tryptophan at the active site of the β subunit is also very rapid and practically irreversible. The reaction of the L-serine in the catalytic site of the β subunit modulates the rate of indole formation at the catalytic site of the α subunit. The communication between the subunits maintains the reactions of the α and β subunits in phase, so that the indole does not accumulate at the active centre of the α subunit.

15.2.2. ENZYME STRUCTURE

The three-dimensional structure of the enzyme from *Salmonella typhimurium* was resolved in 1988 by the DAVIES group (HYDE et al., 1988) at 2.5 Å resolution. The four subunits of the tetramer are arranged in a quasi-linear manner in the order αββα, forming a 150 Å long complex. The two β subunits are closely associated at the centre of the tetramer, the α subunits are positioned at the extremities of the β dimer.

The structure of the α subunit is of the type $(\alpha\beta)_8$, like the triose-phosphate isomerase (TIM barrel); this type of structure is also found for other enzymes (see Sect. 15.1). The active centre of the α subunit was found by structural analysis of the complex that the enzyme forms with a substrate analogue, indole propanolphosphate (Fig. 15.7 below). The inhibitor binds in a hydrophobic cavity, in a region where most of the amino acids are highly conserved in the sequences of α subunits from about ten microorganisms. It is positioned at the interface between neighbouring α and β subunits. The active centres of these two subunits are separated by a distance of about 25–30 Å.

The β subunits interact *via* a large surface that is almost flat. The buried surface is of the order of 1 440 $Å^2$. Dissociation studies under high pressure of the β_2 dimer showed that these two subunits are closely associated, and a dissociation constant of $K_d = 3{,}7 \times 10^{-10}$ M was evaluated at a pressure of one bar. The β subunit is structured in two domains. The N-terminal domain comprises a β sheet made of four antiparallel strands flanked by four helices. The C-terminal domain possesses a β sheet

made of six strands surrounded by helices. These two domains display a high structural homology. The coenzyme, pyridoxal phosphate, is found at the interface between the two domains close to the centre of each β subunit. The active centres of the two β subunits are separated by 25 Å. It seems unlikely that an interaction exists between the sites; the pair αβ seems to be the functional unit.

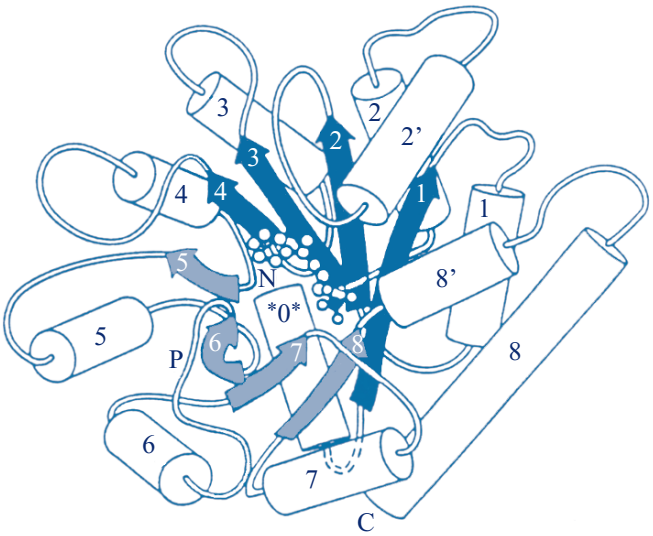


Fig. 15.7 Schematic representation of the structure of the tryptophan synthase α subunit
The active site is indicated by the inhibitor position, indole propanolphosphate, represented by balls and sticks
(From *J. Biol. Chem.*, **263**, HYDE C.C. et al., 17860. © (1988) with permission from The American Society for Biochemistry and Molecular Biology)

The structure clearly shows the existence of a tunnel linking the active centres of the α and β subunits (Fig. 15.8).

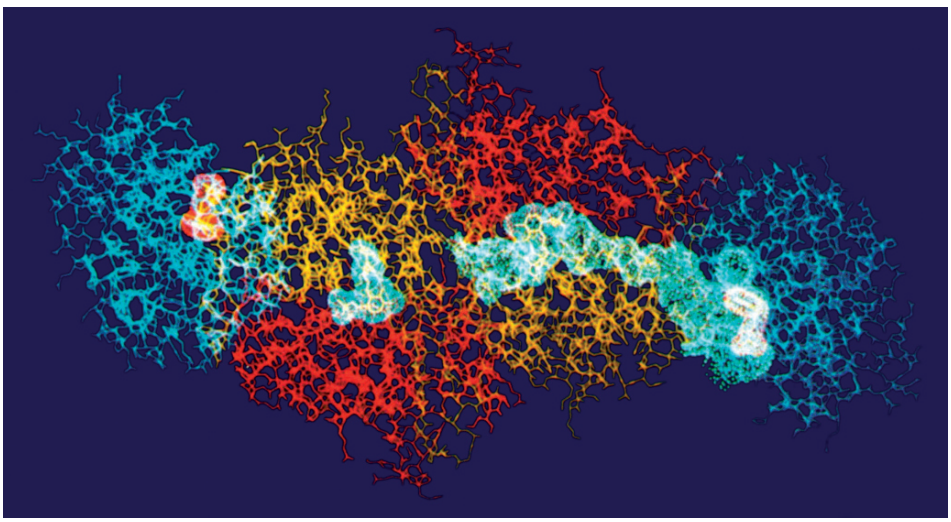


Fig. 15.8 Structure of tryptophan synthase showing the channel that links the active sites of the α and β subunits
(From *J. Biol. Chem.*, **263**, HYDE C.C. et al., 17861. © (1988) with permission from The American Society for Biochemistry and Molecular Biology)

The diameter of this tunnel is large enough to contain the indole. It stretches from the α subunit active centre to the pyridoxal phosphate site of the β subunit and beyond, passing through the centre of the interface between the N- and C-termini of the β subunit. The longest part of the tunnel is through the β subunit; to evaluate its dimensions, four indole molecules, one after the other, can be simultaneously placed inside by modelling. The function of the tunnel appears to be to allow the free diffusion of indole between the α and β subunit active centres, preventing the indole from escaping into the solvent. Figure 15.9 illustrates the functioning of tryptophan synthase.

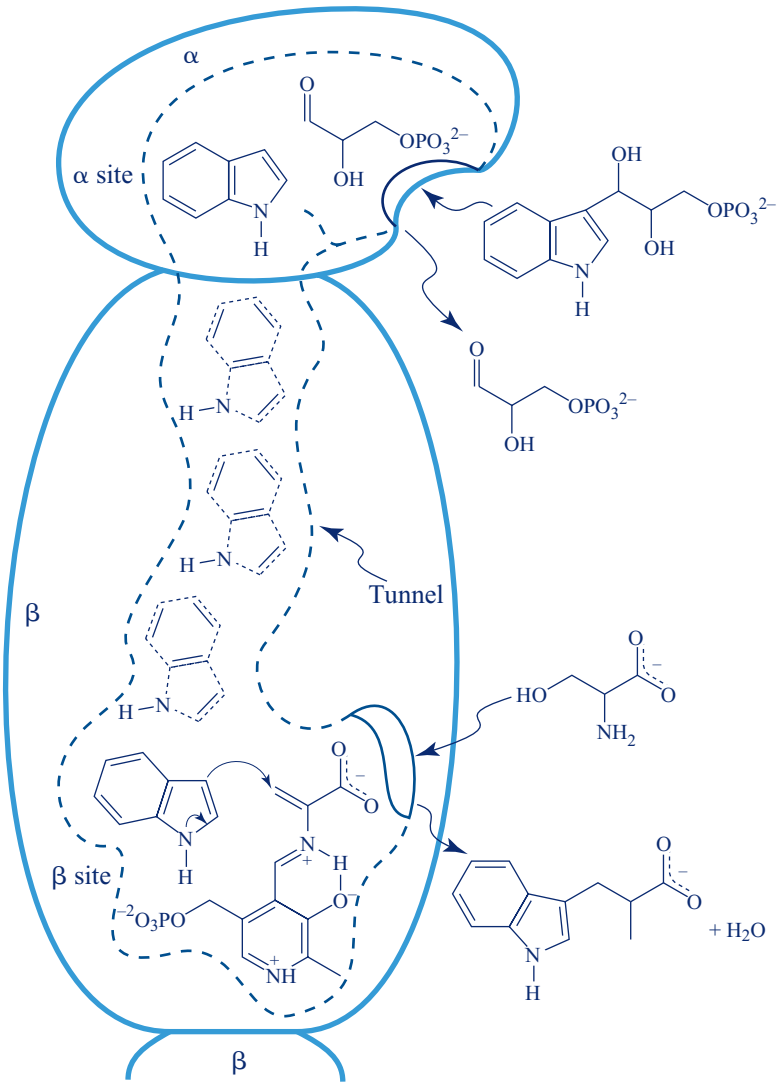


Fig. 15.9 Schematic representation of the reactions catalysed by tryptophan synthase at the different sites of α and β subunits, showing the tunnel by which the indole transits from one site to the other

(Reprinted with permission from *Biochemistry*, **35**, HUR O. et al., 7378. © (1996) American Chemical Society)

Tryptophan synthase is therefore an excellent example of channelling, where the phenomenon has been proved both by kinetic studies and by structural data.

15.2.3. STUDY OF A MUTANT LEADING TO CHANNEL OBSTRUCTION

Knowledge of the tryptophan synthase structure not only brought an additional proof of channelling, but also led to the construction of a mutant in which the channel is obstructed. After considering the nature of the amino acids lining the channel, SCHLICHTING et al. (1994) constructed a mutant in which cysteine 170 of the β subunit was replaced by a much bulkier amino acid, tryptophan. The kinetic analysis was performed by quench-flow methods. An experiment of a single turn over of the enzyme showed that the presence of indole is detectable as an intermediate. In addition, the global reaction rate is decreased by two orders of magnitude compared to the wild type enzyme. The rate constants of the different reaction steps were determined and compared to the "channelled" reaction (Fig. 15.10).

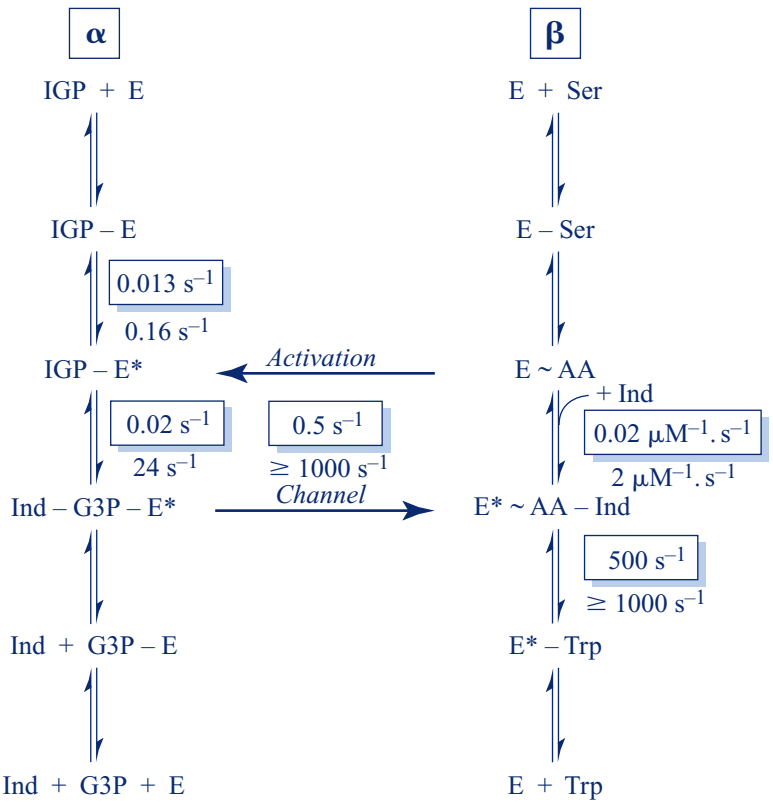


Fig. 15.10 Kinetic scheme of the reaction catalysed by the mutant C170W of tryptophan synthase compared to that of the wild type enzyme

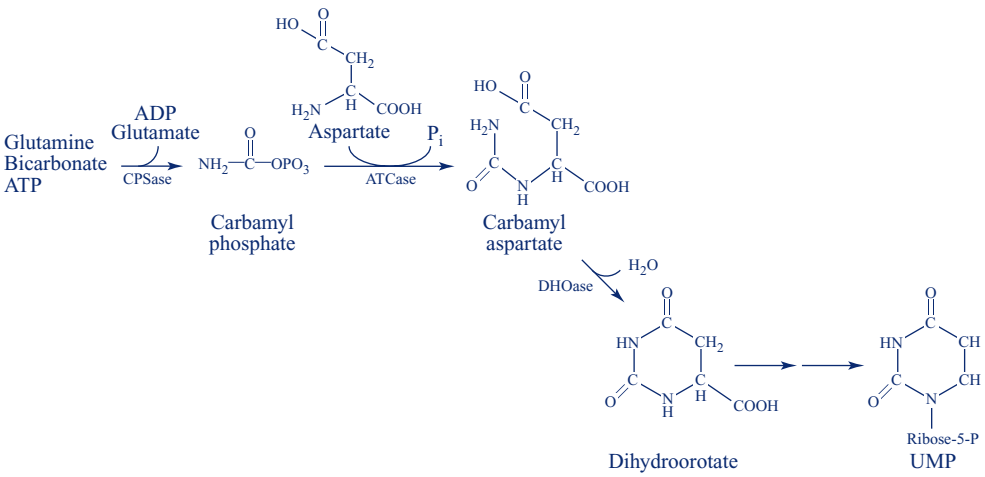
The values of the boxed rate constants refer to the mutant. (From *J. Biol. Chem.*, **269**, SCHLICHTING I. et al., 26591. © (1994) with permission from The American Society for Biochemistry and Molecular Biology)

The fact that the detection of indole is linked to an obstruction of the channel was confirmed by the crystallographic structure. The authors specify however that the crystallographic structure alone does not allow the conclusion that the observed effects come only from the free diffusion of indole in solution, and not from dynamic fluctuations in the mutation region that allow active channelling.

15.3. CAD PROTEIN

15.3.1. THE FIRST ENZYMES IN THE PYRIMIDINE BIOSYNTHESIS PATHWAY

The biosynthesis of pyrimidines from glutamine, bicarbonate and ATP involves a succession of enzymatic reactions. The first three reactions bring sequentially into play carbamyl phosphate synthetase (CPSase), aspartate transcarbamylase (ATCase) and dihydroorotase (DHOase) and are represented below:



These first three steps of biosynthesis are common to all species. In prokaryotes, these enzymatic activities are catalysed by three independent enzymes, whilst in eukaryotes (mammals, drosophila and dictyostelium) they are carried by a trifunctional protein. In these organisms, the trifunctional enzyme is coded by a single gene. The structure and functional properties of this trifunctional oligomeric enzyme are the subject of detailed studies by the EVANS, DAVIDSON and CARREY groups. Carbamyl phosphate synthetase itself displays two activities, a glutamine-dependent amido transferase activity (GLN) and a synthetase activity (SYN) (see Sect. 15.4). In yeast, carbamyl phosphate synthetase and aspartate transcarbamylase are associated in a single polypeptide chain, whilst the dihydroorotase activity is carried by a separate protein (SOUCIET et al., 1989). The latter complex contains however a "pseudo DHOase" sequence, which is devoid of any enzymatic activity. Figure 15.11 below shows the genetic organisation of these three enzymes in different organisms (for a review, see HERVÉ et al., 1993).

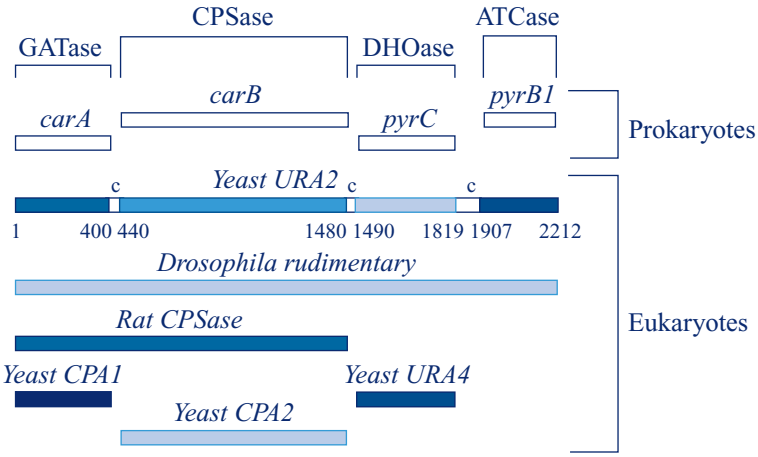


Fig. 15.11 Genetic organisation of pyrimidine biosynthesis in different species
(Reprinted from *Gene*, **79**, SOUCIET J.L. et al., 59. © (1989) with permission from Elsevier)

15.3.2. STRUCTURAL ASPECTS

The CAD protein genes were sequenced in several organisms and it was shown that the domains possessing the different enzymatic activities are organised in the following order: N-terminal-GLN-SYN-DHO-ATC-C-terminal. The protein is a homomultimer composed of three subunits or more. A model of a trimeric protein was proposed, as shown in Fig. 15.12. It was based, amongst other things, on the three-dimensional structure of the homologous ATCase from *E. coli*. In yeast, the URA2 gene codes for a multi-functional protein and the URA4 gene for the single DHO. In *E. coli*, each enzyme is coded by a separate gene (Fig. 15.11). In addition, each enzymatic domain is separated from the other by a loop sensitive to proteases. EVANS and collaborators showed that when the domains are separated by limited proteolysis or by genetic engineering, these enzymes remain active. These results suggested that the CAD protein results from a fusion of genes that occurred at least 100 million years ago.

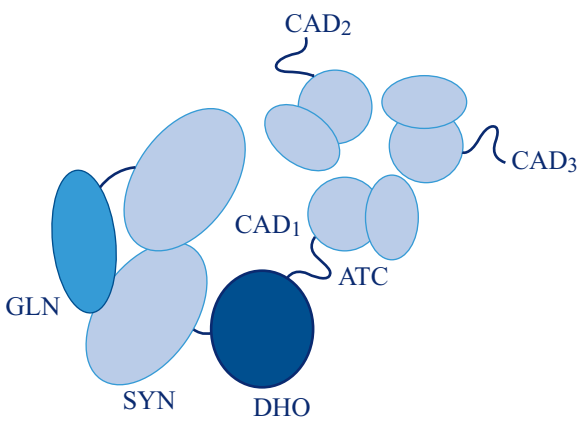


Fig. 15.12 Schematic model of a CAD trimer
The three identical CAD subunits are represented forming a complex centred around a domain of the ATCase
(From *Bioessays*, **15**, DAVIDSON J.N. et al., 157-164. © (1993 Wiley Periodicals, Inc.). Reprinted with permission of John Wiley & Sons, Inc.)

15.3.3. FUNCTIONAL PROPERTIES

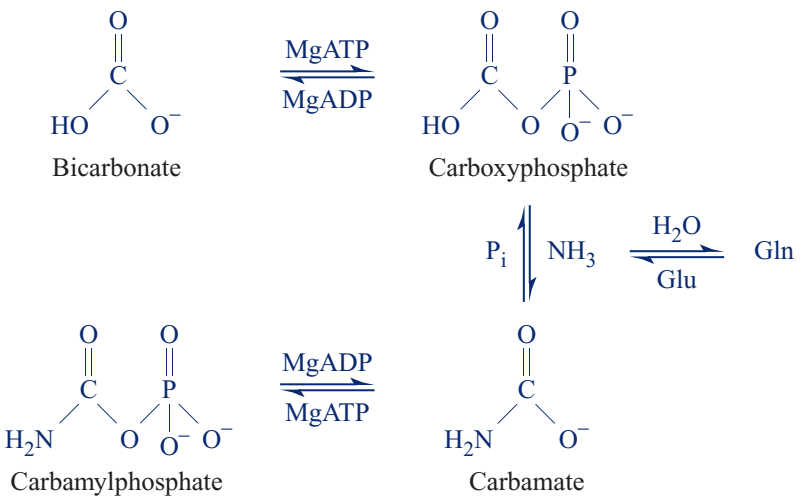
Enzymatic activity studies in vitro and in vivo revealed the existence of channelling, which is less effective in the CAD protein from mammals than in the URA2 protein from yeast. It was shown that the production of dihydroorotate is faster than when the three enzymes are separated (MAILLY et al., 1980). In yeast, very significant channelling was observed for the first two steps of the metabolic pathway (PENVERNE et al., 1994). These different behaviours seem to indicate that channelling is not necessarily the result of a selection pressure in the evolution of multi-functional enzymes. The selective advantage that it represents may reside in the fact that this situation allows a coordinated expression of these enzymes (DAVIDSON et al., 1993).

15.4. CARBAMYL PHOSPHATE SYNTHETASE

The transfer of a metabolite in the interior of an enzyme can occur over long distances. In the case of tryptophan synthase, for which the first structural evidence of channelling was obtained, the presence of a tunnel of 25 Å ensuring the passage of indole was observed. The more recent resolution of carbamyl phosphate synthetase (CPSase), in 1997, by the group of THODEN and collaborators showed the presence of a tunnel of almost 100 Å length.

15.4.1. FUNCTIONAL PROPERTIES

CPSase catalyses the production of carbamyl phosphate from a molecule of bicarbonate, two molecules of $Mg^{++}ATP$ and a glutamine molecule following the reaction:



The enzyme from *E. coli* is constituted of two polypeptide chains, the large and the small subunit. The small subunit catalyses glutamine hydrolysis. The C-terminal part of the small subunit is homologous with the N-terminal domains of amido transferases of trpG type like the GMP synthase. It is on the large subunit that the two phosphorylation reactions occur. The large subunit also contains binding sites for monovalent cations like K^+, which is an enzyme activator, and for effector molecules such as ornithine, which behaves as an activator, and UMP, which is an inhibitor. These effectors act in the reaction by modulating the MICHAELIS constant for MgATP.

The genes that code for the large and small subunits were sequenced in 1983 and 1984 by LUSTY and collaborators. In the large subunit, 40% sequence identity is observed between the sequences Met1-Arg400 and Ala553-Leu933, suggesting that this large subunit derives from gene duplication.

15.4.2. STRUCTURAL PROPERTIES

The enzyme from *E. coli* is a heterotetramer $(\alpha,\beta)_4$, each heteroprotomer being constituted of two different α and β subunits. The tetramer displays a 222 symmetry. The small subunits are found at the extremities of the molecule. The interaction zone between the small and large subunits is extended and contains 35 hydrogen bonds between the two chains. The small subunits are folded into two structural domains. The N-terminal domain (Leu1-Ser153) contains two layers of β sheets perpendicular to one another, one containing four parallel β strands, the other four anti-parallel β strands. The C-terminal domain (Asp154-Lys382) is formed of a ten β strand sheet, some parallel and some anti-parallel, and six α helices. The small subunit carries the glutamine hydrolysis site with Cys269 that acts as a nucleophile catalyst helped by His353. The reaction proceeds *via* the formation of a glutamyl thioester intermediate, formation that was confirmed by crystallographic analysis of a mutant H353N (THODEN et al., 1997). The Ser47 O^γ and the NH hydrogen of the peptide backbone constitute an ideal environment to position the substrate carbonyl and stabilise the oxyanion that forms during the reaction. Another residue, Glu273, interacts with the glutamyl thioester intermediate. These observations represent the first structural evidence of the formation of a glutamyl thioester intermediate in this enzyme family.

The large subunit contains several regions, two that are homologous and contain the formation sites of carboxyphosphate and carbamylphosphate, going respectively from Met1 to Glu403 and from Val404 to Ala553, an oligomerisation region where the contact between the subunits occurs (Asp554-Asn936) and a part called the allosteric motif (Ser937-Lys1073). The two active sites are topologically but not structurally equivalent and, in the interior of the large subunit, they display between them a 2-fold symmetry axis, suggesting that the current activity of CPSase derives from a homodimeric ancestor.

Each of the two homologous regions are constituted of three structural patterns called A, B and C. The A domains contain a β sheet formed of five parallel strands, B domains are constituted of a β sheet of four anti-parallel strands, C domains are formed of a sheet of seven anti-parallel β strands. In the structure, an Mn^{++}ADP molecule was bound to the carboxyphosphate and carbamylphosphate formation sites. In addition, an inorganic phosphate molecule was present on the carboxyphosphate site leading to the closure of the B domain. A similar phosphate position was observed on another enzyme that proceeds *via* the formation of a carboxyphosphate intermediate, biotin carboxylase, suggesting that this region of CPSase is responsible for the stabilisation of the intermediate. The different interactions of ADP with the enzyme and the coordinations of Mn^{++} were determined, as well as the surrounding water molecules. It was possible to determine the K^{+} position.

15.4.3. THE TUNNEL

One of the most remarkable facts to emerge from the structure is that the three active sites are separated by a distance close to 100 Å. Taking into account the instability of the reaction intermediates, the reaction must be well coordinated in time. The careful examination of the structure using a modelling program showed a possible tunnel linking the three active sites in the αβ heterodimer (Fig. 15.13 below). This tunnel extends from the base of the small subunit active site towards the surface of the carboxyphosphate formation site, then until the second site of the large subunit where the carbamylphosphate is formed. The amino acids that line the tunnel wall are essentially non-reactive residues and atoms of the peptide backbone. The amino acids located at 3.5 Å from the tunnel centre in the small subunit are residues Ser35, Met36, Gly293, Ala309, Asn311, *cis*-Pro358 and Gly359, and in the large subunit Glu219 and Cys232; the residues of the tunnel between the two sites of the large subunit are slightly less hydrophobic. Many residues that line the tunnel wall are conserved in the sequences of CPSase molecules from different organisms.

Structural proof of the channelling mechanism is therefore established for several enzymes. Proof of channelling was also furnished with the resolution of the phosphoribosylpyrophosphate amido transferase structure, the first enzyme in the purine biosynthetic pathway, which also uses glutamine as nitrogen source. In this enzyme that contains only two active sites, the tunnel is shorter. It extends over 20 Å and is of a very hydrophobic nature (KRAHN et al., 1997). It is likely that as more structures of enzyme complexes are determined to high resolution, other examples of channelling will be brought to light.

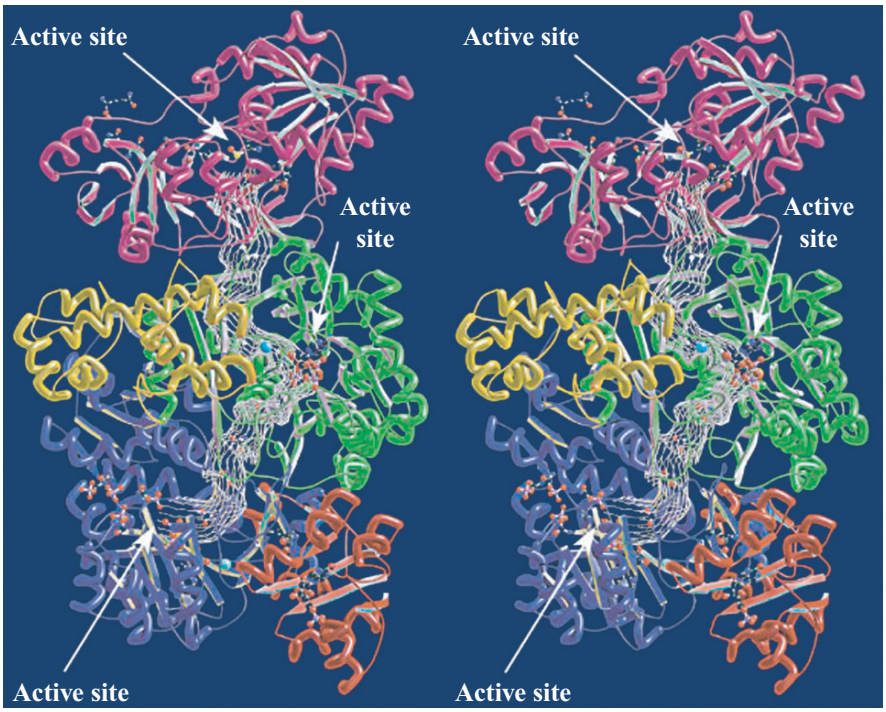


Fig. 15.13 Structure of the αβ heterodimer of CPSase showing the tunnel between the active sites (stereoscopic view)

(Reprinted from *Curr. Op. Struct. Biol.*, **8**, HOLDEN H.M. et al., 679. © (1998) with permission from Elsevier)

15.5. THE PYRUVATE DEHYDROGENASE COMPLEX

At the junction between glycolysis and the citric cycle, the multi-enzymatic complex of pyruvate dehydrogenase catalyses the transformation of pyruvate into acetyl-CoA following the global reaction:

$$\text{pyruvate} + \text{CoA} + \text{NAD}^+ \longrightarrow \text{acetyl-CoA} + \text{CO}_2 + \text{NADH}$$

It is constituted by the assembly of multiple copies of three different kinds of enzymes, the component E1 that is a pyruvate decarboxylase, the component E2 that is a dihydrolipoamide acetyl transferase and the component E3 that is a dihydrolipoamide dehydrogenase.

15.5.1. FUNCTIONAL PROPERTIES

The component E1, which has the pyruvate decarboxylase activity, functions in the presence of a coenzyme, thiamine pyrophosphate (TPP) and catalyses the reaction:

$$\text{pyruvate} + \text{TPP} \longrightarrow \text{hydroxyathyl-TPP} + \text{CO}_2$$

The catalytic mechanism requires the intervention of TPP, whose carbon ionises to give a carbanion to which the pyruvate is added. The addition compound is thus formed. The positive charge of the nitrogen attracts electrons, allowing decarboxylation and the formation of hydroxyethyl-TPP (Fig. 15.14).

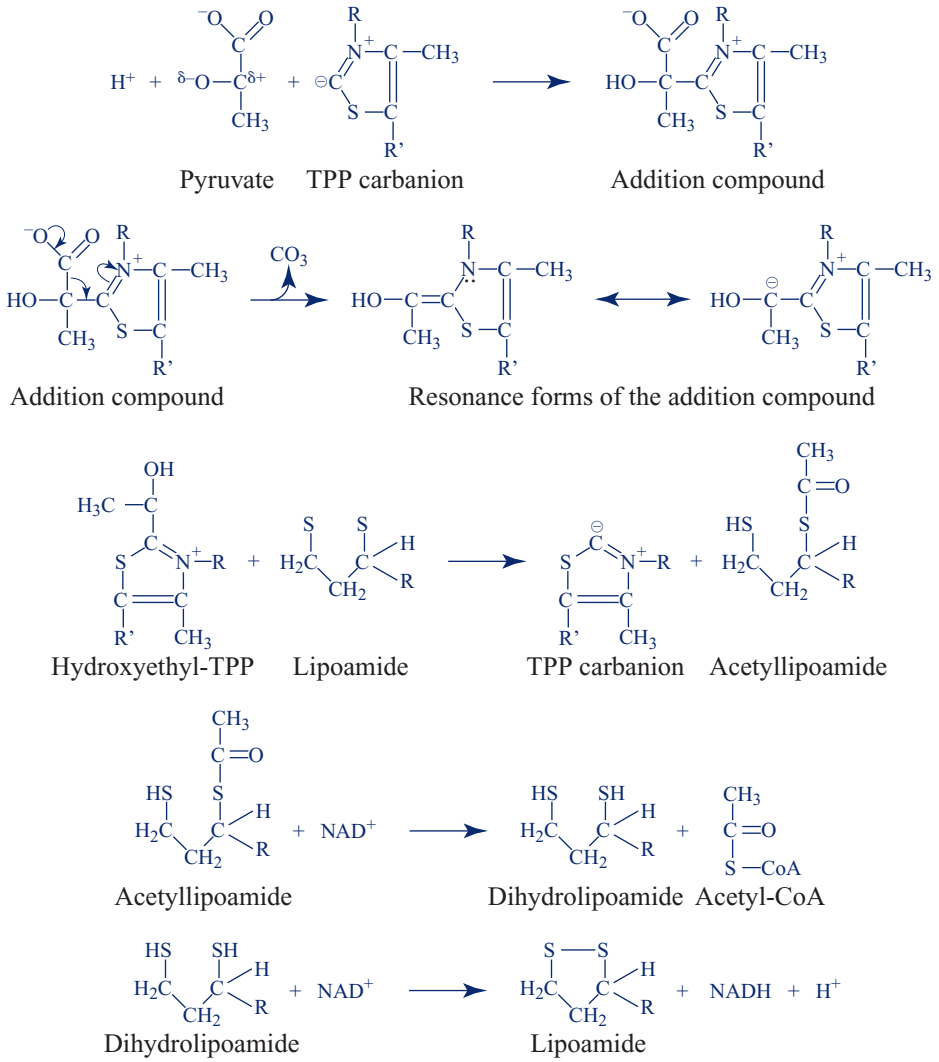


Fig. 15.14 Detail of the reactions catalysed by the three enzymes of pyruvate dehydrogenase

(Reprinted from *Biochemistry,* STRYER L. et al., 5th ed., 516, 2002. © W.H. Freeman and Company/Worth Publishers)

The component E2 is a dihydrolipoamide acetyl transferase, it contains a lipoic acid covalently linked to a lysine; it is lysine 42 in the lipoyl domain of the pyruvate dehydrogenase from *B. stearothermophilus*. The hydroxyethyl group linked to the TPP is oxidised to form an acetyl group, which is transferred to the lipoamide. The

oxidant is the disulphur group of the lipoamide that is transformed into a sulfhydril group (Fig. 15.14). Dihydrolipoamide transacetylase catalyses the transfer of the acetyl group of acetyllipoamide on the coenzyme A to form acetyl-CoA. Dihydrolipoamide dehydrogenase catalyses the regeneration of the oxidised form of the lipoamide (Fig. 15.14). The global diagram of these reactions is summed up in Fig. 15.15.

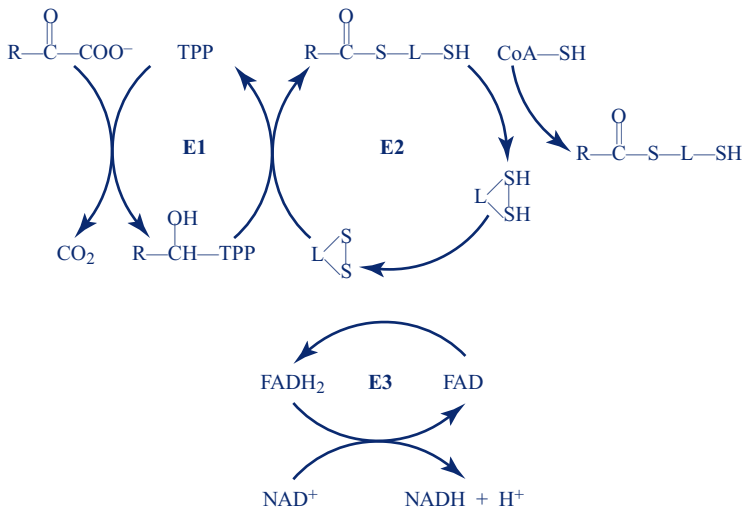


Fig. 15.15 Diagram of the reactions catalysed by the multi-enzymatic complex of pyruvate dehydrogenase
L represents the lipoamide group

15.5.2. STRUCTURAL PROPERTIES

Pyruvate dehydrogenase is an enormous complex, its molecular mass varies between 5 and 10 million Daltons. In *E. coli*, this complex has a polyhedral structure of 30 nm in diameter, as was revealed by electron microscopy studies (REED, 1974). It is constituted of 24 subunits of the enzyme E1, 24 subunits of the enzyme E2 and 12 subunits of lipoamide dehydrogenase (Fig. 15.16).

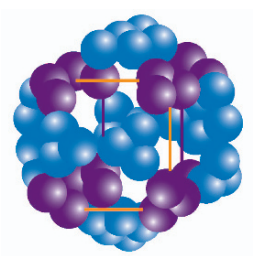


Fig. 15.16 Dodecahedral model of the multi-enzymatic complex of pyruvate dehydrogenase
(Reprinted from *PNAS,* **96**, IZARD T. et al., Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes, 1240. © (1999) National Academy of Sciences, USA)

In all the enzymes studied, the transacetylase subunits (E2) form the centre of the complex. However, the architecture of the enzyme varies depending on the species.

In mammals and in Gram positive bacteria, there are 60 copies of E2, whose association forms an icosahedron, while in Gram negative bacteria, 24 subunits are organised following an octahedral symmetry.

The structure of this complex is currently known to atomic resolution. That of the enzyme E1 was resolved more recently than E2 and E3. The HOL group determined the structure of the enzyme from *Pseudomonas putida* in 1999 (ÆVARSSON et al., 1999), and that of the human enzyme in 2000 (ÆVARSSON et al., 2000) by X-ray crystallography. The enzyme E1 is an $\alpha_2\beta_2$ heterotetramer formed of two types of subunits, α of 410 residues and β of 339 residues. The two α subunits and the two β subunits are linked by a 2-fold symmetry axis. Each of the subunits of the tetramer interacts with the three others by large contact surfaces. At the centre of the tetramer is a small 10 Å cavity, filled with water molecules, that connects the two active sites. Each α subunit is constituted by a large N-terminal domain of α/β type with a central β sheet surrounded by 16 helices, then a few additional helices, one of which has 23 residues, which increase the size of the tetramer. The small C-terminal domain contains two anti-parallel helices. The β subunits are folded into two domains linked by a flexible loop. The N-terminal domain possesses a central β sheet formed of six parallel strands, surrounded by 7 helices. The smaller C-terminal domain is formed of a β sheet of five strands surrounded by four helices.

More recently, the structure of the β subunit isolated from the archaeon *Pyrobaculum aerophilum* was determined (KEIGER et al., 2001). The global structure of the enzymes from the three species is similar.

The multi-enzymatic complex from *Azotobacter vinelandi* is also well-characterised. The three-dimensional structure of the recombination catalytic domain of the subunit E2 was resolved by X-ray diffraction to 12.6 Å by HOL's group (MATTEVI et al., 1992). Each subunit E2 is constituted of several domains, one or several lipoamide domains that contain about one hundred residues, each bearing a lipoamide molecule bound to a lysine residue, a subunit E3 association domain of about 50 residues and a C-terminal catalytic domain of 250 residues that contains the catalytic centre and the association site with another subunit E2. These domains are bound by segments rich in alanine and proline residues which confer a great flexibility, important for enzymatic activity; in particular, the mobility of the lipoamide domains allows them to interact with the catalytic groups of different enzymes of the complex. The structure of the subunit E2 association domain with the peripheral subunit of *B. stearothermophilus* free and complexed with the subunit E3 was resolved by X-ray crystallography and by NMR. The structures of the lipoamide domain from *B. stearothermophilus* and *E. coli* were resolved by NMR.

The structure of the subunit E2 catalytic domain of pyruvate dehydrogenase from *A. vinelandii* obtained by genetic recombination showed that each monomer is tightly associated to the two other subunits forming a kind of cylindrical trimer. The N-terminal part of a monomer stretches over the neighbouring subunit (Fig. 15.17 below).

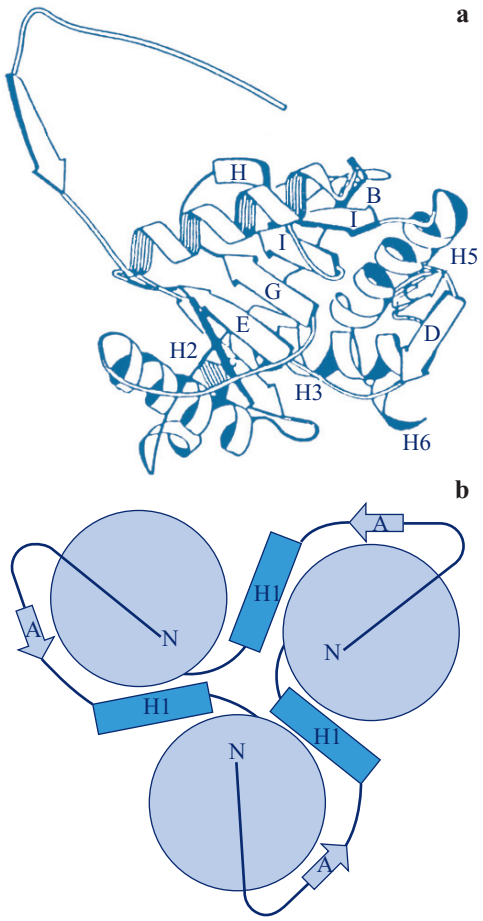


Fig. 15.17 (a) Three-dimensional structure of the subunit E2 catalytic domain of pyruvate dehydrogenase from A. vinelandii – (b) schematic representation of the trimer showing the organisation of the N-terminal part of a subunit interacting with the neighbouring subunit (From *Science*, **255**, MATTEVI A. et al., Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex, 1544. © (1992) reprinted with permission from American Association for the Advancement of Sciences)

This structure displays analogies with that of chloramphenicol acetyl transferase, and certain sequence homologies that were used to identify residues susceptible to be important in catalysis, in particular histidine 610 (Fig. 15.18 opposite). The 24 subunits of the enzyme E2 are organised following a 432 symmetry. While the contacts between subunits in the trimers are very tight, the number of inter-trimer contacts is small. The assembly of the 8 identical trimers generates a sort of cage, whose internal cavity is connected to the exterior by channels.

The structure of the enzyme E3 dimer from *A. vinelandii* was determined to a resolution of 2,2 Å. On this basis with electron microscopy data a model of the dodecahedral oligomer of 60 subunits and of 100 Å diameter was proposed.

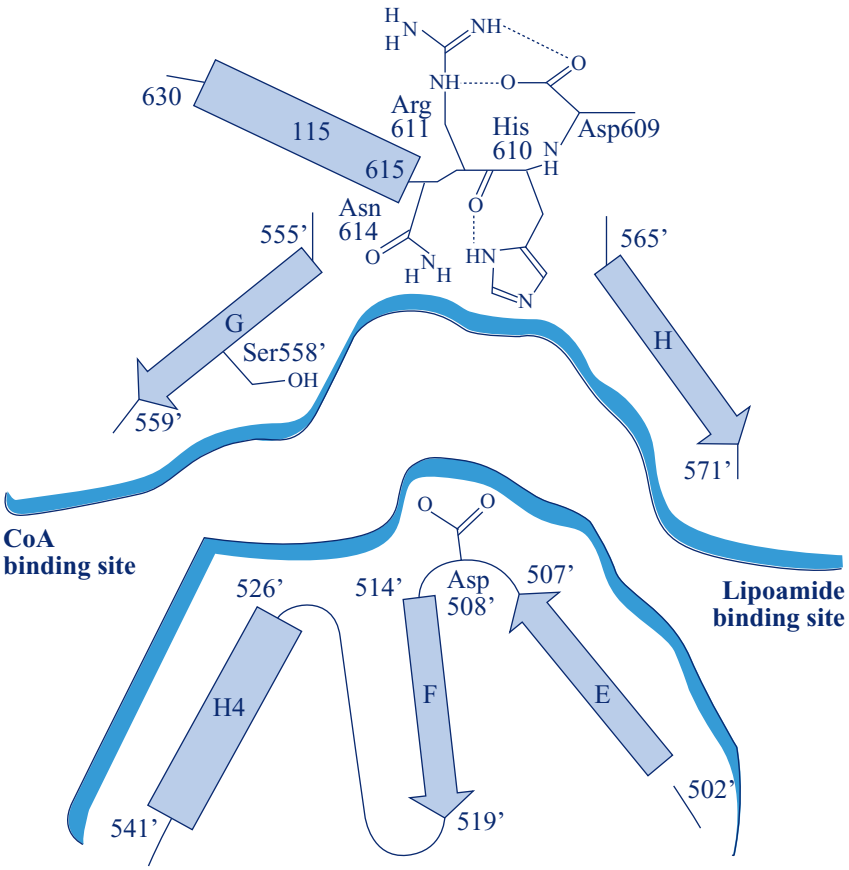


Fig. 15.18 Schematic representation of the catalytic centre of the enzyme E2, showing the 29 Å long channel between the subunits of the trimer with the CoA binding sites and the lipoamide group on the two opposite entrances to the channel

(From *Science*, **255**, MATTEVI A. et al., Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex, 1544. © (1992) reprinted with permission from American Association for the Advancement of Sciences)

15.5.3. ROLE OF LIPOAMIDE DOMAINS IN SUBSTRATE CHANNELLING

The lipoamide groups are bound to small domains of the enzyme E2 which, in all the 2-oxo acid dehydrogenases, have sequence homologies and probably three-dimensional structure analogies. The number of these lipoamide domains varies however according to species (Fig. 15.19 below). Mutagenesis experiments showed that the deletion of one or two lipoamide domains of pyruvate dehydrogenase from *E. coli* does not lead to a variation in the enzymatic activity; however the replacement of the lysine that binds the lipoic acid by a glutamine leads to a loss of activity. It seems then that a single lipoamide domain is sufficient for the enzyme's function (PERHAM, 1991). The conformational flexibility of the lipoyllysine arm allows the

lipoamide group access to the catalytic sites of the different enzymes and ensures the channelling of substrates, which depends on the selective recognition of the E2 lipoamide domain by the enzyme E1 in the corresponding step of the reaction. The interaction between the pyruvate decarboxylase (E1) and the E2 lipoamide domain in the pyruvate dehydrogenase complex from *B. stearothermophilus* was clearly demonstrated by studies combining site-directed mutagenesis and NMR spectroscopy (WALLIS et al., 1996).

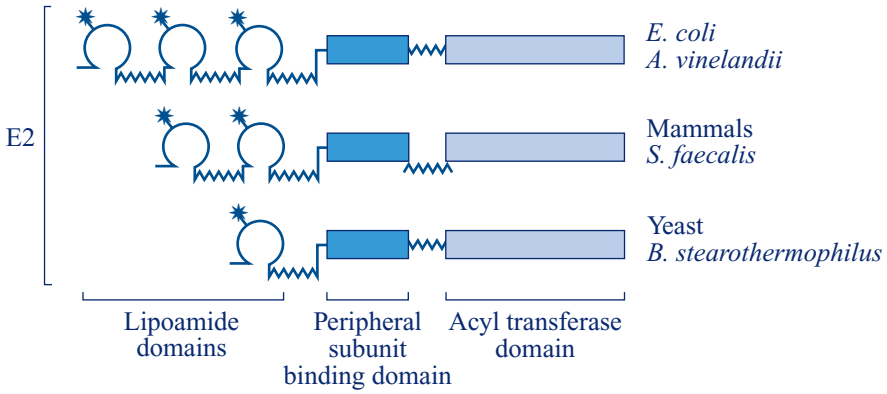


Fig. 15.19 Schematic representation of the domain organisation in pyruvate dehydrogenase from different species

The starred section represents the lipoamide arm (Reprinted with permission from *Biochemistry*, **30**, PERHAM R.N., 8504. © (1991) American Chemical Society)

A fairly similar multi-enzymatic system is found in α-ketoglutarate dehydrogenase, which catalyses the transformation of α-ketoglutarate into acetyl-CoA.

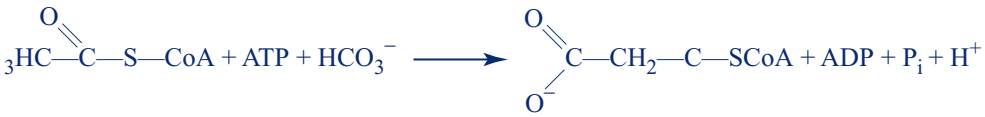
15.6. FATTY ACID SYNTHETASE

Fatty acid synthetase is an enzymatic system that catalyses the sequential reactions involved in the synthesis of long-chain saturated fatty acids. There are two types of fatty acid synthetases: those of type I, found in animal organisms and in mushrooms, are multi-functional proteins comprising several different active sites on the same polypeptide chain; those of type II, which exist in the majority of prokaryotic organisms and in plants, are constituted of mono-functional enzymes and an acyl carrier protein (ACP, see further on). Such a multi-functional complex was isolated for the first time by LYNEN and collaborators in 1969.

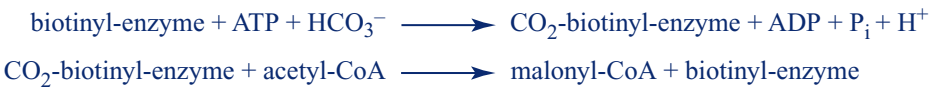
15.6.1. FUNCTIONAL PROPERTIES

The synthesis of saturated fatty acids occurs in the cytosol, unlike degradation that takes place in the mitochondrial matrix. The elongation of the fatty acid chain

occurs by sequential addition of two carbon atom units derived from acetyl-CoA. Malonyl-ACP (ACP for Acyl Carrier Protein) acts as the active donor of the two carbon units in the elongation step. As WAKIL showed, fatty acid synthesis starts by the carboxylation of acetyl-CoA into malonyl-CoA catalysed by acetyl-CoA carboxylase. The reaction is the following:

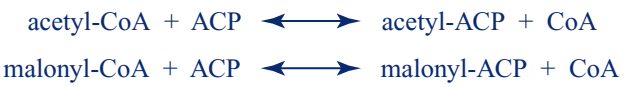


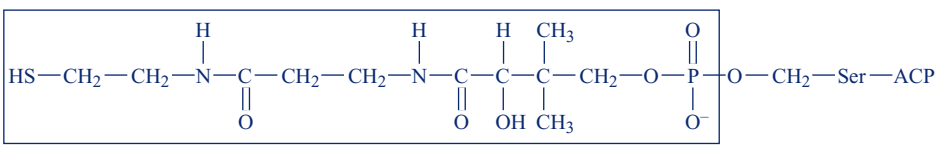
This enzyme which is distinct from the multifunctional complex, has biotin covalently bound to a lysine ε-amino group as its prosthetic group. The reaction occurs in two steps with the formation of an intermediate carboxybiotinyl-enzyme; this first reaction requires the presence of ATP:



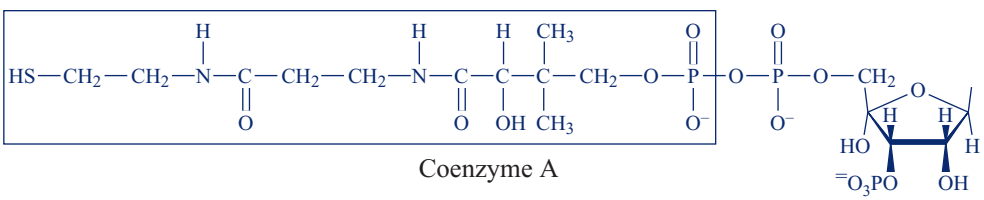
The activity of acetyl-CoA carboxylase is under allosteric control of citrate and isocitrate that are activators.

The intermediates in the synthesis of fatty acids are bound to a protein domain that ensures the acyl transfer, the acyl carrier protein or ACP. The bond is made by the SH terminal group of phosphopantetheine bound to an ACP serine; the phosphopantetheine group represents the active part of ACP and CoA (Fig. 15.20). The flexibility of the phosphopantetheine part, the maximum length of which is 20 Å, appears critical for the formation of the multi-enzymatic complex. The first phase of the elongation resides in the formation of acetyl-ACP and malonyl-ACP, catalysed respectively by acetyl transacylase and malonyl transacylase, following the reactions:





Phosphopantetheine: ACP prosthetic group



Coenzyme A

Fig. 15.20 Phosphopantetheine, reactive unit of ACP and CoA

The condensation of acetyl-ACP and malonyl-ACP then occurs to form a four carbon unit, the acetoacetyl-ACP, catalysed by the condensation enzyme:

$$\text{acetyl-ACP} + \text{malonyl-ACP} \longleftrightarrow \text{acetoacetyl-ACP} + \text{ACP} + CO_2$$

The three following reactions lead to the reduction of the ketone group in C3 to a methylene group. First, the acetoacetyl-ACP is reduced to D-3-hydroxybutyryl-ACP, NADPH being the reducing agent that functions as a coenzyme of β-hydroxyacyl-ACP dehydrogenase. The D-3-hydroxybutyryl-ACP is then dehydrated to form crotonyl-ACP; this reaction is catalysed by enoyl-ACP hydratase. In the last reaction, this substrate is reduced to butyryl-ACP by enoyl-ACP reductase, whose coenzyme NADPH is the reducing agent (Fig. 15.21).

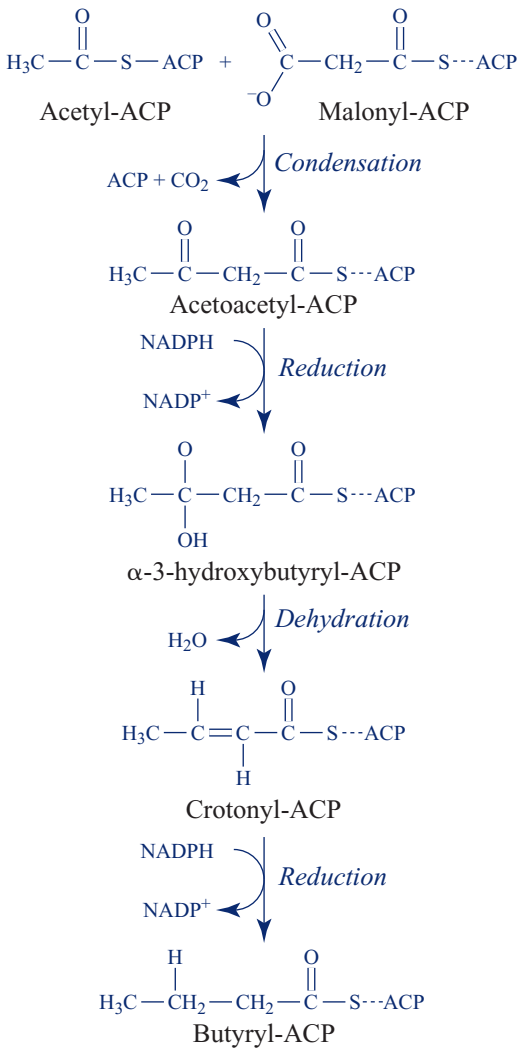


Fig. 15.21 Sequence of reactions in fatty acid synthesis: condensation, reduction, dehydration and reduction

In the yeast synthetase, the enoyl reductase activity requires the presence of FMN; the enoyl reductase acts as a transhydrogenase by transferring the hydride of NADPH to the double bond of the fatty acid *via* the FMN. No flavin free radical has been detected, so it would seem that the reduction implies the transfer of two electrons. The butyryl-ACP so formed can then be condensed with a new malonyl-ACP molecule which, in a second cycle of synthesis implying condensation, reduction, dehydration and reduction, leads to C6-acyl-ACP. The following cycles of elongation end in the formation of C16-acyl-ACP, which is hydrolysed to give palmitate and ACP (Fig. 15.22).

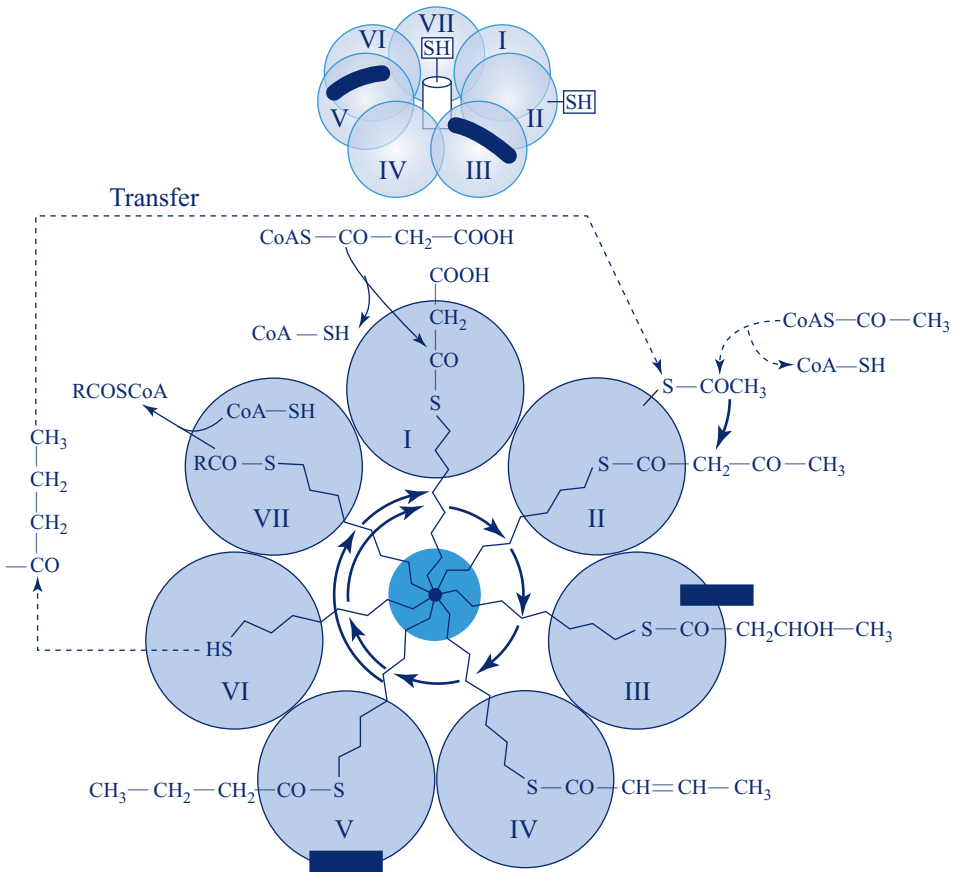


Fig. 15.22 Schematic representation of saturated fatty acid biosynthesis, showing the rotation of the system in several elongation cycles, according to LYNEN
(Reprinted from *PNAS*, **75**, WIELAND F. et al., Distribution of yeast fatty acid synthetase subunits: three-dimensional model of the enzyme, 5792, 1978 with authors' permission)

15.6.2. STRUCTURAL CHARACTERISTICS

The yeast enzymatic complex and that of animal tissues differ slightly in their structural organisation, although they are both type I synthetases. The available data were

obtained by electron microscopy, small angle X-ray scattering and chemical analysis. In contrast, the structures of a few monofunctional enzymes from *E. coli* were resolved by X-ray crystallography. The yeast fatty acid synthetase is a multi-enzymatic complex of molecular weight 2.3×10^6. It is composed of two types of subunits, the α and β chains of respective molecular weights 185 000 and 180 000. Each of these subunits is polyfunctional; the β subunit catalyses 5 out of the 7 partial reactions and the α subunit the two others. The α subunit possesses the enzymatic activities of condensation and β-ketoacyl-reductase. In electron microscopy, the enzyme appears as an ellipsoid, the long axis being 250 Å and the short 210 Å. Small angle X-ray scattering studies showed that the molecule is a hollow ellipsoid comprising a dense part along the short axis. These different data, taken in conjunction with immunochemical studies, led the LYNEN group to propose an $\alpha_6\beta_6$ type structure (WIELAND et al., 1978). Figure 15.23 shows the subunit organisation in the yeast fatty acid synthetase.

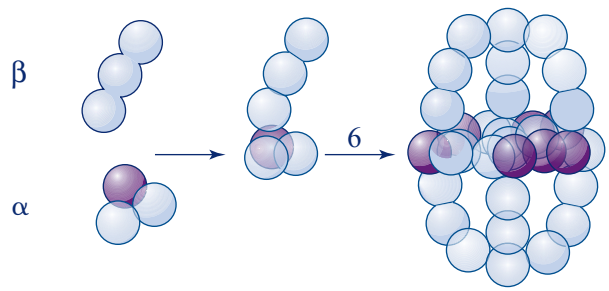


Fig. 15.23 Structure of the fatty acid synthase according to LYNEN
(Reprinted from *PNAS*, **75**, WIELAND F. et al., Distribution of yeast fatty acid synthetase subunits: three-dimensional model of the enzyme, 5792, 1978, with authors' permission)

The synthetase from animal tissues possesses a different structure. The structure is an α_2 homodimer, the subunit has a molecular weight of ~250 000. Electron microscopy studies show that the dimer has a length of 216 Å and a width of 144 Å (KIMOTO et al., 1988). Each protomer contains the ACP site and the seven catalytic activities. However, only the dimer is the active form of this multifunctional protein. The two subunits that contain the same catalytic domains are positioned "head-to-tail", so that the reactive cysteine SH group of the condensation enzyme and the SH group of the ACP phosphopantetheine are found in close proximity (Fig. 15.24 opposite). The fact that the enzymes of the two subunits are simultaneously functional was the subject of controversy. Two models were proposed; in one each subunit plays the role of catalyst and coordinator, in the other the symmetrical dimer possesses independent sites that function simultaneously. It seems that the animal fatty acid synthetase functions as in the second model (SMITH et al., 1985). The necessity of the dimeric structure is linked to the "head-to-tail" arrangement of the two protomers, which allows the thiol group of the reactive cysteine of the condensation enzyme and that of the phosphopantetheine group to come closer together.

In bacteria, the main use of fatty acids is their incorporation in cellular membranes. The different reactions of their synthesis are performed by monofunctional cytoplasmic enzymes, each catalysing a specific reaction. The structures of two enzymes from *E. coli* were resolved by X-ray diffraction, those of malonyl-CoA-ACP trans-

acylase (SERRE et al., 1995) and enoyl reductase bound to NAD^+ (RAFFERTY et al., 1998). Malonyl-CoA-ACP transacylase displays an α/β type architecture, but its fold appears unique. The protein is constituted of two domains. The large domain comprises two non-contiguous segments including residues 3–123 and 206–307. It contains a β sheet formed of four parallel strands and twelve α helices. The small domain (residues 124–205) contains a β sheet consisting of four antiparallel strands capped by two helices. The active site is found in a groove between the two domains. The nucleophile Ser92 group forms a hydrogen bond with His201 in a manner that is reminiscent of the organisation of the serine hydrolase catalytic groups. Instead of a carboxylic group however, it is the carbonyl of the Gln250 main chain that is the acceptor in the hydrogen bond with His201.

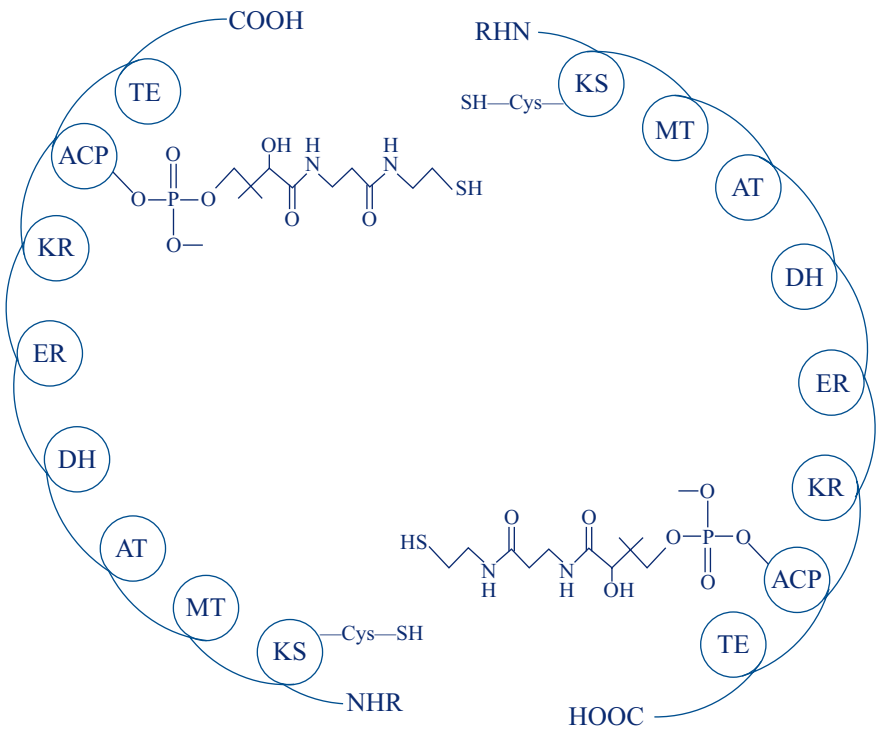


Fig. 15.24 Schematic representation of the organisation of the two chains in the fatty acid synthetase from animal tissues

KS: ketoacyl synthetase (or condensation enzyme); MT: malonyl transacylase; AT: acetyl transacylase; DH: dehydratase; ER: enoyl reductase; KR: ketoacyl reductase; ACP: acyl carrier protein; TE: thioesterase (Reprinted, with permission, from the Annual Review of Biochemistry, **52**, WAKIL S.J. et al., 537. © (1983) by Annual Reviews www.annualreviews.org)

The structure of enoyl-ACP reductase associated with NAD^+ was obtained at a resolution of 2.1 Å. The enzyme is a homotetramer, each of the subunits having bound NAD^+. Each protomer comprises a β sheet constituted of 7 parallel strands flanked on each side by three α helices; an additional helix is also found close to the

C-terminal extremity of the β sheet. The fold is reminiscent of the ROSSMANN fold of NAD^+ dehydrogenases. The comparison with a family of alcohol dehydrogenases allowed the identification of a highly conserved tyrosine and a lysine residue, which are important for catalytic activity. Modelling studies suggested that a region of the protein surface that contains several highly conserved hydrophobic residues and that is close to the nicotinamide centre, could form the fatty acid binding site.

It is interesting to note that, in higher organisms, evolution favoured the association of these enzymes catalysing the seven reactions of the fatty acid biosynthesis as well as ACP into a single multifunctional complex, carried by a single polypeptide chain and sequentially organised. Separated in prokaryotic organisms, these enzymes are found partially condensed in two chains in yeast, each bearing several enzymatic activities. In animal organisms, all the enzymatic activities and ACP are found united in the same chain. It seems that evolution has in this manner proceeded towards the system's stabilisation.

15.7. TRANSITORY MULTIENZYMATIC COMPLEXES AND CHANNELLING

Until now, we have considered the case of enzymes from the same metabolic pathway associated in stable multi-enzymatic complexes. In these complexes, there is a regular organisation that ensures the proximity of active sites. In addition, certain enzymes, in the cellular context, interact more or less strongly with each other. This is the case for the membrane enzymatic systems such as enzymes from the respiratory pathway of mitochondria, but also for certain cytosolic enzymes. The concentration of enzymes in the cytosol is often very high. Specific interactions between different enzymes can have functional consequences. The interactions are often very weak however, which makes it difficult to prove the existence of such complexes. In some cases, only arguments from kinetic studies suggest the existence of interactions between several enzymes. Their existence is therefore still the subject of controversy.

The specific aggregation of glycolytic pathway enzymes was reported by different authors (SRIVASTAVA & BERNHARD, 1985), in particular in muscle and erythrocytes. In muscle sarcoplasm, the glycolysis enzymes are present at concentrations of the order of 0.1 mM and even higher (Table 15.2 opposite), so that the concentration of active sites is in excess compared to that of intermediate metabolites. It is also known that certain glycolysis enzymes associate at erythrocyte membranes in vitro. The association of enzymes that catalyse two consecutive reactions can have the consequence of direct metabolite transfer from one site to the other, which avoids the transfer by free diffusion through the solvent. In this respect, an interaction between aldolase and glyceraldehyde-3-phosphate dehydrogenase of rabbit muscle was shown by the KÉLÉTI group (OVADI et al., 1978). The kinetic study of the

reaction shows the formation of a complex between the two enzymes. In addition, the existence of such a complex was observed by fluorescence polarisation measurements; an apparent dissociation constant of 3×10^{-7} M was determined (OVADI et al., 1978). Even weaker protein-protein interactions ($K_d \sim 10^{-3}$ M) can have functional consequences. However the detection in vitro of such complexes is very difficult, because of the weak affinity of these molecules. Nevertheless, MC GREGOR et al. (1980) showed by gel penetration the formation of a 1:1 stoechiometry complex of 300 kDa between aldolase and fructose 1,6-biphosphatase from rabbit liver; the functional consequences of such a complex remain unknown however.

Table 15.2 Cellular concentrations of glycolysis enzymes

(This article was published in *Curr. Top. Cell. Regul.*, **28**, SRIVASTAVA D.K. & BERNHARD S.A., Enzyme-enzyme interactions and the regulation of metabolic reaction pathways, 1. © Elsevier (1986))

Enzyme	*Active site concentration (mM)*
Phosphoglucomutase	3.9
Aldolase	809.3
α-glycerolphosphate dehydrogenase	61.4
Triose phosphate isomerase	223.8
Glyceraldehyde-3-phosphate dehydrogenase	1398.6
Phosphoglycerate kinase	133.6
Phosphoglycerate mutase	235.9
Enolase	540.7
Pyruvate kinase	172.9
Lactate dehydrogenase	296.0

Amongst glycolysis enzymes, an interaction between glyceraldehyde-3-phosphate dehydrogenase (GPDH) and phosphoglycerate kinase (PGK) from halibut muscle was proposed by the BERNHARD group, based on kinetic data (WEBER & BERNHARD, 1982). The experimental approach consisted in studying the coupled reaction between the two enzymes:

$$NADH + PGK\text{-}1,3\text{-}DPG \longrightarrow G\text{-}3\text{-}P + NAD^{+} + HPO_4^{2-} + PGK$$

1,3-DPG is the 1,3-diphosphoglycerate, 3-PG the 3-phosphoglycerate. In these experiments, the concentration of PGK was equal or superior to two times the concentration of 1,3-DPG and very superior to that of GPDH. The 1,3-DPG has a very strong affinity for PGK, the dissociation constant of the complex is of the order of 1 nM. In these conditions, all the 1,3-DPG is then bound to PGK. The steady state conditions for the 1,3-DPG and the 1,3-DPG-PGK complex are fulfilled. A slow disappearance of NADH is observed. In contrast, in the same conditions, but in the absence of PGK, the reaction is instantly finished (Fig. 15.25 below). The NADH disappearance rate was studied as a function of the concentration of each of the two enzymes. The results obtained indicate that the transfer of the 1,3-diphosphoglycerate between the glyceraldehyde-3-phosphate dehydrogenase and the phosphoglycerate

kinase happens directly by a channelling mechanism, involving a complex between the two enzymes:

$$E_1S + E_2 \longleftrightarrow E_1SE_2 \longleftrightarrow \text{products}$$

and not by transfer through free diffusion:

$$E_1S + E_2 \longleftrightarrow E_1 + S + E_2 \longleftrightarrow E_2S \longleftrightarrow \text{products}$$

E_1 is the phosphoglycerate kinase and E_2 the glyceraldehyde-3-phosphate dehydrogenase; the phosphoglycerate kinase-1,3-diphosphoglycerate complex behaves as the substrate of the glyceraldehyde-3-phosphate dehydrogenase.

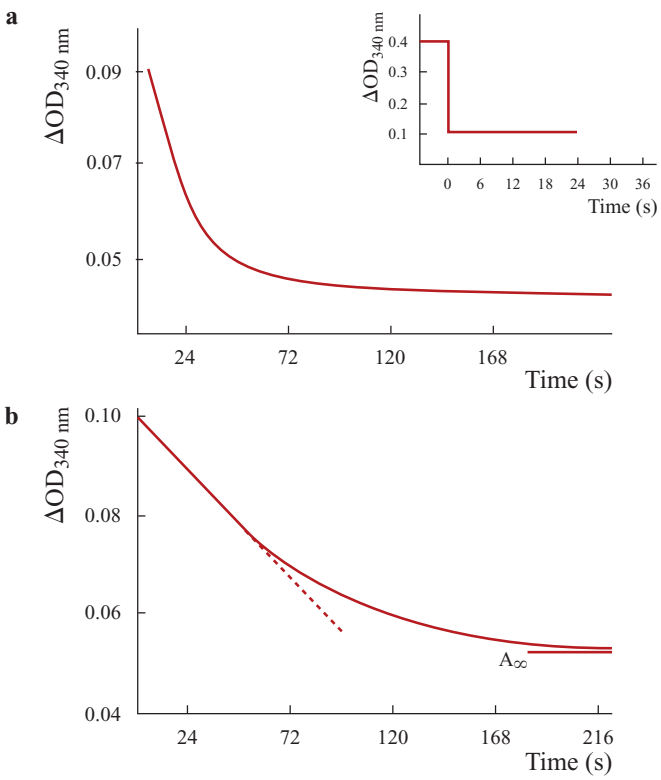


Fig. 15.25 Recording of the NADH disappearance reaction

(a) *[GPDH] = 5 × 10^{-7} M; [PGK-1,3-DPG] = free [PGK] = 8 × 10^{-6} M; [NADH] = 2 × 10^{-4} M, pH 7.4, 25°C. The GPDH is that of rabbit muscle, PGK is from halibut. In inset, the same reaction in the same conditions with free 1,3-DPG and without PGK –* ***(b)*** *GPDH is from halibut. The conditions are the following: [GPDH] = 10^{-7} M; [PGK-1,3-DPG] = 8,4 × 10^{-6} = free [PGK]; NADH] = 2 × 10^{-4} M, pH 7.4, 25°C*

(Reprinted with permission from *Biochemistry*, **21**, WEBER J.P. & BERNHARD S.A., 4189. © (1982) American Chemical Society)

An analogous situation was observed by SRIVASTAVA and BERNHARD (1984) in the case of dehydrogenases. A direct transfer of NADH between glyceraldehyde-

3-phosphate dehydrogenase and alcohol dehydrogenase of horse liver was shown by the same experimental procedure. The rate of benzaldehyde and p-nitrobenzaldehyde reduction by NADH, catalysed by alcohol dehydrogenase is faster when NADH is bound to glyceraldehyde-3-phosphate dehydrogenase. The results of the kinetic studies indicate a direct transfer of the coenzyme from GPDH to the alcohol dehydrogenase. In contrast, in similar conditions, no direct transfer of the coenzyme was observed between lactate dehydrogenase and alcohol dehydrogenase. A direct transfer of NADH was observed only when both dehydrogenases exert different chiral specificities for the hydrogen transfer between the C4 of the nicotinamide centre and the substrate. There are two classes of dehydrogenases, depending on whether NADH in the enzyme has the *anti* or *syn* configuration (Fig. 15.26); they are designated by A and B respectively.

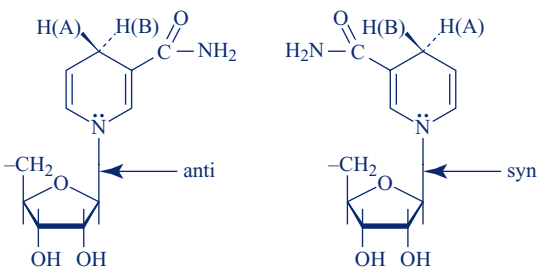


Fig. 15.26 Anti (A) and syn (B) configurations of NADH bound to dehydrogenase A and B respectively
The chirality of the hydrogens H(A) and H(B) at C4 is shown in the two configurations
(From *Science*, **234**, SRIVASTAVA D.K. & BERNHARD S.A., Metabolite transfer *via* enzyme-enzyme complexes, 1081. © (1986) reprinted with permission from American Association for the Advancement of Sciences)

From the results of SRIVASTAVA and BERNHARD (1986), the direct transfer of the coenzyme occurs only between the A–B pairs, and not between A–A or B–B pairs (Table 15.3 below). The results obtained by these authors indicate a strong association of NADH in the complex between the two dehydrogenases compared to the association of the coenzyme in each individual enzyme. The authors concluded that the interactions between pairs of dehydrogenases seem to modulate the dissociation and the transfer of the coenzyme from one site to the other.

The association of enzymes that induces channelling seems then to display specificity. The specific recognition between two enzymes can result in conformational changes upon substrate or coenzyme binding, which results in the appearance of interaction surfaces; these surfaces would be hidden in the free enzyme structure, as proposed by SRIVASTAVA and BERNHARD (1985). It is known that conform-ational changes occur that are induced by phosphogylcerate substrates in phosphoglycerate kinase and by NAD binding in alcohol dehydrogenase. According to the same authors, the interaction between enzymes in a sequential metabolic pathway would facilitate the transfer of metabolites, maintaining equilibrium constants close to unity, and would constitute in this way a regulation system. In particular, in the glycolytic pathway, the system of enzymes would function as a metabolic reserve, allowing the fast interconversion between intermediary metabolites and final

products. In this perspective, the unidirectional character of the metabolic flux is due to the use of the final product by another metabolic pathway.

Table 15.3 Comparison of the reduction rates of the S2 substrate by NADH, catalysed by the enzyme E_2 in the presence of the enzyme E_1, with those predicted assuming free NADH as the only competent coenzyme

(From *Science*, **234**, SRIVASTAVA D.K. & BERNHARD S.A., Metabolite transfer *via* enzyme-enzyme complexes, 1081. © (1986) reprinted with permission from American Association for the Advancement of Sciences)

E_1	E_1-site (m*M*)	[NADH] (m*M*)	E_2 (n*M*)	E_2	Rate (m*M*. min^{-1}) Predicted	Observed	Stereochemistry E_1	E_2
GPDH	220.0	134	LDH	0.029	0.17	1.05	B	A
LDH	170.0	49.5	GPDH	0.108	0.02	0.52	A	B
LDH	141.0	32.6	αGDH	0.138	0.03	0.32	A	B
GPDH	47.0	22.5	αGDH	0.211	0.24	0.24	B	B
αGDH	67.0	49.5	GPDH	0.108	0.23	0.16	B	B
αGDH	168.7	35.8	LDH	0.032	0.02	0.10	B	A
GPDH	104.1	16.7	LADH	0.56	0.03	0.88	B	A
LADH	41.0	24.4	LDH	0.02	0.08	0.08	A	A
LDH	168.7	38.6	LADH	0.45	0.08	0.08	A	A

The work and conclusions of the BERNHARD group were questioned by CHOCK and GUTFREUND (1988) who re-examined the kinetics of the NADH transfer between glycerol-3-phosphate dehydrogenase and lactate dehydrogenase. According to these authors, the dissociation rate of NADH from the NADH-glycerolphosphate dehydrogenase complex determined by SRIVASTAVA and BERNHARD is much to low, because the authors failed to take into account its strong pH dependence which, in their conditions, is strongly reduced by the effect of the high NAD concentrations. CHOCK and GUTFREUND again performed kinetic studies both in steady and pre-steady state conditions. On the basis of their results, they refuted the interpretation of SRIVASTAVA and BERNHARD of a direct coenzyme transfer between the two dehydrogenases and concluded that the transfer occurs by free diffusion.

In response to these criticisms, SRIVASTAVA et al. (1989) performed new kinetic experiments, both in the steady and pre-steady states, with a careful control of pH. As a control, the experiments were independently performed in two different laboratories. The results of both laboratories coincided perfectly, and differed from those of CHOCK and GUTFREUND. Taking into account the CHOCK and GUTFREUND remark on the NADH dissociation rate from the NADH-glycerolphosphate dehydrogenase complex, SRIVASTAVA et al. did find an increased value for the rate. However, the value found is not sufficiently high to reject a direct transfer mechanism of the coenzyme between the two enzymes that their results confirm. In addition, the fact that certain dehydrogenases of A–A or B–B type do not show any direct transfer constitutes a significant control.

In turn, SRIVASTAVA and collaborators raise several criticisms to the GUTFREUND work. In particular, they underlined the incorrect measurement of K_m for NADH in the reaction catalysed by glycerolphosphate dehydrogenase, the small concentration range of NADH making it impossible to distinguish between the two mechanisms. In a later paper the GUTFREUND group (WU et al., 1991) tried to provide direct proof of a complex between the two enzymes in the presence of saturating NADH concentrations, by sedimentation and gel filtration. No complex was observed and, taking into account the concentration range used, the authors considered that if such a complex is formed, its dissociation constant must be > 100 μM. They deduce from their experiments that the substrate channelling in glycolysis is a "phantom phenomenon". However, a simple calculation shows that in the experimental conditions of SRIVASTAVA and BERNHARD (see Table 15.3), when E1 is the glycerolphosphate dehydrogenase and E2 the lactate dehydrogenase, 20% of the latter would be in complexed form at equilibrium if the dissociation constant is 10^{-3} M and 40% if it is 5×10^{-4} M.

The controversy over channelling in the glycolysis enzymes illustrates well the difficulty of proving a direct transfer of substrate or coenzyme between two enzymes when they do not form a stable and well-defined complex with a regular geometry. In this case, the proof of direct transfer rests only on kinetic data, and the conditions to observe it must be carefully chosen.

OVADI (1988) proposed a glycolysis control mechanism by dynamic associations of enzymes modulated by metabolites, from data concerning the formation of complexes between different enzymes and on the oligomerisation states of the glycolysis enzymes in the complexes. The interactions are considered as being specific thus modulating enzyme activity. The interactions are under the control of metabolite concentration. Figure 15.27 below represents the molecular model of glycolytic pathway control according to OVADI. The top figure corresponds to the situation in which the fructose 1,6-biphosphate concentration is high, that of the bottom to the low concentration of this metabolite. OVADI considers that the dynamic assembly of glycolysis enzymes under the effect of metabolites is characterised by several properties. The enzyme associations occur by pairs and are specific. They are controlled by homologous interactions between identical subunits of the same enzyme. Their association/dissociation is modulated by fructose 1,6-biphosphate and other metabolites. The effects of substrate concentration changes and changes in association/dissociation of the enzymatic complexes are mutual. The species involved in the complexes are more active, and the transitory complexes ensure the channelling.

The existence of multi-enzymatic complexes in the metabolic pathways presents obvious advantages. The proximity of the enzymes catalysing the consecutive steps of a metabolic pathway decreases the required time for substrate diffusion and ensures that the substrates are not intercepted by enzymes of other metabolic pathways. By avoiding the passage of labile metabolites in solution, the protein matrix

ensures their protection. Channelling is therefore important to maintain the metabolic flux even with small concentrations of metabolites.

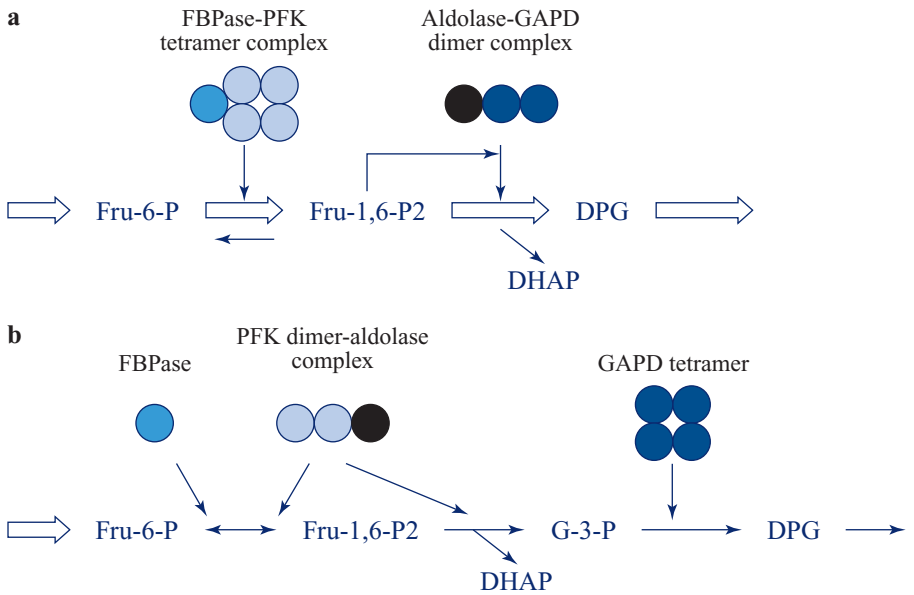


Fig. 15.27 Schematic representation of glycolysis regulation by metabolites
(a) *at a high concentration of fructose 1,6-bisphosphate –* ***(b)*** *at a low concentration of this metabolite. FBPase: fructose bisphosphatase; PFK: phosphofructokinase; GAPD: glyceraldehyde-3-phosphate dehydrogenase; Fru-6-P: fructose 6-phosphate; Fru-1,6-P2: fructose 1,6-bisphosphate; DPG: diphosphoglycerate; DHAP: dihydroxycetone phosphate* (Reprinted from *TIBS*, **13**, OVADI J., Old pathway - new concept: control of glycolysis by metabolite-modulated dynamic enzyme associations, 489. © (1988) with permission from Elsevier)

An example illustrates well the protection of metabolites when channelling exists. A fusion protein was obtained from citrate synthase (CS) and yeast malate dehydrogenase (MDH) and expressed in *E. coli*. The two enzymes were linked by a short Gly-Ser-Gly sequence (LINDBLADH et al., 1994). The enzymatic activity kinetics of the fusion protein and of the two free enzymes were compared. The results showed that the oxaloacetate is effectively channelled between MDH and CS. The transient time of the coupled reaction is decreased. In addition, the presence of another enzyme such as the aspartate amino transferase, which catalyses the transfer of the glutamate amino group to oxaloacetate, is not in competition for the intermediary substrate, although this competition exists when MDH and CS are free; this clearly proves the oxaloacetate confinement (Fig. 15.28 opposite) in the bi-enzymatic complex obtained by genetic recombination.

More recently, several studies showed the role of electrostatic effects in channelling efficiency. In particular, the efficiency of substrate transfer in the bifunctional enzyme dihydrofolate reductase-thymidylate synthase is lowered when the ionic

strength increases (TRUJILLO et al., 1996). In order to evaluate the importance of electrostatic effects in channelling, Brownian dynamics simulations were performed on a modelled structure of the MDH-CS fusion protein (ELCOCK & MC CAMMON, 1996). The calculations showed that in the absence of electrostatic forces, less than 1% of the molecules leaving the MDH active site are transferred to the CS active centre. However, at zero ionic strength, when electrostatic forces are present, 45% of molecules are transferred.

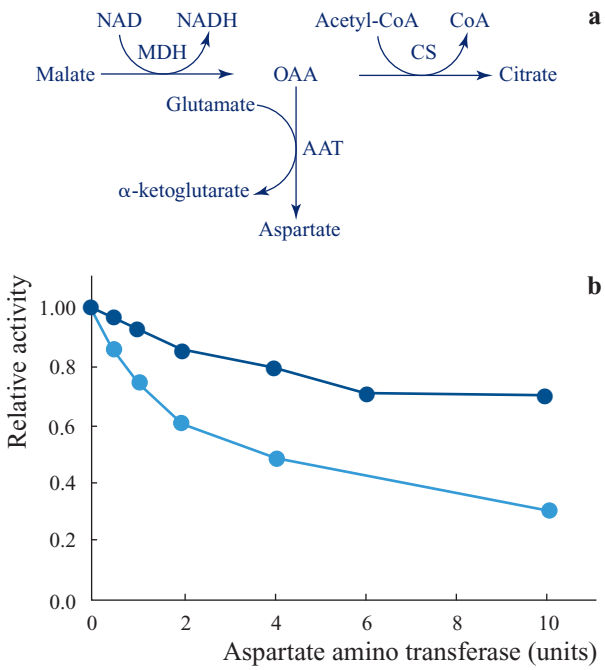


Fig. 15.28 (a) reaction scheme showing the competition between aspartate amino transferase (AAT) and citrate synthase (CS) for the oxaloacetate in the coupled reaction malate dehydrogense (MDH)/CS (b) effect of AAT in the coupled reaction catalysed by MDH and CS

The dark dots correspond to the fusion protein MDH-CS, the pale blue dots refer to an equimolar mix of MDH and CS at the same concentration as in the fusion protein

(Reprinted with permission from *Biochemistry*, **33**, LINDBLADH C. et al., 11692. © (1994) American Chemical Society)

The different examples treated in this chapter, multifunctional enzymes, multi-enzymatic complexes where channelling occurs, which in certain cases is today well characterised, illustrate the different strategies used by enzymes to ensure efficiency in metabolic fluxes necessary to the survival of living organisms.

BIBLIOGRAPHY

SPECIALISED ARTICLES

ÆVARSSON A., CHUANG J.L., WYNN R.M., TURLEY S., CHUANG D.T. & HOL W.H. –2000– *Struct. Fold. Des.* **8**, 277.

ÆVARSSON A., SEGER K., TURLEY S., SOKATCH J.R. & HOL W.H. –1999– *Nat. Struct. Biol.* **6**, 785.

ANDERSON K.S., MILES E.W. & JOHNSON K.A. –1991– *J. Biol. Chem.* **266**, 8020.

BALDOCK C., RAFFERTY J.B., STRUITZE A.R., CABAS A.R. & RICE D.W. –1998– *J. Mol. Biol*. **284**, 1529.

CHOCK P.B. & GUTFREUND H. –1988– *Proc. Natl Acad. Sci. USA* **85**, 8870.

DAVIDSON J.N., CHEN K.C., JAMISON R.S., MUSMANNO L.A. & KERN C.B. –1993– The evolutionary history of the first three enzymes in pyrimidine biosynthesis, in *Bioessays* **15**, 157.

ELCOCK A.H. & MCCAMMON A. –1996– *Biochemistry* **35**, 12652.

HERVE G., NAGY M., LE GOUAR M., PENVERNE B. & LADJIMI M. –1993– *Biochem. Soc. Trans.* **21**, 195.

HOLDEN H.M., THODEN J.B. & RAUSHEL F.M. –1998– Carbamyl phosphate synthetase, in *Curr. Op. Struct. Biol.* **8**, 679.

HOROWITZ N.H. –1945– *Proc. Natl. Acad. Sci. USA* **31**, 153.

HUR O., LEJA C. & DUNN M. –1996– *Biochemistry* **35**, 7378.

HYDE C.C., AHMED A., PADLAN E.A., MILES E.W. & DAVIES D.R. –1988– *J. Biol. Chem.* **263**, 17857.

IZARD T., ÆVARSSON A., ALLEN M.D., WESTPHAL A.H., PERHAM R.N., DE KOK A. & HOL W.G.J. –1999– *Proc. Natl Acad. Sci. USA* **96**, 1240.

KELETI T., BATKE J., OVADI J., JANCSIK V. & BARTHA F. –1977– *Adv. Enzyme Regul.* **15**, 233.

KIMOTO T., NISHIGAI M., SASAKI T. & IKAI A. –1988– *J. Mol. Biol.* **203**, 183.

KIRSCHNER K., SZADKOWSKI H., JARDETZKY T.S. & HAGER V. –1987– *Methods Enzymol.* **142**, 386–397.

KLEIGER G., PERRY J. & EISENBERG D. –2001– *Biochemistry* **40**, 14484.

KRAHN J.M., KIM J.H., BURNS M.R., PARRY J., ZALKIN H. & SMITH J.L. –1997– *Biochemistry* **36**, 11061.

LEHOTSKY J., BEFAKOVA G., KAPLAN P. & RAEYMAEKERS L. –1993– *Gen. Physiol. Biophys.* **12**, 339.

LINDBLADH C., RAULT M., HAGGLUND C., SMALL W.C., MOSBACH K., BÜLOW L., EVANS C. & SRERE P.A. –1994– *Biochemistry* **33**, 11692.

LINN T.C. & SRERE P.A. –1980– *J. Biol. Chem.* **255**, 10676.

LUGER K., HOMMEL U., HEROLD M., HOFSTEENGE J. & KIRSCHNER K. –1989– *Science* **243**, 206.

MACGREGOR J.S., SINGH V.N., DAVOUST S., MELLONI E., PONTREMOLI S. & HORECKER B.L. –1980– *Proc. Natl. Acad. Sci. USA* **77**, 3889.

MATTEVI A., OBMOLOVA G., SCHULZE E., KALK K.H., WESTPHAL A.H., DE KOK A. & HOL W.G.J. –1992– *Science* **255**, 1544.

OVADI J. –1988– *TIBS* **13**, 486.

OVADI J., SALERNO C., KELETI T. & FASELLA P. –1978– *Eur. J. Biochem.* **90**, 499.

PENVERNE B., BELKAÏD M. & HERVÉ G. –1994– *Arch. Biochem. Biophys.* **309**, 85.

PERHAM R.N. –1991– *Biochemistry* **30**, 8501.

PRIESTLE J.P., GRÜTTER M.G., WHITE J.L., VINCENT M.G., KANIA M., WILSON E., JARDETZKY T.S., KISCHNER K. & JANSONIUS J.N. –1987– *Proc. Natl Acad. Sci. USA* **84**, 5680.

REED L.J. –1974– *Accounts Chem. Res.* **7**, 40.

SCHLICHTING I., YANG X.J., MILES E.W., KIM A.Y. & ANDERSON K.S. –1994– *J. Biol. Chem.* **269**, 26591.

SCYLLY J.L. & EVANS D.R. –1991– *Proteins: Struct. Funct. Genet.* **9**, 191.

SERRE L., VERBREE E.C., DAUTER Z., STUITJE A.R. & DEREWENDA Z.S. –1995– *J. Biol. Chem.* **270**, 12961.

SMITH S., STERN A., RANDHAWA Z.I. & KNUDSN J. –1985– *Eur. J. Biochem.* **152**, 547.

SOUCIET J.L., NAGY M., LE GOUAR M., LACROUTE F. & POTIER S. –1989– Organization of the yeast URA2 gene: identification of a defective dihydroorotase-like domain in the multifunctional carbamoylphosphate synthetase-aspartate transcarbamylase complex, in *Gene* **79**, 59.

SRIVASTAVA D.K. & BERNHARD S.A. –1986a– *Curr. Top. Cell. Regul.* **28**, 1.

SRIVASTAVA D.K. & BERNHARD S.A. –1986b– *Science* **234**, 1081.

SRIVASTAVA D.K., SMOLEN P., BETTS G.F., FUKUSHIMA T., SPIVEY H.O. & BERNHARD S.A. –1989– *Proc. Natl Acad. Sci. USA* **86**, 6464.

THODEN J.B., HOLDEN H.M., WESENBERG G., RAUSCHEL F.M. & RAYMENT I. –1997– *Biochemistry* **36**, 6305.

TRUJILLO M., DONALD R.G., ROOS D.S., GREEN P.J. & SANTI D.V. –1996– *Biochemistry* **35**, 6366.

WAKIL S.J., STOOPS J.K. & JOSHI V.C. –1983– *Annu. Rev. Biochem.* **52**, 537.

WALLIS N.G., ALLEN M.D., BROADHURST R.W., LESSARD I.A.D. & PERHAM R.N. –1996– *J. Mol. Biol.* **263**, 464.

WALSH G.R. & EASTERBY J.S. –1994– *Trends Biochem. Sci.* **19**, 193.

WEBER J.P. & BERNHARD S.A. –1982– *Biochemistry* **21**, 4189.

WIELAND F., SIESS E.A., RENNER L., VERFÜRTH C. & LYNEN F. –1978– *Proc. Natl Acad. Sci. USA* **75**, 5792.

WILMANNS M., PRIESTLE J.P., NIERMANN T. & JANSONIUS J.N. –1992– *J. Mol. Biol.* **223**, 477.

WILMANNS M., HYDE C.C., DAVIES D.R., KIRSCHNER K. & JANSONIUS J.N. –1991– *Biochemistry* **30**, 9161.

WU X., GUTFREUND H., LAKATOS S. & COCK P.B. –1991– *Proc. Natl Acad. Sci. USA* **88**, 497.

YOSHIDA S., TAMIYA-KORZYMI K. & KOJIMA K. –1989– *Biochim. Biophys Acta* **1007**, 61.

PART VI

ENZYMOLOGY IN STRUCTURED ENVIRONMENT

INTRODUCTION

Mechanisms of regulation were presented in Part V in order of increasing complexity, and in many of the systems described the cellular context is implicit, sometimes even explicit. In this last part, we elaborate on the importance of cellular compartmentalisation, enzyme localisation and the resulting functional consequences for metabolic control processes and the maintenance of homeostasis in living beings. The refinement of cytological and biochemical methods has allowed both the precise localisation of the diverse cellular constituents and the evaluation of the local concentrations of enzymes and metabolites.

In the cell, enzymes exist in very specific environments, which differ considerably from the homogenous solutions in which studies are carried out in vitro. The cell is a structured environment and enzymatic reactions take place either in well-determined compartments separated from one another by membranes, in the cytosol, or at the membrane when the enzymes are directly associated to them. The role of membranous barriers is extremely important for the regulation of cellular metabolism.

Numerous factors contribute to the modulation of enzyme reactions in situ. **Cellular enzymology** has its own laws which add to those from classical enzymology in vitro, and are not always easy to define due to the complexity of the system. However, important progress has been achieved over the last few years in providing a better understanding of the different mechanisms by which the reactions that occur within organisms are modulated, and thus ensuring their autoregulation.

16 – LOCALISATION AND CELLULAR COMPARTMENTALISATION

16.1. LOCALISATION OF ENZYMES IN CELLULAR COMPARTMENTS

In prokaryotic cells like *E. coli*, most enzymes are contained within the cytosol. In contrast, the enzymes involved in oxidative phosphorylation and the enzymes required for cell-wall synthesis are membrane-bound enzymes (Fig. 16.1 below).

Eukaryotic cells are more complex. They contain diverse compartments and particules: the nucleus, mitochondria, the GOLGI apparatus, lysosomes, the endoplasmic reticulum and the ribosomes (Fig. 16.2 below). The **plasma membrane**, the permeability barrier, contains many enzymes as well as active transport systems for a number of metabolites and ions. In cells such as hepatocytes, this membrane has a thickness of about 90 Å. The proteins in this membrane are predominantly glycoproteins. The **nucleus** contains DNA associated with histones and organised into chromosomes. The enzymes involved in replication and transcription forming mRNA, rRNA and tRNA are localised here. The **mitochondria**, truly vestigial cells, comprise an internal and external membrane, cristae, granules and a matrix. These organelles are the centres of important metabolic reactions of the tricarboxylic acid cycle, or the KREBS cycle, oxidative phosphorylation and electron transport, fatty acid oxidation, the metabolism of amino acids and some steps of urea synthesis. In the **cytosol**, the reactions of glycolysis take place (EMBDEN-MEYERHOF pathway) as do several reactions of gluconeogenesis, fatty-acid synthesis, certain steps in the synthesis of mononucleotides, the synthesis of several amino acids and other steps of urea synthesis. The **lysosomes** possess a large hydrolytic activity as they contain a great many hydrolytic enzymes, including ribonuclease and alkaline phosphatase. The **endoplasmic reticulum** is the place of lipid and steroid synthesis and hydroxylation reactions; it also serves to transport the products of biosynthesis. The **ribosomes** provide the support for protein synthesis. As they are often bound to the rough endoplasmic reticulum (RER), the proteins formed there must cross the RER membrane to travel to their destination. Membrane passage is made possible by the N-terminal part of the protein or signal peptide which is hydrolysed by a protease, signal peptidase, upon exit from the RER (see Chap. 5).

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_17, © Springer-Verlag Berlin Heidelberg 2010

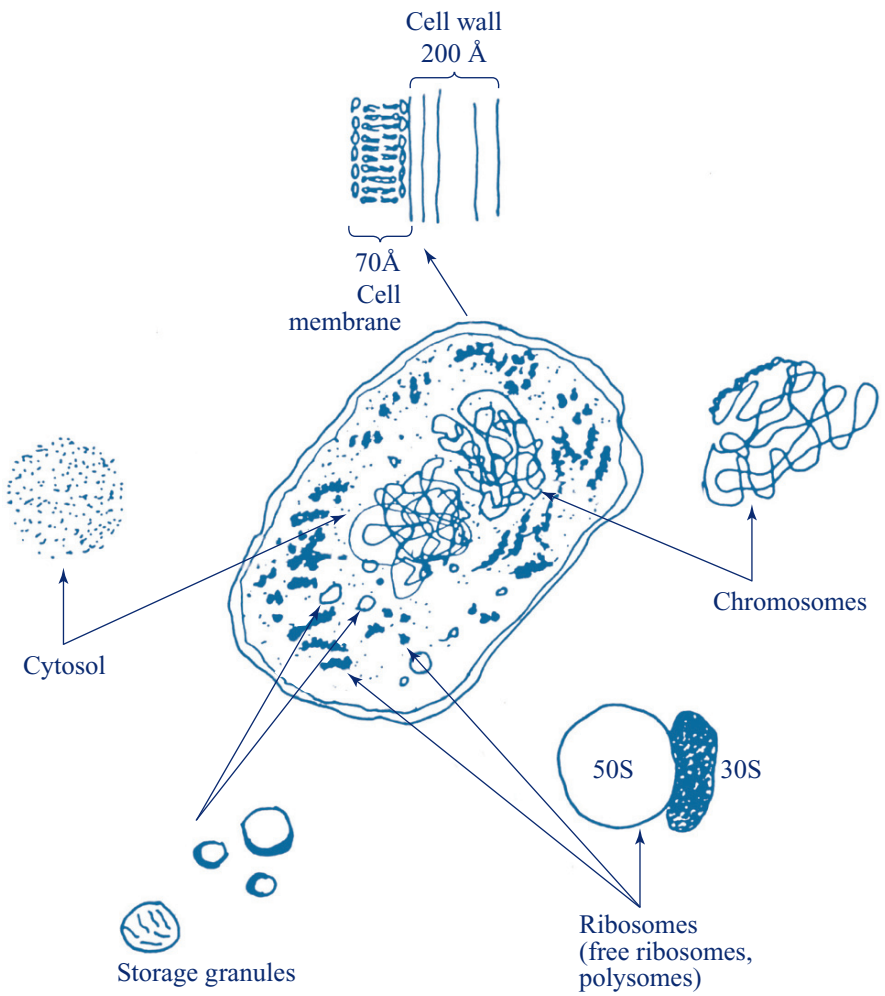


Fig. 16.1 An E. coli cell

The enzymes of the different metabolic pathways are therefore found localised to well-defined cellular compartments. Enzymes from the same pathway are frequently organised sequentially in stable or transient complexes –designated by the term "metabolons"– which facilitates the channelling of metabolites, as shown in Chap. 15. These complexes are often associated to structural elements in the cell. SRERE (1988) distinguished two types of metabolic pathway: those that produce intermediate metabolites usable in several metabolic pathways, e.g. glucose-6-phosphate and dihydroxyacetone phosphate from glycolysis; and others, for which only the final product is usable, for instance in protein synthesis where the intermediates do not fulfil any metabolic function. Eighty per cent of intermediates only participate in a single metabolic pathway.

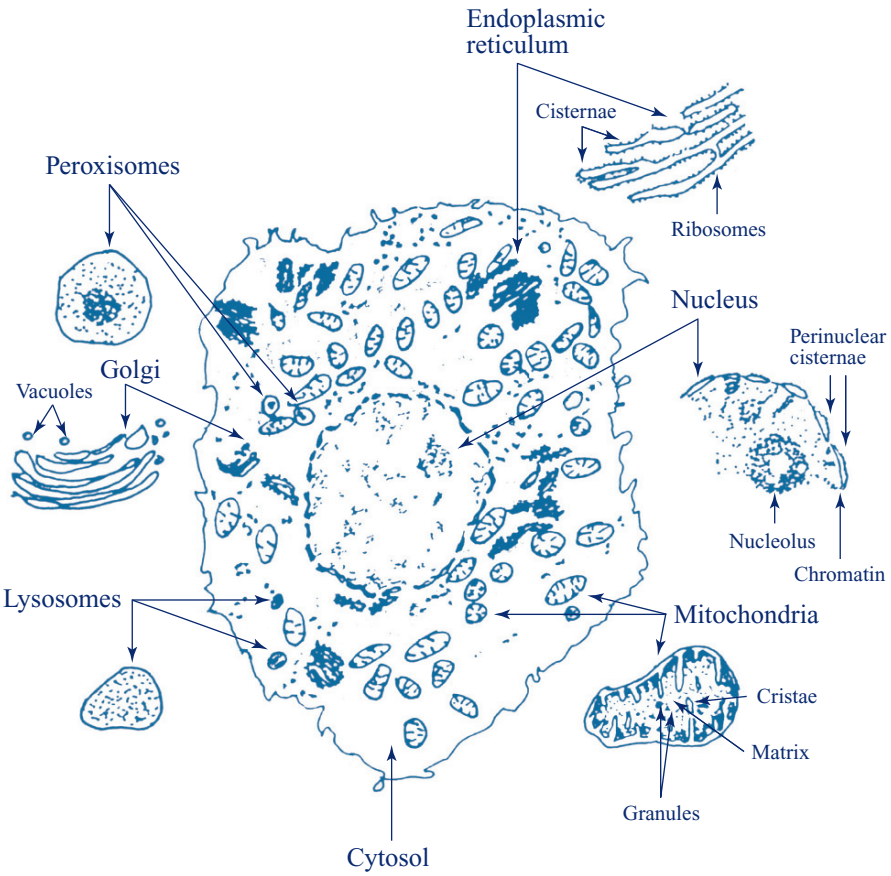


Fig. 16.2 An eukaryotic cell

The DNA replication system provides an example of a stabilised complex in which DNA not only serves as a template but also gives support to the complex. This complex advances along the DNA template driven by the energy from nucleoside triphosphates to synthesise a new DNA molecule. DNA polymerases participate in different types of complex, each having a well-defined function: either the synthesis of a single strand or DNA repair. DNA polymerase III is a complex with a molecular mass of ~ 900 kDa comprising 10 different polypeptide chains. It has been shown in a strain of *E. coli* infected by the T4 bacteriophage that the 10 enzymes responsible for the synthesis of the four deoxyribonucleotide triphosphates necessary for DNA synthesis form a complex associated to the replication complex that comprises several enzymes: a helicase, a primase, DNA gyrase, DNA polymerases I and III, as well as DNA ligase. In eukaryotic cells several enzymes are associated in complexes, called replicons, for DNA replication: DNA polymerase, ribonucleotide reductase, thymidylate synthase, thymidine kinase, nucleoside diphosphate kinase and CMP kinase. It seems that many enzymes in the nucleus are associated to non-soluble structural components of the nuclear matrix. The process of

DNA transcription to RNA presents the same degree of complexity and the polymerases involved exist in the form of complexes. Protein synthesis on ribosomes requires aminoacyl-tRNA synthetases which are associated in complexes consisting of up to eight synthetases in higher eukaryotes, as shown by WALLER and coworkers (MIRANDE et al., 1993).

The enzymes involved in glycogen metabolism are generally associated to particles of this storage product. Elements from the sarcoplasmic reticulum interact with glycogen particles. These particles have various sizes and the distribution of enzymes differs according to particle size. The activity of phosphorylase is higher in small particles; synthetase activity is higher in large particles. All available information suggests that there are two separate systems in the cytosol, one for glycolysis and the other for gluconeogenesis. The system for glycolysis has largely been described in Chap. 15.

As a result of this cellular compartmentalisation, certain enzymes are found in a very specific environment, for example in the vicinity of polyelectrolytes formed by the polar headgroups of membrane phospholipids. Besides, in certain cellular compartments the local enzyme concentrations are often much greater than those of the substrates whose penetration into the compartment is controlled by the membrane. The enzyme reactions thus take place in conditions that differ from those that are routinely employed in experiments in vitro where the enzyme concentrations are negligible compared to those of the substrate – a necessary condition for the MICHAELIS equation to be valid.

For membrane enzymes or for reactions in general that involve crossing a cellular barrier, important effects of diffusion, which modulate the substrate and product concentrations over time, complicate the situation. The rate of an enzyme reaction depends therefore on two terms: one corresponding to the enzyme reaction itself, which may be Michaelian or allosteric; the other takes into account the diffusional effects.

Thus in the cell, enzyme behaviour can frequently be observed to diverge from MICHAELIS' law even when the enzymes are not allosteric. Cellular compartmentalisation and enzyme localisation can involve other mechanisms of metabolic regulation. Besides, the metabolites are transformed by a sequence of reactions; also, in a cellular context it is important to consider the flux as well as the concentration of a given metabolite. Furthermore, some consecutive reactions are catalysed by enzymes associated in multi-enzyme complexes as discussed in Part V. As a result of the proximity of the active sites in enzymes that catalyse two sequential reactions, a channelling effect bestows on the system greater efficacy.

We will examine a few of the cellular factors likely to lead to important functional consequences for enzyme behaviour.

16.2. THE CELLULAR CONCENTRATIONS OF MACROMOLECULES

The organelles therefore delimit the specialised regions of the cell. This partitioning is achieved by membranes or the cell wall in which the enzymes are integrated and their functions are different from those diluted in solution, thus ensuring a means to regulate cellular metabolism. The ensemble is organised in such a way that the different compartments communicate with each other. Due to this organisation, the cellular contents differ notably from a homogenous solution in which macromolecules are dispersed.

The molecular crowding that exists in the cytoplasm of *E. coli*, for example, has been evaluated by GOODSELL (1991). This is illustrated in Fig. 16.3 below. Figure 16.3a represents a cube having sides of 100 nm cut out from the cytoplasm. The volume of the cytoplasm in an *E. coli* cell would fill 600 similar-sized cubes. In this cube up to 30 ribosomes coexist with more than 100 protein factors, 30 aminoacyl-tRNA synthetases, 340 tRNAs, 2-3 molecules of mRNA, 6 RNA polymerases, 300 other molecules, 330 protein molecules among which 130 glycolytic enzymes, 100 enzymes from the KREBS cycle, and 30 000 small molecules such as H_2O, cofactors, precursors and metabolites, as well as approximately 50 000 ions. It is interesting to note that in experiments in vitro, this same cube would only contain a single enzyme molecule, salts from the buffer solution and water molecules. This molecular crowding corresponds to protein concentrations varying between 100 and 500 $mg \,.\, mL^{-1}$, depending on the cellular compartment (LINN & SRERE, 1980; WALSH & EASTERBY, 1984). It follows that the enzyme concentrations in vivo are of the order of 10^{-5} to 10^{-4} M, and therefore much higher than those used in in vitro experiments (which are between 10^{-7} and 10^{-10} M).

SRERE (1982) showed that in a mitochondrion, the distance between two tetrameric proteins having similar molecular masses of 100 kDa is less than their average diameter. The diffusion rate of molecules becomes diminished as a result of the local crowding. GOODSELL evaluated the diffusion rate of a 160 kDa protein: it needs 2 μs to cross 10 nm whereas in an homogenous solution this same process occurs in 2 ns.

Figure 16.3c represents a section of the cell wall from *E. coli*, and Fig. 16.3d the nuclear region whose crowding is dominated by the presence of DNA. The molecular density is therefore high in different regions of the cell; furthermore, these different elements are in constant movement, their rates being limited by the density of the medium.

Molecular crowding promotes association between proteins and other constituents of the cell, the membrane and cytoskeleton. Such interactions may lead to different conformations of the same enzyme, thus modulating its activity. Many examples have been given previously in Part V.

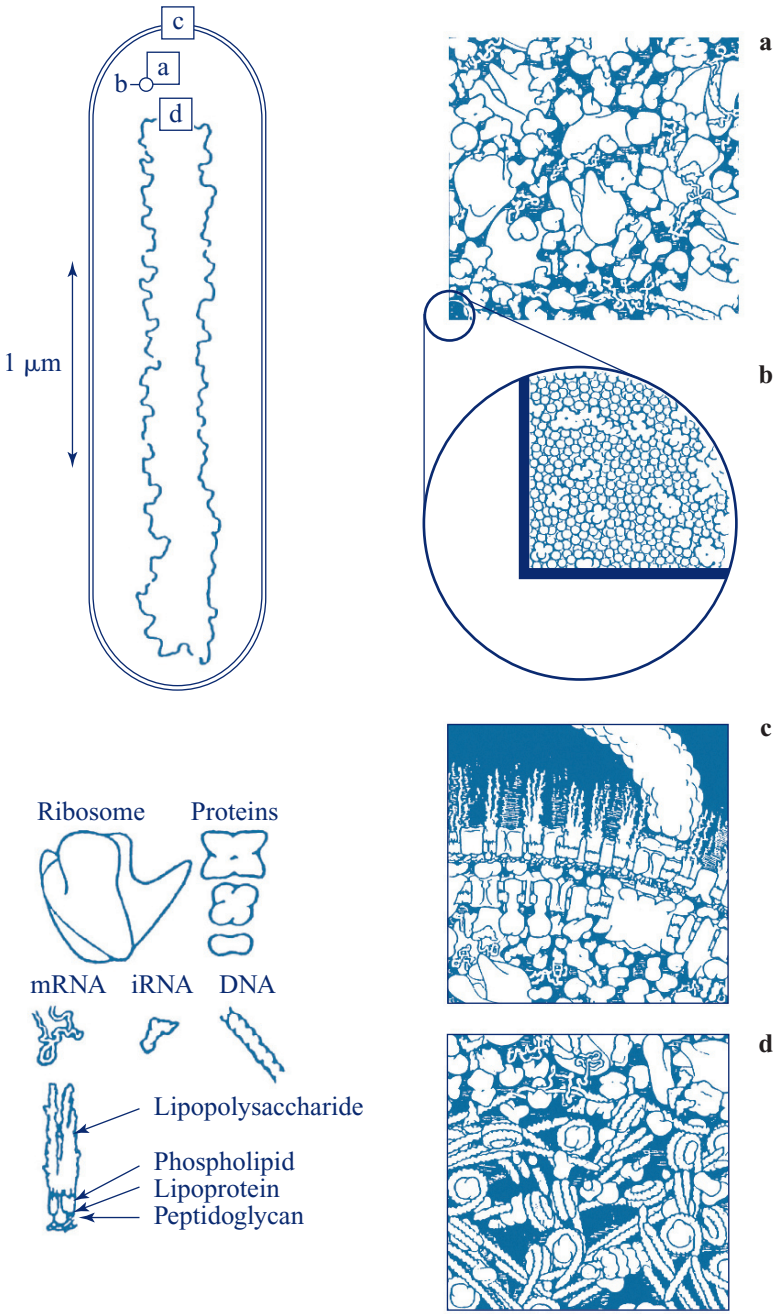


Fig. 16.3 Cellular crowding illustrated by different parts of an E. coli cell

Left, an E. coli cell and a representation of the different macromolecular components. ***(a)*** *the cytoplasm –* ***(b)*** *a portion of the cytoplasm showing the molecules, including water molecules, symbolised by* circles *–* ***(c)*** *the cell wall –* ***(d)*** *the nuclear region dominated by the presence of DNA* (Reprinted from *TIBS*, **16**, GOODSELL D.S., Inside a living cell, 203. © (1991) with permission from Elsevier)

16.3. INTERACTIONS OF ENZYMES WITH CELLULAR CONSTITUENTS

The cellular compartments are delimited by the plasma membrane. The transmission of signals between the different compartments must therefore take place across these physical barriers in order to coordinate cellular functions and to respond to environmental changes. Signal transduction is predominantly ensured by membrane-bound enzymes, although certain signals can cross the cellular barriers by themselves. In response to these signals, there may be association or dissociation of enzymes, which modulates their activities. Proteins also feature in the constitution of membranes and cell walls. They are distributed heterogeneously in cells containing soluble proteins in the cytoplasm, proteins integrated into membranes or the cell wall and proteins situated at the periphery yet interacting with these cellular barriers. These associations have functional consequences; by modulating enzyme activity they contribute to the regulation of cellular metabolism.

16.3.1. MEMBRANE ENZYMES

The structure of the membrane formed from a lipid bilayer comprising essentially phospholipids creates a high charge density on the internal and external sides with consequences for the activity of the enzymes that are bound to it. Two types of membrane association must be considered. Some enzymes are integrated in the membrane and possess a hydrophobic transmembrane region which ensures membrane anchorage. Others are peripheral and interact directly with lipids or with other integral membrane proteins. The interaction of enzymes with membrane lipids sometimes brings about their inactivation. Such is the case for DNA polymerase γ involved in the replication of mitochondrial DNA. This enzyme is inhibited by cardiolipin which represents 20% of the lipids in mitochondrial membranes. YOSHIDA et al. (1989) considered that this dynamic association constitutes a means for the replication of mitochondria. The activity of mammalian glutamate dehydrogenase is partly inhibited by its interaction with cardiolipin and phosphatidylserine. This inhibition, electrostatic in nature, is alleviated by the addition of L-leucine (COUÉE & TIPTON, 1989), which explains why the enzyme is more activated by L-leucine *in vivo* than the isolated enzyme. Glyceraldehyde-3-phosphate dehydrogenase is also inhibited during its temporary association with molecules of phosphatidylinositol and phosphatidylserine; this inhibition can be total in rabbit muscle (SIDEROWICZ et al., 1990).

In contrast, the interactions of protein kinases C (PKCs) with lipids induce activation of these kinases. PKC activity depends on the concentrations of the intracellular calcium and of membrane lipids. In a resting cell, the calcium concentration is less than 10^{-7} M; PKCs are thus cellular and inactive. The PKCs are composed of a single polypeptide chain, having an N-terminal regulatory region of about 20–40 kDa, and a catalytic C-terminal region (~45 kDa). These enzymes are formed from four

domains (C1-C4) that are highly conserved in the different PKCs. The C1 domain contains a binding site for diacylglycerol. The C2 domain harbours a recognition site for phospholipids; C3 and C4 domains possess, respectively, ATP- and substrate-binding sites. The two regulatory and catalytic halves are separated by a hinge region which becomes accessible to proteases when the enzyme is bound to the membrane. Figure 16.4 opposite illustrates the mechanism of PKC regulation, taken from NEWTON (1995). The newly synthesised enzyme is associated with the insoluble fraction in the cells. It undergoes three successive phosphorylations at positions Thr500, Thr641 and Ser660. This triply-phosphorylated form is inactive since the N-terminal part covers up the substrate-binding site. The presence of diacylglycerol, which binds to C1 domain in a stereospecific way, considerably increases the enzyme's affinity for the membrane. This affinity, higher still in the presence of calcium, engenders a conformational change in the protein and increases accessibility to the active site, the translocation of the enzyme to the membrane and the binding of phosphatidylserine to C2 domain. The interaction of diacylglycerol with the C1 domain appears to serve as a hydrophobic anchor, tethering the enzyme to the membrane, with one of the hydrophobic chains of a phosphatidylserine penetrating into the hydrophobic cavity in the C2 domain. The hinge region becomes accessible to proteases during the activation period. PKC activity depends therefore on both the concentration of intracellular calcium and on membrane lipids. Activation is the end result of a series of events. Diacylglycerol is produced by the hydrolysis of phosphatidylinositol 4,5-bisphosphate (PIP2) by phospholipases C. This hydrolysis is the first step in a cascade of events in response to an external signal. The interaction of the C2 domain with phosphatidylserine is electrostatic in nature, involving the polar phospholipid headgroups.

Other soluble enzymes interact with the membrane indirectly through an intermediate association with transmembrane proteins. In erythrocytes, glyceraldehyde-3-phosphate dehydrogenase is associated to the glucose transporter and to band 3 protein. Electrostatic interactions are specific and reinforced by a raised haemoglobin concentration. They are favoured under conditions of low pH and ionic strength and promote inhibition of the enzyme. The presence of ATP and NAD modulates the distribution of the enzyme between the soluble and membrane phases. Phosphofructokinase also interacts via its N-terminal extremity with band 3 protein. As with glyceraldehyde 3-phosphate dehydrogenase, this association is modulated by ATP, ADP, NADH and fructose-6-phosphate. The interaction with band 3 protein leads to the rapid activation of phosphofructokinase due to the fact that it becomes insensitive to allosteric inhibitors, followed by a slow inhibition resulting from the enzyme's dissociation into dimers (JENKINS et al., 1985).

Dynamic association between enzymes and membranes, regardless of the type of interaction, can therefore considerably modify their catalytic behaviour.

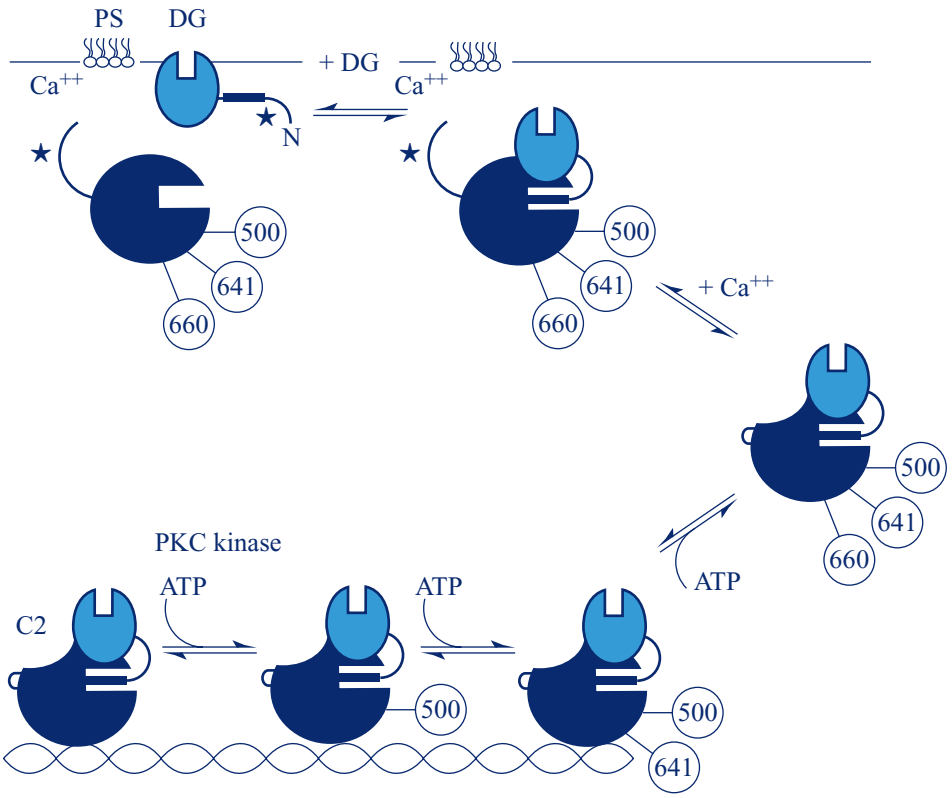


Fig. 16.4 Model for the regulation of protein kinase C activity

(From *J. Biol. Chem.*, **270**, NEWTON A.C., 28497. © (1995) with permission from The American Society for Biochemistry and Molecular Biology)

16.3.2. ENZYMES ASSOCIATED TO THE CYTOSKELETON

The interaction of enzymes with cytoskeletal proteins constitutes another aspect of macromolecular interaction existing at the heart of a cell. Cytoskeletal proteins are abundant in cells. Actin represents about 15% of the total protein in a muscle fibre, and myosin 20%. In other cells the proportion of actin varies between 5 and 15%, and that of myosin represents only 0.5–2% of the total protein. The interactions between enzymes and cytoskeletal proteins are dynamic and regulated by signals such as metabolite concentrations, pH and ionic strength which control the processes of association and dissociation of the different proteins. These temporary associations create a micro-compartment in the vicinity of the cytoskeleton, resulting in reciprocal effects. The association of enzymes to proteins of the cytoskeleton can modulate the degree of organisation of these proteins; reciprocally, these can modify the kinetic behaviour of the enzymes.

Some glycolytic enzymes associate with different proteins of the cytoskeleton. For instance, fructose bisphosphate aldolase interacts with actin and myofibrils, leading to competitive inhibition of the enzyme. This is caused by the masking of its active site by the cytoskeletal proteins. These interactions are specific and electrostatic. A rise in the intracellular calcium concentration tends to promote the dissociation of aldolase from the cytoskeletal proteins, which subsequently augments its activity. As calcium is involved in the transmission of external signals, the dynamics of these associations thus allows coordination of the enzyme activities in different cells. While aldolase activity is controlled by its degree of association to cytoskeletal proteins, reciprocally, the cytoskeleton tends to become structured when in the presence of aldolase; the binding of the enzyme to actin filaments causes them to self-associate and form networks of parallel filaments.

Glyceraldehyde-3-phosphate dehydrogenase is activated by its association to actin and inhibited when bound to microtubules; this drop in activity results from the enzyme's dissociation. Phosphofructokinase is activated when associated to both myofibrils and actin, but is no longer sensitive to allosteric effectors. These interactions depend on the pH, and only exist when the pH is lower than 7–7.5. However, tubulin and microtubules inhibit phosphofructokinase after their association to monomeric and dimeric forms of the enzyme (LEHOTSKY et al., 1993). The effectors – fructose bisphosphate, ATP and ADP – prevent these interactions following the conformational changes that they induce in the enzyme. Reciprocally, phosphofructokinase affects the organisation of microtubules by inducing the appearance of aligned, parallel groups of microtubules connected periodically by side-branches composed of phosphofructokinase tetramers (VERTESSY et al., 1996). Such associations also exist between PKC and actin, and between cellulases and actin (BLOBE et al., 1996).

16.3.3. ENZYMES ASSOCIATED TO PLANT CELL WALLS

Plant cell walls confer on enzymes bound to them properties that differ from those they display in an isolated state. Plant cell walls are composed of complex assemblies of cellulose, hemicellulose and pectins, and appear as an insoluble polyanion due to the presence of the pectins' negative charges. As a result, there is a potential difference between the interior and exterior of the cell wall. The charge density considerably modifies the kinetic behaviour of the enzymes bound to the cell wall, as we shall see in the following chapter.

The behaviour of enzymes associated to cellular membranes, whether directly or indirectly via integral membrane proteins, or to proteins of the cytoskeleton, or to the plant cell wall, differs notably from that of enzymes in solution. These associations represent another means of regulation available to cells.

16.4. COMPARTMENTALISATION OF METABOLITES

Cellular compartments contain enzyme systems and metabolites that ensure their functional specialisation. The concentrations of enzymes are often much higher than those of their substrates. The intermediate metabolites in reaction chains do not exist in a free state when they are produced inside multi-enzyme complexes in which channelling takes place, as is described in Part V. The circulation of metabolites between diverse compartments is ensured by membranes. Thus, external glucose is actively transported into the cytosol via the plasma membrane. The conversion of glucose into pyruvate takes place in the cytosol; pyruvate is then actively transported to the internal compartments of mitochondria, which contain all the enzymes and metabolites of the KREBS cycle. The internal membrane of mitochondria contains all of the enzymes of oxidative phosphorylation, including those that are involved in electron transfer from NADH to oxygen and in ATP synthesis. In higher organisms, the spatial organisation extends far beyond individual cells. The circulation of certain metabolites from one organ to another occurs *via* the blood system.

Over the last few years much research has been dedicated to determining the cellular localisation of enzymes. It has been shown that, frequently, enzymes catalysing different steps in a metabolic pathway are not located in the same cellular compartment. This phenomenon is linked, amongst others, to the fact that numerous metabolic pathways are connected and possess intermediate metabolites in common. One consequence of this organisation is the so-called **"metabolic compartmentalisation"** by which a particular metabolite is sequestered in a given subcellular compartment and cannot diffuse into other compartments. Furthermore, for certain metab-olic pathways the organisation and localisation of enzymes vary from one organism to another. Such is the case for the biosynthetic pathway of pyrimidine nucleotides. Eukaryotic organisms possess two carbamylphosphate synthetases (CPSases): one is specific for the arginine biosynthetic pathway (CPSase-A); the other is specific for the biosynthetic pathway of pyrimidine nucleotides (CPSase-P). In *Neurospora crassa* and *Saccharomyces cerevisiae* the CPSase-P-ATCase (aspartate transcarbamylase) complex is predominantly localised to the nucleus. In contrast, the two first enzymes of the arginine biosynthetic pathway, CPSase-A and ornithine transcarbamylase, are localised to the mitochondria in *N. crassa*, but are found in the cytoplasm of *S. cerevisiae*. One of the consequences of this difference in localisation is that, in *N. crassa*, the two internal reservoirs of carbamylphosphate are strictly compartmentalised and so this metabolite cannot be transferred from one metabolic pathway to the other. On the contrary, in *S. cerevisiae*, a partial exchange of the metabolites between the two metabolic pathways is observed.

Such differences have been evidenced even in very closely related species. Thus, dihydroorotate dehydrogenase (the fifth enzyme in the biosynthetic pathway of pyrimidine nucleotides) is localised to the mitochondria in *S. cerevisiae*, whereas it is situated in the cytoplasm of *Saccharomyces pombe*. This difference in localisation

comes with a difference in the enzymatic mechanism and in the nature of the cofactor.

It is interesting to note that some intermediate metabolites have to migrate from one compartment to another before acting as substrates in the subsequent reaction, and this may happen several times in the same metabolic pathway. This behaviour illustrates well the existence of a complex "metabolic and enzymatic physiology" about which we still know relatively little, though knowledge of which will be essential for therapeutic intervention.

Metabolic compartmentalisation and the existence in the cell of several reservoirs of the same metabolite which do not communicate are well documented, including those for amino acids and nucleotides (HERVÉ, 1996). Taking the amino acids, in *S. cerevisiae*, aspartate, glutamate, leucine and isoleucine are present in the cytoplasm in raised concentrations, whereas the other amino acids are present essentially in vacuoles. The nature of the nitrogen source and the degree of protein synthesis are linked to the dynamic equilibrium of these reservoirs. In some mammals and the protozoan *Tetrahymena pyriformis*, it has been shown that exogenous amino acids are incorporated into proteins in preference to the amino acids generated within the cell either by synthesis or by proteolysis. In rat liver, the amino acids supplied by the lysosomal degradation of proteins are utilised almost exclusively for the biosynthesis of intracellular proteins, whereas amino acids of extracellular origin are used in the biosynthesis of both intra- and extracellular proteins. In cultured muscle cells, the leucine of external origin is used preferentially for protein synthesis whereas leucine of internal origin is used preferentially for oxidation. Even the bacterium *E. coli* is able to maintain two reservoirs of amino acids: one supplied by the external medium, the other by internal metabolism.

These phenomena are even better documented in the case of nucleotides. Two completely independent reservoirs communicating very slowly are used for the synthesis of DNA and RNA. In some cases, it has been shown that the *de novo* biosynthetic pathway of nucleotides and the recycling pathway for bases supply different cellular compartments. The compartmentalisation between uracil originating from external sources and that coming from cellular degradation of RNA has been demonstrated. Elegant isotopic chase experiments have shown that during infection of NOVOKOFF hepatoma cells by the MENGO virus, tritiated uridine from an external origin is incorporated into the viral RNA which is synthesised in the cytoplasm, whereas no labelling of the host RNA, synthesised in the nucleus, was observed. In *Euglena* it has been shown that the *de novo* synthesis of UTP preferentially supplies the synthesis of messenger RNA, whereas the cytoplasmic and nuclear reservoirs of uridine ensure the production of UTP used in the synthesis of ribosomal RNA.

Compartmentalisation appears more important still in the case of DNA precursors. The distribution of deoxyribonucleotides in the cytoplasm and the nucleus is

variable, particularly as a function of the cell cycle. The nuclear content of dCTP and dTTP increases significantly during the phase of DNA replication. A small reservoir of thymidylate is used exclusively for the synthesis of nuclear DNA, while another larger reservoir is involved in nuclear DNA repair and the synthesis of mitochondrial DNA. It has also been shown that mitochondria possess their own nucleotide reservoir which equilibrates very slowly with the cytoplasmic reservoir. It is evident that this complex heterogeneity in the cellular nucleotide content must be taken into account in the implementation of chemotherapeutic protocols, failing which the product used could miss its target.

17 – KINETICS OF ENZYMATIC REACTIONS CATALYSED BY IMMOBILISED ENZYMES

As we have just seen, in living cells some enzymes are bound to membranes, whereas others are bound to the cell wall in plants; they are immobilised in well-defined cellular compartments. Their behaviour differs from that of enzymes in solution; catalysis is no longer homogenous, rather it is heterogenous. For over thirty years diverse enzymes have been immobilised artificially to solid supports for use in biotechnology, and also to serve as models of immobilised enzymes in vivo. Numerous studies have been carried out on these systems in order to determine the kinetic laws that govern their functions (see the reviews by SILMAN & KATCHALSKI, 1966; ENGASSER & HORVATH, 1976; RICARD et al., 1994). Although in most cases artificially immobilised enzymes have proved to be more stable than when in solution, they generally exhibit weaker activities. Various factors are liable to affect the activity of enzymes fixed to a rigid support: steric constraints, a reduction in conformational flexibility, diffusional resistance and the repulsive or attractive electrostatic effects created by the polar headgroups of biological membranes or cell walls. In these conditions, diffusion of substrate and product occurs during the reaction. There is a coupling between diffusion and the enzymatic reaction. We shall first of all consider the consequences of these as they can be evaluated quantitatively. Then we shall examine the effect produced by the electrostatic constraints.

17.1. FUNDAMENTAL COUPLING EQUATION

The notion of coupling between the enzyme reaction and the diffusion of substrate is related to the diffusional resistance in a cell. Let us consider enzyme molecules bound to a surface (Fig. 17.1 below). In proximity to the surface, in an unperturbed environment, the substrate concentration diminishes and the product concentration increases; a concentration gradient is then established. The substrate or product diffusion rate is generally of the same order of magnitude or lower than that of the enzymatic steps.

At steady state the substrate flux J_s is equal to the reaction rate v:

$$J_s = v$$

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_18, © Springer-Verlag Berlin Heidelberg 2010

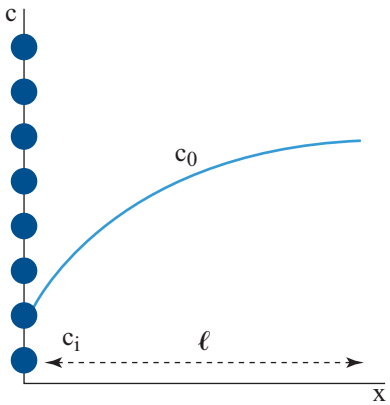


Fig. 17.1 Enzyme molecules **(black circles)** ***fixed to a rigid cell wall*** *The* curve *indicates the gradient substrate concentration*

The flux of molecules that cross a surface element is given by FICK's first law:

$$J_s = \partial c / \partial t = - D \partial c / \partial x$$

D is the diffusion coefficient, c the concentration and x the distance. The flux of substrate molecules that crosses a volume element is given by FICK's second law:

$$\partial c / \partial t = D \partial c^2 / \partial x^2$$

Assuming that at the interior of this unit volume the diffusion of molecules is at steady state, we have the following relationships:

$$\partial^2 c / \partial x^2 = 0 \quad \text{and} \quad \partial c / \partial x = (c_0 - c_i) / \ell$$

c_0 and c_i being, respectively, the concentrations of solute near to the surface and at a distance ℓ (Fig. 17.1). By putting:

$$h_d = D/\ell$$

the expression for the flux becomes:

$$J_s = h_d(c_0 - c_i)$$

Let us consider an enzyme bound to an impermeable surface, membrane or solid support. In the simplest case of a Michaelian enzyme, the rate equation is given by:

$$v = \frac{V_m s_\ell}{K_m + s_\ell}$$

where s_ℓ is the local substrate concentration.

The net flux of substrate, which reaches the immobilised enzyme, is:

$$J_s = h_d(s_0 - s_\ell)$$

s_0 and s_ℓ being, respectively, the solution and local substrate concentrations. Their respective reduced concentrations are:

$$\alpha_0 = s_0/K_m \quad \text{and} \quad \alpha_\ell = s_\ell/K_m$$

The coupling between the diffusion and the reaction is expressed as follows:

$$h_d(s_0 - s_\ell) = V_m\alpha_\ell/(1 + \alpha_\ell)$$

By letting $h^*_d = h_dK_m/V_m$, this equation becomes:

$$h^*_d(\alpha_0 - \alpha_\ell) - \alpha_\ell/(1 + \alpha_\ell) = 0$$

This can also be written in the following form:

$$h^*_d\alpha_\ell^2 + [h^*_d(1 - \alpha_0) + 1)]\alpha_\ell - h^*_d\alpha_0 = 0$$

The positive root of this equation is:

$$\alpha_\ell = \frac{\sqrt{[1+h^*_d(1-\alpha_0)]^2 + 4h^{*2}_d\alpha_0} - [h^*_d(1-\alpha_0)+1]}{2h^*_d}$$

This expresses the variation in the ***reduced local-substrate concentration*** as a function of the ***reduced substrate concentration*** in solution. If α_ℓ is replaced by this value in the rate equation, it is clearly apparent that the kinetics no longer follow the MICHAELIS law. For low substrate concentrations, the global reaction rate is limited by diffusion; for high concentrations, it is limited by the enzyme reaction rate. This behaviour is illustrated in Fig. 17.2, which shows the variation in the intrinsic enzyme reaction rate, V_{kin}, and the diffusion rate, V_{diff}, V being the resulting rate of the reaction catalysed by the enzyme fixed to the rigid surface (ENGASSER & HORVATH, 1976).

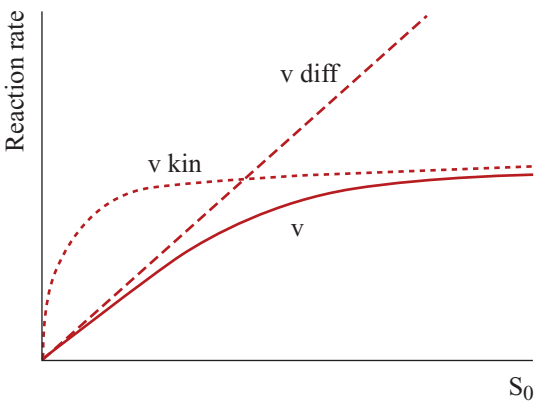


Fig. 17.2 Change in the rate of a reaction catalysed by an immobilised enzyme
(Reprinted from *Applied Biochem. Bioengng.*, **1** , ENGASSER J.M. & HORVÁTH C., Immobilized Enzyme Principles, 127. © (1976) Academic press, with permission from Elsevier)

17.2. HYSTERESIS LOOPS IN REACTIONS INVOLVING DIFFUSION-REACTION COUPLING

◤ The rate equation for the enzyme reaction can contain non-linear terms; this is the case, for example, when there is inhibition by a substrate excess. In these conditions the rate equation is:

$$v = \frac{V_m (s_\ell / K_s)}{1 + (s_\ell / K_s) + (s_\ell^2 / K_s K_2)}$$

K_2 is the dissociation constant for the ternary enzyme complex with two substrate molecules; letting $K_s/K_2 = \lambda$, the coupling equation is:

$$h^*_d(\alpha_0 - \alpha_\ell) - [\alpha_\ell/(1 + \alpha_\ell + \lambda\alpha_\ell^2)] = 0$$

This expression can also be written as:

$$h^*_d\lambda\alpha_\ell^3 - h^*_d(\alpha_0\lambda - 1)\alpha_\ell^2 + [h^*_d(1 - \alpha_0) + 1]\alpha_\ell - h^*_d\alpha_0 = 0$$

This equation has three positive real roots (from DESCARTES' Rule of Signs). If we plot on a graph the change in the reduced local-substrate concentration, α_ℓ, as a function of the reduced substrate concentration in the solution, α_0, we obtain the curve displayed in Fig. 17.3, which shows a hysteresis loop. Such situations were described in Chap. 3. At the same substrate concentration, α_0 can correspond to two stable local-substrate concentrations, α_ℓ. The existence of one or other of these concentrations depends on the rise or fall of the substrate concentration in the solution. This means that, due to coupling between diffusion and reaction, **the fixed enzyme is not only sensitive to the substrate concentration, but also to the direction of its change.** The support or membrane conserves, as it were, *a memory of the preceding events*. This situation is interesting, in so far as it represents a good physical model of a very simple process of "biosensing".

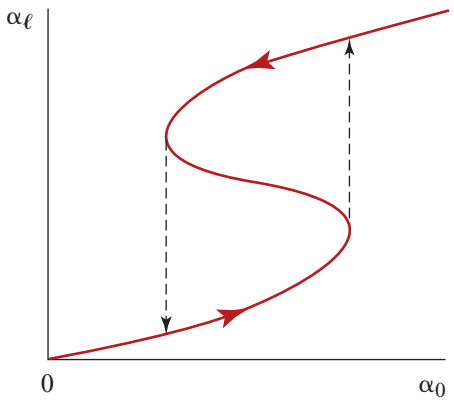


Fig. 17.3 Change in the reduced local-substrate concentration as a function of the reduced substrate concentration in solution, for inhibition by a substrate excess, the enzyme being bound to the surface

Another interesting effect is given by the inhibition of membrane enzymes. Let us consider a competitive inhibitor. The rate equation is:

$$v = \frac{V_m (s_\ell / K_s)}{1 + (s_\ell / K_s) + (I_0 / K_i)}$$

By putting $I_0/K_i = \beta$, we obtain the following relation:

$$h^*_d(\alpha_0 - \alpha_\ell) - [\alpha_\ell/(1 + \alpha_\ell + \beta)] = 0$$

or:

$$h^*_d\alpha_\ell^2 + [h^*_d(1 + \beta - \alpha_0) + 1]\alpha_\ell - h^*_d\alpha_0(1 + \beta) = 0$$

The positive root of this equation is:

$$\alpha_\ell = \frac{\sqrt{[h^*_d(1+\beta-\alpha_0)+1]^2 + 4h^{*2}_d\,\alpha_0(1+\beta)} - [h^*_d(1+\beta-\alpha_0)+1]}{2h^*_d}$$

When the concentration of inhibitor β greatly increases, the following relation holds for low to moderate substrate concentrations:

$$h^*_d(1+\beta) - h^*_d\alpha_0 \gg 1$$

and: $$[h^*_d(1+\beta) - h^*_d\alpha_0]^2 + 4h^*_d\alpha_0(1+\beta) = [h^*_d(1+\beta) + h^*_d\alpha_0]^2$$

Consequently, $\alpha_\ell = \alpha_0$, so the local substrate concentration becomes identical to its concentration in solution. In other words, the rise in inhibitor concentration tends to remove the substrate concentration gradient. THOMAS and collaborators (1974) studied the effect of competitive and non-competitive inhibitors on the behaviour of immobilised enzymes. Figure 17.4a above shows the effect of xanthine, a competitive inhibitor of uricase, when the enzyme is in solution or immobilised. Figure 17.4b demonstrates clearly that the activity of lactate dehydrogenase is much more sensitive to non-competitive inhibition by tartrate when the enzyme is immobilised.

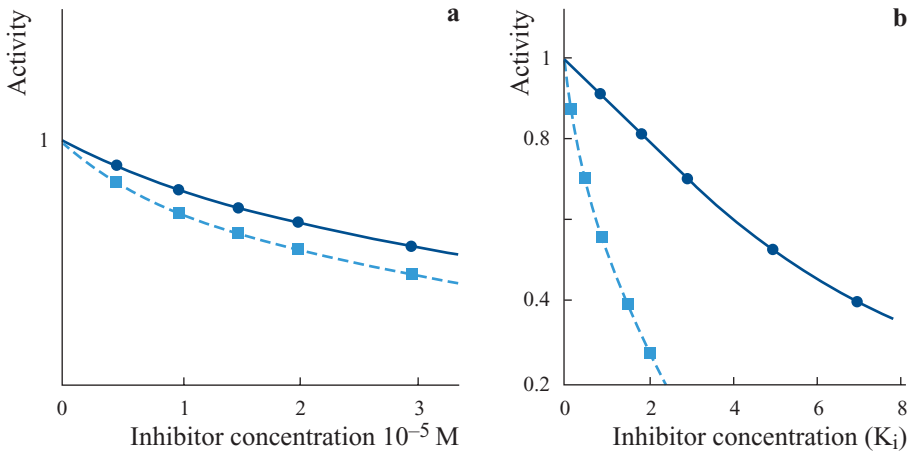


Fig. 17.4 Action of inhibitors on immobilised enzymes
(a) *competitive inhibition of uricase by xanthine; the* dashed curve *represents the enzyme in solution, the* unbroken curve *corresponds to the immobilised enzyme* ***(b)*** *non-competitive inhibition of lactate dehydrogenase by tartrate; same symbols as in* ***(a)*** (Reprinted with permission from *Biochemistry*, **13**, THOMAS D. *et al.*, 2997 and 2998. © (1974) American Chemical Society)

17.3. ELECTROSTATIC CONSTRAINTS ON IMMOBILISED ENZYMES

As well as diffusional constraints, other factors may modulate enzymatic activity within the cell. Without expanding the mathematical formalism, it is appropriate to note that the electrostatic effects may produce such phenomena as cooperativity or anti-cooperativity of membrane enzymes or enzymes bound to charged matrices, creating electrostatic attractions or repulsions depending on the respective charges

on the matrix and substrate. It follows that there are also shifts in the enzyme's activity profile as a function of pH, relative to the enzyme in solution.

At equilibrium the local substrate concentration is related to its concentration in the medium by the relationship:

$$s_\ell = s_0[\exp - (F\Delta\psi/RT)]^z$$

where $\Delta\psi = \psi_0 - \psi_\ell$ and is the electrostatic potential difference between the matrix or the membrane and the external medium, F is the Faraday and z, the valence of the charged substrate. If the enzyme obeys MICHAELIS' law, the relation:

$$v = \frac{V_m s_\ell}{K_m + s_\ell}$$

is written:

$$v = \frac{V_m s_0}{K_m[\exp(F\Delta\Psi / RT)] + s_0}$$

If the membrane or the support matrix is a polyanion, Δ^- is the density of negative charge and if the substrate is a monovalent anion (z = 1), it can be shown that, at low ionic strength:

$$\exp(-F\Delta\Psi / RT) = \frac{(\Delta^-)^2 + s_0^2}{s_0\Delta^-}$$

The MICHAELIS equation becomes:

$$v = \frac{\dfrac{V_m\Delta^-}{K_m + (\Delta^-)}s_0^2}{\dfrac{K_m(\Delta^-)^2}{K_m + \Delta^-} + s_0^2}$$

which indicates cooperativity.

If, on the other hand, the membrane or the matrix is polycationic, we have:

$$\exp(-F\Delta\psi/RT) = s_0/\Delta^+$$

and the MICHAELIS equation becomes:

$$v = \frac{V_m}{1 + \dfrac{K_m}{\Delta^+}}$$

indicating anti-cooperativity. The reaction rate thus becomes practically independent of the substrate concentration. It is worth remarking that the electrostatic repulsion of the substrate by fixed charges from a polyanionic matrix translates as cooperativity, whereas the electrostatic attraction of the substrate by the fixed charges from a polycationic matrix manifests as apparent anti-cooperativity. These phenomena of apparent cooperativity and anti-cooperativity are controlled by the ionic strength. At high ionic strength, the term $\exp - (F\Delta\psi/RT)$ tends to unity and the membrane enzyme or the enzyme fixed to a matrix behaves as an enzyme free in solution. This property enables apparent cooperativity to be distinguished from an enzyme's intrinsic cooperativity.

Enzymes from plant cell walls play important roles in the transport of organic solutes in plants and also in the extension of the cell wall. Their kinetic behaviour generally deviates from MICHAELIS' law, exhibiting apparent cooperativity or anti-cooperativity, yet when in an isolated state in an homogenous solution, the enzyme follows MICHAELIS' law. The difference in kinetic behaviour between the bound enzyme and the enzyme free in solution is due to the effects of electrostatic attraction or repulsion between the negative charges of the cell wall and the mobile charges present in the reaction mix. Acid phosphatases from soybean and sycamore cell walls, studied by RICARD (NOAT et al., 1980; RICARD et al., 1985) provide an illustration of this. The enzyme was solubilised and purified, and in solution it obeyed perfectly Michaelian kinetics. Fixed to the cell wall, it displays apparent anti-cooperativity which varies with the ionic strength of the solution. At high ionic strength, the enzyme bound to the cell wall behaves like the enzyme free in solution as a result of the elimination of electrostatic interactions.

Calcium binds reversibly to plant cell walls leading to an inhibition of cell wall extension. This process results from electrostatic effects exerted on the activity of the cell-wall enzymes. A study of acid phosphatase was carried out by the group of RICARD (CRASNIER et al., 1985) with cell walls from soybean. The calcium had been extracted by acid treatment, causing a considerable reduction in enzyme activity. The reintroduction of calcium led to enzyme activation and a decline in apparent anti-cooperativity, returning to Michaelian behaviour at high ionic strengths (Fig. 17.5), whereas calcium had no effect on the enzyme in solution.

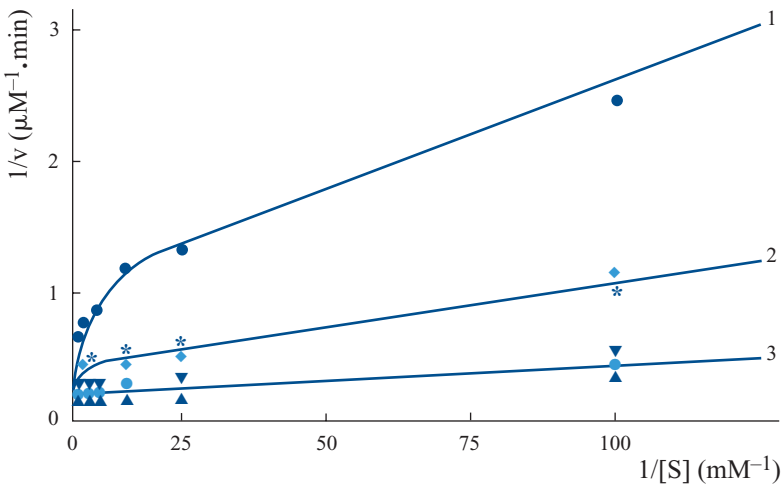


Fig. 17.5 Apparent anti-cooperativity of acid phosphatase from plant cell walls at low ionic strength

Curve 1 *corresponds to cells lacking calcium,* curve 2 *to native (+) cells, as well as to calcium-free cells that have been supplemented with 1 mM* $CaCl_2$*;* curve 3 *shows the suppression of anti-cooperativity by the addition of 0.2 M NaCl*

(Reproduced from *Eur. J. Biochem.*, **151**, CRASNIER M. *et al.*, Electrostatic effects and calcium ion concentration as modulators of acid phosphatase bound to plant cell walls, 187. © (1985) with permission of Blackwell Publishing Ltd)

Based on these charge effects, RICARD and co-workers proposed a mechanism for the growth of plant cell walls. The charge density of the cell walls due to negative charges present on pectins is maintained by different enzymes of the cell wall implicated in its growth. During periods of elongation, the volume of the cell wall increases without alteration in the number of charges, since pectins are introduced in an uncharged methylated form. The electrostatic potential difference of the cell wall diminishes leading to a pH rise and consequently activation of pectin methylesterases, which are much more sensitive to the pH when they are integrated in the cell wall; their optimal pH is about 7.5. These enzymes demethylate pectins and the number of negative charges goes up. There follows an increase in the electrostatic potential difference, and a drop in pH. Enzymes like glucanases, which have an optimal pH of between 4 and 5, become activated and then digest the pectins, thereby enabling elongation of the cell wall. A sort of hypercycle is therefore produced, in which two antagonistic enzyme reactions occur, one ending in a reduction of the charge density, the other leading to its increase, and hence assuring the growth of plant cell walls.

Thus, the spatial organisation within a cell gives rise to new properties and enables modulation of enzyme activity as a function of the cell's demands in response to changes in the environment. ◢

18 – METABOLIC CONTROL THEORY

The life of a cell relies on multiple mechanisms for regulating the diverse reactions that constitute metabolism, as has been previously explained. Some mechanisms are provided by the intrinsic properties of enzymes, such as allosteric enzymes. Others stem from the spatial organisation that allows modulation of enzymatic activity, as with membrane enzymes and multi-enzyme complexes. In these complexes the product of an enzyme reaction is frequently the substrate for the subsequent reaction. In this respect, the metabolic pathway can be considered to be an organised system.

Understanding metabolic processes in the cell requires, on the one hand, a precise knowledge of each reaction, which represents a **local property**, and on the other, the analysis of the **global properties** associated with a metabolic pathway. This is a more or less complex system of coupled enzyme reactions constituting a form of supramolecular organisation. The regulation of metabolic pathways depends on the properties of each of the constituent enzyme reactions. The coupling between these diverse enzyme reactions confers new properties on the system. Some of these enzymes play a **key role** in the metabolic pathway by exerting substantial control over the flow of metabolites that pass along it. Thus phosphofructokinase, an allosteric enzyme, is a **key enzyme** in the glycolytic pathway.

In order to illustrate the behaviour of such systems, we shall consider the control of a linear metabolic pathway at steady state and that of a metabolic cycle.

18.1. CONTROL OF A LINEAR METABOLIC PATHWAY AT STEADY STATE

Let there be a sequence of enzyme reactions:

$$S_0 \xrightarrow[v_1]{E_1} S_1 \xrightarrow[v_2]{E_2} S_2 \ \ldots \xrightarrow[v_n]{E_n} S_n$$

with S_0, $S_1 \ldots S_n$ being the metabolites, E_1, $E_2 \ldots E_n$ the corresponding enzymes, and v_1, $v_2 \ldots v_n$, the individual rates of each reaction. The flux of the metabolic pathway is defined as:

$$J = dS_n/dt$$

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0_19, © Springer-Verlag Berlin Heidelberg 2010

For some metabolic pathways the flux is regulated by an allosteric enzyme which is inhibited or activated through feedback by one of the effectors of the metabolic pathway, as has been described in Part V. But this is not true in all cases. Metabolic control theory aims to discover the general laws that control the behaviour of metabolic pathways at steady state. Many investigators, such as HEINRICH and RAPOPORT (1974), FELL and SAURO (1985), KACSER and BURNS (1979), and KACSER (1987) have analysed this theoretically. They defined two types of parameter: **control coefficients** and **elasticity coefficients**. The study of metabolic control has established two essential relationships which are applicable to all metabolic pathways regardless of their complexity. These are **the summation theorems** and the **connectivity theorems**. The **summation theorems** establish the relations between different coefficients of flux control, and for each metabolite, the relations between the control coefficients of the metabolite concentrations. The **connectivity theorems** describe the relationships between the control coefficients of the system and the elasticity coefficients. The application of these theorems permits an analysis of the enzymatic control of the variables of a metabolic pathway based on the local properties associated with each enzymatic step in the pathway.

18.1.1. CONTROL COEFFICIENTS

In a linear sequence of enzyme relations, the control coefficients are defined by the relationship:

$$C_i = \frac{E_i}{J}\frac{\partial J}{\partial E_i} = \frac{\partial \ln J}{\partial \ln E_i}$$

This relationship, introduced by HEINRICH and RAPOPORT (1974) and then by KACSER and BURNS (1979), expresses the manner in which the enzyme concentration or the reaction rate –assumed to be proportional to the enzyme concentration– affects the global metabolic flux. These expressions in fact comprise several simplifying hypotheses. They are only valid when the enzyme concentration is low compared to that of the substrate, which is not always the case in a cellular context. Furthermore, the enzymes involved in the metabolic pathway are assumed to be independent without their interacting.

Whatever may be the properties of the individual enzymes, we can write:

$$J = f(E_1, E_2 \ldots E_n)$$

The flux can be considered as a function of state:

$$dJ = \sum_{j=1}^{n}\left(\frac{\partial J}{\partial E_i}\right) dE_i$$

Taking into account the definition of the control coefficients, we derive:

$$\frac{dJ}{J} = \sum_{i=1}^{n} C_i \frac{dE_i}{E_i}$$

This relation shows that the relative perturbation of the flux is a linear first-order combination of the perturbations in each enzyme concentration.

The control of a sequence of enzyme reactions to maintain a constant steady state under the effect of a perturbation can be exerted in two ways. When an enzyme is perturbed by an external factor, one way of maintaining the steady state unchanged is to perturb in the same way all of the enzymes in the metabolic pathway. The second method consists of individually modifying the enzymes in order to eliminate the effect of the perturbation. The first mode of control rests on the **summation property**, the second on the **connectivity property**.

The summation property is expressed by the summation theorem. If the concentration or the activity of all metabolic-pathway enzymes is modified by a factor of dE_i/E_i ($i = 1, 2 \ldots n$), the flux will vary in the same manner:

$$dJ/J = dE_i/E_i$$

It follows that:

$$\sum_{i=1}^{n} C_i = 1$$

which is the expression for the **summation theorem**. This can also be written:

$$\sum_{i} E_i \frac{\partial J}{\partial E_i} = J$$

It signifies that all of the enzymes from the same metabolic pathway can contribute to controlling the flux.

18.1.2. ELASTICITY COEFFICIENTS AND THE CONNECTIVITY RELATION

The control coefficients express the manner in which the enzyme concentration or its reaction rate affects the global flux of a metabolic pathway. The elasticity coefficients enable quantification of the effect of an infinitesimal perturbation of an effector metabolite on the rate v_i of a given enzyme reaction in the metabolic pathway. The rate v_i of an enzyme reaction is a function of the concentrations of both the corresponding enzyme E_i and the effector metabolites that control the enzyme activity:

$$v_i = f(E_i, S_i \ldots S_j \ldots)$$

For a linear metabolic pathway, the control coefficients are expressed by the relation:

$$\varepsilon_j^i = \frac{S_i}{v_i} \frac{\partial v_i}{\partial S_j} = \frac{\partial \ln v_i}{\partial \ln S_j}$$

These coefficients convey how a metabolite S_j affects the reaction rate, v_i, of an individual step in the metabolic pathway.

◤ This can be illustrated using two examples. Firstly, let us consider a Michaelian enzyme competitively inhibited by an inhibitor I. In these conditions of quasi-irreversibility, the reaction rate is:

$$v = \frac{V_m s}{K_m\left(1+\frac{I}{K_i}\right)+s}$$

The elasticity coefficient relative to the inhibitor is:

$$\varepsilon_I^v = \frac{I}{v_i}\frac{\partial v}{\partial I}$$

or:
$$\varepsilon_I^v = \left(I\frac{K_m\left[1+\frac{I}{K_i}\right]+s}{V_m s}\right)\left(\frac{-K_m V_m s}{K_I\left[K_m\left(1+\frac{I}{K_i}\right)+s\right]^2}\right)$$

giving the following expression:

$$\varepsilon_I^v = -\frac{\frac{I}{K_I}}{1+\frac{s}{K_m}+\frac{I}{K_I}}$$

The elasticity coefficient is negative; indeed, with an inhibitor, the rate falls as its concentration increases.

We have: $\lim_{I\to\infty}\varepsilon_I^v = -1 \le \varepsilon_I^v \le 0 = \lim_{I\to 0}\varepsilon_I^v$

The absolute value of the elasticity coefficient grows larger as the substrate concentration declines; **the strength of the inhibitor's control is greater at low substrate concentrations.**

The second example is that of an enzyme subjected to the action of an allosteric activator A. To simplify this, the enzyme considered belongs to a K system in which the binding of the substrate and activator occurs exclusively to the R form (see Chap. 13). The principal definitions are recalled here:

$$\alpha = s/K_R \quad ; \quad \gamma = A/K_A \quad ; \quad L_0 = (T)/(R)$$

The reaction rate is given by the expression:

$$v = \frac{\alpha(1+\alpha)^{n-1}}{(1+\alpha)^n + L} \qquad \text{where } L = \frac{L_0}{(1+\gamma)^n}$$

The elasticity coefficient with respect to the activator is:

$$\varepsilon_\gamma^v = \frac{\gamma}{v}\frac{\partial v}{\partial L}\frac{\partial L}{\partial \gamma} = \frac{\gamma[(1+\alpha)^n + L]}{\alpha(1+\alpha)^{n-1}}\left[-\frac{\alpha(1+\alpha)^{n-1}}{[(1+\alpha)^n + L]^2}\right]\left[-\frac{nL_0}{(1+\gamma)^{n+1}}\right]$$

And after rearrangement of the expression:

$$\varepsilon_{\gamma}^{v} = \frac{n\gamma}{(1+\gamma)} \frac{L_0}{L_0 + [(1+\alpha)(1+\gamma)]^n}$$

The elasticity coefficient is positive; its value depends on the concentrations of both substrate and activator. When the allosteric constant is very large, the elasticity coefficient becomes equal to $n\gamma/(1 + \gamma)$. The elasticity coefficients for allosteric enzymes can take higher values than for Michaelian enzymes.

These coefficients express how an intermediate S_j affects the reaction rate v_i of a global metabolic pathway.

For a linear sequence of reactions, the differential of the reaction rate is:

$$dv_i = \frac{\partial v_i}{\partial E_i} dE_i + \sum \frac{\partial v_i}{\partial S_j} dS_j$$

This expression can also be written:

$$dv_i = \frac{\partial v_i}{\partial E_i} \frac{E_i}{v_i} \frac{dE_i}{E_i} v_i + \sum \frac{\partial v_i}{\partial S_j} \frac{S_j}{v_i} \frac{dS_j}{S_j} v_i$$

The rate of each enzyme reaction may be considered as a homogenous function of the corresponding enzyme concentration.

Thus:
$$\frac{E_i}{v_i} \frac{\partial v_i}{\partial E_i} = 1$$

and therefore:
$$\frac{dv_i}{v_i} = \frac{dE_i}{E_i} + \sum_j \varepsilon_j^i \frac{dS_j}{S_j}$$

This expression clearly shows that if the enzyme concentration is altered, the global flux of the reaction will not be affected.

The system thus displays **elasticity** relative to the perturbation in the enzyme concentration without affecting the flux, as this can be maintained constant by changing the concentration of the intermediates. If the concentration of a single intermediate S_j is modified, we have:

$$\frac{dE_i}{E_i} = -\varepsilon_j^i \frac{dS_j}{S_j}$$

In these conditions, the change in the flux is written:

$$\frac{dJ}{J} = \sum_{i=1}^{n} C_i \frac{dE_i}{E_i} = 0$$

or:
$$\frac{dJ}{J} = \frac{dS_j}{S_j} \sum_{i=1}^{n} C_i \varepsilon_j^i$$

The condition whereby the change in the flux $dJ/J = 0$ is therefore:

$$\sum_{i=1}^{n} C_i \varepsilon_j^i = 0$$

This expression is known as the **connectivity relation**.

From the summation relations of the control coefficients and the connectivity relations, and based on knowledge of the properties of each individual step, the control coefficients of the global system may be obtained. This theoretical approach to the control of metabolic pathways shows that the system's global properties differ from the sum of the properties of the individual enzymes. ◢

18.1.3. EXPERIMENTAL APPROACHES

The experimental applications of metabolic control lagged behind theoretical developments. Some experiments were carried out using isolated cells. Since the control coefficients express the effects that result from changes in the concentration or the activity of enzymes on the fluxes and the metabolite concentrations, the most direct approach consists of varying the concentration or modifying the activity of a given enzyme. In order to achieve this, diverse methods have been employed, such as altering the enzyme activity by genetic means or by introducing effectors (FELL, 1992). Metabolic control analysis requires, on the one hand, that the metabolic pathway is well characterised, and on the other, that the concentrations of substrates and allosteric effectors are measured very precisely in vivo, without ruining the system. Thanks to developments in NMR spectrometry in vivo, we currently have available a non-invasive method that enables precise evaluation of the metabolite fluxes in cells, tissues, organs and in whole animals. This explains its increasing use in clinical diagnostics replacing, in certain circumstances, the technique of biopsy.

The first studies began towards the end of the 1970s and were marked by the use of ^{31}P for studying muscle metabolism, which allowed the principal metabolites (ATP, phosphocreatine, inorganic phosphate) to be observed. More recently, NMR spectroscopy with ^{13}C has enabled significant avances in the study of carbohydrate metabolism. In view of the low natural abundance of ^{13}C, it is often necessary to administer components enriched in this isotope in order to follow their cellular metabolism. By this method, the regulation of several metabolic pathways has been studied, including the pentose cycle, the KREBS cycle and the biosynthesis of fatty acids (COHEN, 1989).

The group of SHULMAN at Yale studied the control of glycogen metabolism in muscle and in liver by ^{13}C NMR spectroscopy. In one of these studies, SHULMAN et al. (1995), while analysing their results in the context of the theory of KACSER and BURNS, showed that glycogen synthesis in muscle is not controlled by glycogen synthase, contrary to general belief, but by the glucose/hexokinase (GT/HK) transport system. The pathway for glycogen biosynthesis is schematised in Fig. 18.1 opposite. Glucose is transported into muscle by the glucose transporter (GT) and it is phosphorylated inside the cell by hexokinase to produce glucose-6-phosphate (G-6-P). A glucose molecule is added to glycogen *via* two other intermediates, glucose-1-phosphate and UDP-glucose (UDPG). This addition is catalysed by glycogen synthase, an allosteric enzyme whose activity is dependent on phosphorylation-dephosphorylation. The use of ^{13}C NMR permitted monitoring the

concentration of glycogen and ^{31}P NMR, that of glucose-6-phosphate. From these experiments the investigators found that the rate of glycogen synthesis is proportional to the transport activity of glucose/hexokinase; this step has a control coefficient near to unity at constant insulin concentrations in a healthy subject. From summation theory, this implies that the control coefficients of the other steps in the metabolic pathway are close to zero. The phosphorylation state of glycogen synthase is independent of the plasma glucose concentration. From metabolic control theory, the enzyme displaying the greatest elasticity coefficient exerts the least control over the metabolic flux. Now, glycogen synthase is highly sensitive to the G-6-P concentration, whereas the GT/HK step is insensitive. This ensemble of data led to the conclusion that the GT/HK step controls the rate of glycogen synthesis and that the role of phosphorylation of glycogen synthetase is to adapt the enzyme's activity to the flux, thus controlling the metabolite concentration, but not the flux.

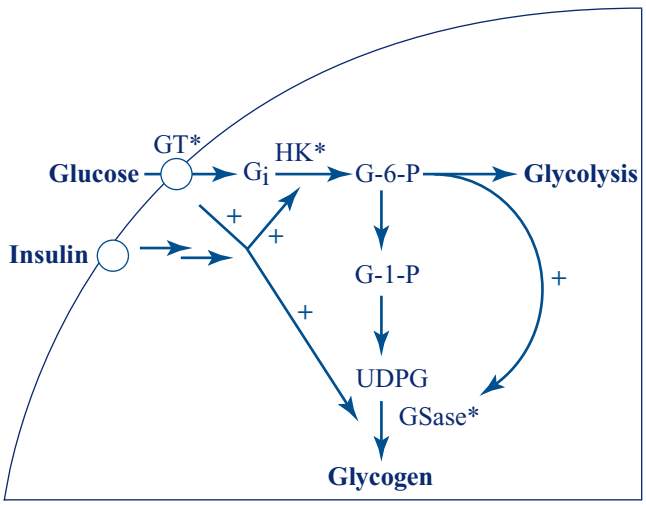


Fig. 18.1 Glycogen synthesis pathway
The + sign indicates the steps that are subjected to regulation by insulin

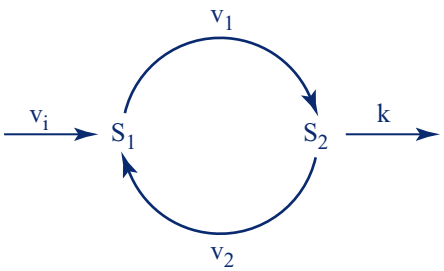
This study illustrates the potential of NMR for studying metabolic pathway control in vivo. At present, noteworthy methodological advances are being made in in vivo proton NMR spectroscopy, which, amongst other things, enables the measurement of lactate in muscle. New developments are to be anticipated in the coming years that will offer huge possibilities for the exploration of cellular metabolism.

18.2. CONTROL OF METABOLIC CYCLES

Cellular metabolism comprises a certain number of cyclical processes. Some, such as futile cycles, can be considered to be models of metabolic cycles and complexes.

These systems are able to produce an amplified reponse to small variations in the signal's intensity.

Let us consider a simplified, open metabolic cycle comprising two antagonistic enzyme reactions:



Let S_1 and S_2 be the concentrations of two metabolites, S_{1ss} and S_{2ss} their concentrations at steady state, and σ_1 and σ_2 the respective deviations with respect to the steady state. When these deviations have small amplitudes, the system can be analysed by the phase-plane technique (NICOLIS & PRIGOGINE, 1979). The appearance rate of the metabolites is given by the following relations:

$$\frac{dS_1}{dt} = \frac{dS_{1ss}}{dt} + \frac{d\sigma_1}{dt} = \frac{d\sigma_1}{dt}$$

$$\frac{dS_2}{dt} = \frac{dS_{2ss}}{dt} + \frac{d\sigma_2}{dt} = \frac{d\sigma_2}{dt}$$

At steady state, the following equations are valid:

$$\frac{dS_{1ss}}{dt} = 0 = f(S_{1ss}, S_{2ss}) = v_i + v_2 - v_1$$

$$\frac{dS_{2ss}}{dt} = 0 = g(S_{1ss}, S_{2ss}) = v_1 - (v_2 + kS_{2ss})$$

If the deviations σ_1 and σ_2 are small enough, it is possible to expand them in a TAYLOR series by conserving only the linear terms:

$$\frac{d}{dt}\begin{bmatrix}\sigma_1 \\ \sigma_2\end{bmatrix} = \begin{bmatrix}\frac{\partial f}{\partial S_1} & \frac{\partial f}{\partial S_2} \\ \frac{\partial g}{\partial S_1} & \frac{\partial g}{\partial S_2}\end{bmatrix}\begin{bmatrix}\sigma_1 \\ \sigma_2\end{bmatrix}$$

The characteristic equation for this system is:

$$D^2 - T_jD + \Delta_j = 0$$

in which D is the differential operator d/dt, Tj the trace and Δj the determinant. The general solution to this system is:

$$\sigma_1 = c_{11}e^{\lambda_1 t} + c_{12}e^{\lambda_2 t}$$
$$\sigma_2 = c_{21}e^{\lambda_1 t} + c_{22}e^{\lambda_2 t}$$

c_{11}, c_{12} and c_{22} are constants of integration; λ_1 and λ_2 are the macroscopic rate constants and have the following relationships:

$$T_j = \lambda_1 + \lambda_2$$
$$\Delta_j = \lambda_1\lambda_2$$

The system's behaviour is defined by the values of T_j, Δ_j and $T_j^2 - 4\Delta_j$. The parabola, corresponding to the relation $T_j^2 - 4\Delta_j = 0$ (Fig. 18.2), defines six regions in the plane. Region I corresponds to $T_j < 0$, $\Delta_j > 0$ and $T_j^2 > 4\Delta_j$. The two roots λ_1 and λ_2 are real and negative, which signifies that the system is stable. When the system is perturbed from its initial steady state, it returns there in a monotonous way. The phase plane corresponds to a stable node (Fig. 18.3a below and see Chap. 3). In region IV, however, $T_j > 0$, $\Delta_j > 0$ and $T_j^2 > 4\Delta_j$, and the two roots are real and positive. The system is unstable (Fig. 18.3b). When it undergoes a perturbation from its initial steady state, it tends in a monotonous way towards a new steady state.

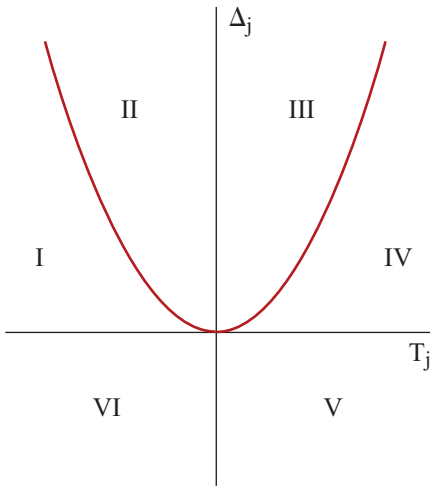


Fig. 18.2 The plane $T_j - \Delta_j$

In regions II and III, the situation is completely different. Region II corresponds to $T_j < 0$, $\Delta_j > 0$ and $4\Delta_j > T_j^2$. The two roots are complex and contain a negative, real component:

$$\lambda_{1,2} = (T_j/2) + i\omega$$

with:
$$\omega = \frac{1}{2}\sqrt{4\Delta_j - T_j^2}$$

The general solution is:
$$\sigma_1 = e^{(Tj/2)t}(c_{11}e^{i\omega t} + c_{12}e^{-i\omega t})$$
$$\sigma_2 = e^{(Tj/2)t}(c_{21}e^{i\omega t} + c_{22}e^{-i\omega t})$$

As T_j is negative, the system is stable. The phase plane is represented by a stable focus (Fig. 18.3c). Perturbed from its steady state, it returns there by a series of damped oscillations. Conversely, for region III, which corresponds to $T_j > 0$, $\Delta_j > 0$ and $4\Delta_j > T^2$, the trajectory in the phase plane is an unstable focus (see Fig. 18.3d). Perturbed from its initial steady state, the system may tend towards a new steady state by amplified oscillations.

Regions V and VI correspond to $T_j > 0$ (region V) and $T_j < 0$ (region VI), with $\Delta_j < 0$ and $T^2_j > 4\Delta_j$. The two roots are real and have opposite signs. The system is unstable and the trajectory is described by a saddle point (Fig. 18.3e). A particularly interesting case arises where $T_j = 0$ ($\lambda_1 = -\lambda_2$), $\Delta_j > 0$ and $T^2_j < 4\Delta_j$. The trajectory in the phase plane decribes a limit cycle (Fig. 18.3f) in which the system's oscillations are sustained.

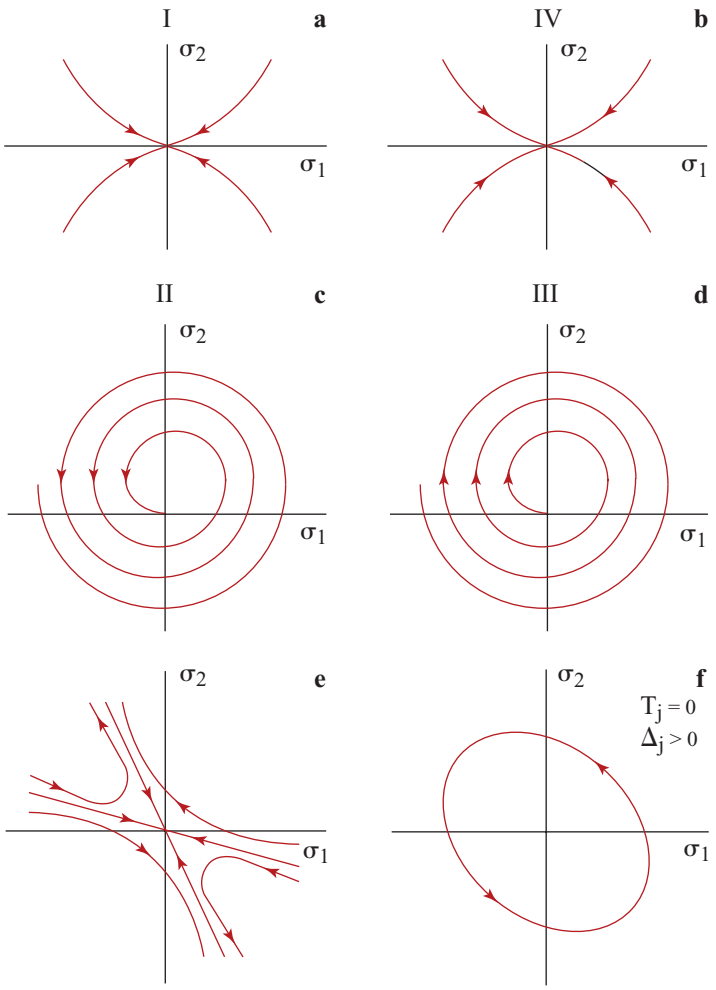


Fig. 18.3 (a) stable node of a metabolic cycle – (b) unstable node – (c) stable focus (d) unstable focus – (e) saddle point – (f) limit cycle

We explored the theoretical bases of these different behaviours in Chap. 3. They apply to open systems that are not at equilibrium. A system whose trajectory is a limit cycle represents a *dissipative structure* according to PRIGOGINE's definition. Such behaviour cannot be achieved with Michaelian enzymes, which lead to linearisable rate equations, unless there is inhibition by a substrate excess causing a deviation from linearity. However, an allosteric enzyme can induce this type of situation. Many biological systems involve oscillations. Among metabolic pathways, glycolysis represents the best studied system. Periodic oscillations in the concentrations of glycolysis-pathway intermediates have been observed in intact yeast cells as well as in extracts. They are obtained by injecting the substrate at a constant flux. Figure 18.4 shows the changes in the metabolite concentrations (from HESS & BOITEUX, 1971). The amplitude of the NADH oscillations is around 10^{-4} M. The amplitude and the frequency vary as a function of the substrate injection rate. Phosphofructokinase and pyruvate kinase –allosteric enzymes– play essential roles in the oscillatory behaviour of glycolysis. The localisation of the control points was demonstrated by HESS and BOITEUX with the help of a "crossover" diagram (Fig. 18.5 below). In this diagram, the phase angle between two adjacent intermediates in the glycolytic pathway is represented as a function of the localisation of intermediate metabolites in the reaction pathway. It appears very clearly that phosphofructokinase and pyruvate kinase are major control points in the system. All control points are coupled to the adenosine phosphate system arising either in the synthesis or in the degradation of ATP; the ATP system is responsible for the propagation of oscillations throughout the glycolytic pathway.

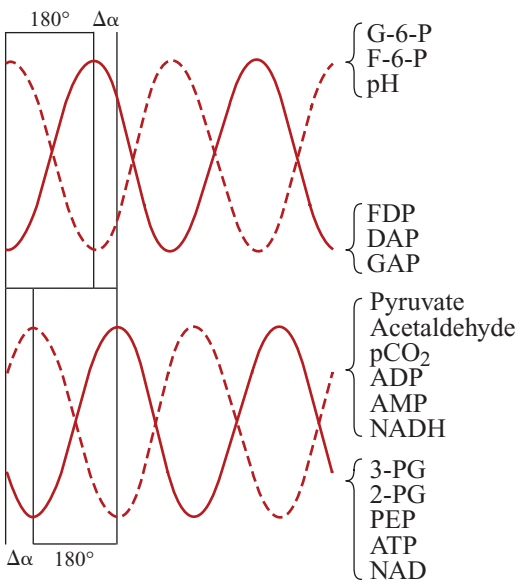


Fig. 18.4 Changes in the metabolite concentrations during glycolysis
(Reprinted, with permission, from the *Annual Review of Biochemistry*, **40**, HESS B. & BOITEUX A., 237. © (1971) Annual Reviews www.annualreviews.org)

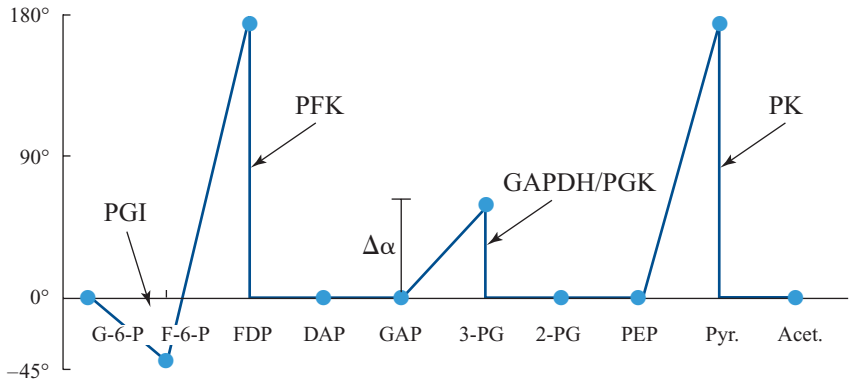


Fig. 18.5 Localisation of the control points of metabolism with the aid of a "crossover" diagram

(Reprinted, with permission, from the *Annual Review of Biochemistry*, **40**, HESS B. & BOITEUX A., 237. © (1971) Annual Reviews www.annualreviews.org)

Glycolytic oscillations have been observed in other systems, in particular in erythrocytes (RAPPOPORT et al., 1977), in ascite tumours and in pig-heart extracts. Systems other than glycolysis give rise to oscillations. The first, soluble oscillatory system to be described was an open system composed of horseradish peroxidase and lactoperoxidase in the presence of a reducing agent and oxygen. The oscillations in the ion fluxes in mitochondrial systems can be induced by adding ionophores. Oscillations in the membrane potential are observed in different membrane systems (Fig. 18.6).

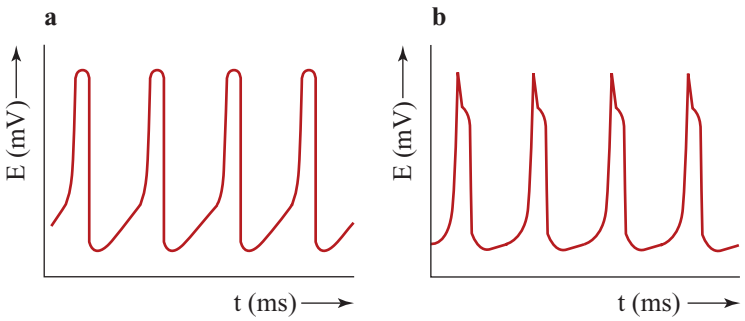


Fig. 18.6 Oscillations in membrane potentials

*(**a**) nerve fibre from octopus – (**b**) cardiac muscle fibre* (From FRANCK U.F.: Chemical Oscillations. *Angewandte Chemie*. 1978. **17**. 8. © Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission)

One of the most spectacular metabolic oscillators occurs during the aggregation of *Dictyostelium discoideum* cells over the course of its differentiation. In *D. discoideum*, cyclic AMP and folic acid induce a chemotactic response; they stimulate the synthesis of GMP and cyclic AMP. The periodic augmentation of cyclic AMP is based on the oscillatory control of the adenylate kinase activity (GERISCH et al., 1977).

These few examples illustrate that numerous oscillatory phenomena exist in living organisms. An exhaustive treatment is not feasible or necessary here. They are observed at the molecular and cellular levels, and in the more integrated context of whole organisms (respiratory and cardiac rhythms). Their physiological significance has been the subject of much discussion. It has been proposed that oscillations, like those in glycolysis, represent a model for the molecular basis of biological rhythms. It has also been suggested that such oscillatory behaviour reflects an evolutionary process for the optimisation of metabolic pathways, to ensure control of a system by allowing it to react and adapt to continual changes in external conditions. The existence of a steady state or a limit cycle may confer a selective advantage and lead to the evolution of a regulatory system. Oscillatory processes can be considered as tools of evolution for the continual selection and optimisation of dynamic structures.

BIBLIOGRAPHY

REVIEWS

HESS B. & BOITEUX A. –1971– Oscillatory phenomena in biochemistry, in *Annu. Rev. Biochem.* **40**, 237–258.

SILMAN I.H. & KATCHALSKI E. –1966– Water-insoluble derivatives of enzymes, antigens and antibodies, in *Annu. Rev. Biochem.* **35**, 873–908.

SRERE P.A. –1988– Complexes of sequential metabolic enzymes, in *Annu. Rev. Biochem.* **56**, 89–124.

SPECIALISED ARTICLES

BLOBE C.G., STRIBLING D.S., FABBRO D., STABEL S. & HANNUN Y.A. –1996– *J. Biol. Chem.* **271**, 15823.

COHEN S.M. –1989– *Methods Enzymol.* **177**, 417.

COUÉE I. & TIPTON K.F. –1989– *Biochem. J.* **261**, 921.

CRASNIER M., MOUSTACAS A.M. & RICARD J. –1985– *Eur. J. Biochem.* **151**, 187.

ENGASSER J.M. & HORVÁTH C. –1976– *Applied Biochem. Bioengng.* **1**, 127.

FELL D.A. –1992– *Biochem. J.* **286**, 313.

FELL D.A. & SAURO H.M. –1985– *Eur. J. Biochem.* **148**, 555.

FRANCK U.F. –1978– *Angewandte Chemie* **17**, 8.

GERISCH G., MAEDA Y., MALCHOW D., ROOS W., WICK U. & WURSTER B. –1977– Cyclic signals and the control of cell aggregation in *Dictyostelium discoideum*, in *Developments and differentiation in the cellular slime moulds*, CAPUCCINELLI & ASHWORTH eds, Elsevier North Holland Biomedical Press, New York.

GOODSELL D.S. –1991– *TIBS* **16**, 203.

HEINRICH R. & RAPOPORT T.A. –1974– *Eur. J. Biochem.* **42**, 89.

HEINRICH R., RAPOPORT S.M. & RAPPOPORT T.A. –1977– *Prog. Biophys. Biol.* **32**, 1.

HERVÉ G. –1996– Structural and functional dynamics of the cell, in *Concepts in Biotechnology*, K. JAYARAMAN, D. BALASUBRAMANIAN & K. DHARMALINGAM eds, Madras University-UNESCO, 21–37.

JENKINS J.D., KEZDY F.J. & STECK T.L. –1985– *J. Biol. Chem.* **260**, 10426.

KACSER H.–1987 *J. Theoret. Biol.* **126,** 505.

KACSER H. & BURNS J.A. –1979– *Biochem. Soc. Trans.* **7**, 1149.

MIRANDE M., LAZARD M., KERJEAN P., BEC G., AGOU F., QUEVILLON S. & WALLER J.P. –1993– in *The translational Apparatus*, NIERHAUS *et al*. eds, Plenum Press, New York, 657–668.

NEWTON A.C. –1995– *J. Biol. Chem.* **270**, 28495.

NICOLIS G. & PRIGOGINE I. –1979– *Proc. Natl. Acad. Sci. USA* **76**, 6060.

NOAT G., CRASNIER M. & RICARD J. –1980– *Plant Cell Environ.* **3**, 225.

RICARD J. –1989– *Catal. Today* **5**, 275.

RICARD J., NOAT G., CRASNIER M. & JOB D. –1981– *Biochem. J.* **195**, 357.

RICARD J., MULLIERT G., KELLERSHOHN N. & GIUDICI-ORTICONI M.T. –1994– Dynamics of enzyme reactions and metabolic networks in living cells. A physico-chemical approach, in *Progress in Molecular and Submolecular Biology*, W.E.G. MÜLLER ed.

SHULMAN R.G., BLOCH G. & ROTHMAN D.L. –1995– *Proc. Natl Acad. Sci. USA* **92**, 8535.

SIDEROWICZ A., MODRZYCKA T., GOLEBIOWSKA J. & SIEMIENIEWSKI H. –1990– *FEBS Lett.* **226**, 175.

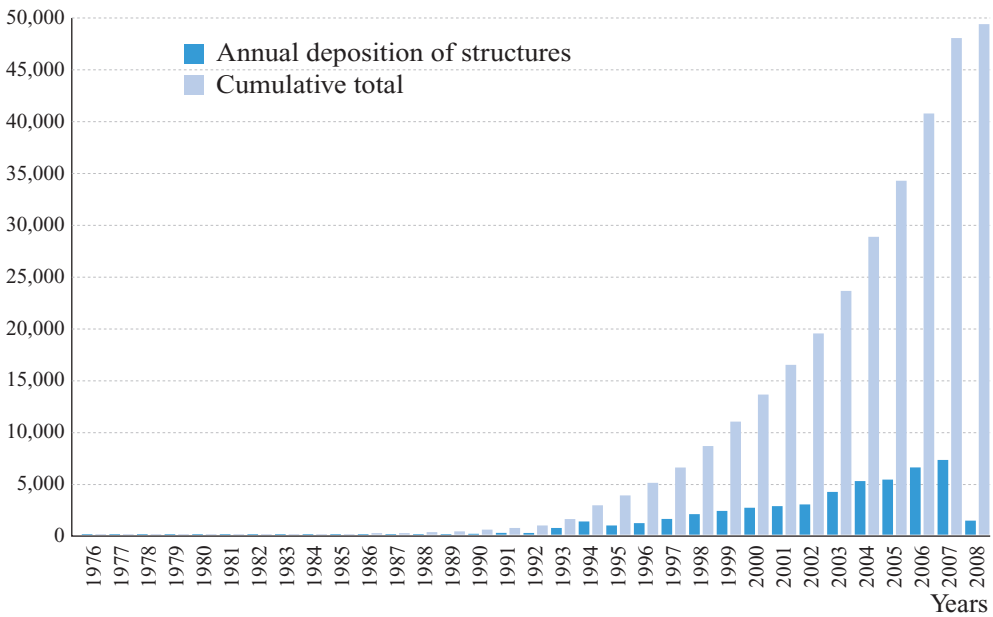
THOMAS D., BOURDILLON C., BROUN G. & KERNEVEZ J.P. –1974– *Eur. J. Biochem.* **13**, 2995.

VERTESSY B.G., KOVACS J. & OVADI J. –1996– *FEBS Lett.* **379**, 191.

WANG J., MORRIS, TOLAN & PAGLIARO L. –1996– *J. Biol. Chem.* **271**, 6861.

CONCLUSIONS AND PERSPECTIVES

In this work we have sought to cover as widely as possible the diverse aspects of enzymology. While it is clear that the fundamental aspects of thermodynamics, enzyme kinetics and the physico-chemical principles of enzymatic catalysis remain unchanged, the data concerning the structure-function relationships of particular enzyme systems are constantly evolving. This is due to the rapid growth in structural knowledge which has resulted from technological progress. It suffices to note that in 1965, the three-dimensional structure of three proteins were solved by X-ray diffraction. These were haemoglobin, myoglobin and lysozyme. In 1970, the three-dimensional structures of 11 proteins were known. In March 2008, nearly 50 000 structures were available in the *Protein Data Bank*. The figure below illustrates this progress of structural knowledge. The number of structural motifs and protein superfamilies remains more limited.



Evolution of the number of protein structures deposited in the Protein Data Bank

Today, we have particularly powerful tools at our disposal, such as synchrotron radiation sources, high-field nuclear magnetic resonance spectrometers and powerful

J. Yon-Kahn, G. Hervé, *Molecular and Cellular Enzymology*,
DOI 10.1007/978-3-642-01228-0, © Springer-Verlag Berlin Heidelberg 2010

computers which permit the resolution of protein structures and multi-molecular structures. Structural knowledge allied to genetic engineering methods and to protein chemistry enables investigation in increasing detail of the active and regulatory sites of enzymes. Even if each day new data appear, the systems that we have chosen to present serve well as useful paradigms, as much in relation to investigative strategies as to the mechanisms of action brought about by these systems.

Although structural knowledge is necessary for understanding the mechanisms involved in the functioning of enzymes and complex enzyme systems, it is, however, insufficient. The basis of their functioning even lies in the structural dynamics of enzyme molecules. Indeed, proteins are not rigid objects; their structure fluctuates around an average position. Their functional properties depend on structural fluctuations and on the flexibility of certain parts of the molecule. The existence of internal movements within proteins is not a new concept. In 1955, LINDERSTRØM-LANG hypothesised that proteins have relatively flexible structures and fluctuate in solution; he described this mobility as the "breathing" of proteins. In order to confirm this experimentally, he developed the methods of hydrogen/deuterium exchange. A little later, as we described in Part IV, KOSHLAND proposed the induced-fit model to describe the conformational adjustments of an enzyme that occur upon binding of its substrate, allowing a better orientation of the atomic groups that react during catalysis. This image of a flexible enzyme has gradually superceded the fixed image of a lock and key as was first proposed. However, this dynamic vision of protein structures was neglected for over a decade and obscured by the static view of structures in three dimensions. Since the 1980s developments in nuclear magnetic resonance spectroscopy have led to a reconsideration of the dynamic aspects of protein structure in solution and have confirmed the hypothesis of LINDERSTRØM-LANG. The existence of movement within proteins has also been recognised by crystallographers. Crystal structures only represent time-averaged structures; the temperature factors analysed during the refinement of crystallographic data have revealed larger structural fluctuations at the surface than in the protein core. Significant fluctuations have also been observed in the active sites of enzymes. More than twenty years of research has underlined the decisive role of these movements in enzymatic activity.

A great variety of internal movements occurs in proteins. These movements, of varying amplitudes, take place over a large time-scale, ranging from 10^{-14} s to several seconds. Vibrational movements correspond to times from 3×10^{-14} s to 3×10^{-12} s. Rotations about dihedral angles in the peptide backbone and protein side-chain have time constants around 10^{-8} s. Concerted movements with higher amplitudes occur over longer time periods. Proteins possess flexible and compact parts which ensure the transmission of movement across the molecule. For example, the relative rigidity of helices allows vectorial transmission of movement. Knowledge of the three-dimensional structure of an enzyme is a necessary but insufficient condition for understanding mechanisms of catalysis; an analysis of the internal dynamics of the molecule is also necessary. Local motion as well as

collective motion of high amplitude play an important role in enzyme catalysis. Their existence and analysis have been the subject of important experimental approaches and theoretical developments over the last few years. Simulation methods using molecular dynamics, which involve applying NEWTON's equations of motion to the different atoms in a molecule, have been adapted to the study of large molecules and been made possible by the increased power and speed of computers.

If molecular enzymology had been for a while obscured by the spectacular progress in molecular genetics, it nowadays experiences a growing interest within the scientific community. On the one hand, knowledge of an increasing number of genomes stimulates research aiming to characterise the function of those proteins encoded by the genes; proteomics has superceded genomics. On the other hand, the development of theoretical studies during the 1990s linking quantum mechanics with molecular mechanics (QM/MM) has seen the study of the catalytic mechanisms of enzymes regain its former stature (FIELD et al., 1990). The strategy consists of applying quantum mechanical methods to those atoms reacting in the active site of an enzyme and molecular mechanical methods to the molecule as a whole. The application of these methodologies represents a very productive approach to understanding catalytic mechanisms. Thus, several enzymes have been the subject of such studies, among which the aspartyl proteases including those from HIV, triose phosphate isomerase, lysozyme, methylene dehydrogenase, alcohol dehydrogenase, citrate synthase, thymidine phosphorylase, protein tyrosine phosphatase, chorismate mutase, β-lactamase and uracil-DNA glycosylase. In most cases, these studies have helped to remove the ambiguity concerning the mechanism and structure of the transition state.

In parallel to the development of theoretical methods, emerging technologies are permitting the evolution of a reaction to be followed over ever shorter time-periods (of nanosecond and even picosecond order). Time-resolved crystallography, made possible by the power of synchrotron radiation, enables conformational kinetics to be monitored over times as short as a few hundred picoseconds. At the moment this method is limited to a few usable systems; its extension to the study of other molecules rests with the possibility to trigger rapidly a synchronous reaction: all the molecules must be in the same state at the same time. Currently, we have at our disposal powerful tools which enable understanding of enzyme function in terms of the structure-dynamics-function relationship.

Besides the molecular aspects, developments in cellular enzymology in vivo are foreseeable in the coming years. The continual improvement in NMR spectroscopy in vivo offers an increasingly useful means to study metabolism and its regulation in cells, tissues or perfused organs and even in complete, living beings. The non-invasive character of the technique means that its use as a diagnostic tool is progressing quickly. Although not yet widespread in fundamental research, *in vivo* NMR constitutes the most appropriate method for analysing metabolic control pathways because it enables a precise measurement of the metabolite concentrations as

a function of time. These approaches, linked to studies in vitro, offer huge potential for enzymatic studies which will be increasingly necessary with the deciphering of genomes, including the human genome. It is not enough simply to know the genes; it is vital to know the proteins for which they code and their functions, of which a large number remain as yet unknown.

An entire field of study is opening up with modern enzymology which will increasingly demand a multi-disciplinary approach. Knowledge progression will result from the synergy of groups of experts in specialities relating to physics, chemistry and most biological fields. Modern enzymology is therefore progressing along two complementary axes. The first is the high-resolution molecular aspect which includes the temporal dimension of structure. The second is the cellular aspect which takes into account the biological context in which catalytic activity manifests itself and its many regulatory aspects in all their complexity.

BIBLIOGRAPHY

SPECIALISED ARTICLES

We provide here some references concerning the studies of motion within proteins in relation to their functional properties.

BRUCCOLERI R.E., KARPLUS M. & MAC CAMMON J.A. –1986– *Biopolymers* **25**, 1767.

BRUCCOLERI R.E. & KARPLUS M. –1985– *Proc. Natl Acad. Sci. USA* **82**, 4995.

CHOTHIA C., HUBBARD T., BRENNER S., BARNS H. & MURZIN A. –1997– *Annu. Rev. Biophys. Biomol. Struct.* **26**, 597.

COLONNA-CESARI F., PERAHIA D., KARPLUS M., EKLUND H., BRÁNDÉN C.I. & TAPIA O. –1986– *J. Biol. Chem.* **261**, 15273.

ECH-CHERIF EL KETTANI M.A., ZACKRZEWSKA K., DURUP J. & LAVERY R. –1993– *Proteins Struct. Funct. Genet.* **16**, 393.

FIELD M.J., BASH P.A. & KARPLUS M. –1990– *J. Comput. Chem.* **11**, 700.

GUILBERT C., PERAHIA D. & MOUAWAD L. –1995– *Comput. Phys. Commun.* **91**, 263.

LINDERTRØM-LANG K.U. –1955– *Chem. Soc. Spec. Pub.* **2**, 1.

LINDERTRØM-LANG K.U. & SCHELLMAN J.A. –1959– Protein structure and enzyme activity, in *The Enzyme*s, 2nd ed., Vol. I, P.D. BOYER ed., Acad. Press, New York, 443–510.

MARQUES O. & SANEJOUAND Y.H. –1995– *Proteins Struct. Funct. Genet.* **23**, 557.

MOUAWAD L. & PERAHIA D. –1996– *J. Mol. Biol.* **258**, 393.

THOMAS A., FIELD M.J., MOUAWAD L. & PERAHIA D. –1996– *J. Mol. Biol.* **257**, 1070.

THOMAS A., FIELD M.J. & PERAHIA D. –1996– *J. Mol. Biol.* **261**, 490.

GENERAL BIBLIOGRAPHY

SPECIALISED ARTICLES

Part I

DODE P. –1956– *Bases fondamentales et applications de la thermodynamique chimique*, Sedes, Paris.

GLANSDORFF G. & PRIGOGINE I. –1971– *Structure, stabilité, fluctuations*, Masson, Paris.

STRYER L., BERG J.M. & TYMOCZKO J.L. –2002– *Biochemistry*, 5th ed., Freeman Pub., San Francisco.

Part II

BERTRANDIAS F. & BERTRANDIAS J.P. –1997– *Mathématiques pour les sciences de la vie, de la nature et de la santé*, Collection Grenoble Sciences, EDP Sciences, Paris.

PROTASSOV K. –2002– *Analyse statistique des données expérimentales*, Collection Grenoble Sciences, EDP Sciences, Paris.

RICARD J. –1973– *Cinétique et mécanismes d'action des enzymes. I- Cinétique enzymatique*, Doin, Paris.

SEGEL I.W. –1975– *Enzyme kinetics*, John Wiley & Sons, New York.

Part III

CANTOR C.R. & SCHIMMEL P.R. –1980– *Biophysical chemistry*, Freeman Pub., San Francisco.

GHELIS C. & YON J.M. –1982– *Protein folding*, Acad. Press, New York.

JANIN J. & DELEPIERRE M. –1994– *Biologie structurale*, Hermann, Paris.

LUNDBLAD R.L. & NOYES C.M. –1985– *Chemical reagents for protein modifications*, Vol. I and II, C.R.C. Press, Boca Raton.

MC PHERSON M.J. –1990– *Directed mutagenesis: a practical approach*, Oxford Univ. Press.

Part IV

CHRISTEN P. & METYZLER D.E. –1985– *Transaminases*, John Wiley & Sons, New York.

FERSHT A.R. –1985– *Enzyme, structure and function*, 2nd ed., Freeman Pub., San Francisco.

JENCKS W.P. –1969– *Catalysis in chemistry and enzymology*, McGraw Hill, New York.

PELMONT J. –1996– *Enzymes: catalyseurs du vivant*, Collection Grenoble Sciences, EDP Sciences, Paris.

WALSH C. –1979– *Enzymatic reaction mechanisms*, Freeman Pub., San Francisco.

PHYSICAL CONSTANTS

Force: 1 dyne $= 10^{-5}$ Newton $= g \,.\, cm \,.\, s^{-2}$

Energy: 1 erg $= 10^{-7}$ J $= g \,.\, cm^2 \,.\, s^{-2}$
1 J $= 4.18$ kcal

AVOGADRO' number: $N = 6.023 \times 10^{23}$ mol

BOLTZMANN' constant: $k_B = 1.381 \times 10^{-23} \, J \,.\, K^{-1} = 3.298 \times 10^{-24} \, cal \,.\, K^{-1}$

PLANCK' constant: $h = 6.626 \times 10^{-34} \, J \,.\, s = 1.58 \,.\, 10^{-34} \, cal \,.\, s$

FARADAY' constant: $F = 9.649 \times 10^4 \, C \,.\, mol^{-1} = 2.306 \times 10^4 \, cal \,.\, vol^{-1} \,.\, eq.^{-1}$

ABBREVIATIONS

C = coulomb

cal = calorie

K = Kelvin

eq. = equivalent

J = joule

s = second

g = gram

mol = mole

INDEX

Q

R

S

T