PRINCIPLES OF MEDICAL BIOLOGY

Edited by

E. EDWARD BITTAR

NEVILLE BITTAR

CELLULAR ORGANELLES AND THE EXTRACELLULAR MATRIX

Cellular Organelles and the Extracellular Matrix

PRINCIPLES OF MEDICAL BIOLOGY
A Multi-Volume Work, Volume 3

Editors: **E. EDWARD BITTAR,** *Department of Physiology, University of Wisconsin, Madison*
NEVILLE BITTAR, *Department of Medicine, University of Wisconsin, Madison*

This Page Intentionally Left Blank

Principles of Medical Biology
A Multi-Volume Work

Edited by **E. Edward Bittar,** *Department of Physiology, University of Wisconsin, Madison* and
Neville Bittar, *Department of Medicine, University of Wisconsin, Madison*

This work in 25 modules provides:

- A holistic treatment of the main medical disciplines. The basic sciences including most of the achievements in cell and molecular biology have been blended with pathology and clinical medicine. Thus, a special feature is that departmental barriers have been overcome.

- The subject matter covered in preclinical and clinical courses has been reduced by almost one-third without sacrificing any of the essentials of a sound medical education. This information base thus represents an integrated core curriculum.

- The movement toward reform in medical teaching calls for the adoption of an integrated core curriculum involving small-group teaching and the recognition of the student as an active learner.

- There are increasing indications that the traditional education system in which the teacher plays the role of expert and the student that of a passive learner is undergoing reform in many medical schools. The trend can only grow.

- Medical biology as the new profession has the power to simplify the problem of reductionism.

- Over 700 internationally acclaimed medical scientists, pathologists, clinical investigators, clinicians and bioethicists are participants in this undertaking.

The next seven physical volumes are planned for Fall 1996.

This Page Intentionally Left Blank

Cellular Organelles and the Extracellular Matrix

Edited by **E. EDWARD BITTAR**
Department of Physiology
University of Wisconsin
Madison, Wisconsin

NEVILLE BITTAR
Department of Medicine
University of Wisconsin
Madison, Wisconsin

JAI PRESS, INC.

Greenwich, Connecticut *London, England*

Library of Congress Cataloging-in-Publication Data

Cellular organelles and the extracellular matrix / edited by E. Edward Bittar, Neville Bittar.
p. cm. -- (Principles of medical biololgy ; v. 3)
Includes index.
ISBN 1-55938-804-8
1. Cell organelles. 2. Lysosomal storage diseases. 3. Peroxisomal diseases. 4. Extracellular matrix. I. Bittar, E. Edward. II. Bittar, Neville. III. Series.
[DNLM: 1. Organelles. 2. Extracellular Matrix. QH 591' C3935 1995]
QH581.2.C453 1995
574.87'34--dc20
DNLM/DLC
for Library of Congress 95-24345
CIP

55 Old Post Road, No. 2
Greenwich, Connecticut 06836

JAI PRESS LTD.
The Courtyard
29 High Street
Hampton Hill, Middlesex TW12 1PD
England

ISBN: 1-55938-804-8

Manufactured in the United States of America

CONTENTS

LIST OF CONTRIBUTORS

Denis Crane	Faculty of Science and Technology Griffith University Nathan, Brisbane, Australia
A. J. Freemont	Department of Pathological Sciences University of Manchester Manchester, England
David S. Goodsell	Department of Molecular Biology The Scripps Research Institute La Jolla, California
D. W. L. Hukins	Department of Biomedical Physics and Bioengineering University of Aberdeen Aberdeen, England
M. J. Humphries	School of Biological Sciences University of Manchester Manchester, England
Karl I. Kivirikko	Biocenter and Department of Medical Biochemistry University of Oulu Oulu, Finland
Grazia M. S. Mancini	Department of Clinical Genetics Rotterdam Academic Hospital and Medical School Rotterdam, The Netherlands
Colin Masters	Faculty of Science and Technology Griffith University Nathan, Brisbane, Australia

Rajendra Raghow
Department of Pharmacology
School of Medicine
University of Tennessee
Department of Veterans Affairs
Medical Center
Memphis, Tennessee

Brian Storrie
Department of Biochemistry and
Anaerobic Microbiology
Virginia Polytechnic Institute and
State University
Blacksburg, Virginia

Joseph M. Tager
E.C. Slater Institute
University of Amsterdam
Amsterdam, The Netherlands

Alan M. Tartakoff
Institute of Pathology
Case Western Reserve University
Cleveland, Ohio

Jerrold R. Turner
Institute of Pathology
Case Western Reserve University
Cleveland, Ohio

David Drake Tyler
Royal Veterinary College
University of London
London, England

Frans W. Verheijen
Department of Clinical Genetics
Erasmus University
Rotterdam, The Netherlands

Ronald J.A. Wanders
Department of Pediatrics
Academic Hospital of the
University of Amsterdam
Amsterdam, The Netherlands

S. A. Weston
School of Biological Sciences
University of Manchester
Manchester, England

PREFACE

This volume is in two parts. The first contains the remaining chapters on cellular organelles and several chapters relating to organelle disorders. An account of mitochondriopathies is given in the chapter on the mitochondrion rather than in a separate one. The subject matter of this part of the volume shows quite clearly that the interdisciplinary approach to the study of organelles has shed considerable light on the nature of the mechanisms underlying the etiology and pathobiology of many of these disorders. As an example, mutations in the genes encoding integral membrane proteins are found to lead to disturbances in peroxisome assembly. It is also interesting and significant that mistargeting of protein is now thought to be another cause. It will be revealing to see whether mistargeting is the result of mutations in the genes encoding chaperones.

The second part of the volume is concerned with the extracellular matrix. It sets out to show that a vast body of new knowledge of the extracellular matrix is available to us. Take for example the integrin family of cell adhesion receptors. It turns out that integrins play a key role not only in adhesion but also in coupling signals to the nucleus via the cytoskeleton. As for fibronectins, they seem to link the matrix with the cytoskeleton by interacting with integrins.

Collagen molecules are dealt with in the last two chapters. The boundaries of collagen in disease are defined by drawing a clear line of demarcation between systemic connnective tissue disorders (e.g., scleroderma), better known as autoimmune diseases, and the heritable diseases such as osteogenesis imperfecta and the

Marfan syndrome. This classification takes into account a second group of acquired disorders of collagen forming tissues in which regional fibrosis is the hallmark. Liver cirrhosis and pulmonary fibrosis are prime examples.

The decision to place Volumes 2 and 3 before those dealing with cell chemistry was not easily made. It was based on the view that most students will have had an undergraduate course in biochemistry or cell biology or both courses, and that they could go to Volumes 4–7 in which the subject of cell chemistry is covered, and then return to Volumes 2 and 3.

Our most grateful thanks are due to the authors who have made this volume possible and to the staff of JAI Press for their courtesy and skill.

E. EDWARD BITTAR
NEVILLE BITTAR

Chapter 1

The Lysosome: Its Role in the Biology of the Cell and Organism

BRIAN STORRIE

INTRODUCTION

Each of us possess within our cells the organellar equivalent of the gastrointestinal tract. This is the vacuolar system which, like the gastrointestinal system, consists of a series of compartments. Of these, the lysosome is the final digestive compart-

Principles of Medical Biology, Volume 3
Cellular Organelles and the Extracellular Matrix, pages 1–18

ISBN: 1-55938-804-8

ment and has a special significance because it was the first described organelle within the system. In this membrane limited organelle, proteins, carbohydrates, and lipids of both extracellular and intracellular origin are digested to their constituent building blocks. These are then utilized for cellular metabolism. Nutritionally the lysosome is responsible for the release of dietary cholesterol from internalized serum proteins. In times of starvation, cells respond by increased intracellular digestion of proteins in the lysosome. Our cytosolic pools of free amino acids must be replenished. Like the stomach, the lysosome is an acidic pH compartment containing multiple acid pH optima hydrolytic enzymes. As with all subcellular organelles, the lysosomal system is marked by the presence of distinct proteins in its membrane and lumen.

In this chapter, attention centers on the lysosome. I discuss the routes of delivery of molecules, bacteria, and organelles to the lysosome and the biological role of the various compartments of the vacuolar system: endosomes, phagosomes, autophagosomes, and lysosomes. Each of these compartments, although functionally and perhaps physically interconnected, plays a distinct physiological role in delivering material to the lysosome. An emphasis on the structure-function interrelations between these compartments leads to a discussion of the biosynthesis of lysosomes. To illustrate the physiological roles of the lysosomal system, three pathways will be described for the delivery of material to lysosomes for digestion. Through heterophagic route(s) the lysosome functions in the digestion of internalized extracellular molecules with accrued benefit to the cell and the human body. Through the autophagic and the KFERQ (lys-phe-glu-arg-gln) amino acid sequence-dependent routes the lysosome functions in the turnover of intracellular molecules. Finally, I focus on the essentiality of functional lysosomes to human growth and development.

DISCOVERY OF LYSOSOMES

In the early 1950s, lysosomes were discovered inadvertently by the laboratory of Christian de Duve (Louvain, Belgium) in the course of studies initiated to understand the role of insulin in regulation of liver cell metabolism (Bainton, 1981). de Duve wished to establish the subcellular localization of the enzyme, glucose 6-phosphatase, which is involved in sugar metabolism. As a control for his subcellular fractionation work, he assayed a second phosphatase, acid phosphatase, using β-glycerol phosphate as substrate. In the course of these studies, de Duve's group found that acid phosphatase activity was present, in gently prepared cell homogenates, in a form inaccessible to substrate. Only upon further disrupting the preparation by mechanical sheer, freeze-thawing, or detergent treatment did the enzyme become readily accessible to substrate. To de Duve, this suggested the existence of a hitherto undescribed organelle in which acid pH optima enzymes were rendered inaccessible to substrate by the organelle delimiting membrane. As

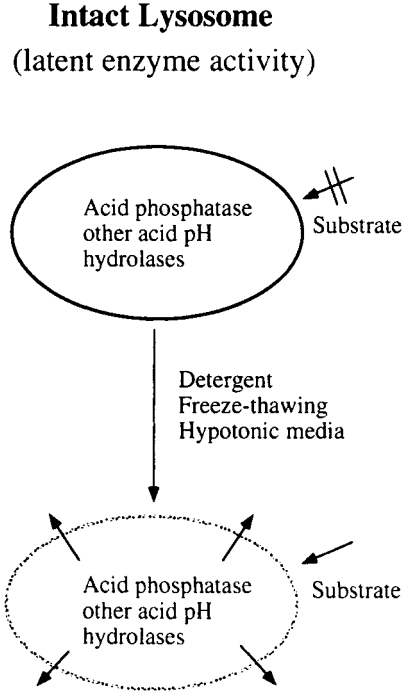


Figure 1. An intact lysosomal membrane limits the access of substrate to enzymes enclosed within the organelle. Exposure of lysosomes to detergent, freeze-thawing, or hypotonic media disrupts the lysosomal membrane and results in a highly permeable membrane through which enzymes and substrate readily move. Before exposure the lysosomal enzymes are said to be latent (i.e., inaccessible).

conceived by de Duve, the addition of detergent, for example, removed the barrier to substrate access by dissolving the membrane (Figure 1).

This concept that "latent" or sequestered enzyme activities for a series of acid pH optima enzymes are present within a distinct organelle termed the lysosome has been abundantly confirmed and extended by subsequent work. To date fifty or more enzymes have been found to be present, as a group, within the lumen of the lysosome. Moreover, all these enzymes catalyze the same general reaction:

$$A\text{–}B + H_2O \rightarrow A\text{–}H + B\text{–}OH$$

Because the reaction results in the hydrolysis of water, the enzymes are termed hydrolases. Collectively these acid pH optima hydrolases are capable of degrading almost any cellular macromolecule to its constituent monomers. These monomers are then transported from the lysosome by transport enzymes and reutilized by the cell.

In the mid-1950s, lysosomes were identified morphologically in electron micrographs of cell fractions by Alex Novikoff (University of Vermont and later Albert Einstein College of Medicine). The organelle appeared to be heterogeneous in its morphology and on the whole could be described as a small dense body of a few tenths of a micron in diameter surrounded by a single limiting membrane. The key step in showing that these dense bodies were indeed rich in acid hydrolases involved cytochemistry. In a cytochemical assay, enzyme activity is used to generate *in situ* localized deposits which can be subsequently identified under a microscope. Novikoff used the acid phosphatase catalyzed hydrolysis of β-glycerol phosphate to precipitate lead as lead phosphate. Lead deposits are electron dense and can be readily recognized under the electron microscope. With this approach, he was able to show both in cell fractions and in intact cells that the small dense bodies were indeed rich in acid phosphatase.

The localization of proteins to lysosomes by virtue of their cytochemical activity has been extended, in recent years, to the ability to localize proteins to lysosomes by their antigenicity. In the most developed form of this approach, immunospecific gold particles are used to localize the proteins themselves, without any reference to their possible enzyme activities. Moreover, using different sized immunospecific gold particles, up to three to four proteins may be localized in the same electron micrograph (Figure 2). This technique has been very important in recent work on lysosomes. Hans Geuze and Jan Slot (Utrecht, Netherlands) and Gareth Griffiths (European Molecular Biology Laboratory, Heidelberg, Germany) pioneered this approach for the study of lysosomes and the vacuolar apparatus (see Griffiths, 1993).

From the beginning, de Duve suspected that the functional role of lysosomes was intracellular digestion of macromolecules. In fact, this role was suggested in the laboratory's first paper on latent acid phosphatase activity. The first direct evidence for the digestive role of lysosomes came from the work of Werner Straus (University of North Carolina, Chapel Hill) on endocytosis in the kidney. During the mid-1950s, Straus found that the droplets responsible for the degradation of

Figure 2. Immunogold labeling of lysosomes using four different sized immunogold particles (5 nm, white filled stars; 6 nm, black arrowheads; 9 nm, white arrowheads; 16 nm, white open stars). **A**, **B**, and **C** show three examples of lysosomes (Ly) at different magnifications (bar = 0.2 μm in all cases).

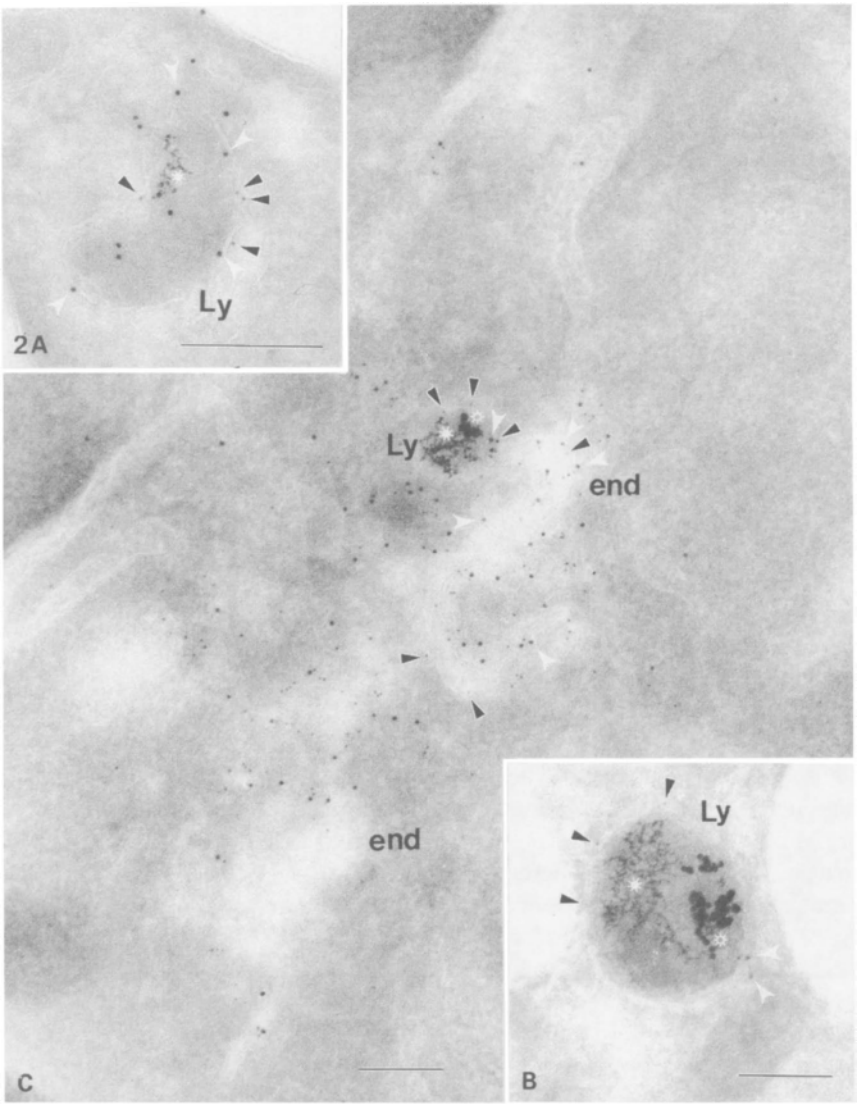


Figure 2. The lysosomes are surrounded by a single membrane (white double lines) and have an electron dense lumen. Also shown in **C** are two late endosomes (end). The late endosome is a larger structure with an electron lucent lumen. Both the lysosomes and late endosomes are positive for lysosomal membrane proteins and lumenal hydrolases (gold markers). In these preparations, the membranes are negatively stained and appear as two closely apposed, parallel white lines. Reproduced from the Journal of Cell Science with permission (Deng, Y., Griffiths, G., and Storrie, B. (1991). Comparative behavior of lysosomes and the pre-lysosome compartment (PLC) in *in vivo* cell fusion experiments. J. Cell Sci. 99, 571–582).

reabsorbed proteins by the kidney were lysosomes; they were rich in the same acid pH hydrolases described for rat liver lysosomes. Following upon that Hirsch and Cohn (Rockefeller University, New York) in the 1960s established a role for lysosomes in the digestion of internalized bacteria and cells and later in endocytosis. At about the same time, lysosomes were shown to play a central role in the digestion of cellular organelles, a process termed autophagy, by Farquhar then at the University of San Francisco Medical Center. Building upon these accomplishments, Brown and Goldstein (University of Texas Southwestern Medical School, Dallas) showed in the mid-1970s that lysosomes, as the site of digestion of cholesterol esters internalized from the blood stream by receptor mediated endocytosis, were a major participant in human cholesterol metabolism (Goldstein et al., 1985).

In summary, over a period of 25 years, the lysosome went from an unknown organelle to one known to have diverse roles in cellular digestion. Moreover, during the same time period, the digestive role of lysosomes proved to be essential to human health. As first shown by Hers (Belgium) in 1963, a single gene defect in a lysosomal enzyme, α-glycosidase, resulted in an inborn metabolic disease, an example of what is now categorized as a lysosomal storage disease because of the accumulation of undigested macromolecules within the cell's lysosomes (Bainton, 1981). This and other examples of lysosomal storage diseases are covered in depth in Chapter 5.

"MEALS ON WHEELS:" DELIVERY OF MACROMOLECULES TO THE LYSOSOME

Cells use three major routes to deliver macromolecules to the lysosome for digestion (Figure 3). These are heterophagy, autophagy, and KFERQ (lys-phe-glu-arg-gln) sequence-dependent import. In heterophagy, foreign molecules and particulates such as bacteria or viruses are internalized into the cell by encapsulation in membrane-limited vesicles and typically delivered to the lysosome. Within the lysosome, the fate of many molecular complexes is degradation. However, in other cases, the molecules or particulates are adapted for intracellular trafficking and are not degraded but rather are recycled to the cell surface and reutilized or transferred across the membranes of the vacuolar system into the cytosol. In some cases, an internalized bacterium may even make itself at home and multiply within the confines of the vacuolar system (Silverstein et al., 1989). Enveloped viruses frequently use the vacuolar system as their entry portal into the cell. In autophagy, the cell delivers parts of itself via the vacuolar system to the lysosome (Dunn, 1993). In well-nourished cells, small portions of the cytosol are internalized into late endosomes by invaginations of the organelle membrane to give rise to what is often called, on the basis of its morphology, a multivesicular body. This is a non-specific process that contributes to general protein turnover. Under conditions of starvation, mitochondria and other organelles, larger portions of the cytosol, and secretory granules (a special case termed crinophagy) are delivered in bulk and later digested

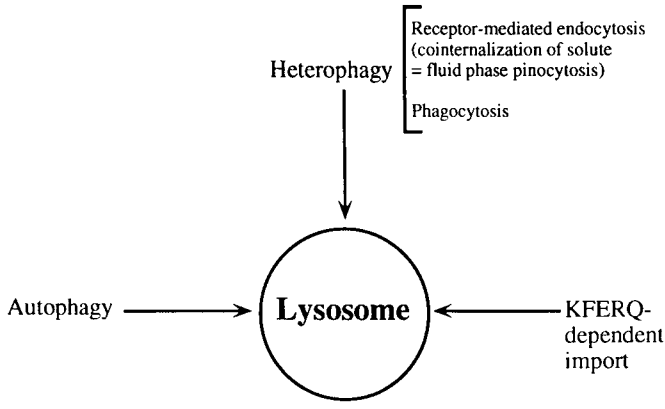


Figure 3. Routes of transport to lysosomes. Material is delivered to lysosomes for digestion by autophagy, heterophagy and KFERQ-dependent import.

within the lysosome. Finally, in the case of KFERQ-dependent import, cytosolic proteins are transported selectively as individual molecules into the lysosome system for degradation (Dice et al., 1990). In the next three sections heterophagy, autophagy, and KFERQ-dependent import are considered separately. Description of the transport of newly synthesized macromolecules and lipids to the lysosome is treated at the end of the chapter.

Heterophagy

The internalization of particulates by animal cells was first described by Metchnikoff of the Pasteur Institute, Paris in the early 1890s (Metchnikoff, 1887). Metchnikoff studied the feeding of protozoa and the uptake of particulates by defensive ameboid cells in multicellular organisms using simple light microscopy. The first description of the uptake of soluble macromolecules came some 50 years later in the mid 1930s and is due to the work of Lewis at Johns Hopkins Medical School, Baltimore. A little bit more than 40 years after that, primarily as an outcome of the work of Brown and Goldstein at University of Texas Southwestern Medical School, Dallas, and Griffins, Silverstein, and Cohn at the Rockefeller University, New York, the involvement of membrane proteins termed receptors in much of macromolecule internalization, including that of particulates, came to be appreciated.

Heterophagy, the cellular engulfment of extracellular macromolecules (endocytosis) and particulates (phagocytosis) into membrane limited vesicles, is principally a process in which receptor bound molecules or particulates are concentrated at the cell surface and then internalized (Figure 4). Internalization may be accompanied

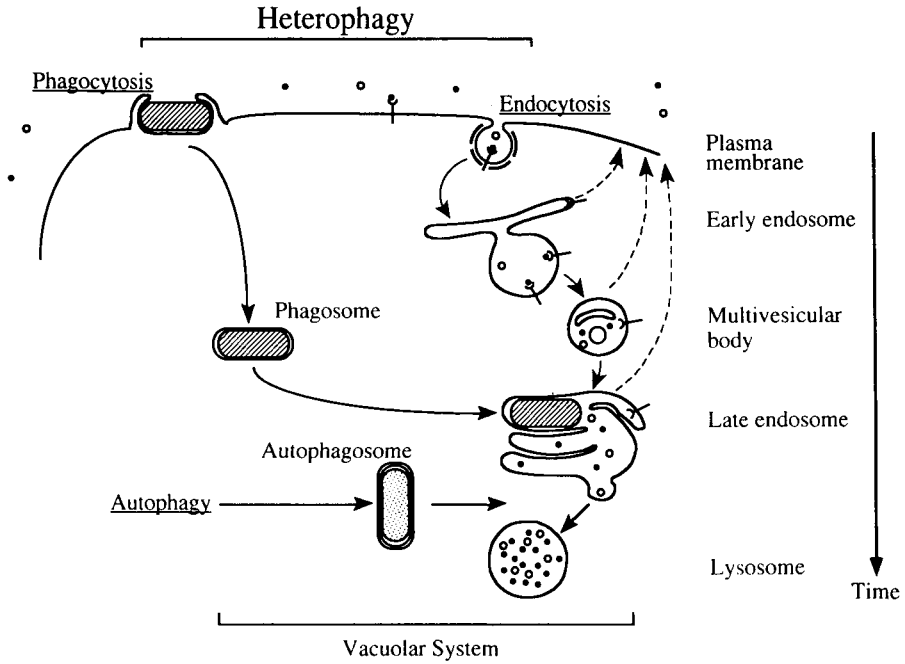


Figure 4. Schematic depiction of organelles involved in autophagic, phagocytic, and endocytic delivery routes to lysosomes. ● ligand molecule, ○ solute molecule, ⊁ endocytic receptor. Binding of ligand to receptor is a concentrative process. Adapted from Hoflack and Lobel, 1993.

by the inclusion of free solute molecules (fluid phase pinocytosis). The engulfment of bacteria is initiated by the binding of antibody coated bacteria to specific cell receptors (Silverstein et al., 1989). These are the Fc receptors of human macrophages. The Fc receptor recognizes the Fc (constant) domain of the heavy chain of immunoglobulin molecules. Because we respond to bacteria by the production of circulating antibodies to their surface components, the macrophage Fc receptor is a general receptor for the internalization of antibody coated bacteria. Receptor binding, as already indicated, is a concentrative mechanism for the engulfment of bacteria. It should be noted that, in addition to the Fc receptor, other phagocytic receptors play an important role in the turnover of such cells as the red blood cell. Aged red blood cells are engulfed and digested by macrophages in the spleen and liver. Similarly soluble macromolecules in the human blood stream, such as low density lipoprotein (LDL), a major cholesterol carrier, or transferrin, the major iron carrier, are concentrated by binding to specific endocytic receptors on the cell surface (Goldstein et al., 1985). The LDL and transferrin receptors are common proteins on the surface of many cell types in the human body while the Fc receptor is restricted to such cell types as the macrophage. All human cells require choles-

terol for their membranes and iron as a co-factor for proteins while the ability to internalize antibody coated bacteria is a specialized cell function. The role of the lysosome in these processes will be described later in the chapter.

Phagocytosis

The cellular engulfment of bacteria and other large particles such as whole cells is termed phagocytosis (Figure 4). It is an actin dependent process restricted to the immediate area of the plasma membrane to which the particle is bound and is inhibited by drugs such as cytochalasin B or D which depolymerize actin (Silverstein et al., 1989). In phagocytosis the particulate is progressively enveloped by the cellular plasma membrane in a zippering-like process in which plasma membrane receptor proteins progressively bind ligand molecules distributed over the surface of the particulate. As a result, the plasma membrane progressively envelopes the particulate. Finally, there is a membrane fusion event which leads to the complete enclosure of the particulate within what is now termed a phagosome. Initially the phagosome is devoid of acid hydrolases and its delimiting membrane is virtually identical in composition to the cellular plasma membrane. For digestion to happen, the phagosome must be transformed into a phagolysosome, a lysosome-like organelle containing the ingested particulate. The phagolysosome is an acidic compartment due to the presence of an ATP-dependent proton pump. The transformation of a phagosome into a phagolysosome is a maturation process and is the subject of current research. Maturation appears to be a progressive process in which plasma membrane derived components are conserved and recycled back to the plasma membrane at the same time that lysosomal hydrolases and membrane proteins are inserted into the organelle. In all cases, proteins are inserted and retrieved from the phagosome by membrane fusion events with vesicular carriers. The source of many of the newly inserted components may well be endosomal rather than lysosomal. Endosomes contain lysosomal hydrolases and an ATP-dependent proton pump which is capable of acidifying the organelle. Fusion of phagosomes with endosomes has been readily demonstrated *in vitro* but fusion of phagosomes with lysosomes has not. Some microorganisms, for example, *Toxoplasma gondii*, *Chlamydia psittaci*, *Legionella pneumophilia*, *Mycobacterium tuberculosis*, and *Mycobacterium microti*, when included within a phagosome, are capable of inhibiting the transformation of phagosomes into phagolysosomes. This is an important but little understood process.

Receptor-mediated endocytosis

The concentrative uptake of soluble macromolecules and small particulates such as viruses occurs by receptor-mediated endocytosis (Figure 4). Examples include low density lipoprotein and transferrin, both of which have nutritive roles as

cholesterol and iron carriers, respectively, epidermal growth factor and insulin, both of which are hormones whose internalization leads to down regulation of their receptors, and galactose terminated blood proteins which are scavenged from the blood stream by galactose receptors on the surface of liver cells. Many enveloped viruses are also internalized after binding to sugar specific receptors. Each of these is an example of a ligand in the context of ligand-receptor binding.

Receptor-mediated endocytosis is a selective process in which the ligand-receptor complex interacts with cytoplasmic proteins that coat the plasma membrane at certain places. These are indented regions of the plasma membrane known as coated pits. They occupy a small amount of the plasma membrane surface area, ~2% in fibroblasts. The coat complex consists of several polypeptides including the heavy and light chains of clathrin, the major coat protein, and adaptor proteins that interact in an as yet unknown manner with the cytoplasmic tail of transmembrane endocytic receptor proteins. The cytoplasmic tail of endocytic receptors all possess a common structural motif that contains a tyrosine residue and a three-dimensional structural feature known as a tight turn. This motif interacts with adaptor proteins which in turn react with clathrin to cause receptor accumulation in the coated pit. Accumulation is followed by internalization and the formation of a coated vesicle. This vesicle rapidly loses its coat and fuses with other newly formed vesicles to generate an early endosome. A second type of coated structure, the caveola, concentrates small molecules before their uptake into cells by transporter complexes. This process is ATP-dependent and termed *potocytosis* (Anderson, 1993). It does not involve lysosomes.

Transport between early endosomes and lysosomes is a selective process in which many plasma membrane components are conserved and recycled back to the plasma membrane while others are transported to lysosomes. In this process, the fate of ligand and receptor is frequently different. This is best illustrated by examining three examples: 1) low density lipoprotein (LDL) and its receptor, 2) transferrin and its receptor, and 3) epidermal growth factor (EGF) and its receptor.

In the case of LDL, the LDL receptor is recycled while LDL is delivered to the lysosome and degraded (Goldstein et al., 1985). Receptor conservation is achieved here as it also is in the transferrin case by virtue of the effect of acidification on ligand binding. For LDL, early endosome acidification results in the endosomal dissociation of LDL from its receptor. LDL receptor is returned to the cell surface and used for repeated rounds of internalization of LDL from the blood stream. LDL itself is delivered as solute to late endosomes and subsequently to lysosomes. LDL is degraded in the lysosome to release cholesterol. The free cholesterol is transported across the lysosomal membrane and utilized by the cell for membrane biosynthesis. Cholesterol is an important membrane lipid. Transferrin is an iron chelating protein, and at neutral pH in the blood stream it is iron laden. At acidic pH within endosomes, the iron ligand dissociates and the iron is delivered to other organelles. The resulting iron-free apotransferrin is returned to the cell surface as

an apotransferrin-transferrin receptor complex. There, at the neutral pH of the blood stream, apotransferrin dissociates from the receptor and both receptor and apotransferrin may be reutilized for another round of iron transport into the cell. In striking contrast, binding of EGF and its receptor is pH insensitive; the pH stable ligand receptor complex is delivered to lysosomes and both components degraded. Early endosomes have a pH of about 6 while lysosomes have a pH of about 4.5–5.0. The blood stream has a pH of about 7.3.

Autophagy

Cells respond to starvation by sequestering segments of their cytosol or organelles such as mitochondria within a double membrane system termed an autophagosome (Figure 4). Autophagosomes appear to be non-selective with respect to included material. As shown by the recent immunoelectron microscopic studies of Dunn (1993), the protein antigens present on the enveloping double membrane system are characteristic of the rough endoplasmic reticulum (RER). Hence, the source of the enveloping double membrane must be ribosome-free portions of the RER. This newly formed autophagosome subsequently matures into a late autophagosome by the insertion of lysosomal membrane proteins into its outer membrane and acquisition of a proton pump. These proteins may be transported to the autophagosome from the Golgi apparatus or elsewhere. At the same time a large portion of the RER proteins are lost. This acid pH organelle then acquires acid hydrolases along with additional lysosomal membrane protein from lysosomes, endosomes, or both and becomes an autolysosome. During this process, the inner enveloping membrane of the organelle is lost along with further RER proteins. Whether these RER derived proteins are lost by digestion or are recycled to the RER remains an open question. The autolysosome is a digestion competent compartment and the enveloped material is digested to its constituent monomers. The resulting monomers are transported from the lysosome and utilized by the cell for protein biosynthesis. Protein digestion within autolysosomes has a major role in maintaining cellular amino acid levels during starvation.

KFERQ (Lys-Phe-Glu-Arg-Gln)-Dependent Import

As shown by Dice et al. (1990) over the last 15 years, human cells respond to growth factor deprivation by KFERQ-sequence dependent import of proteins into lysosomes. Unlike autophagy this is a selective process which can lead to the large scale degradation of proteins containing the KFERQ motif. About 30% of cytosolic proteins in human fibroblasts contain an exposed KFERQ-sequence motif and upon growth factor deprivation up to four fifths of these proteins are degraded. The KFERQ motif is recognized by hsc73, a heat shock cognate protein of 73 kD. Hsc73 is a member of the family of hsp70 heat shock proteins. Heat shock proteins are

ATPases capable of folding and unfolding proteins. Binding of hsc73 to the KFERQ motif is a specific recognition event that leads to the subsequent import of the target protein into the lysosomal apparatus. *In vitro* studies with isolated lysosomes suggest that this import is into lysosomes rather than endosomes although this has not been directly shown. Presumably there must exist a hsc73-mediated protein import system on the surface of lysosomes or endosomes. At present the physiological benefit of KFERQ-dependent protein import is unknown.

ENDOSOMES AS WELL AS LYSOSOMES PLAY A ROLE IN PROTEIN DEGRADATION

In emphasizing the lysosome as the end site of "meals-on-wheels" degradation of proteins, we should not lose sight of the fact that proteinases are found in almost all subcellular compartments including the cytosol, mitochondria, endosomes, plasma membrane, and endoplasmic reticulum. The lysosome is only one of several sites of protein degradation in cells. Nevertheless, it is often the major site of protein degradation. In liver cells, for example, lysosomal protein degradation accounts for as much as 70% of total protein degradation (Dunn, 1993). Although protein degradation has been emphasized here, attention should be drawn to the fact that lysosomal hydrolases represent a wide range of enzymes including glycosidases and lipases. The lysosome plays a major role in lipid degradation as well as protein degradation. Finally, the reader should realize that although lysosomal hydrolases are concentrated in the lysosome, degradation may occur in earlier compartments such as one or another variety of endosome or phagosome (Brodsky, 1992). These compartments are acidic and do contain small amounts of lysosomal acid hydrolases. For molecules which are readily degraded, much of the degradation may occur as the molecule is being transported to the lysosome.

ROLE OF ENDOSOMES AND LYSOSOMES IN ANTIGEN PRESENTATION

Cytotoxic and helper T-cells are activated in response to peptides bound to class I and class II molecules coded by the major histocompatibility complex (MHC) (Figure 5, Brodsky, 1992; Teyton and Peterson, 1992; van Bleek and Nathenson, 1992; Kosacsovics-Bankowski et al., 1993; Pfeifer et al., 1993). Class I molecules during their export to the cell surface typically bind peptides generated from proteins synthesized by the same cell. When these peptides are non-self, for example, a cellularly synthesized, but virally coded peptide, binding is to class I molecules and cytotoxic T-cells are activated when the peptide is presented as a cell surface complex with class I molecules. Likewise class II molecules also bind peptide during their export. Here binding occurs within the endocytic apparatus. When the peptide is the product of an endocytosed foreign molecule helper, T-cells are activated by peptide-class II molecules complexes at the cell surface. Overall

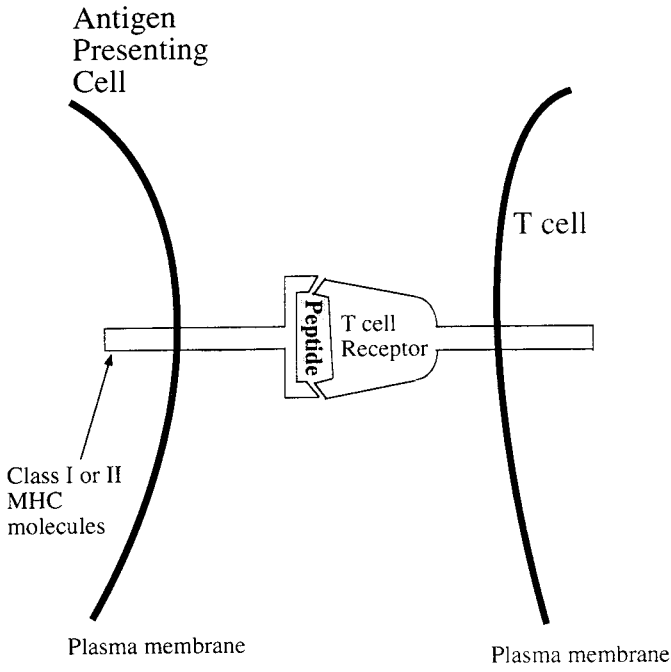


Figure 5. Schematic depiction of peptide-MHC class I or II molecular complex binding with the T cell receptor. This is an essential step in antigen presentation. The peptide is generated intracellularly within the antigen presenting cell. Presentation of antigen, i.e., peptide, occurs at the cell surface.

these processes are referred to as antigen presentation. For a detailed description of T-cells and cellular immunity, the student is referred to an immunology text.

The peptides bound to class II molecules are generated by lysosomal hydrolases within acid pH compartments. Drugs such as chloroquine and ammonium chloride, which neutralize endosomal and lysosomal pH, inhibit antigen presentation by class II molecules. Whether the peptides are generated within endosomes or lysosomes is a point of current controversy. Some data, such as the inhibition of presentation by mutations which affect endosomal but not lysosomal pH, point to the major site being endosomal. Other data from experiments using liposomes as carriers to deliver antigen to late compartments (pH = 4.5–5.0) suggest a major role for lysosomes in peptide generation. The question of whether the generation of peptide and its binding to class II molecules must occur within the same endosome or lysosome is a second example of a major unanswered question with respect to the class II presentation pathway. Recent evidence suggests that phagocytosis, another heterophagic process, can play a role in the presentation of exogenous foreign antigens by class I molecules.

In conclusion, antigen presentation and the role of the endosomal/lysosomal system in this process is an important area in which the general outlines of the process are only beginning to emerge.

LYSOSOMAL BIOGENESIS

The biochemical pathway by which lysosomes are formed may be drawn as a Y-shaped figure in which input streams from both the secretory and the endocytic pathways converge to contribute to lysosomal biogenesis (Figure 6, Kornfeld and Mellman, 1989; Hoflack and Lobel, 1993; Storrie, 1993). Proteins and lipids destined for the lysosome are synthesized and/or modified in the endoplasmic reticulum and the Golgi apparatus. From there they are transported to a prelysosomal/endosomal compartment by vesicles which bud from the *trans* Golgi network, a sorting compartment on the exit side of the Golgi apparatus. Both membrane and water soluble, lumenal proteins of the lysosome are synthesized on membrane bound polysomes of the RER, initially glycosylated within the RER, and from there transported to the Golgi apparatus (Figure 7).

N-linked oligosaccharide chains of glycoproteins are further modified in the Golgi apparatus. Newly synthesized lysosomal membrane proteins become highly sialylated in the *trans* Golgi apparatus but not phosphorylated. Newly synthesized lysosomal acid hydrolases become phosphorylated on mannose residues in the *cis* Golgi apparatus but not sialylated in the *trans* Golgi apparatus. The *cis* and *trans* Golgi compartments are located on the entry and exit sides, respectively, of the organelle. Phosphorylation is initiated by a specific enzyme in response to a non-contiguous set of amino acids located on the surface of newly synthesized acid hydrolases (Figure 8). These amino acids form a recognition feature known as a

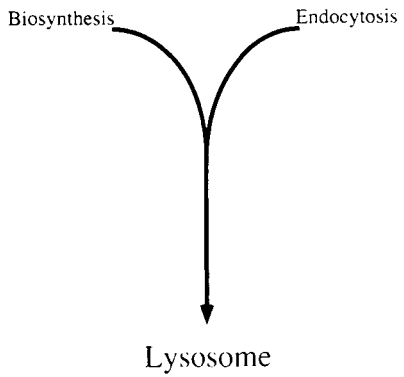


Figure 6. Schematic depiction of the convergence of biosynthetic and endocytic contributions to lysosomal biogenesis.

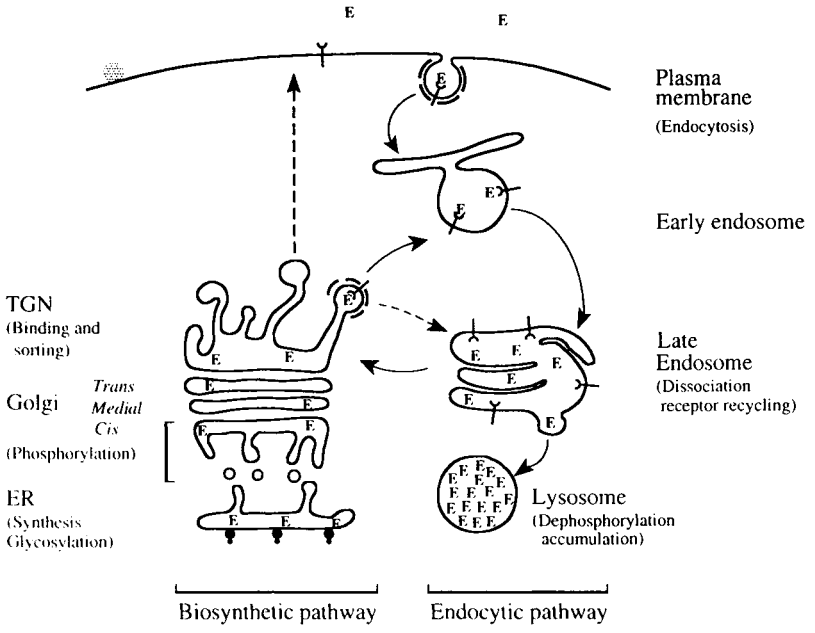


Figure 7. Schematic depiction of the trafficking of newly synthesized lysosomal enzymes from their site of synthesis in the rough endoplasmic reticulum to the lysosome. E, enzyme; ⋎ mannose 6-phosphate receptor. Small amounts of lysosomal enzymes secreted from cells are captured by cell surface mannose 6-phosphate receptors. Adapted from Hoflack and Lobel, 1993.

targeting patch. The mannose 6-phosphate residues generated play an important role in the sorting of these glycoproteins to lysosomes.

Two different mechanisms exist for the targeting of proteins to lysosomes. The first is mannose 6-phosphate dependent and targets newly synthesized lysosomal hydrolases to the lumen of the lysosome. In this mechanism lysosomal hydrolases bind to mannose 6-phosphate receptors within the Golgi apparatus. In the *trans* Golgi network, the ligand receptor complex is sorted into clathrin coated vesicles and transferred from the Golgi to a prelysosomal endosomal compartment. There the hydrolase-receptor complex dissociates to generate free mannose 6-phosphate receptors. These are recycled back to the Golgi apparatus for another round of hydrolase transport and the free hydrolases are transported to lysosomes via endosomes. The second mechanism is dependent on sequence features of the cytoplasmic domain, the cytoplasmic tail, of newly synthesized lysosomal membrane proteins. This domain is both *necessary* and *sufficient* for the targeting of membrane proteins to lysosomes. When deleted, targeting to the lysosome ceases and the protein is delivered to the cell surface. When placed as the cytoplasmic

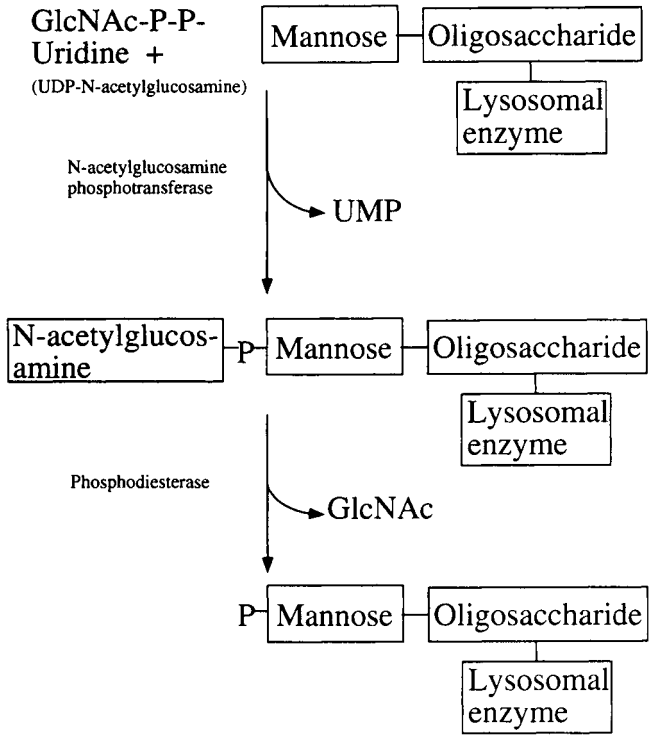


Figure 8. Mechanism of mannose 6-phosphate phosphorylation of lysosomal enzymes.

domain of another protein, the protein is now targeted to the lysosome. As with hydrolase transport, membrane protein transport to the lysosome converges with the endocytic pathway. For both soluble hydrolases and membrane proteins the exact compartment of convergence is a point of current controversy. For some newly synthesized proteins, convergence may be at or near early endosomes. For others, convergence may well be at a late endosomal compartment. This late endosomal compartment, also referred to as the prelysosomal compartment (PLC), is rich in mannose 6-phosphate receptors.

ESSENTIALITY OF LYSOSOMES TO THE CELL AND ORGANISM

Lysosomes and associated heterophagic and autophagic organelles have a wide range of roles which are important to both the individual cell and the organism. Genetic defects in lysosomal enzymes result in the premature death of humans. Lysosomes are essential for our health. Lysosomes, however, do not play an

essential role to the cell itself. Human cells with multienzyme lysosomal defects do grow and divide. Indeed in culture, these cells are healthy with the exception of an enlarged lysosomal compartment due to the accumulation of molecules that would normally be digested. Humans with I-cell disease in which lysosomal enzymes as a class are secreted are born and live for several years. In the end, the essentiality of the lysosome is to the organism, to us as complex multicellular beings. We must indeed degrade foreign molecules and molecules which our cells produce to grow and develop into healthy adults.

SUMMARY

The lysosome is a degradative organelle located in the cytoplasm of eukaryotic cells. It is marked by the presence of distinctive acid pH optima hydrolases in its lumen and highly sialylated membrane proteins. Together these hydrolases are capable of digesting most molecules synthesized by cells to their constitutive monomers. Lysosomal enzymes and membrane proteins were originally thought to be restricted to the lysosome itself. However, these proteins may also be found in related organelles such as endosomes, phagosomes, and autophagosomes which deliver material to lysosomes for its final digestion. Three major routes exist for the transport of molecules to lysosomes for digestion. These are heterophagy (phagocytosis and endocytosis), autophagy, and KFERQ-dependent import of cytosol proteins into the lysosomal/endosomal apparatus. The vacuolar system (i.e., lysosomes, endosomes, phagosomes, and autophagosomes) plays an essential role in the protection of cells from bacterial invasion, antigen presentation, and in the turnover of cellular molecules. Defects in lysosomal enzymes result in lysosomal storage diseases and the premature death of humans.

REFERENCES

Anderson, R. G. W. (1993). Potocytosis of small molecules and ions by caveolae. Trends Cell Biol. 3, 69–72.

Bainton, D. F. (1981). The discovery of lysosomes. J. Cell Biol. 91, 66s–76s.

Brodsky, F. M. (1992). Antigen processing and presentation: close encounters in the endocytic pathway. Trends Cell Biol. 2, 109–115.

Dice, J. F., Terlecky, S. R., Chiang, H.-L., Olson, T. S., Isenman, L. D., Short-Russell, S. R., Freundlieb, S., & Terlecky, L. J. (1990). A selective pathway for degradation of cytosolic proteins by lysosomes. Sem. Cell Biol. 1, 449–455.

Dunn, W. A., Jr. (1993). Mechanisms and regulation of autophagic degradation of cellular proteins. In: Endosomes and Lysosomes: A Dynamic Relationship (Storrie, B. & Murphy, R. F., eds.), Advances in Cell and Molecular Biology of Membranes (Tartakoff, A. M., ed.), 1, pp. 117–138, JAI Press, Greenwich.

Goldstein, J. L., Brown, M. S., Anderson, R. G. W., Russell, D. W., & Schneider, W. J. (1985). Receptor-mediated endocytosis: concepts emerging from the LDL receptor system. Ann. Rev. Cell Biol. 1, 1–39.

Griffiths, G. (1993). Fine Structure Immunocytochemistry. Springer, Heidelberg.

Hoflack, B. & Lobel, P. (1993). Functions of the mannose 6-phosphate receptors. In: Endosomes and Lysosomes: A Dynamic Relationship (Storrie, B. & Murphy, R. F., eds.), Advances in Cell and Molecular Biology of Membranes (Tartakoff, A. M., ed.), 1, pp. 51–80, JAI Press, Greenwich.

Kovacsovics-Bankowski, M., Clark, K., Benacerraf, B., & Rock, K. L. (1993). Efficient major histocompatibility complex class I presentation of exogenous antigen upon phagocytosis by macrophages. Proc. Natl. Acad. Sci. USA 90, 4942–4946.

Kornfeld, S. & Mellman, I. (1989). The biogenesis of lysosomes. Ann. Rev. Cell Biol. 5, 483–525.

Metchnikoff, E. (1887). Sur la lutte des cellules de l'organisme contre l'invasion des microbes. Ann. Inst. Pasteur 1, 321.

Pfeifer, J. D., Wick, M. J., Roberts, R. L., Findlay, K., Normark, S. J., & Harding, C. V. (1993). Phagocytic processing of bacterial antigens for class I MHC presentation to T cells. Nature 361, 359–362.

Silverstein, S. C., Greenberg, S., Di Virgilio, F., & Steinberg, T. H. (1989). Phagocytosis. In: Fundamental Immunology, 2nd ed. (Paul, W. E., ed.), pp. 703–720, Raven Press, New York.

Storrie, B. (1993). Towards an understanding of the inheritance of mammalian lysosomes and yeast vacuoles. In: Endosomes and Lysosomes: A Dynamic Relationship (Storrie, B. & Murphy, R. F., eds.), Advances in Cell and Molecular Biology of Membranes (Tartakoff, A. M., ed.), 1, pp. 391–406, JAI Press, Greenwich.

Teyton, L. & Peterson, P. A. (1992). Invariant chain—a regulator of antigen presentation. Trends Cell Biol. 2, 52–56.

van Bleek, G. M. & Nathenson, S. G. (1992). Presentation of antigenic peptides by MHC class I molecules. Trends Cell Biol. 2, 202–207.

RECOMMENDED READINGS

Holtzman, E. (1989). Lysosomes. Plenum, New York.

Storrie, B. & Murphy, R. F., eds. (1993). Endosomes and Lysosomes: A Dynamic Relationship. In: Advances in Cell and Molecular Biology of Membranes, Vol. 1 (Tartakoff, A. M., ed.), JAI Press, Greenwich.

Chapter 2

The Golgi Complex

ALAN M. TARTAKOFF and JERROLD R. TURNER

DISTRIBUTION AND BASIC STRUCTURAL CHARACTERISTICS

Camillo Golgi's unexpected staining of an "internal reticular apparatus" of spinal neurons with silver nitrate made the Golgi Complex (GC) first visible as a "black reaction" at the very end of the last century. Ironically, neither the basis of this reduction of heavy metal salts, nor the reason why the GC is so conspicuous in neurons, is adequately understood. What is well-established, however, is that the GC exists in all animal, plant, and fungal cells which contain rough endoplasmic reticulum (RER). The GC is now known as a station along both the secretory and endocytic paths where an impressive number of covalent and non-covalent post-

Principles of Medical Biology, Volume 3
Cellular Organelles and the Extracellular Matrix, pages 19–38

ISBN: 1-55938-804-8

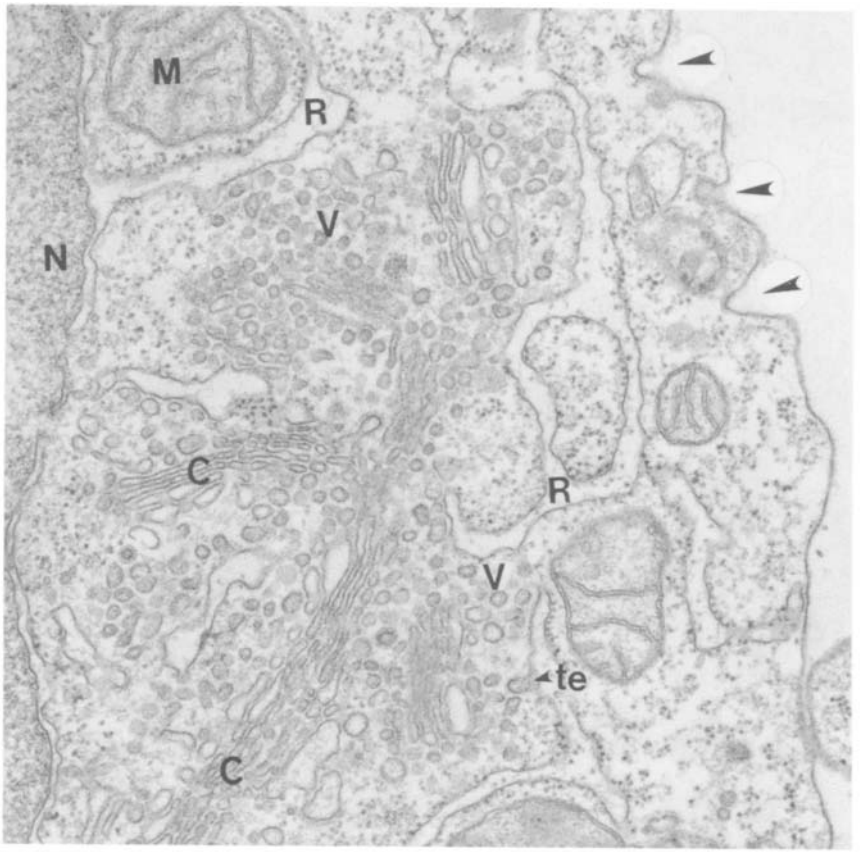


Figure 1. Thin section of the Golgi region of a myeloma cell fixed with glutaraldehyde followed by osmium tetroxide. Note the rough endoplasmic reticulum (R), transitional elements (te), conspicuous stacked cisternae (C) and associated vesicles (V). (N) nucleus, (M) mitochondrion. Arrowheads designate coated pits at the cell surface.

translational modifications occur and where macromolecules are sorted to a variety of destinations.

The ultrastructural hallmark of the GC is the set of 3–30 elongated, closely-apposed cisternae (Figure 1) which in three dimensions constitute a pile of somewhat saucer-shaped closed membrane-bounded saccules. Individual cisternae are often perforated by pores, and small smooth-surfaced vesicles are found in the vicinity of the cisternae. The cisternae closest to the RER are generally referred to as "cis," "proximal" or "forming," while cisternae further removed are called "trans," "distal" or "mature." Beyond the most distal of the closely apposed cisternae are immature secretion granules (condensing vacuoles, Figure 2) and additional anas-

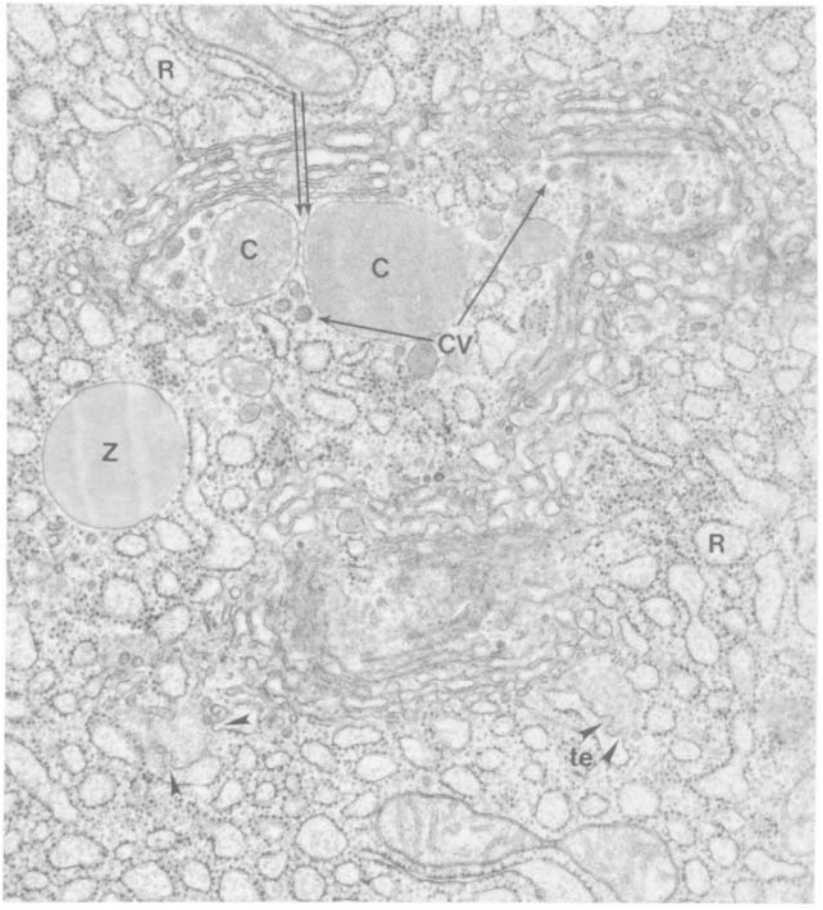


Figure 2. Thin section of a guinea pig exocrine pancreatic cell. Mature secretion granules [zymogen granules] (Z) and immature granules [condensing vacuoles] (C) are adjacent to the Golgi cisternae, some of which contain visibly concentrated content (arrowhead). CV: coated vesicles. (te) designates transitional elements of the RER. The large double arrow indicates the proximal-to-distal axis across the stack of Golgi cisternae. Unlike the cell types illustrated in Figure 1, exocrine pancreatic cells secrete primarily via the "regulated" secretory path, i.e., after concentration of their secretory products in the secretion granules.

tomosing and tubular smooth cisternae, known as GERL, Trans-Golgi Network, or Trans-Golgi Reticulum (Figure 3). These latter structures are best characterized in cells, such as fibroblasts, which do not store secretory products in granules. Their exact identification must await the development of suitable reagents (e.g., antibod-

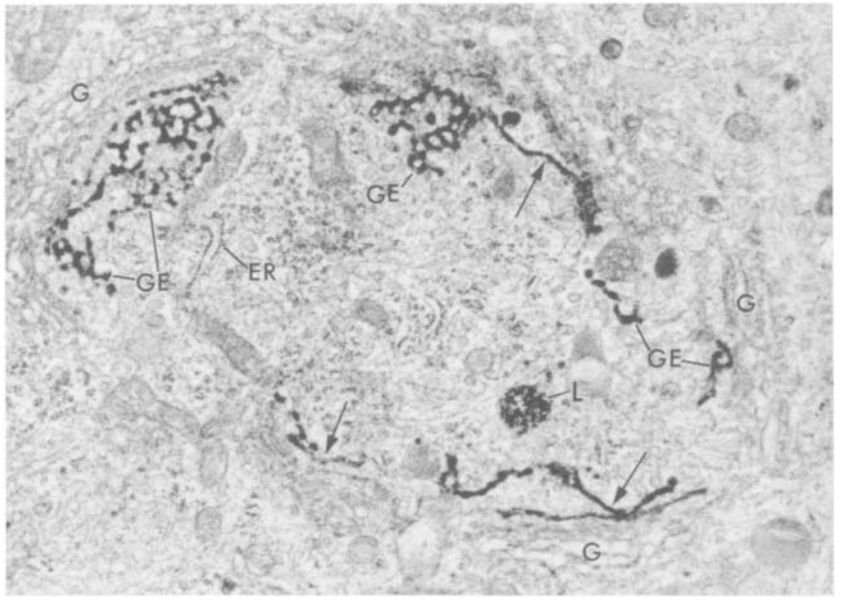


Figure 3. Electron micrograph of a rat dorsal root ganglia neuron incubated for acid phosphatase activity. Electron-opaque reaction product delineates the anastomosing and cisternal portions of GERL (GE and arrows, respectively). Acid phosphatase activity is not present in the cisternal stacks of the Golgi (G). (From P. Novikoff, with permission.)

ies) for their systematic recognition in distinct cell types; however, available information indicates that key sorting events occur in these structures (see below).

The GC of animal cells is found during interphase near the centrosome, on the apical side of the nucleus with its distal face (which is usually convex) toward the apex of the cell. The proximal face is usually closely associated with specialized areas of the RER (transitional elements) from which small smooth-surfaced vesicles appear to bud. Systematic examination of sections with the electron microscope indicates that most animal cells contain one—or certainly no more than a few—stacks of Golgi cisternae. By contrast, in plant cells and in certain insect cells many discrete small Golgi stacks are found scattered throughout the cytoplasm without obvious relation to the RER.

Subcellular fractions enriched in Golgi-derived vesicles and cisternae have been obtained from several cells and tissues. Since the membranes of the GC have a density which is similar to that of other smooth membranes (plasma membrane, outer mitochondrial membrane, etc.), it is only in the case of cells whose GC

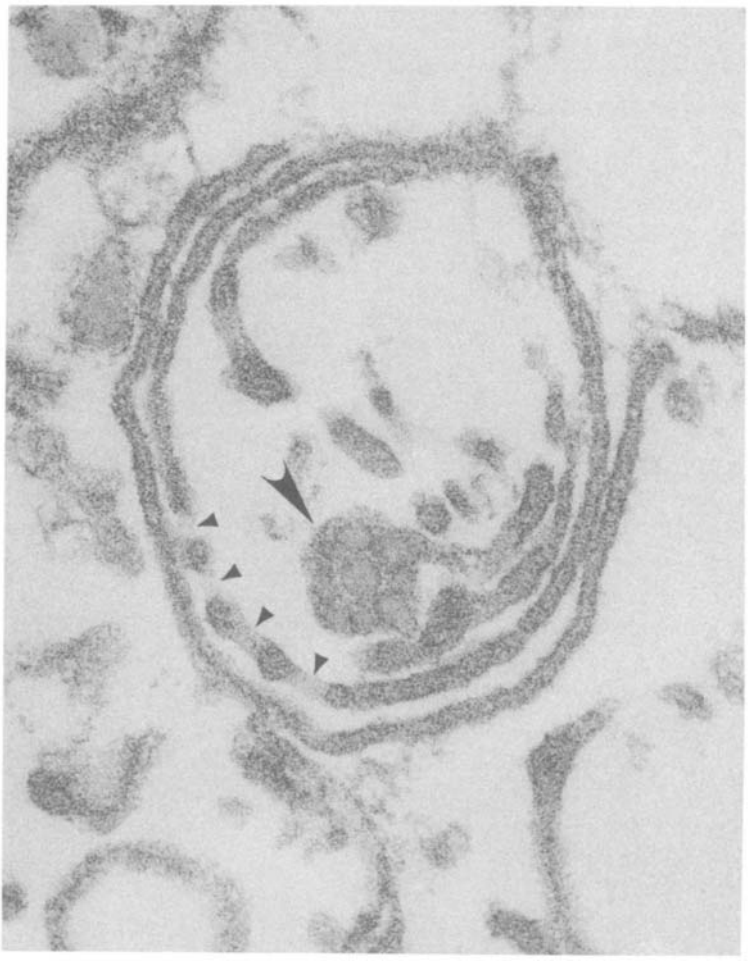


Figure 4. Thin section of an isolated "stacked" Golgi-rich fraction recovered from a rat liver homogenate. Note the abundant secretory lipoprotein content (large arrowhead). The small arrowheads indicate pores in cisternae.

contains secretory products of very low density (e.g., the abundant secretory lipoproteins of the hepatocyte) that the isolated fractions are especially pure (Figure 4). The prospect for obtaining comparably enriched fractions from other cell types is sure to improve as immunoaffinity methods are increasingly applied to isolation of organelles. In the liver, and in a few other tissues, isolated Golgi-enriched

fractions retain a characteristic stacked organization of their cisternae. It is not known what "glue" holds the individual cisternae so firmly together that they withstand the trauma of homogenization. Related materials may be important for the targeting of vesicles to, and their fusion with, Golgi cisternae.

STRUCTURE

No resident proteins are known to be present in both the membranes of the GC and the membranes of other organelles. No soluble macromolecules of the cisternal space of the GC have been identified except for macromolecules which are undergoing transport. The membrane lipids of the GC have not been studied in great detail; however, cholesterol and phospholipids are major components (cholesterol is essentially missing from the RER), with glycosphingolipids also being readily detected, possibly because this is their site of synthesis.

Although the composition of the cisternal space of the GC cannot at present be studied directly, indirect evidence suggests that the ionic composition is distinct from that of the ground substance of the cytoplasm and cisternal space of the ER. For example, when carboxylic ionophores such as monensin are added to living cells, Golgi cisternae dilate within seconds, while the RER is unaffected (Figure 5). This ionophore is known to promote equilibration of sodium, potassium, and proton concentrations across membranes. There is also cytochemical evidence suggesting that the cisternal pH of distal cisternae is somewhat acid.

It is not known why the GC has its characteristic centrosomal location in animal cells; however, an intimate relation exists between the GC and the tubulin-based cytoskeleton. Thus, when cells are treated with drugs which block tubulin polymerization, the GC fragments and becomes widely distributed throughout the cytoplasm as miniature stacks of cisternae (Figure 6). Macromolecular transport through the GC persists, although certain sorting operations may be disturbed.

A striking feature of the GC which has been intensively studied by electron microscopists is its cytochemical heterogeneity (Figure 7). For example, in many cells: 1) intense osmication produces a metallic deposit within proximal cisternae; 2) light fixation followed by incubation with substrates of nucleoside diphosphatase (uridine diphosphate, thiamine pyrophosphate) in the presence of heavy metal salts which precipitate inorganic phosphates ("capture reagents") produces a deposit in distal cisternae; and 3) incubation with substrates of acid phosphatase and comparable capture reagents produces a deposit in the GERL. These cytochemical staining patterns, which are curiously somewhat variable from one cell type and one cell to the next, were identified long before the complexity of the participation of the GC in glycosylation reactions was known. It is now clear, as mentioned below, that the enzymes responsible for glycosylation are also subcompartmentalized across the GC, and that the above-mentioned phosphatases are thought to be important for allowing glycosylation to proceed.

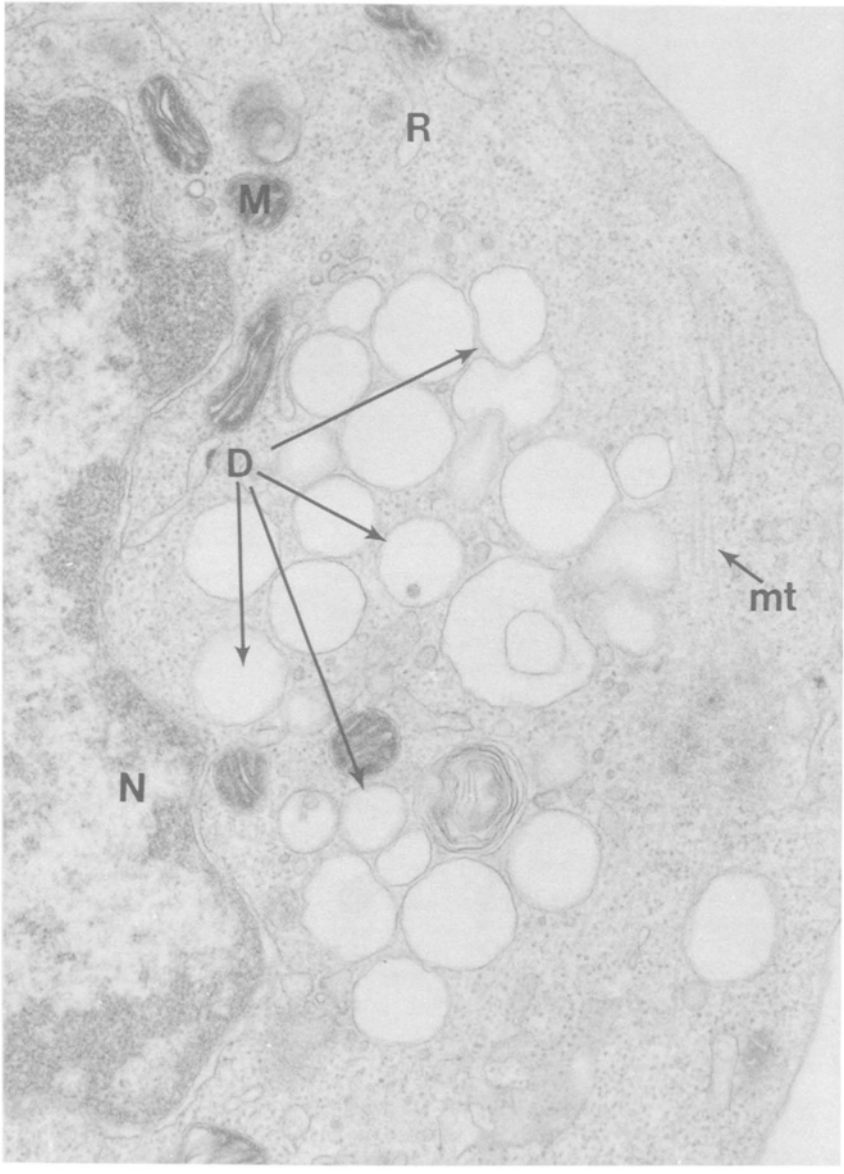


Figure 5. Thin section of a mouse thymoma cell treated for 1 hr with the carboxylic ionophore, monensin. Note the dilated Golgi elements (D), microtubules (mt), mitochondria (M), the RER (R) and the nucleus (N). The dilation is accompanied by major slowing of transport across the Golgi. Both effects are slowly reversible. (Micrograph courtesy of M. Detraz.)

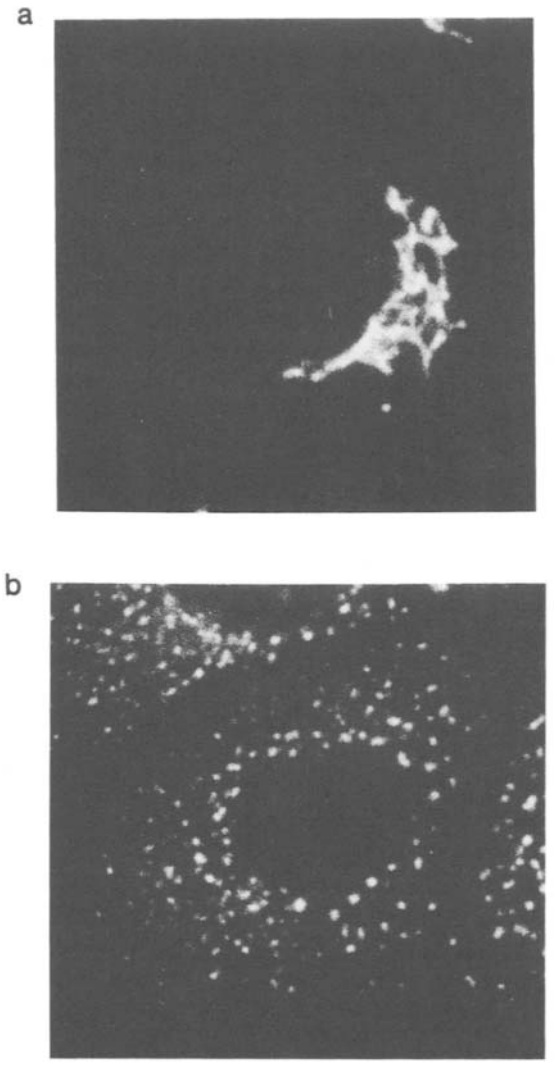


Figure 6. Depolymerization of microtubules results in scattering of the Golgi Complex. Bovine kidney cells in culture were treated with nocodazole for 2 hr at 37 °C and fixed with formaldehyde. Galactosyl transferase was revealed in both control (**a**) and nocodazole-treated (**b**) cells, using immunofluorescent reagents in the presence of detergent to permeabilize all membranes. All staining in the control is juxtanuclear. The gross dispersal of Golgi elements seen in (**b**) is totally reversible upon withdrawal of the drug.

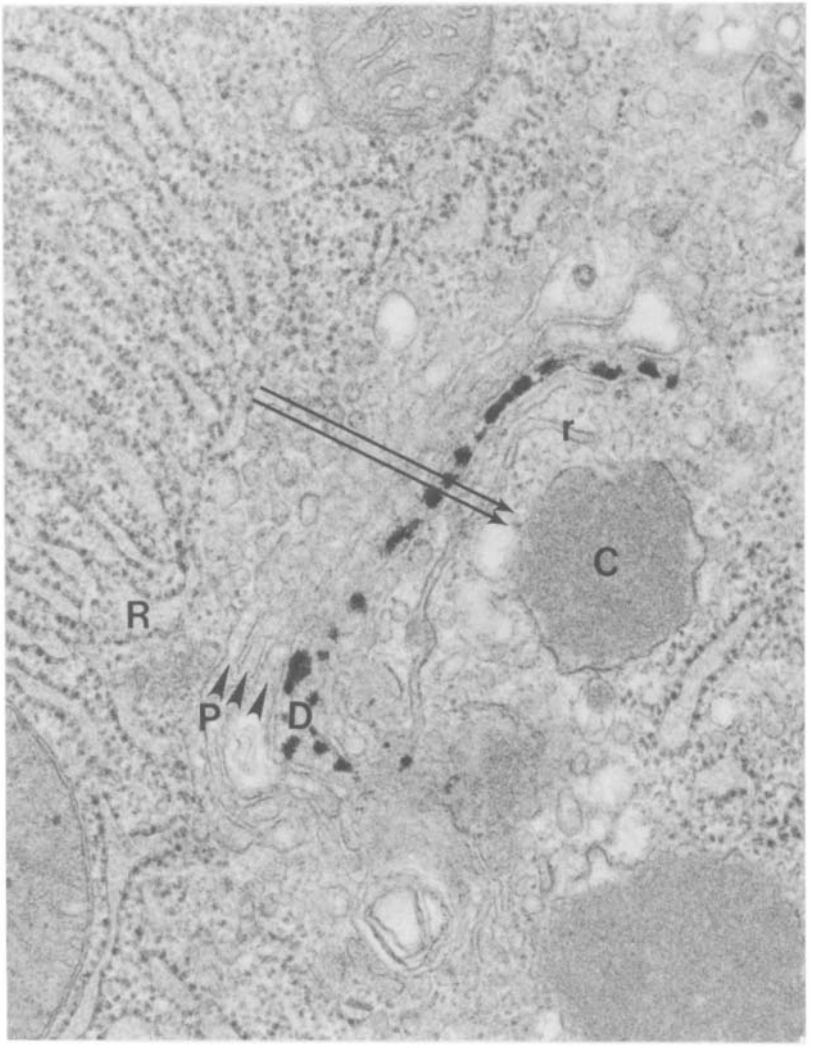


Figure 7. Thin section of the guinea pig exocrine pancreas after cytochemical detection of nucleoside diphosphatase activity. The reaction product is confined to distal cisternae (D) and appears as a black (electron-scattering) deposit. Rough endoplasmic reticulum (R), proximal Golgi cisternae (P) and condensing vacuole (C). The detection procedure involves light fixation with aldehydes, incubation with uridine diphosphate in the presence of lead nitrate (which produces the lead phosphate precipitate at sites of enzyme activity), post-fixation with osmium tetroxide and routine dehydration, embedding, sectioning and examination.

POSTTRANSLATIONAL MODIFICATIONS

Table 1 lists modifications of macromolecules which occur while they pass through the GC along the secretory path. The identified enzymes responsible for these modifications are integral membrane proteins. In a few cases genetic engineering experiments have made it possible to identify the peptide sequences which account for these enzymes residing in the GC.

The best-characterized covalent modifications are involved with glycan addition and maturation. In higher eukaryotes, especially the biosynthesis of N-glycans (whose core is added to asparagine residues in the RER) has been extensively studied. For example, typical N-glycans (Figure 8) are radically remodeled to "complex" structures due to both sugar removal by glycosidases and "terminal" sugar addition in the GC. The presence of immature *vs.* complex units can be readily judged from their differential sensitivity to endoglycosidases.

The biological roles of N-glycans can be studied by use of: a) a sugar analogue (tunicamycin) to interrupt addition of the glycans in the RER, b) sugar analogues which block glycosidase action (nojirimycins, etc.), or c) cell mutants which lack individual oligosaccharide processing activities.

Certain of the enzymes responsible for N-glycan maturation have been localized at the electron microscopic level by immunocytochemistry. In general, the anatomic distribution of the enzymes matches the order of their enzymologic function (the first of the GlcNAc transferases has been detected in medial cisternae, galactosyl and sialyl transferase have been detected in distal and postcisternal structures etc.). Nevertheless, judging from a still small number of cell types which have been studied, as with the distribution of GC phosphatases, transferase and glycosidase distributions do vary among cell types. In yeast there is a further class of oligosaccharide added to N-glycans of many secretory and membrane proteins during Golgi traversal. These are massive phosphate-containing polymannose units. Mutants are available which fail to add complete polymannose units.

For many O-glycans (which are linked to the polypeptide via the hydroxyl of serine or threonine residues), even the addition of the first sugar, N-acetyl galac-

Table 1. Posttranslational Modifications which Occur During Golgi Traversal

Trimming and elongation of N-glycans
Initiation and elongation of O-glycans
Initiation and elongation of glycosaminoglycans
Elongation of glycolipids
Sulfation of glycans
Phosphorylation of glycans
Proteolysis at dibasic amino acids
Acylation?
Concentration of content
Non-covalent addition of lipid to secretory lipoproteins

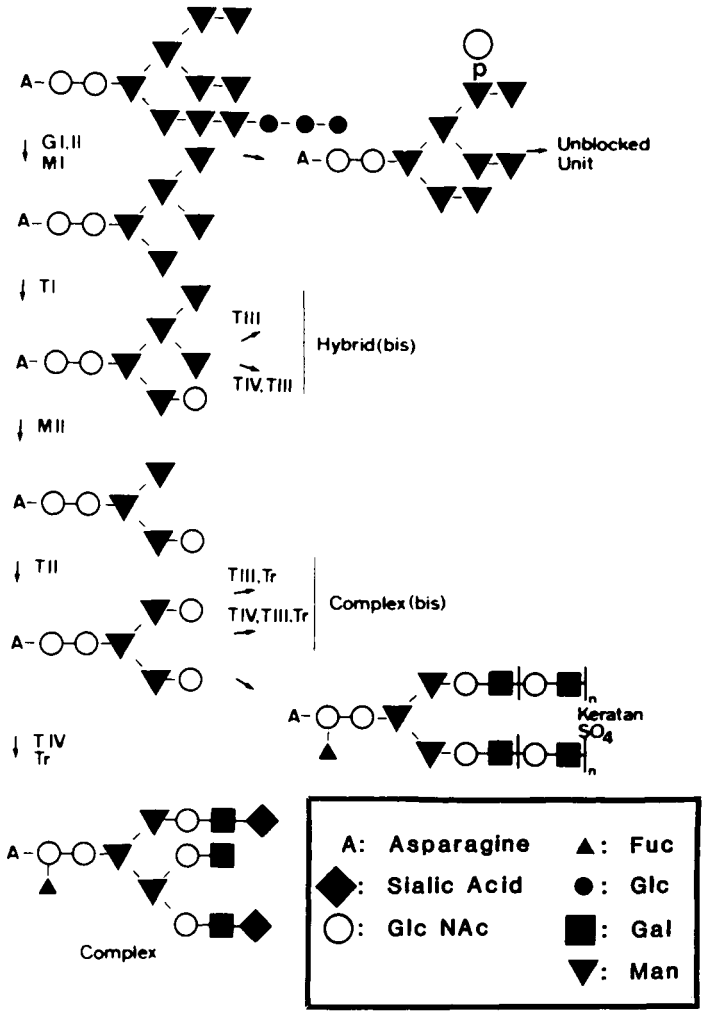


Figure 8. Structural and biosynthetic relations among asparagine-linked oligosaccharides. The parent oligosaccharide (upper left) is donated from a dolichol-linked precursor. Only a few of the processing options are indicated. G, glucosidase; M, mannosidase; T, GlcNAc transferase; Tr, terminal sugar (galactose, fucose, sialic acid) transferases. In hybrid structures, one branch retains terminal nonreducing mannose; in complex units, bi-, tri- or tetra-antennary elongation occurs; in bisected units (bis) additional GlcNAc is added to the innermost mannose. The "unblocked unit" bears the mannose-6-phosphate signal which is responsible for targeting of many acid hydrolases to lysosomes. All steps after G1, II and M1 occur in the Golgi complex. Reproduced with permission from Tartakoff, A., The Secretory and Endocytic Paths, J. Wiley Interscience, 1987.

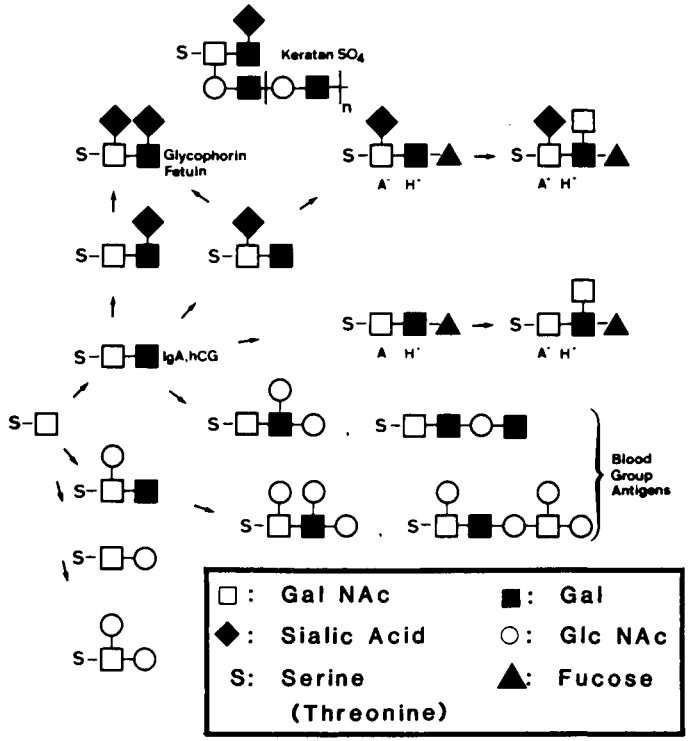


Figure 9. Structural and biosynthetic relations among N-acetylgalactosamine-linked oligosaccharides. A limited number of the many processing options are indicated. A and H blood group specificities are noted. Reproduced with permission from Tartakoff, A., The Secretory and Endocytic Paths, J. Wiley Interscience, 1987.

tosamine, occurs in the GC. Such units (Figure 9) are found on many cell surface proteins (e.g., glycophorin and the low density lipoprotein receptor) and secretory proteins (e.g., chorionogonadotrophin, mucins). The role of O-glycans is difficult to study, since specific inhibitors comparable to those used for analysis of N-glycan structure are not available. One option is to make use of CHO cell mutants which fail to add the initiating of GalNAc.

Initiation of O-linked glycan chains of glycosaminoglycans (Figure 10) is also thought to occur in the GC. In this case a characteristic tetrasaccharide composed of xylose, two residues of galactose, and glucuronic acid is added to selected serine and threonine residues. The more distally located repeating disaccharides, which, along with their polypeptide backbones, distinguish the several classes of proteoglycans, are also added during traversal of the GC. Before exit from the GC, these units undergo sulfation and, in the case of heparin sulfate, epimerization of glucuronic acid to iduronic acid. Sulfation of many O-glycans, N-glycans, and

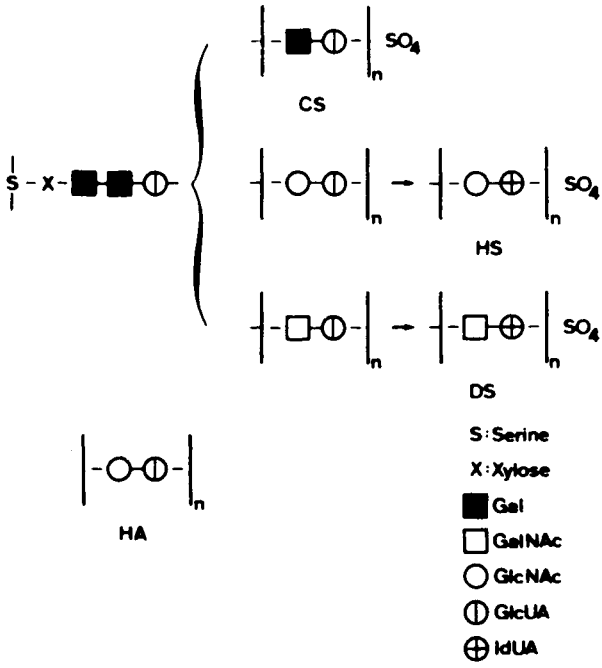


Figure 10. Structural and biosynthetic relations among xylose-linked glycosaminoglycans and hyaluronic acid. CS, chondroitin sulfate; HS, heparin sulfate; DS, dermatan sulfate; HA, hyaluronic acid. Reproduced with permission from Tartakoff, A., The Secretory and Endocytic Paths, J. Wiley Interscience, 1987.

glycolipids also occurs during Golgi traversal. As with O-glycans, the construction of glycosaminoglycan chains fails to occur in selected CHO cell mutants.

Phosphorylation of selected N-glycan chains also occurs during GC traversal. This is of known importance for the lysosomal acid hydrolases of fibroblasts and a number of other cell types. In these cases, one or more mannose residues per glycan chain (there may be many glycan chains per polypeptide and their structures may be different) acquires a blocked phosphate unit composed of phosphodiester-linked N-Acetylglucosamine. A specific phosphodiesterase then releases the terminal GlcNAc, leaving a mannose-6-phosphate unit(s). The significance of this terminal unit is that it can interact with the mannose phosphate receptors (two are known: 215,000 and 46,000 Daltons), which are concentrated in the GC. This interaction leads to the segregation of many such hydrolases from the secretory path and their delivery to lysosomes (see Figure 14B). In the fatal human disease, Mucolipidosis II ("I-Cell Syndrome"), which is characterized by major disorders of connective tissue and neurological function, the enzyme activity responsible for addition of the blocked phosphate unit is missing, massive secretion of acid

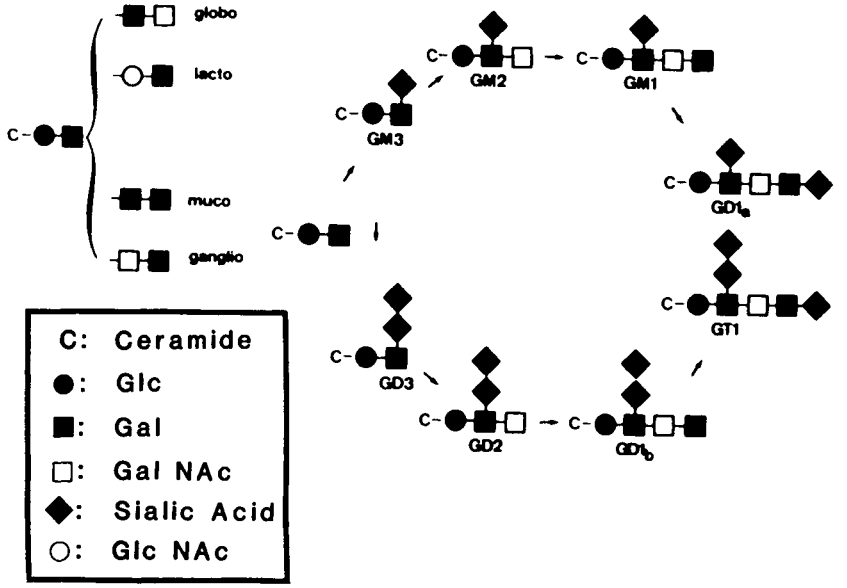


Figure 11. Structural and biosynthetic relations among simpler glycolipids. The ganglio series (characterized by sialic acid) is given in detail. Reproduced with permission from Tartakoff, A., The Secretory and Endocytic Paths, J. Wiley Interscience, 1987.

hydrolases occurs, and lysosomes become engorged by the accumulation of indigestible substrates.

The major glycolipids of animal cells are derivatives of sphingosine. Unlike phospholipids, which are built on glycerophosphate and contain two acyl chains, sphingolipids (Figure 11) bear only one acyl substituent and lack phosphate. The addition of their terminal sugars, and possibly their first sugar glucose/galactose, occurs during Golgi traversal. When the sugar residues bear sulfate or terminal sialic acid, they are known as sulfatides or gangliosides, respectively. Both residues are added during GC transit. The numerous sugar transferases in the GC function effectively because Golgi membranes are equipped with specific ports which allow entry of activated sugars (sugar nucleotides such as uridinediphospho-galactose, cytidinemonophospho-sialic acid, etc.), which are synthesized in the cytoplasm or, in the latter case, in the nucleus. These ports in fact accomplish 1-to-1 exchange of sugar nucleotides for nucleoside phosphates (uridinemonophosphate, cytidinemonophosphate, etc.), which are among the products of the sugar transferase reactions. A port has also been identified for entry of the activated sulfate donor, phosphoadenosinephosphosulfate. Additional ports presumably exist which allow exit of monosaccharides produced by glycosidase action.

Major proteolytic events occur during and after traversal of the stack of Golgi cisternae. For example, many hormones (insulin, glucagon, ACTH, etc.) are

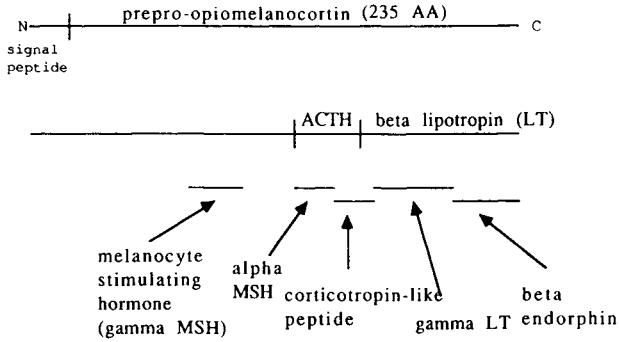


Figure 12. A dramatic example of the extent of proteolysis which can occur during transit of hormones through the Golgi Complex and condensing vacuoles. Proopiomelanocortin is synthesized in the anterior and intermediary lobes of the pituitary and undergoes extensive cleavage at sites marked by pairs of basic amino acids. The upper processing scheme (including ACTH) is characteristic of the anterior pituitary, while the lower scheme pertains to the intermediary lobe.

synthesized at the level of the RER as pro-hormones (proinsulin, proglucagon, proopiomelanocortin, etc.). While in the GC and/or condensing vacuoles they are cut by endoproteases (which have been best identified in yeast) at sites of single or (more often) paired basic amino acids (Figure 12). Exopeptidases also participate in liberating the final product.

The Golgi is responsible for the massive non-covalent addition of lipids to secretory lipoproteins; for example, the low and high density lipoproteins secreted by hepatocytes. The extent of addition of cores of cholesterol and phospholipids to these lipoproteins is so great that they become visible within Golgi cisternae after osmium fixation (Figure 4). Although the situation is not well understood, this lipid addition implies that RER and/or Golgi membranes must flip-flop the component lipids with great efficiency from the endodomain (cytoplasmic face), where lipid synthesis occurs, to the ectodomain and beyond. Apart from sugar addition to glycolipids, other steps of lipid synthesis are not known to occur in the GC.

The Golgi posttranslational modification which first caught the attention of electron microscopists is the impressive ability of distal Golgi elements to concentrate their macromolecular content to the point of producing essentially solid protein. These events of concentration are conspicuous in cells engaged in genesis of secretion granules or specialized lysosomes (endocrine and exocrine cells of the pancreas, pituitary cell types, granulocytes, etc.). The final stages of concentration occur in condensing vacuoles, downstream from Golgi cisternae themselves (Figure 2). This thermodynamically improbable event is not understood, but may involve extensive charge neutralization of secretory protein content (by small and macromolecular ions) and active ion extrusion, followed by water efflux. Judging

from indirect cytochemical measurements, the condensing vacuoles, as well as secretion granules, endosomes and lysosomes, have an acidic interior.

MECHANISMS OF TRANSPORT ALONG THE SECRETORY PATH

Golgi-associated vesicular transport along the secretory path involves: 1) receiving soluble and membrane proteins from the RER, 2) transit from proximal to distal cisternae, 3) exit to the cell surface(s), and 4) exit to lysosomes. The minimum time required for delivery of newly-synthesized proteins to their destinations is about 30 minutes; however, many proteins require as much as several hours.

Passage of RER content and membrane proteins to the GC is thought to be mediated by smooth surfaced "transit vesicles" which bud from the transitional elements of the RER. This event requires ongoing ATP production, GTP and a temperature in excess of 10 °C, but is not affected by inhibitors of protein synthesis. Although direct evidence is lacking, it is likely that a vesicular shuttle is involved and that "empty" vesicles return repeatedly from the GC to encapsulate and transport successive quanta of RER materials. This return leg of transport requires intact microtubules. In the presence of brefeldin A it rapidly causes Golgi membranes to relocate to the RER.

Exit from the RER is selective. Many RER membrane proteins (such as those involved in protein synthesis and drug detoxification), as well as certain proteins which appear to be soluble in the RER content, do not exit efficiently. These include the heavy chain binding protein (also known as BiP) which associates with free heavy chains of immunoglobulin and other incompletely folded polypeptides. Moreover, proteins do not exit at a uniform rate: half-times for RER exit vary from minutes to hours. It has been argued that all proteins exit except for those that bear retention signals. A deeper analysis of RER exit should be forthcoming with the availability of *in vitro* models (cell-free or using perforated cells) as well as multiple yeast mutants in which exit is reversibly blocked at 37 °C.

The mechanism of transit from proximal to distal cisternae has been much discussed. A trivial explanation—direct cisternal continuity—is ruled out by electron microscopy. Two further models (Figure 13) which are able to account for transit are: 1) a "dissociative" model in which vesicles bud from and target hierarchically to successive cisternae, and 2) a model of "cisternal progression," according to which entire cisternae move in a proximal-to-distal direction. The first model has received considerable experimental support from both cell fusion studies and from cell-free analysis of N-glycan maturation of membrane glycoproteins in transit across the Golgi stack. For the cell-free studies, pairs of stacked Golgi fractions are mixed with appropriate supernatant factors. A set of soluble and membrane-associated proteins (coatamers) have been identified and implicated in vesicle formation, targeting, and fusion. This model requires an impressive degree of precision of vesicle targeting in successive rounds of membrane fission and

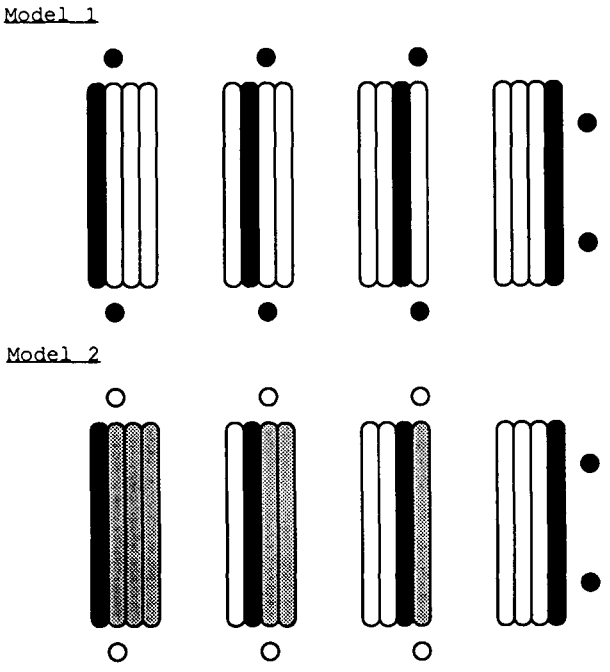


Figure 13. Two models which seek to explain how secretory protein content passes from the proximal face (left) of the Golgi stack to the distal face (right). Each model is illustrated as a sequence of time frames to describe the passage of a cohort of secretory content across the Golgi stack. In Model 1, passage from one cisterna to the next is accomplished by small vesicles which depart from a given cisternae carrying quanta of content and target specifically to the ext cisternae. In Model 2, the entire cisternae move in a proximal-to-distal direction, finally releasing vesicles at the distal face. According to this model, proximal cisternae must continually be regenerated. These newly-created cisternae are white.

fusion. The budding and targeting is thought to occur at the rims (extremities) of cisternae. The second model has the advantage of avoiding this issue of exquisite vesicle targeting specificities, but has difficulty accounting for the progressive modification of the composition of successive cisternae (see above) and must postulate that the most proximal cisternae are continually being regenerated. The strongest evidence for this model comes from study of a specialized (and possibly atypical) system: the genesis of surface scales by certain marine algae. The scales which mature within Golgi cisternae are so large that they could not be encapsulated by small transit vesicles.

Transit across the GC requires ATP, GTP, and fatty acyl CoA and proceeds effectively only above 20 °C. It is dramatically slowed, for unknown reasons, by the carboxylic ionophore, monensin (See Figure 5).

MACROMOLECULAR SORTING

Figure 14 gives an overview of the macromolecular sorting options which are executed by the GC.

Exit from the GC of constitutively secreted proteins and membrane proteins which are constitutively delivered to the plasma membrane occurs after their transit through the most distal cisternae, which house the terminal sugar transferases, judging from the extent of maturation of their N-glycans. The putative carrier vesicles responsible for their delivery to the cell surface have been only partly characterized.

Secretory glycoproteins which are massively stored in secretion granules also bear "matured" N-glycans. As mentioned above, their concentration begins in distal Golgi cisternae and is completed as condensing vacuoles are converted to secretion granules. In the case of cells which produce secretion granules, both constitutively secreted proteins and those destined for storage may be present in condensing vacuoles, but constitutively transported proteins are not stored in secretion granules. As with RER exit, it is thought that a given protein will be constitutively secreted unless it bears a critical (yet undefined) signal which leads to its being retained and concentrated. This sorting decision is upset by agents such as chloroquine or ammonium chloride which dissipate the relative acidity of membrane-bounded compartments. It is therefore thought that the acidity of condensing vacuoles is essential.

In the best-studied animal cells, the exit of acid hydrolases from the GC occurs after addition of the mannose-phosphate units (described above) and involves mannose-phosphate receptors. Subcellular fractionation studies of the liver indicate that the enzymes responsible for mannose-phosphate addition are in relatively proximal Golgi cisternae. Immunocytochemical studies of the larger receptor show that its sub-GC location varies from one cell type to the next and that it can be detected in endosomes but not in lysosomes. It is therefore thought that the cytologic site of acid hydrolase exit from the GC may be variable and that hydrolase arrival in lysosomes follows passage through endosomes. As with many cell surface receptors, the affinity of the mannose-phosphate receptors for mannose-phosphate bearing ligands is high at neutral pH and low at acidic pH. Thus, the acid hydrolase ligand should be released upon arrival in the endosome.

This entire trajectory of acid hydrolase delivery is upset in I-Cell Disease, since mannose-phosphate units are not added to the hydrolases (see above). The result is that they are secreted.

In selected cell types, such as the hepatocyte and in yeast, accurate targeting of acid hydrolases can be accomplished in the absence of mannose phosphate units. Polypeptide determinants are presumably responsible. This is also true for lysosomal membrane proteins, which do not bear mannose-phosphate units.

Constitutive exit to the cell surface may lead to more than a single plasma membrane domain. In hepatocytes and other polarized epithelial cells (e.g., kidney

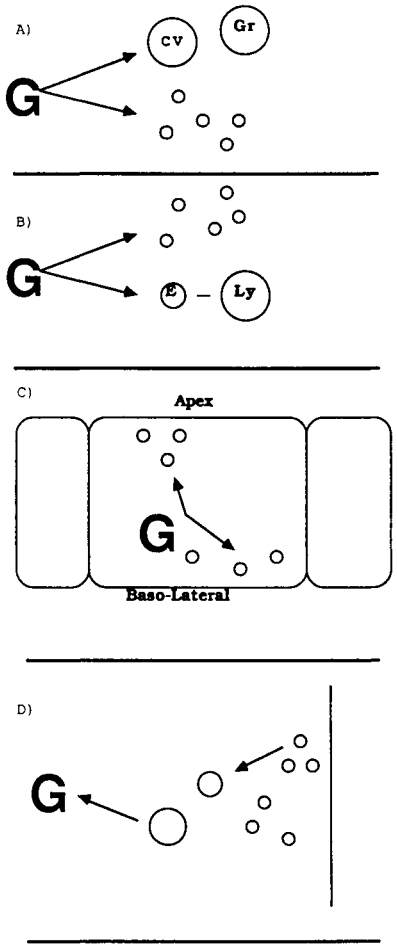


Figure 14. Enumeration of multiple sorting/transport options accomplished by the Golgi Complex. **A)** Delivery of secretory products to both constitutively discharged secretion vesicles and, via condensing vacuoles (CV), to secretion granules (Gr). **B)** Delivery of products to both secretion vesicles and to lysosomes (Ly) via endosomes (E). For simplicity, only one of the secretory options indicated in (**A**) is included. **C)** Delivery of membrane or secretory proteins to both the apical and baso-lateral surfaces of polarized cells. **D)** Receipt of vesicular traffic from the cell surface(s), possibly followed by return to the cell surface.

or intestinal epithelial cells), the functions and composition of the apical plasma membrane are distinctly different from those of the basolateral surface. These differences are critical for many functions of epithelial cells; for example in transepithelial ion transport. Much effort has been devoted to elucidating the path(s) taken by membrane glycoproteins and glycolipids as they exit from post-GC

elements toward these two domains. For experimental simplicity, much of the analysis has been based on following viral envelope glycoproteins (of Vesicular Stomatitis Virus, Semliki Forest Virus, Influenza Virus, etc.). The site at which sorting of apical vs. basolateral membrane proteins occurs varies according to cell type: in kidney epithelial cells, the VSV glycoprotein (G) proceeds directly from the GC to the basolateral surface (where this virus buds) and the hemagglutinin of Influenza proceeds directly from the GC to the apical surface (where this virus buds). In hepatocytes, by contrast, all membrane and secretory proteins appear to move from the GC to the sinusoidal surface (equivalent of the basolateral surface), from which some membrane proteins subsequently pass to the bile canalicular front.

GOLGI FUNCTIONS IN ENDOCYTOSIS

Morphologic studies of endocytosis of some (but certainly not all) soluble tracers and biochemical studies of labeled cell surface glycoproteins have documented a route of vesicular transport from the cell surface to Golgi cisternae (See Figure 14). The data indicate that this endocytic route is followed less often than transit from the cell surface to lysosomes or return to the cell surface (diacytosis); nevertheless, cumulatively, the plasma membrane-to-Golgi transport of both proteins and lipids may account for a large fraction of GC membrane traffic and be responsible for important remodeling of surface glycoconjugates. The diacytotic route is believed to be essential for antigen presentation by cells of the immune system and for recycling of cell surface receptors.

RECOMMENDED READINGS

Balch, W. (1990). Small GTP-binding proteins in vesicular transport. TIBS 15, 473–477.

Bomsell, M. & Mostov, K. (1992). Role of heterotrimeric G proteins in membrane traffic. Mol. Biol. of the Cell 3, 1317–1328.

Bowen, R. (1928). The methods for demonstration of the Golgi apparatus, I–IV. Anat. Rec. 28, 2983; 39, 231; 39, 365; 40, 103.

Farquhar, M. & Palade, G. (1981). The Golgi apparatus (Complex)-1954–1981-from artifact to center stage. J. Cell Biol. 91, 77s–106s.

Inferrerra, C. & Carrozza, G. (1975). The discovery of the Golgi apparatus by the black reaction and its present fine structural visualization. In Golgi Centennial Proceedings (M. Santini, ed.), pp. 13–38, Raven Press, New York.

Kelly, R. (1985). Pathways of protein secretion in eukaryotes. Science 230, 25–32.

Kornfeld, S. (1986). Trafficking of lysosomal enzymes in normal and disease states. J. Clin. Invest. 77, 1–6.

Rothman, J. & Orci, L. (1992). Molecular dissection of the secretory pathway. Nature 355, 409–415.

Simons, K. & Fuller, S. (1985). Cell surface polarity in epithelia. Ann. Rev. Cell Biol. 1, 243–288.

Simons, K. & van Meer, G. (1988). Lipid sorting in epithelial cells. Biochemistry 27, 6197–6202.

Tartakoff, A. (1983). Perturbation of vesicular traffic with the carboxylic ionophore monensin. Cell 32, 1026–1028.

Tartakoff, A. (1987). The secretory and endocytic paths, J. Wiley Interscience, New York.

Whaley, G. (1975). The Golgi Apparatus, Springer-Verlag, New York.

Chapter 3

The Peroxisome

COLIN MASTERS and DENIS CRANE

INTRODUCTION

Peroxisomes are cellular organelles which are widely distributed in eukaryotic cells and play a major role in oxidative metabolism. These organelles are characterized by their content of catalase and several oxidases, and as the name was intended to

Principles of Medical Biology, Volume 3
Cellular Organelles and the Extracellular Matrix, pages 39–58

suggest, the functions of this particle generally involve the production and degradation of hydrogen peroxide.

Because of this broad specification, an overlap of terminology is possible between such commonly applied descriptions of subcellular structures as microbodies, peroxisomes, microperoxisomes, and glyoxysomes. However, de Duve (1969) has made it clear that the term microbody is intended to describe the morphological appearance only, and the term glyoxysome is restrictive in that it refers to a special class of peroxisomes in plants that stress the metabolism of glyoxylate. Also microperoxisomes are now considered to be *bona fide* peroxisomes so that the current usage tends to use the term peroxisome to cover all these descriptions.

Peroxisomes carry out a variety of metabolic functions which vary with cell type. In animal cells, these functions include respiration based on the hydrogen peroxide-forming oxidases and catalase, fatty acid oxidation, plasmalogen biosynthesis, alcohol oxidation, transaminations and the metabolism of purines, polyamines, bile acids and other substrates.

In addition to these important metabolic roles in mammalian tissues, it should be noted that peroxisomes also play significant roles in both catabolic and anabolic pathways of plants (Tolbert, 1971). These additional roles are possible because of the distinctive enzyme compositions in these different life forms. In seeds rich in lipids, peroxisomes (glyoxysomes) are the sites for the breakdown of fatty acids to succinate, via the glyoxylate cycle, and participate in gluconeogenesis in this way. In leaf tissues, peroxisomes serve as sites of photorespiration in the leaf cell. This process involves the oxidation of glycolic acid (a product of photosynthetic CO_2 fixation) to CO_2 and H_2O_2.

Unlike mitochondria, peroxisomes have no separate capacity for the synthesis of their protein components, and no energy-coupled electron transport system. Electron shuttles, however, may provide a system for the oxidation of reduced pyridine nucleotides in this organelle.

MORPHOLOGY

In cell sections, peroxisomes generally appear as spherical organelles that are bounded by a single trilaminar membrane. They contain matrix material that appears slightly more dense than the cytoplasm, and vary in their abundance from hundreds per cell in tissues like liver and kidney, to one or two in other cell types.

The presence of an electron-dense central core, also called a nucleoid or crystalloid, came to be quoted as the most distinctive structural feature of these organelles. The core was later shown to be a polytubular structure which, in most cases, is closely related to the occurrence of urate oxidase.

The limiting membrane of peroxisomes is a single membrane (Figure 1), consisting of a triple-layered structure of varying thickness, from 4.5 nm to 6–8 nm. The peroxisomal membrane was found to be thinner than that of lysosomes,

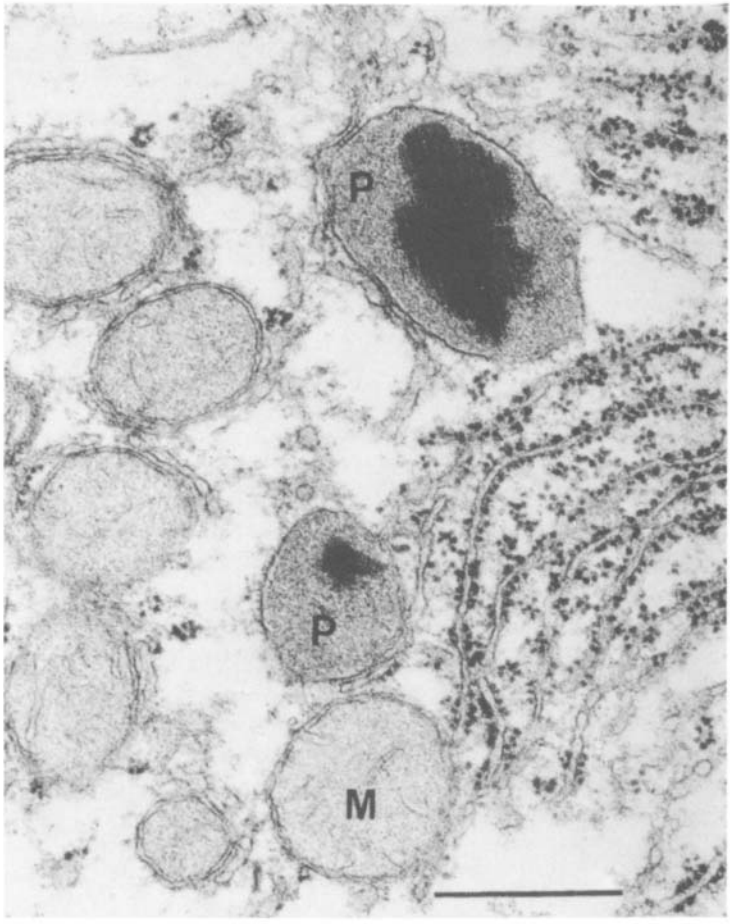


Figure 1. Electron micrograph of peroxisomes (**P**) in rat liver (X 50,000). Note the single limiting membrane, the electron-dense crystalline core, and the finely granular matrix in each case. **M**, mitochondrion. In the right of the figure are shown membranes of the endoplasmic reticulum, with associated ribosomes. The bar represents 0.5 μm. This print was kindly contributed by J. L. Hughes, Adelaide Children's Hospital.

the plasma membrane, and most other single-membrane-bound cell inclusions, but is of comparable thickness to the endoplasmic reticulum (see Masters and Holmes, 1977).

In some species of animals, a marginal plate has been described in peroxisomes. This plate resembles a flat, uniformly thick, straight, or slightly bent structure and is located at the periphery of the peroxisome, separated from the inner surface of

the membrane by a narrow space of lower density than the surrounding matrix. The biological function of these plates is unknown still, although it is now established that the enzyme L-α-hydroxyacid oxidase B is localized in these structures.

Numerous early studies reported that peroxisomes may exhibit direct connections with the endoplasmic reticulum (ER). The presence of membranous projections from peroxisomes, and continuities between the membranes of peroxisomes and ER in several species were reported. Later studies have thrown doubt on the significance of these membranous connections and protrusions in regard to the mechanisms of biogenesis of these organelles (see later discussion on peroxisome biogenesis). However, the possibility that these structures represent potential sites of metabolic interaction between these two organellar compartments is still a consideration.

In some tissues, serial sectioning of peroxisomes has shown a remarkable elongation and branching of tubular peroxisomal profiles. The complexity and morphological diversity of these structures, along with the changes during differentiation, have contributed significantly towards the recent view that peroxisomes may occur as a peroxisomal reticulum rather than as discrete globular organelles (Lazarow and Fujiki, 1985). This matter also bears significantly on concepts of peroxisome biogenesis, and is referred to again in that section.

MAJOR BIOCHEMICAL COMPONENTS

In seeking to define the biochemical characteristics of peroxisomes, comment needs to be directed first towards a characterization of the individual enzymes in these organelles, and some of the major enzymic components discovered to date are listed in the accompanying table (Table 1).

Catalase

The enzyme catalase is widely distributed in a variety of life forms, animal, plant, and bacterial, and is widely regarded as the single component most closely synonymous with the peroxisome. Peroxisomal catalase is localized for the main part in the peroxisomal matrix, but significant amounts of this activity also occur in the cytoplasm of many mammalian species.

Catalase may perform both catalatic and peroxidatic functions with hydrogen peroxide, depending on conditions, as follows:

$$\text{catalatic reaction: } H_2O_2 + H_2O_2 \rightarrow 2H_2O + O_2$$

$$\text{peroxidatic reaction: } H_2O_2 + RH_2 \rightarrow 2H_2O + R$$

Both reactions begin with the formation of a primary complex between hydrogen peroxide and the iron of the hematin prosthetic group, which either reacts with another mole of hydrogen peroxide for the catalatic reaction, or a hydrogen donor

Table 1. Enzymes in Hepatic Peroxisomes

Enzyme	*Substrate*
Catalase	H_2O_2 (catalatic); ethanol, methanol, formate, nitrite (peroxidatic)
Oxidases (H_2O_2-generating)	
α-hydroxyacid	α-hydroxyacids
Urate	Uric acid
D-Amino acid	D-Amino acids
Fatty acyl-CoA	Long-chain fatty acyl-CoA derivatives (> 10C)
Polyamine	Spermidine and spermine
Pipecolic acid	Pipecolic acid
Acyl transferases	
Carnitine acetyl-CoA	Acetyl-CoA
Carnitine octanoyl-CoA	Octanoyl-CoA (8C)
Acyl-CoA:DHAP	DHAP, palmitoyl-CoA
Dehydrogenases	
Glycerol-phosphate	Glycerol-phosphate
Isocitrate	Isocitrate
3-hydroxy fatty acyl-CoA	3-hydroxy fatty acyl-CoA
Glucose 6-Phosphate	Glucose 6-Phosphate
Other	
Enoyl-CoA hydratase	Crotonyl–CoA, enoyl– CoA (6-16C)
Thiolase	3-oxo-fatty acyl-CoA (6-12C)
Fatty acyl-CoA synthetases	Long- and very-long-chain fatty acids
Alkyl-DHAP synthase	Acyl-DHAP, alcohol
Hydroxymethylglutaryl-CoA reductase	Hydroxymethylglutaryl-CoA
Epoxide hydrolase	Arene and alkene oxides
Alanine: glyoxylate aminotransferase	Alanine, glyoxylate

for the peroxidatic reaction. The catalatic activity predominates when hydrogen peroxide is present in high concentrations, while at low concentrations of hydrogen peroxide alternative hydrogen donors may be used as substrate in peroxidatic activities. The chemical nature of the hydrogen donor for the latter reaction may vary greatly and may include such compounds as phenols, formate, alcohols, nitrite, and primary amines.

Catalase from various animal sources has been shown to consist of four identical subunits with a molecular weight of approximately 60,000. Thus native catalase is a tetrameric molecule of 240 kD, and contains four heme prosthetic groups.

A number of studies provide evidence for the existence of multiple forms of catalase in mammalian organisms, and the well-established tetrameric structure of catalase and the single genetic locus for the enzyme (Masters and Crane, 1992) argue strongly for the differences between catalase molecules due to epigenetic modifications of the individual subunits. Mechanisms that have been put forward in this regard include, for example, the attachment or removal of sialic acid residues, which may give rise to a set of five multiple forms because of the tetrameric structure of the enzyme (Masters et al., 1986).

L-α-Hydroxyacid Oxidase

L-α-Hydroxyacid oxidase is a flavin-containing protein that catalyzes the oxidation of L-α-hydroxyacids to produce the corresponding α-ketoacids and hydrogen peroxide:

This enzyme exists as two isozymic forms in mammalian tissues (Masters and Holmes, 1977). The isozyme HAOX-A preferentially oxidizes short-chain aliphatic hydroxyacids and exhibits no activity with aromatic hydroxyacids and is sometimes referred to as glycolate oxidase, because it exhibits maximal activity with this substrate. In contrast, HAOX-B exhibits no activity with glycolate, but oxidizes a variety of other L-α-hydroxyacids including aromatic substrates.

Investigations with purified isozymes of hydroxyacid oxidase from rat liver (A) and kidney (B) have shown similarities in molecular sizes, coenzyme contents, and biochemical properties (Duley and Holmes, 1976). Sodium dodecyl sulfate– (SDS) gel electrophoretic studies yielded subunit sizes of 43,000 and 40,000 for the A and B isozymes, respectively, whereas spectral studies revealed one flavin residue per polypeptide subunit for each isozyme. Amino acid composition studies show that the isozymes are similar in this respect, too, and immunological studies have confirmed a degree of sequence homology between these isozymes, so that the A and B loci encoding these isozymes are indicated as having a common evolutionary origin.

D-Amino Acid Oxidase

D-amino acid oxidase is a flavoprotein, with a prosthetic group of flavin adenine dinucleotide (FAD), which catalyzes the oxidative deamination of D-amino acids. D-amino acid oxidase exhibits an absolute stereospecificity for the unnatural D-enantiomorph, but has a broad specificity for different amino acids.

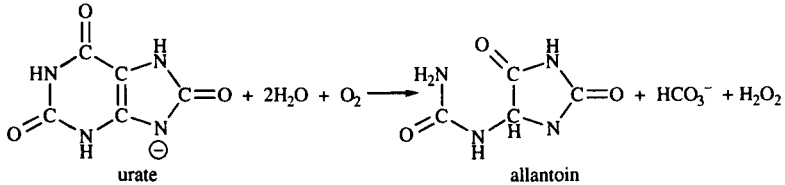
D-Amino acid oxidase is widely distributed among vertebrate species, but has been found to be localized predominantly in kidney tissue and to a lesser extent in liver. The enzyme has been reported to occur in both the cytosol and peroxisomal fractions of these tissues. Within the peroxisome, D-amino acid oxidase is apparently contained in a "compartment" of the peroxisome that is resistant to solubilization (see Masters and Crane, 1995).

The presence of an enzyme that utilizes D-amino acids has presented problems in the assessment of its biological role. As is well known, D-amino acids are neither synthesized by nor incorporated into mammalian cells. The biological activity of D-amino acid oxidase *in vivo* has been demonstrated convincingly, however.

Substitution of D-amino acids for their natural L-enantiomorphs in the diets of rats readily promotes growth, and the conversion of amino acids to the L form has been shown to occur by the oxidative deamination of D-amino acids followed by reamination. D-pipecolic acid is among the proposed physiological substrates of this enzyme. Additionally, it has been suggested that glycolate adducts of thiols may be physiological substrates (Hamilton, 1985).

Urate Oxidase

Urate oxidase (also called uricase) catalyzes the oxidation of urate, with molecular oxygen as the only known oxygen acceptor. In alkaline solution the following overall reaction has been demonstrated:



urate allantoin

This represents the final step in the metabolic degradation of the purine ring in many mammalian species, resulting in the formation of the excretory product allantoin.

The native molecular weight of pig liver urate oxidase has been determined to be 125,000, with the enzyme consisting of four identical subunits. Intracellular distribution studies have shown urate oxidase to be localized exclusively within peroxisomes, being predominantly associated with the nucleoid or crystalline core of the organelle.

Comparative studies on the distribution and properties of urate oxidase are of considerable phylogenetic interest. It has long been known that man and his immediate evolutionary ancestors, the anthropoid apes, are devoid of urate oxidase activity and exhibit high uric acid levels in serum and urine, and recent investigations have extended this condition to several genera of New World monkeys. More primitive primates such as the Old World monkeys and the prosimians, have low levels of uric acid in their body fluids and exhibit significant liver urate oxidase activities. Moreover, these enzyme activities differ in their stabilities, with the phylogenetically more recent animals having an unstable form of this enzyme. Thus, a gradual degeneration of urate oxidase during the phylogenesis of primates has been proposed.

β-Oxidation Enzymes

Although a β-oxidation pathway had been known to occur in plant peroxisomes for some years, the demonstration of peroxisomal β-oxidation in mammals by Lazarow and de Duve (1976) profoundly influenced the biological views of fatty

acid degradation in vertebrate animals. Peroxisomes have been shown to possess an active β-oxidation system which is capable of oxidizing both very-long (> C22) and long-chain (C16–C22) fatty acids. Short- and medium-chain fatty acids (i.e., less than twelve carbons) are not oxidized by the peroxisomes, but may be rapidly utilized by the mitochondria; hence setting up an interesting metabolic interplay between these organelles (see later section).

The enzymes of the peroxisomal β-oxidation spiral are illustrated in the accompanying diagram (Figure 2). They differ from those in the mitochondria in that the first step in the peroxisomal spiral is catalyzed by a flavoprotein acyl-CoA oxidase, which in addition to producing an α,β-unsaturated acyl-CoA as a product, also forms hydrogen peroxide.

The unsaturated acyl-CoA is converted to β-hydroxyacyl-CoA by an enoyl-CoA hydratase and this substrate is further oxidized by a β-hydroxyacyl-CoA dehydrogenase. These last two enzyme activities are associated in a single bifunctional protein in the peroxisome, in contrast to the two separate proteins of the mitochondrial situation. The oxoacyl-CoA is, in turn, acted upon by thiolase to produce a fatty acyl-CoA plus acetyl-CoA.

The products of oxidation in the peroxisome (medium-chain fatty acyl-CoA and acetyl-CoA) may interact with their respective carnitine acyltransferases to be transported out of the peroxisome into the cytoplasm, where they are available for further oxidation in the mitochondria (the medium chain acyl-CoA) or the formation of ketone bodies or the synthesis of complex lipids (acetyl-CoA).

The acyl-CoA oxidase consists of three subunits (A, B, and C) with respective sizes of 72, 52, and 21 kD. It has been proposed that the enzyme is synthesized as the larger subunit (72 kD) with the greater proportion of these molecules being cleaved into the smaller B and C subunits, with the result that a number of isozymes may be formed (e.g., A_2, ABC, and B_2C_2).

The bifunctional protein, which expresses both the enoyl-CoA hydratase and 3-hydroxyacyl-CoA dehydrogenase activities, has been shown to possess a molecular weight of about 78 kD, while the peroxisomal thiolase is a dimeric enzyme of 89 kD.

In regard to the oxidation of unsaturated fatty acids, a number of auxiliary enzymes are required, in addition to those already listed. A 2,4-dienoyl-CoA reductase is required for the complete peroxisomal degradation of unsaturated fatty acids with cis-double bonds on even-numbered carbon atoms. This peroxisomal reductase is a distinct protein from the mitochondrial enzyme. Similarly, an enoyl-CoA isomerase and a 3-hydroxyacyl-CoA epimerase may also be required for the peroxisomal oxidation of unsaturated fatty acids, but firm information on the molecular and catalytic properties of these enzymes is not yet available. Clearly, however, the peroxisome has a superior ability to oxidize many long-chain saturated and unsaturated fatty acids (see Masters and Crane, 1992, 1995).

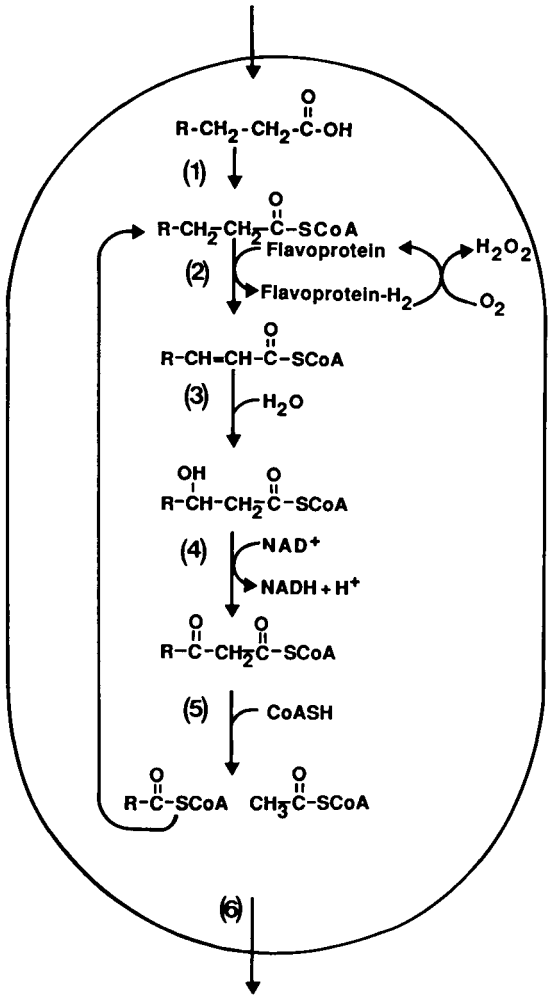


Figure 2. The peroxisomal β-oxidation system. Fatty acids enter the peroxisome (top of diagram) and are converted to their CoA derivatives. These derivatives undergo β-oxidation as indicated, with the release of a 2-carbon product in the form of acetyl-CoA. The other product, a chain shortened fatty acyl-CoA, is able to recycle through the pathway. The numbers in the diagram represent the enzymes which participate in this pathway: (1) fatty acyl-CoA synthetase; (2) fatty acyl-CoA oxidase; (3) enoyl-CoA hydratase; (4) hydroxyacyl-CoA dehydrogenase, and (5) thiolase. The products of this pathway are also able to be transported out of the peroxisome (bottom of diagram) by way of short-chain carnitine acyltransferases (enzyme 6).

A final point of difference between β-oxidation in the peroxisomes and mitochondria is the induciblity of the peroxisomal β-oxidation system (see section on peroxisome proliferators), whereas the mitochondrial system is constitutive (i.e., not modified in activity by dietary treatments).

The Peroxisomal Membrane

Knowledge of the properties of the peroxisomal membrane is also pivotal to the understanding of this organelle's function. The membrane separates the peroxisomal enzymes and peroxisomal metabolites from the cytosol, and acts to define the peroxisomal interior as a distinct intracellular space. Hence the permeability properties of the peroxisomal membrane determine in large part the extent to which the peroxisome functions as a separate metabolic compartment.

The membrane of isolated peroxisomes had been earlier suggested to be quite porous, allowing free passage of a number of small molecules such as sucrose, lactate, and amino acids (de Duve and Baudhuin, 1966). The molecular basis of this porosity is not clear, although Mannaerts and van Veldhoven (1987) have suggested that the permeabilizing activity resides within a 22 kDa integral membrane protein. More recent data, however, indicating a membrane potential across the peroxisomal membrane of yeast cells, is seemingly inconsistent with this concept of a ready porosity of the membrane. These opposing findings may be a reflection on the problems inherent in studying peroxisomes *in vitro* (which may have been damaged by the procedures of preparation), and point to the need for a reexamination of some of these issues.

As can be seen from the accompanying table (Table 2), the peroxisomal membrane of liver is involved in specific reactions of lipid metabolism. For example, the membrane contains acyl-CoA synthetase, which initiates β-oxidation of fatty acids via the activation of fatty acids to CoA derivatives, as well as DHAP acyltransferase, the first enzyme of ether lipid synthesis (Lazarow and Fujiki, 1985). In addition, the reported presence of small numbers of electron transport enzymes such as cytochrome b_5, and associated reductases, has caused speculation

Table 2. Membrane Components of Hepatic Peroxisomes

Constituent	*Location*
Acyl-CoA synthetase	outer side
Cytochrome b_5	outer side
NADH-cytochrome b_5 reductase	outer side
Acyl-CoA reductase	outer side
DHAP-acyltransferase	inner side
Alkyl-DHAP synthase	inner side
22-,35-,68-,70- kD proteins	membrane proteins
phosphatidylcholine, phosphatidylethanolamine	major phospholipid components (PC > PE)

that these proteins might participate in disposing of the reducing equivalents generated during β-oxidation.

The phospholipid composition of the peroxisomal membrane is also worthy of comment. Several authors have noted the relatively low phospholipid/protein ratios of peroxisomes in comparison to other organelles (Crane and Masters, 1986). The composition of the peroxisomal phospholipids is similar to that of the ER, consisting mainly of phosphatidylcholine and phosphatidylethanolamine. Cholesterol is present at only very low levels.

PEROXISOMAL METABOLISM

Scope of Metabolism

The peroxisome is emerging in its own right as an important site of carbohydrate, lipid, and nitrogen metabolism; this organelle would appear to possess significant regulatory functions as well (Masters and Crane, 1992). For some decades, it has commonly been envisaged that peroxisomes are involved in the production and degradation of hydrogen peroxide—the peroxisomal oxidases being active in the formation of this metabolite and catalase in its destruction. These relationships are illustrated in Figure 3 and indicate that fatty acids, D-amino acids, hydroxy acids, and uric acid are common substrates for the peroxisomal oxidases, and that catalase

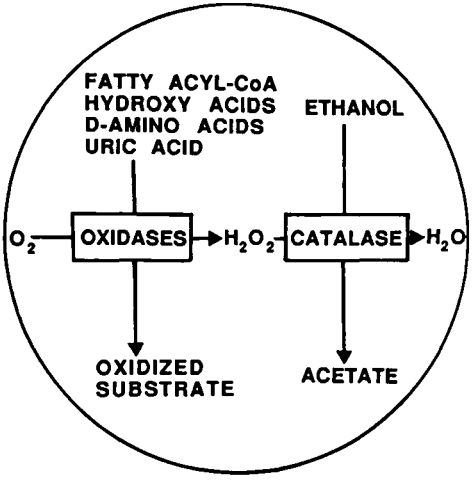


Figure 3. Peroxisomal metabolism. Substrates of the peroxisomal oxidases may be metabolized with the production of hydrogen peroxide. Hydrogen peroxide, in turn, may be catalytically decomposed, with or without the attendant oxidation of substrates such as ethanol.

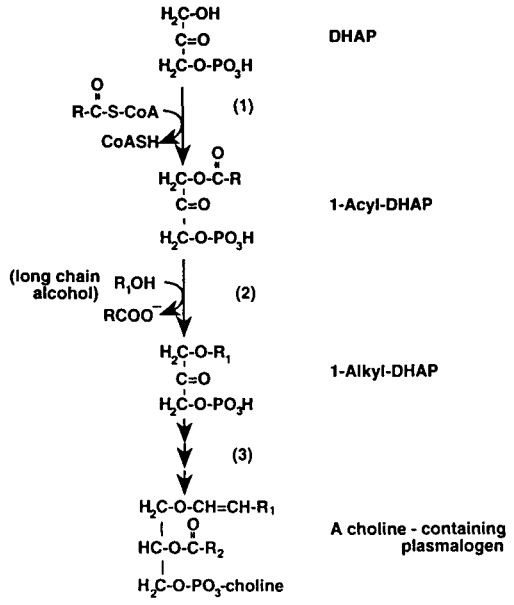


Figure 4. Pathway of glycerol-ether lipid (plasmalogen) synthesis. The numbers in the diagram refer to the enzymes participating in this pathway. (1) DHAP acyltransferase; (2) alkyl-DHAP synthase; and (3) reactions catalyzed by enzymes of the endoplasmic reticulum. DHAP, dihydroxyacetonephosphate.

may act peroxidatically to further oxidize substrates such as ethanol, methanol, formate, or nitrite.

Additionally, the metabolic role of this organelle has been broadened in recent years to include the catabolism of very-long chain fatty acids, the biosynthesis of ether-phospholipids and bile acids, and metabolic conversions of cholesterol, polyamines, oxalate, pipecolic acid, phytanic acid, dicarboxylic acids, and several drugs.

While many of these functions occur by variations of the oxidative procedures already described, special mention may be given to the biosynthesis of ether-phospholipids and bile acids.

Peroxisomes play an important role in catalyzing some of the initial reactions involved in the biosynthesis of ether-phospholipids (Hajra and Bishop, 1982), and the pathways of formation of these significant membrane constituents are shown in the accompanying figure (Figure 4).

In the first stage of ether-lipid synthesis dihydroxyacetone-phosphate (DHAP) is converted to acyl-DHAP by means of the enzyme DHAP-acyltransferase. The acyl-DHAP is then converted to alkyl-DHAP by means of alkyl-DHAP synthase. Alkyl-DHAP is subsequently subjected to a series of oxidation and reduction

reactions by enzymes of the ER, to finally yield a plasmalogen, or glycerol-ether lipid. The enzymes dihydroxyacetone phosphate acyltransferase and alkyl-dihydroxyacetone phosphate synthase are located primarily in the peroxisome.

A number of reports in the literature demonstrate that peroxisomes are also intimately involved in the metabolism of cholesterol and bile acids (Pedersen et al., 1987). A key regulatory enzyme in cholesterol biosynthesis, β-hydroxy-β-methylglutaryl-CoA reductase, has been shown to occur in peroxisomes, as well as in mitochondria (see, for example, Keller et al., 1985), and the peroxisome is also involved in the conversion of mevalonic acid to cholesterol, the conversion of trihydroxycoprostanoic acid (THCA) into cholic acid, and the conversion of 26-hydroxy cholesterol to 3-β-hydroxy-5-cholenoic acid. Many of these steps in bile acid synthesis involve reactions of the β-oxidation sequence.

Metabolic Flux

Of central relevance to the establishment of the metabolic significance of peroxisomes is the evidence provided by direct measurement of peroxidatic flux in cells containing these organelles. Using formate as a substrate, for example, Aebi and Fortwich (1959) found that peroxidatic reactions are indeed a substantial contributor to metabolic flux in mammalian tissues. In liver homogenates, about one-third of normal oxygen utilization occurred via the peroxisome. Other workers have been able to quantitatively assess hydrogen peroxide generation by exploiting the spectral characteristics of the first intermediate formed by enzyme and substrate, i.e., catalase compound 1 (Figure 5), thus demonstrating that a major portion of exogenous ethanol is oxidized by this pathway in rat liver.

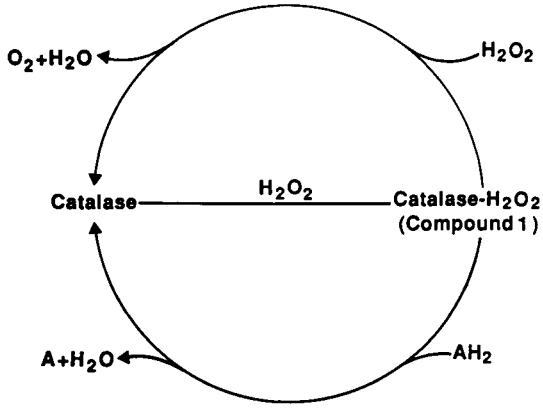


Figure 5. Reactions of catalase and hydrogen peroxide. Compound 1 is characterized by its absorbances at 405 and 665 nm. AH_2 is a generalized representation of a hydrogen donor.

The catalase-H_2O_2 intermediate (catalase compound I) is also observable in perfused liver and intact animals. Consequently, through measurement of this intermediate, both ethanol and methanol have been shown to be metabolized via catalase in living organisms, and the metabolic activity of liver perfused with urate and glycolate was seen to constitute nearly half of the total liver respiration. Thus, biological oxidations of considerable physiological significance have been demonstrated as occurring through the catalase system (Oshino et al., 1973).

Again, using experiments which studied metabolic flux in whole animals, Masters and Crane (1984) provided support for the significant physiological status of peroxisomal metabolism. They used agents such as clofibrate to cause induction of the peroxisomal pathways of β-oxidation by means of an increase in oxidase and catalase, glycolate and ethanol to cause activation of this pathway, albeit at different loci; and aminotriazole to inhibit the pathway at the catalase step. All of these treatments caused significant and widespread changes in lipid flux which *in toto* provide a strong indication of the significant variability of peroxisomal metabolism, and the cell's ability to increase activities to cope with abnormal substrate concentrations. This "inducibility" may well be an important aspect of the metabolism of peroxisomes and the metabolic relationships between peroxisomes, mitochondria, and other subcellular compartments. At the same time, the results suggest that peroxisomes also have an appreciable role in lipid metabolism under normal conditions *in vivo*.

PEROXISOME BIOGENESIS

Although peroxisomes have been recognized as significant structures in animal and plant cells for several years now, the mechanism of formation of these organelles has remained problematical and ill-defined in many respects. As might be expected, the initial views on the mechanism of peroxisome biogenesis arose mainly from morphological studies. The suggestions from these studies were that these organelles grow as buds from the ER, then split off to form discrete particles that mature and are eventually degraded by lysosomal hydrolysis. In this model, peroxisomal proteins are viewed as being synthesized on ribosomes that are bound to the ER, released into the lumen of the ER, and pass through channels to the growing peroxisome. Subsequently, a number of investigators raised serious questions about the acceptance of these previous proposals and reexamination of the biogenetic events on a much shorter time scale provided no support for the above model (see Lazarow and Fujiki, 1985).

The current understanding of peroxisome biogenesis favors a fission model for peroxisomal biogenesis. In this model, catalase and other peroxisomal proteins are synthesized on free polyribosomes and imported posttranslationally into preexisting peroxisomes. In this way, peroxisomes are allowed to grow and undergo fission to form new peroxisomes (Figure 6). One of the interesting implications of this model is that peroxisomes never form *de novo*, and that most animal cells, including

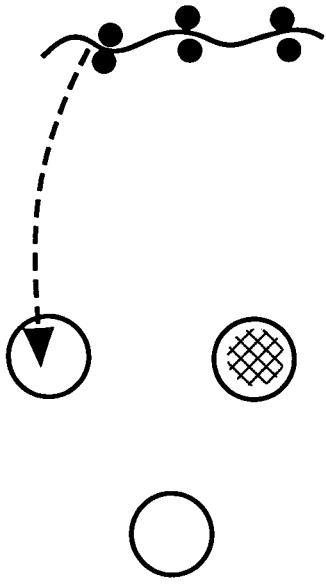


Figure 6. Proposed model for peroxisome biogenesis. Peroxisomal proteins of the membrane, matrix, and core are synthesized on free polysomes and imported into preexisting peroxisomes.

germ cells, must contain at least one peroxisome in order to allow subsequent biogenesis and proliferation. This situation is not yet fully confirmed.

The recent concept of a peroxisome reticulum further extends this fission model to include a series of peroxisome compartments which are interconnected by tubular structures (Figure 7). Some recent studies are consistent with different parts of these reticula housing different subsets of peroxisomal proteins, indicating a heterogeneity of structure and providing a possible insight into the processes of their biogenesis.

A topogenic signal, which directs many proteins to the peroxisome, has been identified as the amino acid sequence, serine-lysine-leucine at the carboxyterminus, enabling these components to be incorporated without accompanying proteolysis (Gould et al., 1989). There is evidence, though, for a degree of variation in the route of import, along with a general requirement for ATP in this process.

A novel and potentially rewarding approach to the definition of peroxisome biogenetic processes is the genetic characterization of peroxisome assembly mutants of yeast and Chinese hamster ovary cells. In particular, the rapid progress in the identification of peroxisome assembly genes from yeast mutants (see Subramani, 1993) holds the promise of a clearer understanding of the molecular mechanisms involved.

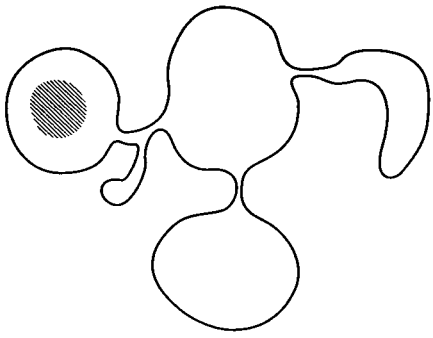


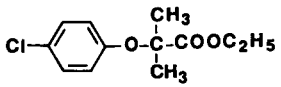
Figure 7. Diagrammatic representation of the "peroxisomal reticulum" concept. In this model new peroxisomes are envisaged as forming by fission of an existing peroxisomal reticulum, with peroxisomal proteins being imported posttranslationally.

PEROXISOME PROLIFERATORS

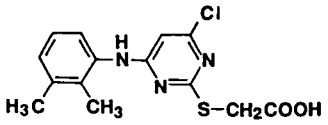
One of the most interesting aspects of peroxisomal function is their ability to be induced by a number of agents called peroxisome proliferators. An example of such a peroxisome proliferator is clofibrate (ethyl-p-chloro-phenoxy-isobutyrate), a well known hypolipidemic drug. Clofibrate treatment has been shown to increase the peroxisomal volume in hepatocytes from about 2% in normal animals to more than 10%.

A number of other hypolipidemic drugs have also been shown to be as effective, or more effective, in causing peroxisome proliferation in rodent liver; and the list has widened to include agents which are not hypolipidemic (see Figure 8). Peroxisome proliferators are now known to encompass a range of xenobiotics as well as compounds that are either analogs of clofibrate or structurally unrelated to clofibrate (e.g., Wy-14,643), as well as completely unrelated chemicals like the drug aspirin (acetylsalicylic acid). Another group of chemicals which have been shown to be potent peroxisome proliferators include industrial plasticizers, such as di-(2-ethylhexyl)phthalate, which are used in plastic manufacture. Again, proliferation of peroxisomes can result as well from physiological manipulations, such as those induced by high fat diets.

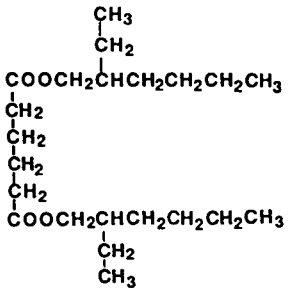
The peroxisomes which are induced by peroxisome proliferators are different in many ways from the peroxisomes present in livers of normal animals. The content of enzymes within these organelles, in particular, is different; thus, the enzymes of the fatty acid β-oxidation pathway may increase some ten-fold following drug induction, and come to represent predominant protein components in these organelles. Again, peroxisome proliferators also appear to significantly influence the permeability characteristic of isolated peroxisomes (Klucis et al., 1985), allowing the release of even such a large molecule as catalase. This effect may reflect



Clofibrate



Wy–14,643



DEHP

Figure 8. Structures of some peroxisome proliferators. Clofibrate, ethyl p-chloro-phenoxyisobutyrate; Wy-14,643, [4-choro-6-(2,3-xylidino)-2-pyrimidinylthio] acetic acid; DEHP, di-(2-ethylhexyl)phthalate.

changes in the phospholipid composition of induced organelles (Crane and Masters, 1986).

The mechanism of the induction of specific peroxisomal proteins has been investigated extensively. Such analyses have indicated that peroxisome proliferators increase the content of these proteins by acting at the level of gene transcription. In addition to this specific receptor-mediated process, evidence also exists in favor of two other mechanisms of proliferation, namely, changes in the permeability of the peroxisomal membrane and the activities of key enzymes, and substrate overload (Masters and Crane, 1995).

One other aspect of the cellular effects of peroxisome proliferators which has attracted much attention is their role in hepatocarcinogenesis. Long-term exposure of rodents to peroxisome proliferators like clofibrate and Wy-14,643, for example, has resulted in the formation of hepatocellular carcinomas (see Goldfischer and Reddy, 1984). The mechanism of tumor production does not relate to chromosomal DNA damage but instead has been proposed to involve the levels of extra-perox-

isomal hydrogen peroxide which are believed to be increased under such circumstances. The increased concentration of hydrogen peroxide outside the peroxisomes has been suggested to result from a combination of factors, viz. the increased formation of this metabolite in proliferated peroxisomes and an unchanged capacity of peroxisomal catalase to detoxify this metabolite. Once having escaped the peroxisome's detoxifying pathways, H_2O_2, being a powerful oxidizing agent, is thought to be responsible for oxidative damage to lipid and protein components of membranes, and for the generation of other destructive oxygen species.

SUMMARY

As this review has indicated, the peroxisome has assumed a position of considerable interest and challenge to cell biologists. It is now recognized that this organelle is widely distributed in body tissues and possesses major metabolic involvements. In addition to the traditional role of the peroxisome in peroxide metabolism (involving substrates such as hydroxyacids, D-amino acids, and urate), the known metabolic capacities have broadened to include the oxidation of fatty acids (including very-long chain fatty acids, the common long-chain fatty acids, and dicarboxylic acids), prostaglandins, alcohols, polyamines, and xenobiotics, as well as significant involvements in transaminations and the synthesis of cholesterol, bile acids, ether lipids, and dolichol.

In addition, significant differences have emerged in relation to the nature of the oxidative function between the mitochondrion and the peroxisome. One such feature is the inducibility of the peroxisomal oxidation as compared to the relatively constitutive nature of mitochondrial β-oxidation. Another is the incomplete oxidation of fatty acid substrates in the peroxisome (as compared to complete oxidation in mitochondria); typically in the peroxisome, long-chain fatty acids are oxidized to medium-chain fatty acids, which may then undergo complete oxidation in mitochondria. These characteristic features of peroxisomal oxidation (namely, its inducible nature and the chain shortening) appear to play an important role in the regulation of lipid metabolism.

In relation to peroxisome biogenesis, understanding has changed considerably in recent years. While it was earlier thought that individual peroxisomes were formed by budding off from the ER, the evidence now appears to favor the presence of interconnections between many of the peroxisomes in cells, the synthesis of peroxisomal proteins on free polysomes, and the production of new peroxisomes by fission from this peroxisomal reticulum. The topogenic signal for many peroxisomal proteins has been shown to be the sequence serine-lysine-leucine at the carboxy-terminus.

An intriguing and characteristic aspect of peroxisomal function is the proliferation caused by a number of hypolipidemic agents, xenobiotics, and physiological treatments. The oxidative capacity of the organelle may be increased up to twenty-fold by such regimens, allowing a massive redistribution of metabolic capacity

Table 3. Essential Cellular Functions Fulfilled by Peroxisomes

1	Oxidation of very-long-chain fatty acids
2	Biosynthesis of bile acids
3	Catabolism of dicarboxylic acids
4	Biosynthesis of ether-phospholipids
5	Catabolism of pipecolic acid
6	Oxidation of polyamines

within cells, and the molecular mechanisms for these changes are the focus of much topical research.

Also attracting increasing interest in recent years is the expanding list of biomedical correlations of peroxisomal diseases associated with abnormalities of the essential cellular functions normally performed by peroxisomes in humans (Table 3). These dysfunctions are referred to at greater length in Chapter 6.

REFERENCES

Aebi, H. & Fortwich, F. (1959). Foriatoxydation and katalase-activitat. Helv. Physiol. Acta 17, 180–201.

Crane, D. I. & Masters, C. J. (1986). The effect of clofibrate on the phospholipid composition of the peroxisomal membranes in mouse liver. Biochim. Biophys. Acta 876, 256–263.

de Duve, C. (1969). Evolution of the peroxisome. Ann. N.Y. Acad. Sci. 168, 369–381.

de Duve, C. & Baudhuin, P. (1966). Peroxisomes (microbodies and related particles). Physiol. Rev. 46, 323–357.

Duley, J. A. & Holmes, R. S. (1976). L-α-hydroxyacid Oxidase Isozymes: purification and molecular properties. Eur. J. Biochem. 63, 163–173.

Hajra, A. K. & Bishop, J. E. (1982). Glycerolipid biosynthesis in peroxisomes via the acyl dihydroxyacetone phosphate pathway. Ann. N.Y. Acad. Sci. 368, 170–182.

Goldfischer, S. L. & Reddy, J. K. (1984). Peroxisomes (microbodies) in cell pathology. Intl. Rev. Exp. Pathol. 26, 45–84.

Gould, S. J., Keller, G.-A., Hosken, N., Wilkinson, J., & Subramani, S. (1989). A conserved tripeptide sorts proteins to peroxisomes. J. Cell Biol. 108, 1657–1664.

Hamilton, G. A. (1985). Peroxisomal oxidases and suggestions for the mechanism of action of insulin and other hormones. Adv. Enzymol. 57, 85–178.

Keller, G.-A., Barton, M. C., Shapiro, D. J., & Singer, S. J. (1985). 3-Hydroxy-3-methylglutaryl coenzyme A reductase is present in peroxisomes in normal rat liver cells. Proc. Natl. Acad. Sci. USA. 82, 770–774.

Klucis, E., Crane, D., & Masters, C. (1985). Sequential alterations in the microlocalization of catalase in mouse liver after treatment with hypolipidemic drugs. Mol. Cell. Biochem. 65, 73–82.

Lazarow, P. B. & de Duve, C. (1976). A fatty acyl-CoA oxidizing system in rat liver peroxisomes; enhancement by clofibrate, a hypolipidemic drug. Proc. Natl. Acad. Sci. USA. 73, 2043–2046.

Lazarow, P. B. & Fujiki, Y. (1985). Biogenesis of peroxisomes. Ann. Rev. Cell Biol. 1, 489–530.

Mannaerts, G. P. & van Veldhoven, P. O. (1987). Permeability of the peroxisomal membrane, In: Peroxisomes in Biology and Medicine, (Fahimi, H. D. & Sies, H., eds.), pp. 169–176. Springer Verlag, Berlin.

Masters, C. & Crane, D. (1984). The role of peroxisomes in lipid metabolism. Trends Biochem. Sci. 9, 314–317.

Masters, C. & Crane, D. (1992). The peroxisome: organization and dynamics. Adv. Mol. Cell. Biol. 4, 133–160.

Masters, C. & Crane, D. (1995). The Peroxisome—A Vital Organelle. Cambridge University Press, Cambridge, U.K.

Masters, C. & Holmes, R. (1977). Peroxisomes: new aspects of cell physiology and biochemistry. Physiol. Rev. 57, 816–882.

Masters, C., Pegg, M., & Crane, D. (1986). On the multiplicity of the enzyme catalase in mammalian liver. Mol. Cell. Biochem. 70, 113–120.

Oshino, N., Chance, B., Sies, H., & Bucher, T. (1973). The role of H_2O_2 generation in perfused rat liver and the reaction of catalase compound I and hydrogen donors. Arch. Biochem. Biophys. 154, 117–131.

Pederson, J. J., Kase, B.-F., Prydz, K., & Bjorkhem, I. (1987). Liver peroxisomes and bile acid formation. In: Peroxisomes in Biology and Medicine (Fahimi, H. D. & Sies, H., eds.), pp. 67–77, Springer Verlag, Berlin.

Subramani, S. (1993). Protein import into peroxisomes and biogenesis of the organelle. Ann. Rev. Cell Biol. 9, 445–475.

Tolbert, N. E. (1971). Microbodies, peroxisomes and glyoxysomes. Annu. Rev. Plant Physiol. 22, 45–74.

RECOMMENDED READINGS

Masters, C. & Holmes, R. (1977). Peroxisomes: new aspects of cell physiology and biochemistry. Physiol. Rev. 57, 816–882.

Böck, P., Kramar, R., & Pavelka, M. (1980). Peroxisomes and related particles in animal tissues. Springer Verlag, Berlin.

de Duve, C. (1983). Microbodies in the living cell. Sci. Am. 248, 74–84.

Fahimi, H. D. & Sies, H. (1987). Peroxisomes in Biology and Medicine. Springer Verlag, Berlin.

van den Bosch, H., Schutgens, R. B. H., Wanders, R. J. A., & Tager, J. M. (1992). Biochemistry of peroxisomes. Ann. Rev. Biochem. 61, 157–197.

Masters, C. & Crane, D. (1995). The Peroxisome—A Vital Organelle. Cambridge University Press, Cambridge, U.K.

Chapter 4

The Mitochondrion

DAVID DRAKE TYLER

Principles of Medical Biology, Volume 3
Cellular Organelles and the Extracellular Matrix, pages 59–132

ISBN: 1-55938-804-8

INTRODUCTION

Mitochondria [Gr. *mitos* = thread + *chondros* = granule] are lipoprotein particles of characteristic shape, size, structure, and staining properties. They are present in the cytoplasm of all eukaryotic cells of higher animals and plants, and also in certain microorganisms, including algae, protozoa, and fungi. Under microscopic examination, mitochondria often appear in fixed tissue sections as spheres, ovoids, or rods, approximately 1–5 μm long and 0.3–1.0 μm wide, but many variations are found even within a single tissue. In contrast, their width is remarkably constant in a wide variety of cell types, a feature which limits the maximum diffusion path required for metabolites to reach the center of the organelle from the surrounding cytoplasm. Mitochondria are believed to have arisen by the symbiotic association between aerobic bacteria and anaerobic hosts cells during the evolution of the eukaryotic cell. The present account is restricted to the properties and physiological functions of mitochondria in mammalian cells and tissues.

For the past forty years, it has been known that mitochondria are the chief site of intracellular respiration and energy metabolism, and that their metabolic activities are responsible for more than 90% of the total adenosine triphosphate (ATP) synthesis in the animal body (Tyler, 1992). The magnitude of this process is illustrated dramatically by the fact that the human body daily synthesizes a quantity of ATP equal to or greater than the total body weight. Because the human body contains only about 50g of adenine nucleotides, it is obvious that on average each molecule of ATP is synthesized in mitochondria and used as an energy source many times each day. At first, mitochondria were regarded as relatively independent power plants of the cell, but they are now recognized as dynamic centers of cell metabolism which share an incessant and profound interdependence with their surrounding cytoplasm (Bereiter-Hahn, 1990). At the molecular level, a substantial traffic of numerous kinds of cell metabolites (including the adenine nucleotides ATP and ADP) occurs across the mitochondrial membranes, often via specific carrier or transporter proteins present in the mitochondrial inner membrane. The carriers play an essential role in the import of cell metabolites for oxidation and ATP synthesis, and they also function (especially in liver cells) during processes taking place partly in the matrix space and partly in the cytosol, including ureogenesis, gluconeogenesis, and lipogenesis. At the structural level, there is often a dynamic balance taking place between small individual mitochondria and mitochondrial networks, which could provide a membranous path for power transmission within the cell (Skulachev, 1988). In some tissues, including liver, a close contact or continuum exists between the mitochondrial outer membrane and the endoplasmic reticulum, and this arrangement probably provides a path for the transport between the organelles of proteins, phospholipids, heme, and intermediates in steroid synthesis. The outer membrane also contains receptor sites which interact with the cell cytoskeleton, the system of fibers responsible for determining and maintaining the distribution of mitochondria within the cell cytoplasm. Most importantly of all, the small quantity of DNA (mt-DNA) in the mitochondrial

matrix space (about 0.2–1.0% of the total cell DNA) functions in partnership with nuclear DNA during the intricate process of mitochondrial biogenesis. Each mitochondrion contains about 10 mt-DNA molecules. The complete nucleotide sequence of the 16,569-base pair human mt-DNA was elucidated in 1981. The function of mt-DNA is to encode the synthesis of the small and large mitochondrial ribosome (mitoribosome) RNA subunits, mitochondrial transfer RNAs (tRNAs), and the messenger RNA required for the synthesis of a small number of polypeptides (probably 13), which function as subunits of respiratory-chain complexes I, III, and IV, and the ATP synthase (Complex V). The synthesis of all these mtDNA gene products uses the enzymic transcription and translation processes present in the matrix space, including DNA polymerase, RNA polymerase, and amino-acid activating enzymes. The remaining much larger number of mitochondrial proteins are encoded by nuclear DNA, synthesized outside the mitochondria on cytosolic ribosomes, and then transported to their correct mitochondrial location via contact sites between the outer and inner mitochondrial membranes (Pfanner and Neupert, 1990).

Disorders of Mitochondria

Patients suffering from disorders of mitochondrial metabolism often show abnormalities in tissues dependent on a high rate of ATP synthesis, including brain and skeletal muscle (Scholte, 1988). Primary disorders are due mainly to mutation, which can affect mitochondrial functions in a number of different ways. Defects can occur in the primary transcript during protein synthesis, and in the enzymes and receptors functional during posttranslational modification of the protein or the process of protein transport into mitochondria. Mutations can affect the catalytic activity of the protein, result in failure of the protein to become incorporated correctly into the mitochondrial structure, or perhaps increase or retard its degradation. A mutation in the presequence of the precursor form of a mitochondrial protein can result in its failure to be transported into mitochondria or in its abnormal routing to another region of the cell. Some mitochondrial disorders are caused by specific deletions in the mitochondrial genome. Secondary defects arise from a number of causes including a block in mitochondrial metabolism due to lack of components (including vitamins) required for normal function, or to the action of inhibitory metabolites accumulating in mitochondria as a consequence of the primary defect. Because mitochondria are the main centers of oxygen reduction in living cells, they are especially vulnerable to the toxic actions of oxygen radicals formed during metabolism. At present, there is much interest in the idea that oxygen radicals are implicated in many cell disorders and also in the process of aging.

STRUCTURE OF MITOCHONDRIA

The characteristic feature of mitochondrial structure revealed by electron microscopy is the presence of a double-membrane system which divides the organelle into

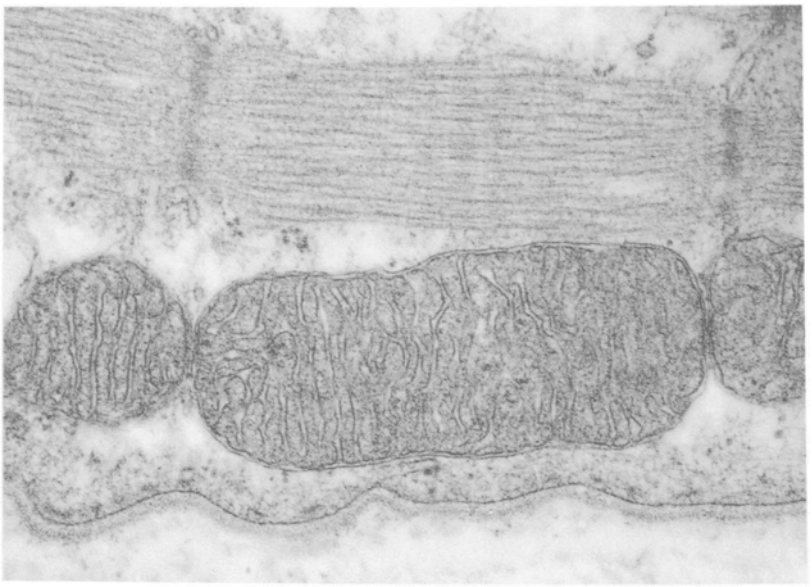


Figure 1. Electron micrograph showing mitochondria in the cranial tibia muscle of the dog *Canis domesticus* (Great Dane). Magnification × 60,000. Courtesy Dr. Jenifer Plummer and John Bredl.

two compartments, an outer compartment or intermembrane space, and an inner compartment or matrix space (Figure 1). Pores are present in the outer membrane, and contact sites occur between the outer and inner membranes. The pores in the outer membrane, about 2–4nm in diameter, are formed by the protein porin. These pores exclude protein molecules but allow the free passage of most, if not all, metabolite molecules including the adenine nucleotides and nicotinamide nucleotides. (The molecular weight of ATP is 507 Da.) The inner-membrane system seems to consist of two regions which have separate functions, namely the inner peripheral membrane lying roughly parallel to the outer membrane; and the cristae, invaginated regions of the inner membrane which more or less penetrate the matrix space (Figure 2). The inner-peripheral membrane probably contains the metabolite carrier proteins. The cristae contain four large enzyme complexes (Complexes I, II, III, and IV) which together with small mobile ubiquinone (Q) and cytochrome *c* molecules constitute the respiratory chain system; and the ATP synthase complex (Complex V) which catalyzes the synthesis of ATP from ADP and inorganic phosphate (P_i). The folding or invagination of the inner membrane greatly increases the membrane surface area available to accommodate the five complexes of the respiratory-chain phosphorylation system. In rat liver, for example, the mitochondria occupy about 17% of the cell volume, but the combined surface area of the

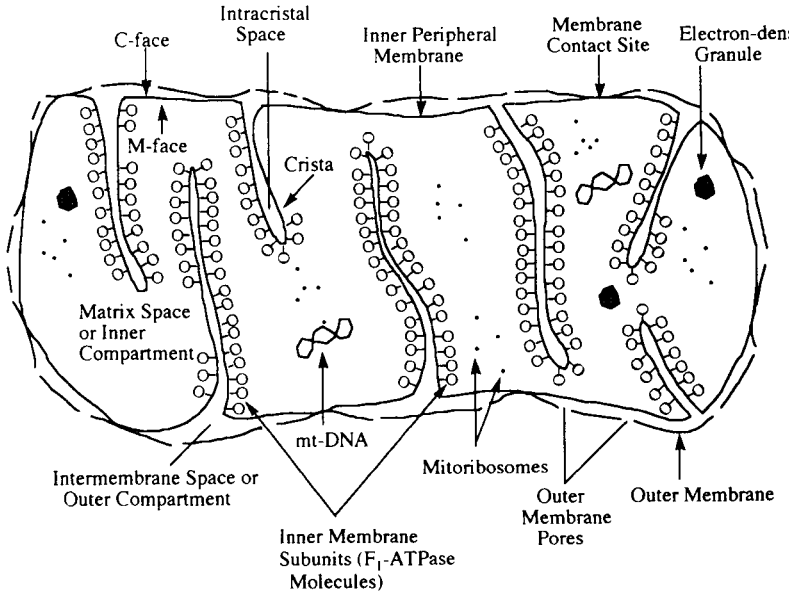


Figure 2. Diagram summarizing the main features of mitochondrial structure revealed by using a variety of positive-staining and negative-staining techniques. In most tissues, electron-dense granules are thought to be assembly sites for the formation of new inner membrane, but in bone-forming tissues they are probably calcium phosphate granules.

cristae is approximately 11 times the surface area of the cell. Various structures are present in the matrix space, including DNA fibers, ribosomes, and electron-dense granules.

For many years, the study of mitochondrial structure by conventional fixation and staining methods suggested that mitochondria contain mainly lamellar or plate-like cristae, but more recent techniques of freeze-fracture and scanning electron microscopy have called into question both the shape of the cristae and the extent of their connections to the inner peripheral membrane. These features differ among the mitochondria of various tissues. Whatever the outcome of these present uncertainties, the primary function of the inner-membrane system is to provide a proton-impermeable barrier, so that the proton electrochemical potential, or proton gradient, established across the membrane by Complexes I, III, and IV of the proton-translocating respiratory chain can be used to drive the synthesis of ATP by Complex V (Nicholls, 1982). An increase in proton permeability, even at one localized region of the membrane in a mitochondrion, causes the collapse of the potential across the whole inner-membrane system, and results in the failure of ATP synthesis. Mitochondria often lie close to energy-requiring sites in cells, so that a

minimal diffusion path is taken by molecules of ATP, synthesized mainly in the mitochondria, to energy-requiring sites which lie mainly outside the mitochondria. An excellent example is provided by the sperm cell, in which mitochondria form a spiral in the mid-piece, adjacent to dynein ATPase in the axoneme responsible for flagellar motility.

Structure-Function Relationships

Knowledge of the precise location of mitochondrial enzymes and other components has been obtained mainly by using systematic methods for the subfractiona-

Table 1. Distribution of Enzymes and Other Components in Mitochondria

Region	*Number of Polypeptide Components*	*Important Components*
Outer membrane	About 50	Porin (pore-forming protein) Redox enzymes, including NADH-cytochrome b_5 reductase; cytochrome b_5; amine oxidase(monoamine oxidase) Enzymes of lipid metabolism including long-chain fatty acyl-CoA synthetase; carnitine palmitoyl-CoA transferase; enzymes of phospholipid metabolism
Intermembrane space	About 15	Adenylate kinase; nucleoside diphosphate kinase; sulfite oxidase
Inner membrane (cristae and inner-peripheral membrane)	At least 100	Respiratory-chain complexes I–IV, including succinate dehydrogenase (part of Complex II) ATP synthase ($F_O.F_1$-ATPase; F_O sector, membrane; F_1-ATPase, M-face) Metabolite carrier proteins (? inner-peripheral membrane) Cytochrome P_{450} monooxygenase system (adrenal, liver, kidney)
Matrix space	At least 200	Citric-acid-cycle enzymes, except succinate dehydrogenase Fatty acid oxidation enzymes (β-oxidation system; short- and medium-chain fatty acyl-CoA synthetases; enzymes of ketone-body metabolism) Enzymes of amino acid and nitrogen metabolism (amino-transferases, glutamate dehydrogenase; some urea cycle enzymes Enzymes of nucleic acid and protein synthesis; DNA (mt-DNA), ribosomes Some enzymes of heme synthesis Superoxide dismutase (Mn enzyme)
Contact sites between the inner and outer membranes	Several	Protein and phospholipid receptors for cytosolic precursor and proteins imported into mitochondria

Note: Information is based largely on data obtained from rat-liver mitochondria.

tion of isolated mitochondria. This information is summarized in Table 1, which shows that the main enzymic activities of mitochondria are located in the inner membrane and the matrix space. In heart mitochondria, which function largely as centers of ATP synthesis, the respiratory chain components, together with the ATP synthase complexes and the metabolite carrier proteins for adenine nucleotides and P_i, constitute about 80% of the total protein in the inner membrane. In liver mitochondria, which have numerous functions in addition to ATP synthesis, the value is about 20%. The wider spacing between cristae in liver compared with heart mitochondria provides a proportionately larger volume of matrix space to accommodate additional enzymes for urea synthesis and other functions, and extra inner peripheral membrane to accommodate the greater variety of metabolite carriers needed to support the more complex metabolism of liver mitochondria.

BIOCHEMISTRY OF MITOCHONDRIA

Mitochondria are the site of oxidation by molecular oxygen of two-carbon acetyl units, in the form of acetyl-coenzyme A (acetyl-CoA), to yield CO_2 and H_2O. Acetyl units are liberated during the catabolism of cell fuels, namely carbohydrate, lipid, and protein. These compounds, obtained from the diet or from tissue stores, are first hydrolyzed to release their constituent units: the hexoses, amino acids, fatty acids, and glycerol. The release of energy during the oxidation of these fuels occurs in three main stages. In the first stage, their carbon skeletons are partially oxidized to yield CO_2, H_2O, and either acetyl-CoA or the citric-acid-cycle intermediates 2-oxoglutarate, succinyl-CoA, fumarate, and oxaloacetate. These intermediates can be converted to pyruvate and then to acetyl-CoA by the action of the citric acid cycle and associated enzymes. (The citric acid cycle is also called the tricarboxylic acid or TCA cycle, the Krebs cycle, and the citrate cycle.) During the second stage, acetyl groups are oxidized by the cycle to form CO_2, H_2O, and the dehydrogenase cofactors nicotinamide adenine dinucleotide (NAD) and flavin adenine dinucleotide (FAD) are reduced by pairs of hydrogen atoms removed by dehydrogenases from cycle intermediates. Finally, in the third stage these reduced cofactors (NADH and $FADH_2$) are reoxidized by the respiratory chain, thus renewing the supply of oxidized cofactors so that the cycle can continue to operate. Nearly all the free energy available in cell fuels is liberated in the third stage, when NADH and $FADH_2$ are oxidized by the respiratory chain to form water, and the oxidoreduction reactions are linked or coupled to ATP synthesis (Figure 3). This process, called oxidative phosphorylation, provides the bulk of ATP used in cells for energy-requiring reactions. A notable exception is found in brown adipose tissue, where much of the oxidative energy is 'uncoupled' from ATP synthesis by a regulated mechanism, so that when necessary the energy of cell fuels can be released directly as heat during non-shivering thermogenesis.

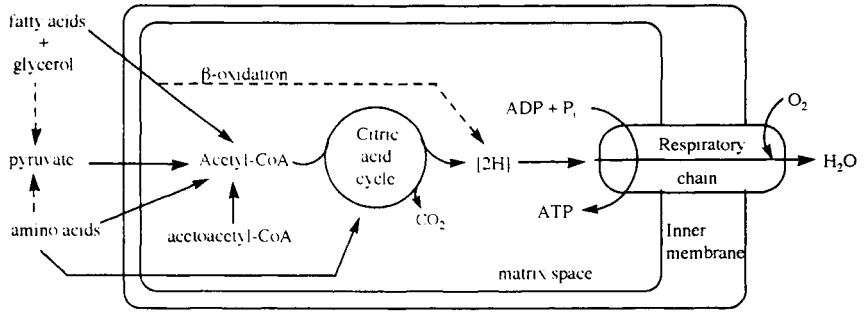


Figure 3. Diagram showing the main pathways for the oxidation of cell fuels by mitochondria, and the role of the respiratory chain in ATP synthesis. Reducing equivalents [2H] released during the oxidation of acetyl-units by the citric acid cycle and during the β-oxidation of long-chain fatty acids are passed to the respiratory chain for oxidation linked or coupled to ATP synthesis.

Sources of Acetyl-CoA

The role of mitochondria in acetyl-CoA formation and oxidation is confined largely to the metabolism of pyruvate, fatty acids, and acetoacetyl-CoA. Glucose is the primary fuel of intracellular respiration. The dehydrogenation or carboxylation of pyruvate formed by glycolysis can provide the acetyl groups (as acetyl-CoA) and oxaloacetate necessary for the complete oxidation of pyruvate to CO_2 and H_2O independently from the metabolism of other cell fuels. The vital importance of this process is illustrated by the fact that the adult human brain constitutes only about 2% of the body weight but is responsible for about 60% of the total body glucose usage and 20% of the total body oxygen usage at rest. Liver glycogen can maintain the blood glucose concentration for periods of up to 24 hours of fasting. During brief periods of starvation or during sustained exercise glycogen stores are largely depleted, and the vital glucose requirement (especially for the brain) must be provided by glucose synthesis (gluconeogenesis) from non-carbohydrate sources. Triglyceride is hydrolyzed in adipose tissue stores, and nonesterified long-chain fatty acids and glycerol are released into the bloodstream to serve as a fuel for other tissues, including muscle and liver. Much of the glycerol released during lipolysis is metabolized by tissues containing glycerol kinase, such as liver and kidney, and is converted to glucose to maintain the blood sugar level. A major portion of the fatty acids liberated are taken up by the liver and partially oxidized to form ketone bodies (acetoacetate and 3-hydroxybutyrate) which then join the blood circulation as fuel for other tissues, the extrahepatic tissues. Ketone bodies provide a fuel for respiration that is available immediately to extrahepatic tissues. The uptake and oxidation of ketone bodies occurs in most, if not all, aerobic tissues except the liver, for example, in muscle, kidney, mammary gland, and small intestine. Under normal

circumstances, the concentration of ketone bodies in blood is too low to make a significant contribution to fuel supplies for the brain. During prolonged starvation, ketone bodies can replace up to 75% of the brain glucose requirement as an energy source, whereas long-chain fatty acids are not taken up to a significant extent by this tissue. Some glucose oxidation (via pyruvate) is necessary in extrahepatic tissues to provide essential citric-acid-cycle intermediates needed for the complete oxidation of fatty acids and ketone bodies. In liver, the function of glycolysis is mainly to provide precursors for biosynthetic reactions rather than simply to supply pyruvate for oxidation. In the fed state, the oxidation of carbon skeletons derived from amino acid degradation can serve as the main fuel for ATP synthesis, but in the starved state a larger proportion of the energy is derived from the β-oxidation of fatty acids, which occurs during the formation of ketone bodies.

Pyruvate Oxidation

Pyruvate formed during glycolysis is transported from the cytosol to the mitochondrial matrix space by a specific carrier protein, the monocarboxylate carrier, which is present in the inner membrane. In the matrix space, pyruvate undergoes oxidative decarboxylation catalyzed by the pyruvate dehydrogenase complex, according to the overall equation:

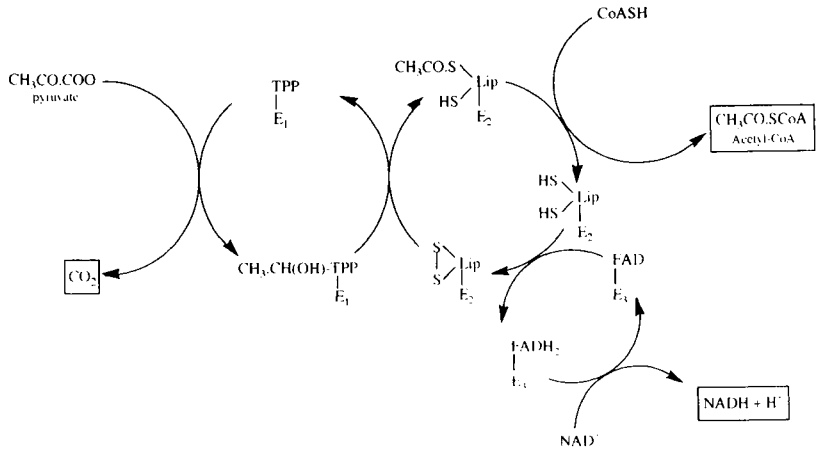
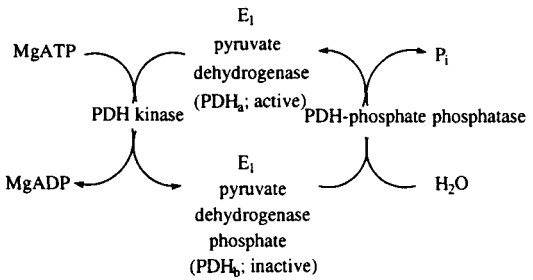


Figure 4. Reaction mechanism of the pyruvate dehydrogenase complex. The enzymes of the complex are as follows: E_1, pyruvate dehydrogenase (decarboxylating); E_2, dihydrolipoamide acetyltransferase; E_3, dihydrolipoamide dehydrogenase. Products are shown in boxes. Lip represents the lipoamide group attached to enzyme E_2. Four enzyme cofactors of the complex, namely thiamine pyrophosphate (TPP), coenzyme A, FAD, and NAD, are derived from members of the vitamin-B group.

$$CH_3CO{-}COO^- + NAD^+ + CoASH \xrightarrow{E_1,E_2,E_3} CH_3CO{-}SCoA + CO_2 + NADH + H^+$$

pyruvate — acetyl–CoA

This important reaction links the glycolytic pathway with the citric acid cycle and the respiratory chain, which together provide a mechanism for the complete oxidation of glucose to CO_2 and H_2O. The reaction pathway is shown in Figure 4. The overall reaction is catalyzed by three different enzymes, E_1, E_2, and E_3, which are physically associated in large numbers to form an efficient catalytic mechanism (Table 2). A single mammalian enzyme complex, one of the largest known enzyme complexes, consists of 60 E_1 subunits, 60 E_2 subunits, and 12 E_3 subunits. Their combined molecular weight is about 8,500 kDa, and the diameter of the roughly spherical complex is 45nm. (In comparison, the hydrocarbon core of a membrane lipid bilayer is about 3nm wide.) In mammalian cells, the complex catalyzes the first reaction in glucose metabolism for which no means of reversal is known. Regulation of its activity therefore has an important influence on intermediary metabolism. The complex also contains two regulatory enzymes, a kinase and a phosphatase, which regulate the activity of the complex by catalyzing phosphorylation and dephosphorylation of the first enzyme (E_1):



The PDH kinase is activated by a high intracellular ATP/ADP ratio, and by increases in the acetyl-CoA/CoASH ratio and the NADH/NAD^+ ratio. These ratios are high during fatty acid oxidation, which therefore has a sparing action on the oxidation of carbohydrate, stored to a very limited extent (as glycogen) compared with the large fat stores in the well-nourished individual.

The simplest disorder of pyruvate metabolism is caused by dietary thiamine deficiency, which results in the accumulation in blood of pyruvate and metabolically-related compounds (lactate, alanine). The effects of thiamine deficiency are seen in the conditions of beri-beri and Wernicke's encephalopathy. Other disorders are caused by defects in the catalytic or regulatory subunits of the pyruvate dehydrogenase complex and in pyruvate carboxylase. In severe cases, the defects are associated with developmental disorders and with death in infancy. Dichloroacetate ($CHCl_2.COO^-$) has been used in the treatment of some of these conditions. It inhibits pyruvate dehydrogenase kinase and consequently it increases the rate of pyruvate oxidation. A defect in dihydrolipoamide dehydrogenase also diminishes

Table 2. The Pyruvate Dehydrogenase Complex

Enzyme	*Symbol*	*Prosthetic Group*	*Reaction Catalyzed*	*Properties*
Pyruvate dehydrogenase (decarboxylating)	E_1	Thiamine (vitamin B1) pyrophosphate (TPP)	Pyruvate decarboxylation	Site of regulation of the complex by phosphorylation and dephosphorylation
Dihydrolipoamide acetyl-transferase	E_2	1 lipoamide per subunit	Oxidation of the 2-carbon unit to an acetyl residue and transfer to CoASH	Site of inhibition of the complex by arsenite
Dihydrolipoamide dehydro-genase	E_3	1 FAD per subunit	Reoxidation of reduced lipoamide	Identical in the 2-oxoglutarate dehydrogenase complex
Pyruvate dehydrogenase ki-nase				Catalyze phosphorylation and dephosphorylation of subunit E_1, to regulate the activity of the complex. The kinase is activated by Mg^{2+}, and the phosphatase is activated by Mg^{2+} and Ca^{2+}.
Pyruvate dehydrogenase phosphate phosphatase				

the activity of the 2-oxoglutarate and the branched-chain 2-oxoacid dehydrogenase complexes. The E_2 subunits of the three oxoacid dehydrogenase complexes are important antigens responsible for the presence of mitochondrial antibodies in the serum of patients suffering from primary biliary cirrhosis. This condition is a chronic liver disease, which affects middle-aged women and results in liver failure and death. The etiology of the disease is unknown and no treatment is available.

Fatty Acid Oxidation

Two main processes are responsible for the mitochondrial oxidation of saturated long-chain fatty acids to acetyl-CoA, namely fatty-acid activation and β-oxidation (Schulz, 1991). Long-chain fatty acid oxidation is also dependent on the activity of the carnitine carrier in the inner membrane, which transports activated fatty acids from their site of formation at the outer mitochondrial membrane to their site of oxidation by the β-oxidation system in the matrix space. Medium-chain-length and short-chain-length fatty acids and the volatile fatty acids (acetic, propionic, and butyric acids) can permeate the inner membrane freely and do not require a carrier system. Their activation and oxidation is catalyzed by enzymes in the matrix space. Medium-chain-length fatty acids are present in the maternal milk provided for neonatal animals.

Several different fatty acyl-CoA synthetases are responsible for catalyzing fatty-acid activation, according to the general equation:

$$(\text{fatty acid})^- + \text{CoASH} + \text{ATP}^{4-} \xrightarrow{Mg^{2+}} \text{fatty acyl-CoA} + \text{ATP}^{2-} + \text{PP}_i^{3-}$$

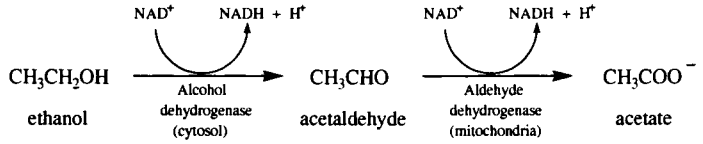
The reactions catalyzed by fatty acyl-CoA synthetases are readily reversible because the product, fatty acyl-CoA, is an energy-rich compound. They are driven to the right by removal of the other reaction product, inorganic pyrophosphate (PP_i), which is hydrolyzed by the enzyme inorganic pyrophosphatase ($PP_i + H_2O \rightarrow 2\ P_i$). Three distinct synthetase enzymes are present in mitochondria. The first of these, which acts on long-chain fatty acids (C_6–C_{20}) is located in the outer membrane with its catalytic center facing the intermembrane space. Long-chain fatty acyl-CoA serves as a substrate mainly for the overt carnitine palmitoyltransferase (CPT 1) situated in the outer membrane. It is also a substrate for the outer-membrane enzyme glycerol 3-phosphate acyltransferase (GPAT), which functions in phospholipid biosynthesis. The second and third mitochondrial synthetase enzymes for fatty acids, acetyl-CoA synthetase and butyryl-CoA synthetase, are located in the matrix space. Acetyl-CoA synthetase catalyzes the reaction:

$$\underset{\text{acetate}}{CH_3COO^-} + \text{CoASH} + \text{ATP}^{4-} \xrightarrow{Mg^{2+}} \underset{\text{acetyl-CoA}}{CH_3CO{-}SCoA} + \text{AMP}^{2-} + \text{PP}_i^{3-}$$

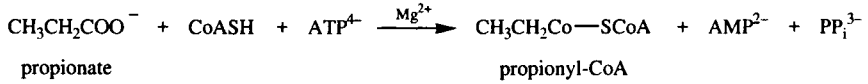
It also activates fluoroacetate, an important step in the "lethal synthesis" of fluorocitrate. Butyryl-CoA synthetase activates short-chain-length and medium-

chain-length fatty acids (C_4–C_{11}; optimum activity with C_4) and the corresponding 3-hydroxy and 2,3- or 3,4-unsaturated fatty acids.

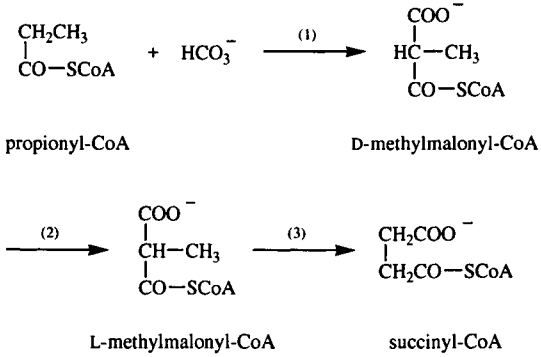
Volatile fatty acids (in the form of acetate, propionate, and butyrate) are formed in the lower digestive tract by the bacterial fermentation of carbohydrates and endogenous substrates such as mucus. They can contribute up to 10% of the calorie requirement in man. Propionate is also formed in the liver during the degradation of isoleucine, methionine, threonine, and valine, and during the β-oxidation of fatty acids containing an odd number of carbon atoms. Aromatic acids, including benzoic acid and phenylacetic acid, also serve as substrates for liver butyryl-CoA synthetase, and this enzyme can therefore function in the oxidation and detoxification of acids other than aliphatic acids. Acetate is also formed in the liver during ethanol oxidation:



Acetyl-CoA synthetase is absent from liver mitochondria but present in the mitochondria of other tissues. This distribution of the enzyme explains why acetate is transported from the liver by the blood, and taken up by other tissues for oxidation or for lipid synthesis. In many mammalian tissues, acetyl-CoA and butyryl-CoA synthetases also activate propionate to form propionyl-CoA:



Liver mitochondria contain an additional propionyl-CoA synthetase. The specialized location of this enzyme is related to the role of the liver as the main organ for gluconeogenesis, for which propionate serves as a precursor by undergoing conversion to succinyl-CoA in a sequence of three reactions catalyzed by mitochondrial enzymes:



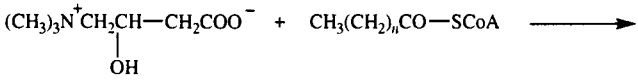
The first reaction (step 1), a carboxylation step requiring ATP and Mg^{2+}, is catalyzed by the biotin-containing enzyme propionyl-CoA carboxylase. In the second reaction (step 2), D-methylmalonyl-CoA undergoes an epimerization catalyzed by methylmalonyl-CoA racemase. During the third reaction (step 3), L-methylmalonyl-CoA is isomerized to form succinyl-CoA. This reaction is catalyzed by methylmalonyl-CoA mutase, one of a small number of enzymes requiring 5′-deoxyadenosylcobalamin, a cofactor derived from vitamin B_{12}. The product of the pathway, succinyl-CoA, can then enter the citric acid cycle and become available for complete oxidation or, in liver, it can be used for gluconeogenesis after conversion to oxaloacetate.

A defect in propionyl-CoA carboxylation (step 1) causes the accumulation of propionate in body fluids, together with other compounds including methylcitrate formed by the action of citrate synthase on propionyl-CoA and oxaloacetate. Methylmalonic acidemia is caused by any one of a number of defects in the third reaction (step 3) including: abnormal vitamin B_{12} function, which is responsive to massive doses of the vitamin; a mutation in the enzyme methylmalonyl-CoA mutase; and impaired transport into mitochondria of the enzyme precursor protein containing a presequence mutation.

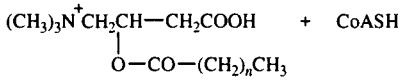
The Role of Carnitine

Carnitine is required for the transport of activated long-chain fatty acids across the mitochondrial inner membrane. A specialized transport system is necessary because, although the membrane is permeable to long-chain fatty acids, no enzyme for their activation is present in the matrix space. The membrane is impermeable to long-chain fatty acyl-CoA esters synthesized by enzymes located outside the inner-membrane barrier. By contrast, medium-chain and short-chain fatty acids can cross the inner membrane by simple diffusion and become available for activation by other fatty acyl-CoA synthetases situated exclusively within the matrix space. Oxidation of these fatty acids is therefore not dependent on carnitine.

Long-chain fatty acid transport requires the formation of a fatty acid ester with carnitine, catalyzed by external carnitine palmitoyltransferase (CPT I), an enzyme which spans the outer membrane:



carnitine fatty acyl-CoA



acylcarnitine

This enzyme commits long-chain fatty-acyl groups to translocation across the inner membrane and their entry into the β-oxidation system, and it also functions as an important site where their rate of oxidation is controlled. Regulation is achieved through the allosteric inhibition of CPT I by malonyl-CoA (Zammit, 1984). When fatty acid synthesis is taking place in the cytosol, the concentration of the pathway intermediate malonyl-CoA is increased, and it causes a decrease in the rate of fatty acid oxidation by inhibiting CPT I. This ensures that newly-formed fatty acid is not immediately degraded. In the liver, the malonyl-CoA concentration acts as an important metabolic signal to ensure that fatty acid oxidation (and hence ketogenesis) is low in the fed state (when fatty acid synthesis is high), and high in the starved state when synthesis is decreased and fatty acid serves as the main fuel for oxidation. After the transport of long-chain acyl-carnitine across the inner membrane (a process catalyzed by the carnitine/acyl-carnitine exchange carrier), fatty acyl-CoA is regenerated in the matrix space by a second internal carnitine palmitoyltransferase (CPT II) situated at the M-face of the inner membrane (Figure 5). The outer-membrane and inner-membrane forms of carntine palmitoyltransferase differ in molecular size (M_r = 86–90 kDa and 70 kDa, respectively) and in their physical and immunological properties. Mitochondria also contain a second acyltransferase enzyme, carnitine acetyltransferase. The apparent presence of octanoyltransferase

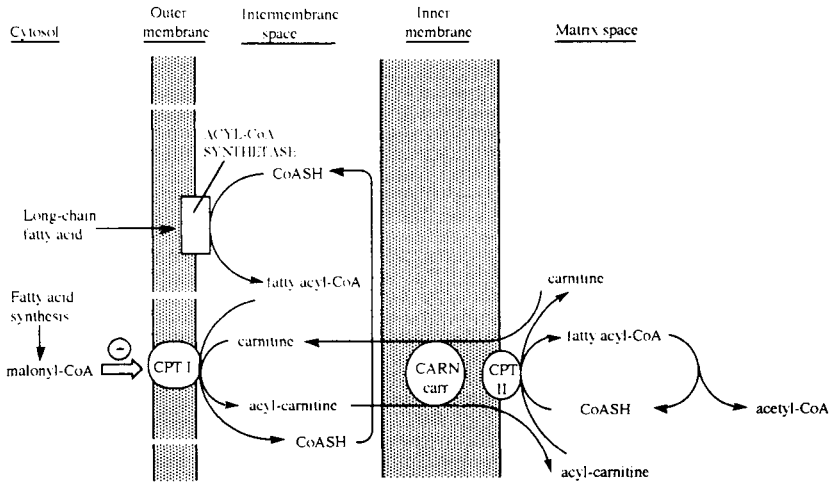


Figure 5. Role of carnitine in the transport of long-chain fatty-acyl groups across the inner mitochondrial membrane. The enzyme carnitine palmitoyltransferase I (CPT I) is a transmembrane protein containing regulatory and catalytic sites, which lie at the outer and inner faces of the mitochondrial outer membrane, respectively. The overall regulation of long-chain fatty acid oxidation is achieved by the binding of malonyl-CoA (an allosteric inhibitor) to the regulatory site of CPT I (see text).

activity in mitochondria is explained by the broad overlapping chain-length specificity of the palmitoyl- and acetyltransferases.

The β-oxidation Pathway

This pathway is catalyzed by four enzymes (Figure 6). They are probably organized as a metabolon bound to the M-face of the inner membrane. (A metabolon is a functional complex of enzymes which act sequentially in a metabolic pathway.) This arrangement would ensure the efficient sequential transfer of intermediates between enzymes of the pathway, and at the same time prevent undesirable competition between the numerous homologue forms of the coenzyme A esters at the catalytic centers. The effect of the four reaction steps is to transfer

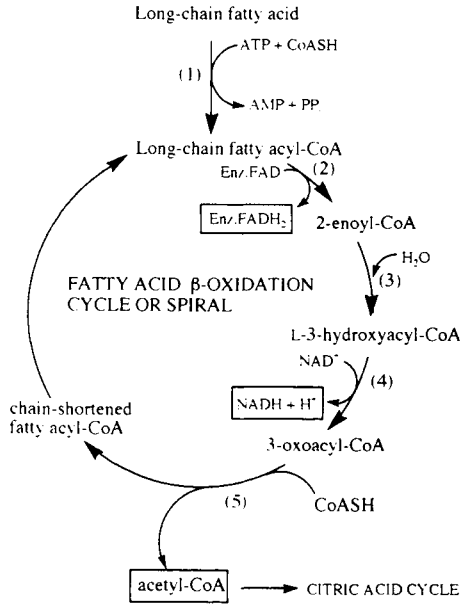
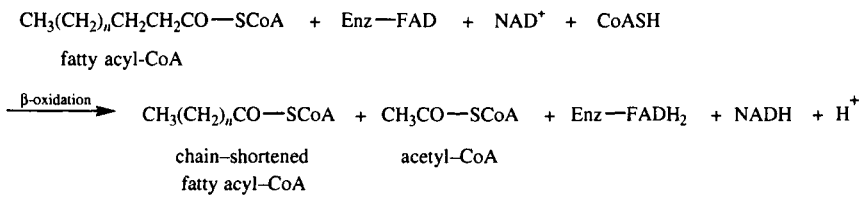
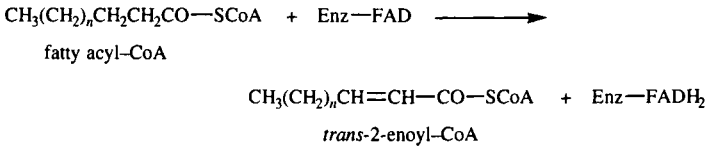


Figure 6. The β-oxidation pathway for the catabolism of straight-chain saturated fatty acids. Products are shown in boxes. The reduced enzyme cofactors produced in the pathway, namely $FADH_2$ in the form of an enzyme-bound prosthetic group ($Enz.FADH_2$), and NADH, are reoxidized by the respiratory chain. When fatty acids are oxidized, hydrogen atoms removed from the substrate molecules are transferred to the respiratory chain in two stages: during the β-oxidation cycle shown in the diagram, and during the further oxidation of the β-oxidation product, acetyl-CoA, by the citric acid cycle. Numbers in parentheses indicate enzyme-catalyzed steps as follows: (1) fatty acyl-CoA synthetase; (2) acyl-CoA dehydrogenase; (3) enoyl-CoA hydratase; (4) L-3-hydroxyacyl-CoA dehydrogenase; and (4) acetyl-CoA acyltransferase (β-ketothiolase).

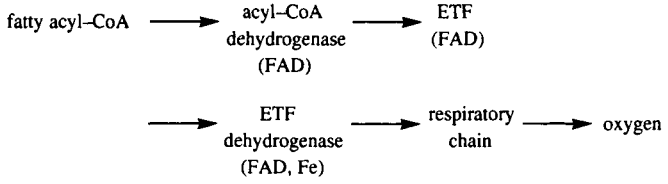
two pairs of hydrogen atoms from the substrate to form reduced cofactors (NADH + H^+ and Enz.$FADH_2$) and to regenerate a fatty acyl-CoA molecule chain-shortened by the removal of a two-carbon fragment in the form of acetyl-CoA:



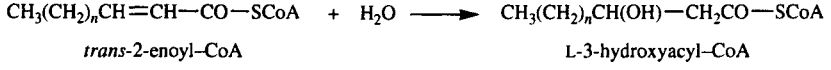
Step 1. This step is catalyzed by flavin-linked acyl-CoA dehydrogenase (Enz.FAD), which forms an unsaturated double bond at the α,β position of the substrate:



Animal tissues contain at least two distinct acyl-CoA dehydrogenases, and multiple forms of the other β-oxidation enzymes seem to be present. They differ in their chain-length specificity, and the variety of enzymes is probably required to accommodate the different physical properties of long-chain, medium-chain, and short-chain fatty acids. The reducing equivalents removed by acyl-CoA dehydrogenases are passed to the electron-transferring flavoprotein (ETF), which is in turn reoxidized by ETF dehydrogenase, a ferriflavoprotein linked to the respiratory chain:



Step 2. The enoyl-CoA formed by the first step is hydrated to form 3-hydroxyacyl-CoA in a reaction catalyzed by enoyl-CoA hydratase:



Step 3. In this step, hydroxyacyl-CoA is oxidized to the corresponding oxo compound in a reaction catalyzed by NAD-linked L-3-hydroxyacyl-CoA dehydrogenase:

$$CH_3(CH_2)_nCH(OH)—CH_2CO—SCoA + NAD^+ \longrightarrow CH_3(CH_2)_nCO—CH_2CO—SCoA + NADH + H^+$$

L-hydroxyacyl–CoA

3-oxoacyl–CoA

Step 4. In the final step, the oxoester is split by acetyl-CoA acyltransferase (β-ketothiolase) to form acetyl-CoA and a new shortened fatty acyl-CoA substrate is made available for further β-oxidation:

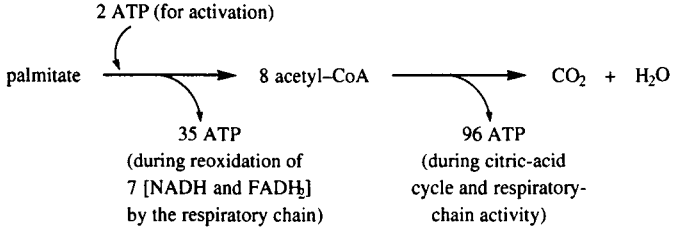
$$CH_3(CH_2)_nCO—CH_2CO—SCoA + CoASH \longrightarrow CH_3(CH_2)_nCO—SCoA + CH_3CO—SCoA$$

3-oxoacyl–CoA — fatty acyl–CoA — acetyl–CoA

The repeated β-oxidation cycle or spiral catalyzes the oxidation of saturated straight-chain fatty acids containing an even number of carbon atoms to acetyl-CoA, with the formation of reduced cofactors (NADH + H^+ and Enz.$FADH_2$). The complete oxidation of one palmitate molecule can be summarized:

$$C_{16}H_{31}COO^- + 8\ CoASH + ATP^{4-} + 7\ FAD + 7\ NAD^+ + 7\ H_2O \xrightarrow{\beta\text{-oxidation}} 8\ acetyl\text{–}CoA + AMP^{2-} + PP_i^{3-} + 7\ FADH_2 + 7\ (NADH + H^+)$$

The sustained activity of the β-oxidation pathway depends on the availability of the oxidized forms of the dehydrogenase cofactors NAD and Enz.FAD. This is achieved by reoxidation of the reduced cofactors via the respiratory chain. Three molecules of ATP are synthesized when one molecule of NADH is oxidized by the respiratory chain, and two molecules of ATP when $FADH_2$ is oxidized (see the section on ATP synthesis). Using this data, it can be calculated that the complete oxidation of one palmitate molecule gives rise to a maximum net synthesis of 35 + 96 – 2 = 129 molecules of ATP:

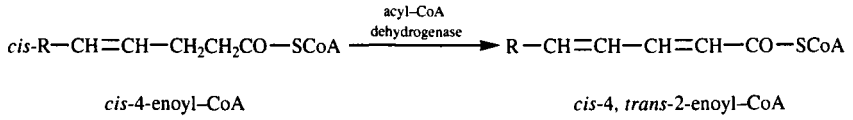


Oxidation of Unsaturated Fatty Acids

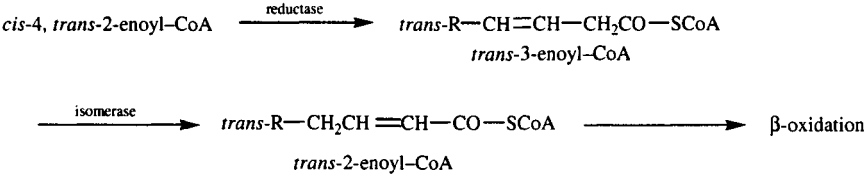
The oxidation of long-chain unsaturated fatty acids is an important energy source. Two properties of their chemical structure prevent their complete metabolism by the β-oxidation pathway, namely the presence of double bonds in the *cis* configuration, and the spacing of double bonds at 3-carbon intervals in the fatty

acid hydrocarbon chain. At least two additional enzymes are required to convert intermediates formed during unsaturated fatty acid oxidation into normal intermediates of the β-oxidation system (Schulz and Kunau, 1987).

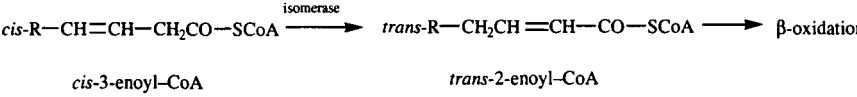
Unsaturated fatty acids with *cis* double bonds extending from even-numbered carbon atoms are chain-shortened to *cis*-4-enoyl-CoA, which undergoes dehydrogenation:



The product is reduced to *trans*-3-enoyl-CoA by the NADP-linked enzyme 2,4-dienoyl-CoA reductase. This reaction is followed by an isomerization step catalyzed by *cis*-3,*trans*-2-enoyl-CoA isomerase, to form *trans*-enoyl-CoA, leading to further β-oxidation:



Chain shortening of unsaturated fatty acids with *cis* bonds extending from odd-numbered carbon atoms yield *cis*-3-enoyl-CoA. The same enoyl-CoA isomerase catalyzes the conversion of the *cis*-3 isomer to the *trans*-2 isomer, which can then enter the β-oxidation spiral:



Defects in Fatty Acid Oxidation

The consequences of any defect in fatty acid oxidation in human patients can be dramatic, especially when the availability of glucose is limited. They are characterized by a lack of energy for vital tissues, and episodes of hypoketotic hypoglycemia. The diminished production of acetyl-CoA causes both a block in gluconeogenesis (because acetyl-CoA normally acts as an allosteric activator of pyruvate carboxylase), and an impairment in ketone body formation. The action of other fatty-acid metabolizing enzymes (located in the endoplasmic reticulum and the peroxisomes) on accumulated β-oxidation intermediates produces a massive excretion of dicarboxylic acids in the urine (dicarboxylic aciduria). Defects in fatty acid oxidation are caused by disorders in the transport of fatty acids into mitochondria, or in their oxidation (Angelini et al., 1987). In patients suffering from a

deficiency of carnitine or carnitine palmitoyltransferase (CPT), the main histological finding is the presence of excess muscle fat. In these patients, the provision of energy from carbohydrate oxidation is sufficient to prevent the morphological changes seen in other mitochondrial myopathies. The primary gene-linked carnitine deficiency syndromes include a systemic form (systemic carnitine deficiency or SCD) and a muscle or myopathic form (muscle carnitine deficiency syndrome or MCD). In systemic carnitine deficiency, excessive amounts of carnitine and acylcarnitine are excreted in the urine, and this causes a decrease in the carnitine content of the blood and body tissues. In muscle carnitine deficiency, the carnitine content of liver and blood is normal and the condition is thought to arise from defective carnitine transport into muscle. Carnitine deficiency also occurs as a secondary consequence of other inborn errors of metabolism, including the organic acidemias. Two forms of CPT deficiency have been recognized, a CPT I deficiency in liver, and a CPT II deficiency in muscle.

Five inborn errors in the mitochondrial β-oxidation system have been identified. The first three of these are deficiencies in the activity of long-chain, medium-chain, and short-chain acyl-CoA dehydrogenases, referred to as LCAD, MCAD, and SCAD deficiency, respectively. The other two errors are variant forms of multiple acyl-CoA dehydrogenase deficiency (MAD), probably caused by defects in the electron-transferring flavoprotein (ETF) and in ETF dehydrogenase. MCAD is the most frequent error found, and it is one of the most common metabolic disorders known. Estimates suggest that its occurrence is about one in 5,000–10,000 individuals. It is seen within the first two years of life after a period of fasting, when body tissues become strongly dependent on fatty acid oxidation as a source of energy. Typical symptoms include vomiting, lethargy, and coma. This deficiency, and possibly other defects in fatty acid oxidation, seem to be present in some cases of sudden infant death syndrome (SIDS; also called cot death or crib death). A number of inherited metabolic disorders, including disorders of fatty acid metabolism, have been associated with Reye's syndrome, a severe and often fatal illness of children and teenagers. The syndrome is characterized by encephalopathy and fatty degeneration of the viscera, especially the liver. Brain and liver mitochondria show abnormalities to a degree which seem to parallel the severity of the illness. Normal mitochondrial structure and function are restored in patients responding to treatment, which includes blood transfusion and intensive hospital care. The condition, therefore, actually seems to be due to specific mitochondrial damage which is acquired and transient in nature.

Ketone Body Formation and Oxidation

The enzymes responsible for ketone-body metabolism occur mainly in mitochondria. The relationships between fatty acid oxidation and ketone-body formation are shown in Figure 7. Acetoacetyl-CoA, the starting material for ketone-body

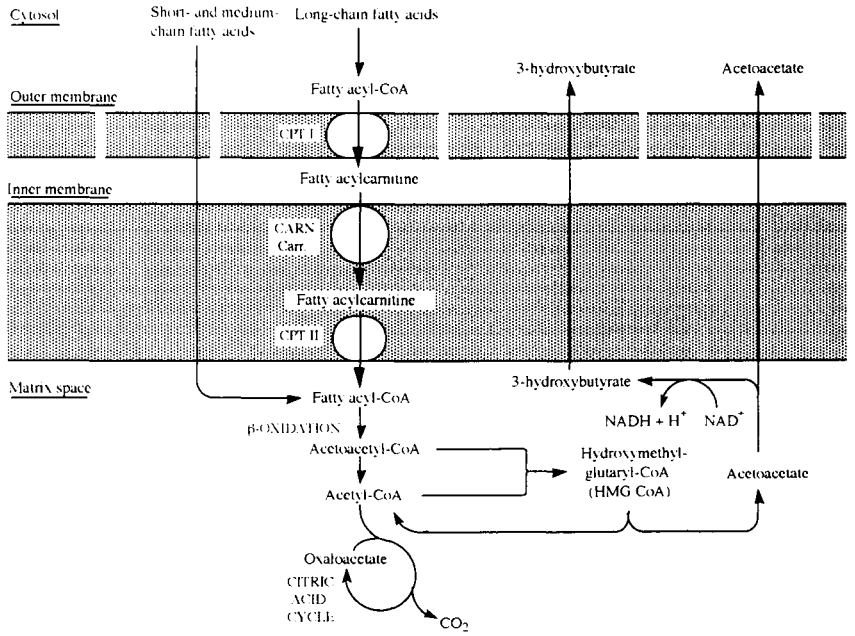
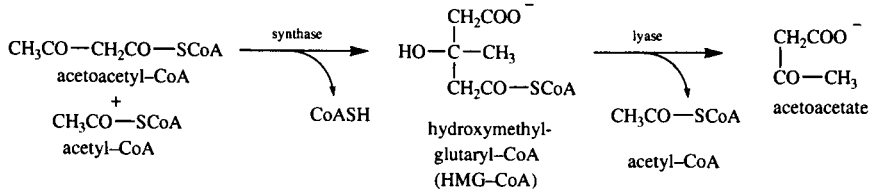


Figure 7. Relationship between fatty acid oxidation and ketone body formation in the matrix space of liver mitochondria. CPT I, CPT II, the two forms of carnitine palmitoyltransferase; CARN Carr., the carnitine/acylcarnitine exchange carrier.

formation, arises during β-oxidation and also by the condensation of two molecules of acetyl-CoA, in a reaction catalyzed by acetyl-CoA acetyltransferase:

$$\underset{\text{acetyl–CoA}}{2\ CH_3CO{-}SCoA} \longrightarrow \underset{\text{acetoacetyl–CoA}}{CH_3CO{-}CH_2CO{-}SCoA} + CoASH$$

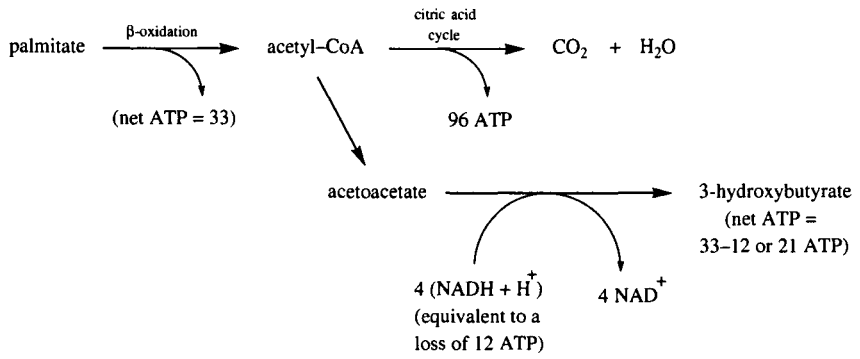
The conversion of acetoacetyl-CoA to acetoacetate occurs mainly via the intermediate formation of hydroxymethylglutaryl-CoA, the HMG-CoA pathway. In the first reaction of the pathway, catalyzed by hydroxymethylglutaryl-CoA synthase, acetoacetyl-CoA condenses with a molecule of acetyl-CoA to form hydroxymethylglutaryl-CoA. The second reaction, catalyzed by hydroxymethylglutaryl-CoA lyase, splits the substrate into acetyl-CoA and free acetoacetate:



Although the acetyl group released by the lyase is not the same as that incorporated by the synthase in the first reaction, in effect acetyl-CoA plays a catalytic role in acetoacetate formation. Liver mitochondria also contain the enzyme fatty acyl-CoA hydrolase, which catalyzes the direct hydrolysis of acetoacetyl-CoA to form acetoacetate. The activity of this enzyme is much lower than that of the HMG-CoA pathway enzymes. The function of the hydrolase could be to prevent the accumulation of high concentrations of fatty acyl-CoA esters in the matrix space, and to ensure that some free CoASH is always available for metabolism. Acyl-CoA esters are potent inhibitors of several enzymes and metabolite carriers, including the adenine nucleotide carrier. Part of the acetoacetate formed in the liver is reduced to 3-hydroxybutyrate by an NAD-linked D-3-hydroxybutyrate dehydrogenase situated on the M-face of the inner membrane:

$$\underset{\text{acetoacetate}}{CH_3CO{-}CH_2COO^-} + NADH + H^+ \rightleftharpoons \underset{\text{D-3-hydroxybutyrate}}{CH_3CH(OH){-}CH_2COO^-}$$

Low rates of fatty acid oxidation are accompanied by the formation mainly of acetoacetate, whereas higher rates result in an increased formation of 3-hydroxybutyrate. Acetoacetate reduction provides a means for the oxidation of reducing equivalents generated in excess of those required for ATP synthesis. For example, when 3-hydroxybutyrate is the product of palmitate oxidation, theoretically about 6 times more fatty acid (129/21) must be oxidized, to obtain an equivalent ATP yield, than is obtained by complete β-oxidation:



Ketone bodies formed in the matrix space of liver mitochondria pass to the cytosol mainly by simple diffusion across the inner membrane in the form of undissociated acids. Some acetoacetate is thought to leave mitochondria by exchange with pyruvate on the monocarboxylate carrier. Ketone bodies are carried in the blood to the extrahepatic tissues, where they are oxidized by the concerted action of three mitochondrial enzymes (Figure 8). The first enzyme, D-3-hydroxybutyrate dehydrogenase, catalyzes the oxidation of 3-hydroxybutyrate to acetoacetate, a direct reversal of the reaction in liver responsible for its formation. Acetoacetate is

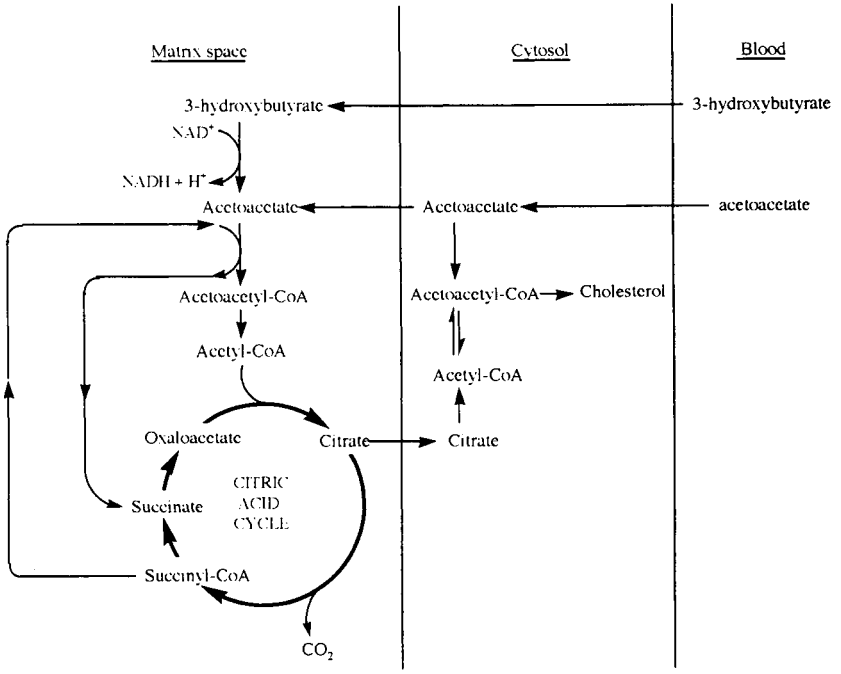
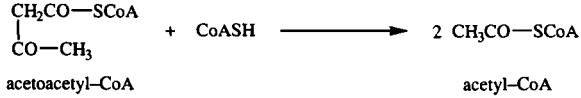


Figure 8. Pathways for the oxidation of ketone bodies carried in the blood to the peripheral (extrahepatic) tissues.

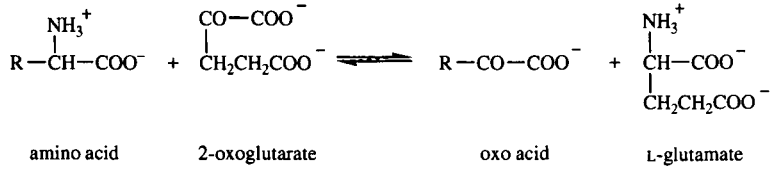
activated by reaction with the energy-rich compound succinyl-CoA to form acetoacetyl-CoA and succinate, in a reaction catalyzed by 3-oxoacid-CoA transferase:

Succinyl-CoA is regenerated from succinate by the action of GTP-specific succinyl-CoA synthetase. In the third and final step, acetoacetyl-CoA is split to form two molecules of acetyl-CoA (which can enter the citric acid cycle), in a reaction catalyzed by acetyl-CoA acetyltransferase, the same enzyme as that responsible for acetoacetyl-CoA synthesis in the liver:



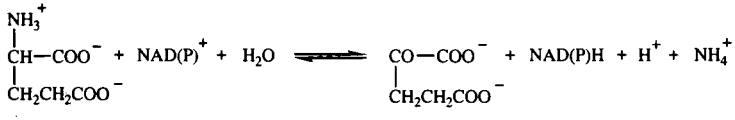
Amino Acid Catabolism as a Source of Acetyl-CoA

In nearly all cases, an early step in the catabolism of individual amino acids is the transfer of the α-amino group of one amino acid to the oxoacid of another, usually 2-oxoglutarate, in a reaction catalyzed by an aminotransferase (transaminase):



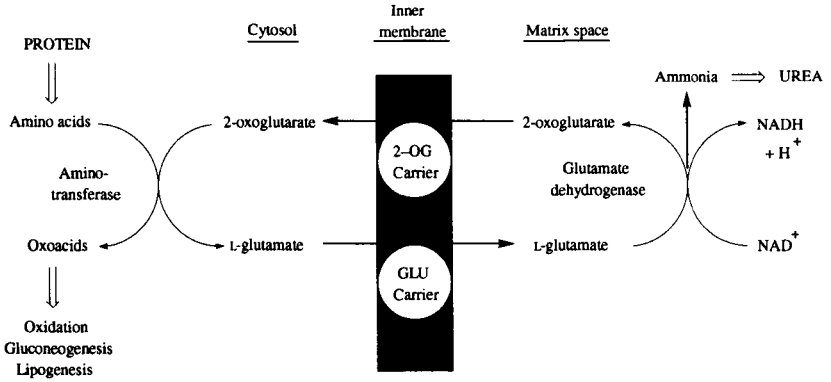
amino acid 2-oxoglutarate oxo acid L-glutamate

The aminotransferase enzymes are specific for a single pair of α-amino acids and α-oxoacids. They all require the cofactor pyridoxal phosphate, which is the active form of vitamin B_6 (pyridoxine, pyridoxal, and pyridoxamine). The glutamate formed during aminotransferase reactions can then undergo oxidative deamination to release ammonia and regenerate 2-oxoglutarate, in a reaction catalyzed by glutamate dehydrogenase:



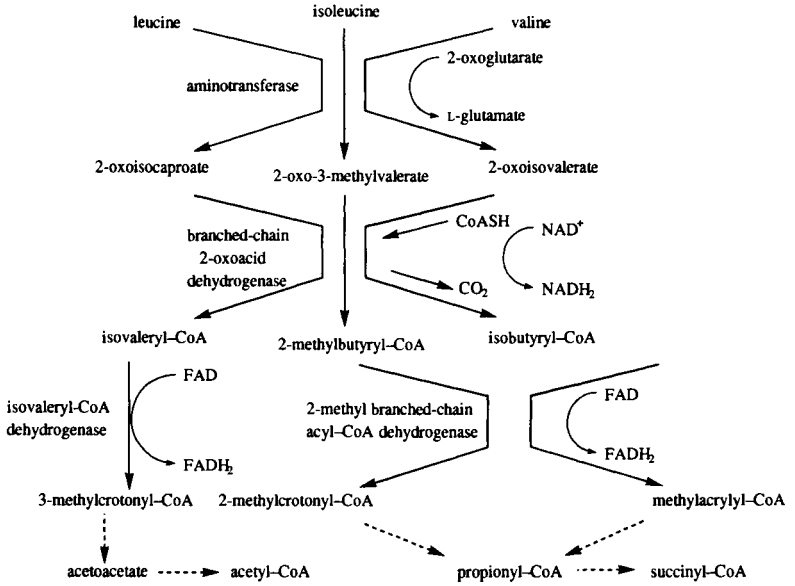
L-glutamate 2-oxoglutarate

The coupled action of the aminotransferases and glutamate dehydrogenase, a process called transdeamination, provides a major route for the conversion of amino acids into oxoacids and ammonia. Most of the aminotransferase activity of mammalian cells occurs in the cytosol, but glutamate dehydrogenase is located exclusively in the matrix space of mitochondria. Glutamate formed in the cytosol by aminotransferase activity is transported to its dehydrogenase in the matrix space by the glutamate carrier present in the inner membrane, and 2-oxoglutarate formed by the dehydrogenase reaction can return to the cytosol on the oxoglutarate carrier to continue the process:



Because the reactions catalyzed by these enzymes are close to equilibrium, the overall process can be reversed readily when oxoacids are available, so that amino acids can be synthesized as well as degraded. In well-fed animals, glutamate dehydrogenase is involved mainly in the reactions of transdeamination, with the formation of 2-oxoglutarate. During amino acid synthesis, the oxoacids commonly involved in aminotransferase reactions, namely 2-oxoglutarate, oxaloacetate, and pyruvate, can be supplied by the metabolism of glucose.

The first two steps in the catabolism of the branched-chain amino acids (leucine, isoleucine, and valine) are catalyzed by an aminotransferase reaction, especially in muscle, followed by the transport in blood of the oxoacid formed to the liver, where it undergoes oxidative decarboxylation:



The irreversible oxidative decarboxylation reaction is catalyzed by the branched-chain 2-oxoacid dehydrogenase complex, which shows many structural and functional similarities to the pyruvate dehydrogenase complex (Yeaman, 1989). Both complexes consist of three enzyme components and are regulated by a phosphorylation-dephosphorylation mechanism. This specialized regulation is related to the fact that the three branched-chain amino acids are essential requirements in the diets of mammalian species. They can be synthesized from their corresponding oxoacids by a reversal of the aminotransferase reactions shown in the diagram above. Regulation of the branched-chain 2-oxoacid dehydrogenase complex protects the oxoacids against further degradation when other cell fuels are available, and it enables essential amino acids to be regenerated. The products of the dehydrogenase reaction undergo further oxidation in reactions catalyzed by two distinct acyl-CoA dehydrogenases, one specific for isovaleryl-CoA and the other specific for 2-methyl

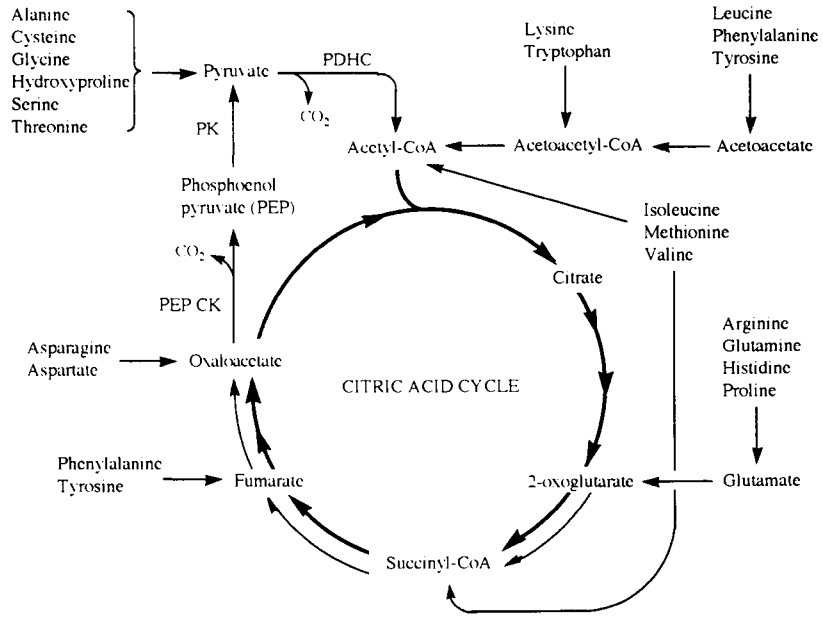


Figure 9. Diagram showing the pathways used for oxidation of the carbon skeletons of amino acids. Skeletons which yield citric-acid-cycle intermediates are converted into acetyl-CoA via pyruvate, along the pathway shown by the thin arrows. PEP CK, phosphoenolpyruvate carboxykinase; PK, pyruvate kinase; PDHC, pyruvate dehydrogenase complex.

branched-chain acyl-CoA. Mammalian mitochondria, therefore, contain at least five different acyl-CoA dehydrogenases, three functional in the β-oxidation of straight-chain fatty acids, and two involved in the metabolism of branched-chain amino acids.

The amino acids can be grouped into two main classes, according to the site of entry by their oxoacid carbon atoms into the citric acid cycle. The catabolism of about one-half of the common amino acids yields acetyl-CoA, after conversion to pyruvate or acetoacetyl-CoA. Their carbon atoms are thus made available directly for oxidation by the citric acid cycle. The carbon atoms of the remaining amino acids are converted to citric-acid-cycle intermediates (Figure 9). In two cases (glutamate and aspartate) the oxoacid products of transamination (2-oxoglutarate and oxaloacetate) are themselves cycle intermediates, but in other cases the oxoacids undergo further metabolism in preparation for their oxidation. The entry of the carbon skeletons derived from amino acids into the cycle at various points does not in itself ensure their complete oxidation, because cycle intermediates act in a catalytic manner and are regenerated during the oxidation of the acetyl unit of acetyl-CoA. In order to undergo net oxidation, intermediates must first be converted

to acetyl-CoA by a pathway in which the enzyme phosphoenolpyruvate carboxykinase plays a key role (Figure 9).

Organic Acidemias

A variety of defects in the mitochondrial metabolism of fatty acids and amino acids are associated with an increase in content of organic acids in the blood and in the urine (organic acidemia) (Devlin, 1992). In these syndromes, organic acids undergo chain-shortening and accumulate at the point of the metabolic defect in the form of acyl-CoA esters, which have potentially toxic effects against a number of important metabolic pathways including the citric acid cycle. For example, propionyl-CoA is an inhibitor of citrate synthase, 2-oxoglutarate dehydrogenase, and succinyl-CoA synthetase. Carnitine protects the cell against these effects by accepting acyl residues from coenzyme A, and the excess acylcarnitine esters formed in this way are excreted in the urine and cause secondary carnitine deficiency. This condition can often be alleviated by oral carnitine therapy. Patients with multiple acyl-CoA dehydrogenase deficiency accumulate medium-chain monocarboxylic acids within the mitochondria, which then undergo further metabolism in the surrounding cytoplasm before excretion in the urine as C_6–C_{10} dicarboxylic acids, including adipic (C_6) and suberic (C_8) acids. Patients with deficiencies in the branched-chain 2-oxoacid dehydrogenase complex, functional in the catabolism of valine, leucine, and isoleucine (see the section on Amino Acid Catabolism), excrete oxoacid intermediates of the pathways and their corresponding hydroxyacids. An unidentified excretory product is responsible for a characteristic odor which gives the name maple-syrup urine disease to this condition. A large proportion of the cases show serious mental retardation, ketoacidosis, and a short life span. Two other errors of leucine metabolism, which give rise to specific patterns of organic acids in the urine, are due to a deficiency of isovaleryl-CoA dehydrogenase (isovaleryl acidemia) and a deficiency of HMG-CoA lyase. Glutaric acidemia is caused by an inherited deficiency of glutaryl-CoA dehydrogenase or a component in the ETF oxidation pathway. Organic acidemia is also present in patients suffering from defects in the pathway responsible for the conversion of propionyl-CoA to succinyl-CoA (see the section on Fatty Acid Oxidation).

Oxidation of Acetyl-CoA by the Citric Acid Cycle

Isoenzyme forms of some citric-acid-cycle enzymes, including fumarase and malate dehydrogenase, have a dual location in the cell, with one form present in the matrix space of mitochondria and the other form in the cytosol. Other enzymes, such as citrate synthase and the 2-oxoglutarate dehydrogenase complex, are located exclusively in mitochondria. The complete citric acid cycle therefore functions in the matrix space, the only region of the cell where all eight enzymes of the cycle

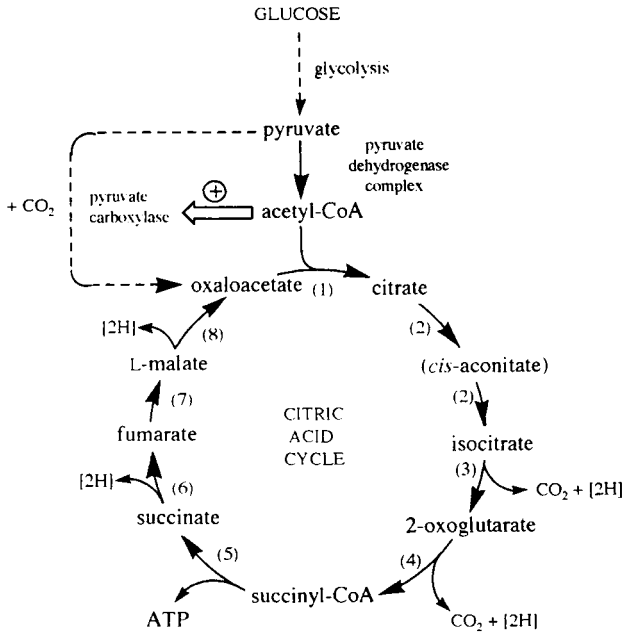


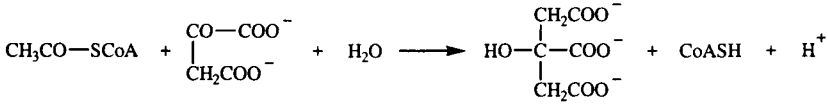
Figure 10. The reaction steps of the citric acid cycle, showing how it functions during the oxidation of glucose. The metabolism of pyruvate produced by glycolysis can provide both acetyl-CoA and oxaloacetate required for the synthesis of citrate, the first reaction step in the cycle. The catabolism of fatty acids and some amino acids also provides important sources of acetyl-CoA for oxidation by the cycle. During each turn of the cycle, the entry of two acetyl-carbon atoms is balanced exactly by the loss of two carbon atoms in the form of CO_2. Numbers in parentheses indicate the eight enzyme-catalyzed steps of the cycle as follows: (1) citrate synthase; (2) aconitase; (3) isocitrate dehydrogenase; (4) 2-oxoglutarate dehydrogenase complex; (5) succinyl-CoA synthetase (thiokinase); (6) succinate dehydrogenase (part of respiratory-chain Complex II); (7) fumarase; and (8) malate dehydrogenase.

are located. The cycle consists of two metabolic pathways, one for the metabolism of tricarboxylic acids and the other for the metabolism of dicarboxylic acids (Newsholme and Leech, 1983). (At physiological pH, these compounds exist in the cell largely in their anion form, as tricarboxylates and dicarboxylates.) The eight reaction steps of the cycle are summarized in Figure 10. With the exception of succinate dehydrogenase, the enzymes can be released in non-sedimentable form when the mitochondrial membranes are disrupted. Succinate dehydrogenase is tightly-bound to the inner membrane, where it forms part of the respiratory-chain Complex II. The catalytic center of the dehydrogenase lies at the M-face of the membrane, and is therefore accessible to succinate generated in the matrix space during citric-acid-cycle activity. The 2-oxoglutarate dehydrogenase complex, with

a molecular weight of 2.7×10^6 Da, is very much larger than the other citric acid cycle enzymes, whose individual molecular weights do not exceed 2×10^5 Da. The dehydrogenase complex is believed to provide a structural nucleus around which the remaining cycle enzymes are assembled to form a metabolon. This arrangement increases the efficiency of the cycle by providing microenvironments where locally high concentrations of pathway intermediates can be maintained. It also ensures the rapid vectorial transfer of intermediates between catalytic centers of the cycle enzymes.

Metabolism of Tricarboxylates

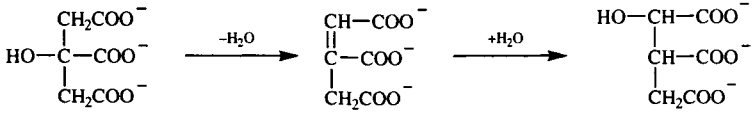
Step 1. The sequence begins with the combination between acetyl-CoA and oxaloacetate to form citrate. The reaction, catalyzed by citrate synthase (formerly called citrate condensing enzyme) occurs with a large change in free energy. This ensures that the reaction is irreversible and proceeds to completion, despite the very low concentration of oxaloacetate in the matrix space:



acetyl-CoA oxaloacetate citrate

For each molecule of citrate metabolized by the subsequent reactions of the cycle, one molecule of oxaloacetate is regenerated to enable additional molecules of acetyl-CoA to be oxidized.

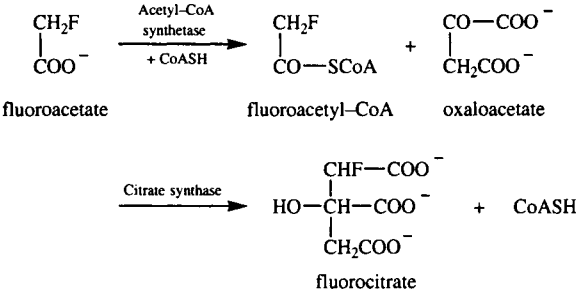
Step 2. Citrate is converted to *threo*-D_S-isocitrate by the isomerase enzyme aconitase (aconitate hydratase). The reaction results in the transfer of the -OH group from carbon 3 to carbon 2 of the citrate molecule. The mechanism is thought to involve the dehydration of citrate to form enzyme-bound *cis*-aconitate, which is then rehydrated to yield isocitrate:



citrate *cis*-aconitate *threo*-D_S-isocitrate

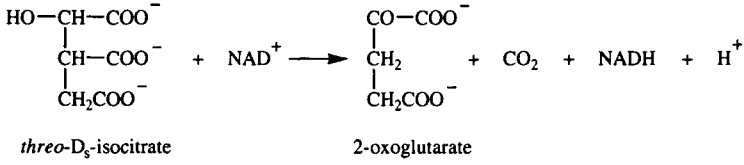
The isocitrate molecule contains two chiral atoms, but only one of the four possible isomers of isocitrate is formed in the stereospecific reaction catalyzed by aconitase. The enzyme contains an iron-sulfur center [4Fe-4S]. Unlike the iron-sulfur centers in the respiratory-chain complexes, this center does not undergo oxidoreduction during catalytic activity, but functions instead as a substrate-binding site at the catalytic center of the enzyme.

The toxic citrate analogue fluorocitrate is a powerful competitive inhibitor of aconitase. Cell membranes are relatively impermeable to fluorocitrate, because of its size and multiple electrical charges. It can however be formed within the cell from fluoroacetate, which permeates cell membranes much more freely. Once inside the matrix space, fluoroacetate is converted first to fluoroacetyl-CoA, and then to fluorocitrate:



Fluoroacetate, which is used as a rodenticide, is therefore more toxic than fluorocitrate when given orally. In several animals it has a similar toxicity to cyanide. The minimum estimated oral dose causing death in the human adult is 2 mg/kg body weight.

Step 3. In the third reaction step, isocitrate is converted into 2-oxoglutarate (also called α-ketoglutarate). The oxidative decarboxylation of isocitrate, catalyzed by isocitrate dehydrogenase, achieves both the essential functions of the cycle, namely the transfer of reducing equivalents to a cofactor (NAD), and the release of CO_2. The conversion of isocitrate to 2-oxoglutarate is catalyzed by two isocitrate dehydrogenases, one specific for NAD and the other specific for nicotinamide adenine dinucleotide phosphate (NADP). In animal tissues, the NAD-linked enzyme is located exclusively in the matrix space of mitochondria, whereas the NADP-linked enzyme is present in both the matrix space and the cytosol. Both types of dehydrogenase require Mg^{2+} or Mn^{2+} for activity. The NAD-linked enzyme, which functions in the citric acid cycle, catalyzes a regulated non-equilibrium reaction:

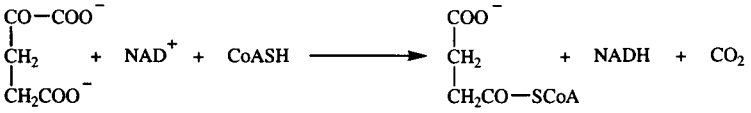


It is an allosteric enzyme regulated by ADP, which activates by lowering the K_m for isocitrate without changing the maximal velocity of the reaction. The NADP-linked enzyme could also be involved in cycle activity, at least when it approaches the maximal rate. In liver mitochondria, the NADP-linked enzyme provides reducing equivalents during the reductive amination of 2-oxoglutarate (glutamate dehy-

drogenase reaction) when urea is synthesized from ammonia. The involvement of oxalosuccinate as an enzyme-bound intermediate of the reaction has been established only for the NADP-linked enzyme.

Metabolism of 2-Oxoglutarate

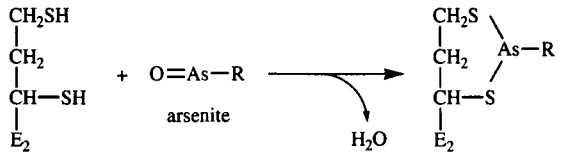
Step 4. The tricarboxylate and dicarboxylate pathways of the cycle are linked by the oxidative decarboxylation of 2-oxoglutarate to form the energy-rich thioester compound succinyl-CoA, in a reaction catalyzed by the 2-oxoglutarate dehydrogenase complex:



2-oxoglutarate succinyl-CoA

This complex shows many similarities to the pyruvate dehydrogenase complex, but unlike the pyruvate complex it is not regulated by phosphorylation and dephosphorylation. The complex contains three physically-associated enzymes, analogous to those of the pyruvate system and also denoted E_1, E_2, and E_3. In the 2-oxoglutarate dehydrogenase complex, the enzymes are E_1 or 2-oxoglutarate dehydrogenase (decarboxylating); E_2 or dihydrolipoate-succinyl transferase; and E_3 or dihydrolipoamide dehydrogenase. Enzymes E_1 and E_2 are different proteins from the E_1 and E_2 enzymes of the pyruvate dehydrogenase complex, but enzyme E_3 is identical in both complexes.

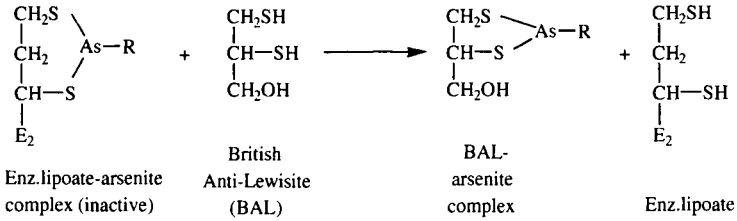
Arsenites and other trivalent arsenic compounds (represented by O=As-R) block the activity of 2-oxoacid dehydrogenase complexes by combining with the adjacent sulfhydryl groups of the lipoate cofactor of dihydrolipoamide acetyltransferase (E_2):



Enz.lipoate Enz.lipoate-arsenite complex (inactive)

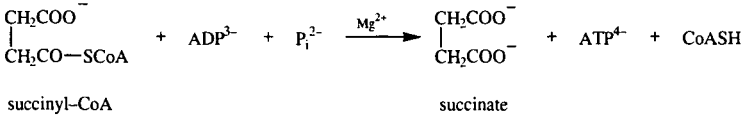
The action of arsenites on the 2-oxoglutarate dehydrogenase complex blocks the activity of the citric acid cycle and therefore abolishes mitochondrial ATP synthesis. Because the animal body contains very small amounts of lipoate cofactor, trivalent arsenical compounds are very toxic. The lethal dose of sodium arsenite in mammalian species is about 8 mg/kg body weight, compared with 1–2 mg/kg body weight for cyanide. Toxic arsendichlorides such as Lewisite ($Cl.CH{=}CH.AsCl_2$)

have been developed for use in chemical warfare. The inhibition of oxoacid dehydrogenase activity by arsenicals cannot be reversed by monothiol compounds. It is however reversed by dithiol compounds, including 2,3-dimercaptopropanol (British Anti-Lewisite or BAL), which is an effective antidote to arsenical poisoning:



Metabolism of Dicarboxylates

Step 5. Succinyl-CoA is converted to succinate by two enzymes, succinyl-CoA synthetase (succinate thiokinase), which is specific for adenine nucleotides, and 3-oxoacid CoA-transferase (succinyl-CoA-acetoacetate transferase) (Jenkins and Weitzman, 1988). The transferase reaction is important during the oxidation of ketone bodies (see the section on Defects in Fatty Acid Oxidation). In many tissues, the conversion of succinyl-CoA is catalyzed mainly by succinyl-CoA synthetase. The large free-energy change occurring during splitting of the thioester bond of succinyl-CoA is coupled to the substrate-level phosphorylation of ADP to form ATP:

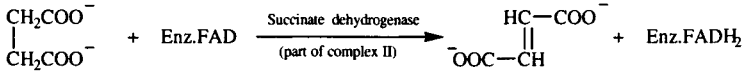


The reaction proceeds via the formation of enzyme-bound succinyl-phosphate and histidine-phosphate. In the citric acid cycle the energy released during acetyl-CoA metabolism by citrate synthase is used to drive citrate synthesis to completion, while in succinyl-CoA metabolism it is conserved as ATP. The matrix space contains a second succinyl-CoA synthetase specific for guanine nucleotides, which functions during ketone-body metabolism and in heme biosynthesis. The substrate-level phosphorylation reaction is an important source of ATP in brown adipose tissue. During nonshivering thermogenesis in this tissue, most or all of the respiratory-chain-linked ATP synthesis is uncoupled, and the tissue relies on substrate-level phosphorylation reactions in glycolysis and the citric acid cycle to supply ATP.

In the matrix space, GTP is synthesized from ATP in a reaction catalyzed by nucleosidediphosphate kinase (GDP + ATP → GTP + ADP). The inner membrane does not contain a carrier system for guanine nucleotides, and GTP generated by the forward reaction is therefore confined to the matrix space, where it is consumed in several GTP-dependent reactions, including the phosphoenolpyruvate car-

boxykinase reaction. It is also used in the conversion to ADP of AMP formed during the activation of medium-chain and short-chain fatty acids (AMP + GTP → ADP + GDP), a reaction catalyzed by Mg^{2+}-dependent GTP-AMP phosphotransferase, during mitochondrial protein synthesis, and in succinyl-CoA synthesis during ketone body oxidation.

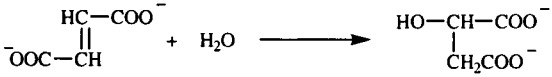
Step 6. Succinate is converted to fumarate (*trans*-isomer) by succinate dehydrogenase, a ferriflavoprotein containing covalently-bound FAD as prosthetic group (represented as Enz.FAD):



succinate fumarate

The dehydrogenase is tightly-bound to the inner membrane with the catalytic center of the enzyme facing the matrix space. Under physiological conditions, the enzyme forms part of the respiratory-chain Complex II (succinate-Q reductase). The reduced flavin of the dehydrogenase ($Enz.FADH_2$) is reoxidized in the respiratory chain by way of the iron-sulfur centers in Complex II and the ubiquinone pool.

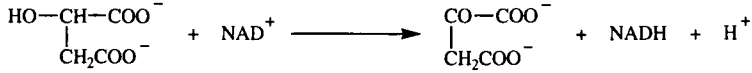
Step 7. Fumarate is hydrated by fumarase (fumarate hydratase) to form L-malate:



fumarate L-malate

Fumarase is inactive towards the *cis*-isomer of fumarate (maleate) and D-malate and therefore it shows both geometric-isomer and optical-isomer specificity.

Step 8. In the final reaction step of the cycle, oxaloacetate consumed in step 1 is regenerated by the action of L-malate dehydrogenase:



L-malate oxaloacetate

This final reaction step regenerates a molecule of oxaloacetate, so that a further molecule of acetyl-CoA can enter the cycle for oxidation. By repetition of the reaction sequence, one molecule of oxaloacetate (or any other cycle intermediate) can act catalytically to stimulate the oxidation of many 2-carbon acetyl groups. During a single turn of the cycle, two carbon atoms are released as CO_2, four pairs of hydrogen atoms are transferred to cofactors (as 3 [NADH + H^+] and $Enz.FADH_2$), and one molecule of ATP is synthesized by substrate-level phosphorylation. Further ATP is synthesized during reoxidation of the reduced cofactors by the respiratory chain. The equations representing pyruvate oxidation by way of the pyruvate dehydrogenase complex and the citric acid cycle can be written:

$$CH_3CO\text{–}COOH + CoASH + NAD^+ \rightarrow CH_3CO\text{–}SCoA + CO_2 + NADH + H^+ \quad (1)$$

$$CH_3CO\text{–}SCoA + 3\,NAD^+ + Enz.FAD + ADP^{3-} + P_i^{2-} + 2\,H_2O \rightarrow 2\,CO_2 + 3\,NADH + Enz.FADH_2 + ATP^{4-} + 2\,H^+ + CoASH \quad (2)$$

The reoxidation of the reduced cofactors NADH and $FADH_2$ by the respiratory chain occurs according to the equation:

$$4\,(NADH + H^+) + Enz.FADH_2 + 21/2\,O_2 \rightarrow 4\,NAD^+ + Enz.FAD + 5\,H_2O \quad (3)$$

To balance the equations, ATP must be hydrolyzed to ADP and P_i thus:

$$ATP^{4-} + H_2O \rightarrow ADP^{3-} + P_i^{2-} + H^+ \quad (4)$$

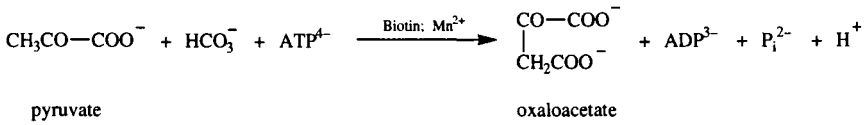
The sum of reactions (1) to (4) gives a balanced equation for the oxidation of pyruvate to CO_2 and H_2O:

$$CH_3CO\text{–}COOH + 21/2\,O_2 = 3\,CO_2 + 2\,H_2O \quad (5)$$

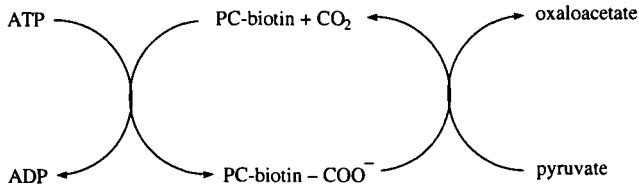
During reaction 3, three ATP molecules are synthesized per molecule of NADH oxidized, and two ATP molecules per molecule of $FADH_2$ oxidized. In addition, one molecule of ATP is synthesized by substrate-level phosphorylation catalyzed by succinyl-CoA synthetase (reaction 2). Therefore during the oxidation of one molecule of pyruvate to CO_2 and H_2O, a total of $[4 \times 3] + [1 \times 2] + 1 = 15$ molecules of ATP are formed (see also Table 5).

Carboxylation and Decarboxylation Reactions

During one turn of the citric acid cycle, for each molecule of oxaloacetate consumed by citrate synthase, one molecule of oxaloacetate is regenerated by the subsequent reaction steps, provided that all the intermediates are retained in the pathway. Although this can occur to a large extent in muscle, where intermediates are largely retained for oxidation and ATP synthesis, in other tissues (especially the liver) many of the intermediates are removed from the cycle for use in biosynthetic reactions. On the other hand, excessive amounts of citric-acid-cycle intermediates can be formed, especially during amino acid catabolism. Regulation of the concentration of cycle intermediates in mitochondria is achieved by the action of enzymes in the matrix space and cytosol which catalyze carboxylation and decarboxylation reactions. Their relationship to the citric acid cycle is shown in Figures 9 and 10. The mitochondrial enzyme pyruvate carboxylase catalyzes the ATP-dependent synthesis of oxaloacetate from pyruvate and bicarbonate:

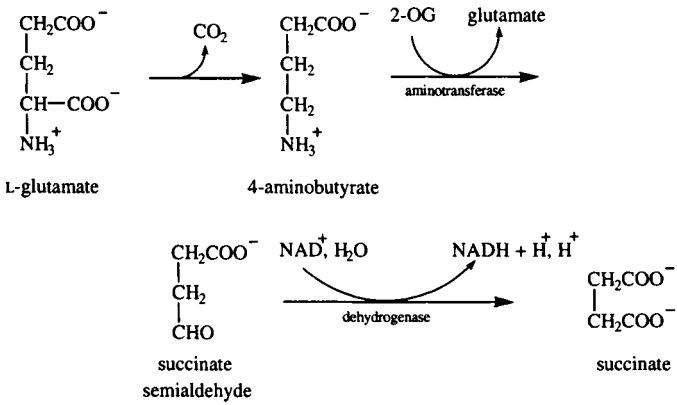


Pyruvate carboxylase (PC) is an allosteric enzyme containing four subunits, four molecules of biotin bound to lysyl residues, and four manganese atoms per molecule of enzyme. In common with other biotinylated enzymes, ATP hydrolysis is coupled to the carboxylation of biotin to form N-1′-carboxybiotin, which then carboxylates the substrate:



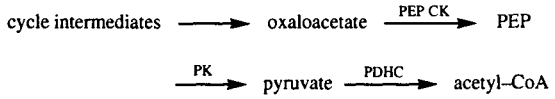
Acetyl-CoA functions as an allosteric activator of the enzyme. This important property provides a rapid and flexible regulation in provision of the substrates acetyl-CoA and oxaloacetate for the citrate synthase reaction. When acetyl-CoA accumulates because oxaloacetate or other intermediates have been removed from the cycle, the activation of pyruvate carboxylase by acetyl-CoA stimulates the formation of additional oxaloacetate, and the oxidation of acetyl units is maintained (Figure 10).

In the brain, up to one-third of the 2-oxoglutarate formed in the citric acid cycle is removed by conversion to glutamate, which passes to the cytosol before undergoing decarboxylation to form the neurotransmitter 4-aminobutyrate (also called γ-aminobutyric acid or GABA). These reactions, together with the uptake of 4-aminobutyrate by mitochondria and further metabolism to yield succinate, constitute the GABA shunt pathway:

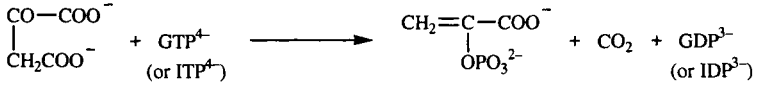


The pathway spans two different compartments, one responsible for the synthesis and release of 4-aminobutyrate at presynaptic terminals and the other for its uptake and metabolism in glial cells. When dehydrogenase activity is defective, succinate semialdehyde becomes reduced to form 4-hydroxybutyrate, and causes 4-hydroxybutyric acidemia.

When excessive amounts of cycle intermediates become available for oxidation, they must first be converted to acetyl-CoA. As shown in Figure 9, this is achieved by the sequential actions of phosphoenolpyruvate carboxykinase (PEP CK), pyruvate kinase (PK) and the pyruvate dehydrogenase complex (PDHC):

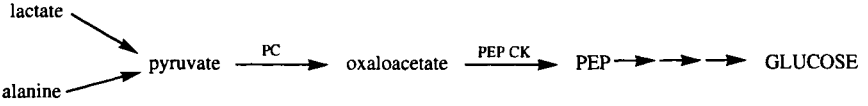


In the first reaction, oxaloacetate is decarboxylated by phosphoenolpyruvate carboxykinase to form phosphoenolpyruvate (PEP):



oxaloacetate phosphoenolpyruvate

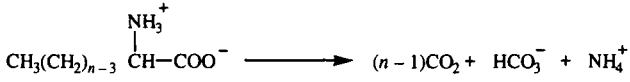
In human liver, this enzyme is distributed equally between the matrix space and the cytosol. The sequential action of pyruvate carboxylase and phosphoenolphosphate carboxykinase is responsible for the conversion of pyruvate to phosphoenolpyruvate in the gluconeogenic pathway of the liver cell, and by-passes the pyruvate kinase reaction of glycolysis, which is irreversible:



The activation of pyruvate carboxylase and inhibition of pyruvate dehydrogenase by acetyl-CoA helps to explain the sparing action of fatty acid oxidation on glucose usage. At the same time, fatty acid oxidation stimulates gluconeogenesis by supplying ATP and GTP required for the pathway.

Relationship of the Cycle to Urea Synthesis

The oxidation of amino acids yields one mole each of NH_4^+ and HCO_3^- from each mole of amino acid oxidized, according to the general equation:



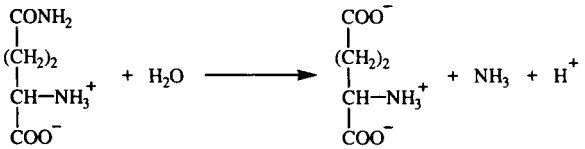
The ammonia released during amino acid catabolism is extremely toxic, especially to the brain. Perhaps the main cause of this toxicity is the effect of ammonia on the

equilibrium of the glutamate dehydrogenase reaction. An increase in ammonium-ion concentration would decrease the concentration of 2-oxoglutarate and thereby lower the concentration of citric-acid-cycle intermediates, which in turn would inhibit the rate of glucose oxidation, the major fuel for ATP synthesis in brain:

$$\text{2-oxoglutarate}^{2-} + NH_4^+ + \text{NADPH} + H^+ \longrightarrow \text{L-glutamate}^- + \text{NADP}^+ + H_2O$$

Other possible explanations are based on the direct inhibition of the 2-oxoglutarate dehydrogenase complex by ammonia; the inhibition of glutaminase by ammonia, resulting in the depletion of glutamate, a brain neurotransmitter; and the stimulation of glycolysis by ammonia at the 6-phosphofructokinase step, causing increased lactate formation which in turn alters the electrical activity of brain cells.

In man and other mammals, both ammonia and bicarbonate are removed principally by conversion to urea in the urea cycle (also called the ornithine cycle) (Meijer et al., 1990). In man, urea synthesis occurs almost exclusively in the liver, and is responsible for 60–90% of the total nitrogen excreted. A large proportion of the ammonia nitrogen used for urea synthesis is formed during amino acid catabolism in extrahepatic tissues, and it is released into the blood mainly in the form of alanine and glutamine. Alanine taken up from the blood by the liver undergoes transamination with 2-oxoglutarate mainly in the cytosol, to form glutamate and pyruvate. Glutamine is hydrolyzed to glutamate and ammonia by the action of the mitochondrial enzyme glutaminase:



L-glutamine L-glutamate

Although this reaction can occur to some extent in liver cells, much of the blood glutamine is taken up and hydrolyzed in the kidney (where ammonia released by the enzyme is used to maintain acid-base balance) and in the intestine. Ammonia released by intestinal glutaminase diffuses to hepatic portal blood and is taken up by the liver for urea synthesis, together with any ammonia absorbed by the intestine after release in the gut during bacterial metabolism.

In each molecule of urea synthesized, one of the two nitrogen atoms is derived from carbamoyl phosphate and the other from aspartate. In man, up to one-third of the urea nitrogen can be supplied by ammonia released during intestinal glutaminase activity. The remainder of the ammonia requirement is supplied by glutamate dehydrogenase activity in the mitochondrial matrix space. The aspartate required for urea synthesis is formed from glutamate by transamination. Glutamate, which functions as a common carrier for the nitrogen of amino acids after transamination, can therefore serve as the source of both urea nitrogen atoms.

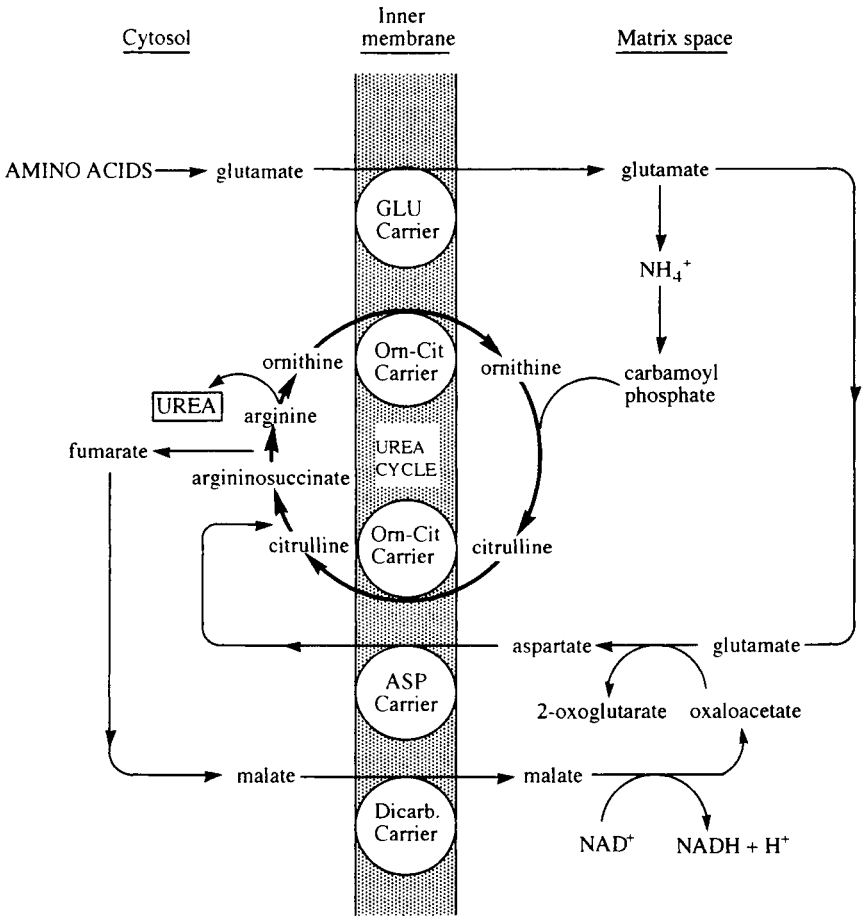
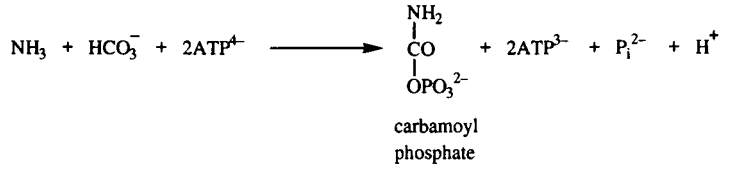


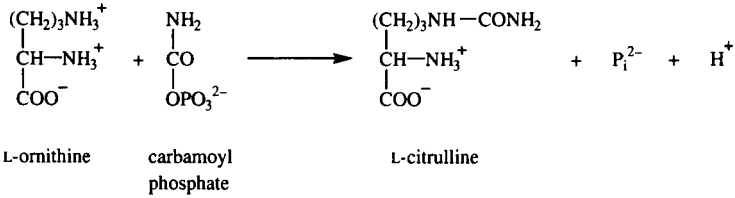
Figure 11. Role of liver mitochondrial enzymes and metabolite carriers during urea synthesis. The urea cycle is shown by thick arrows. The carriers ensure the necessary metabolic cooperation between the mitochondrial and cytosolic enzymes of the pathway. The following abbreviations are used to label metabolite carriers: GLU, glutamate; Orn-Cit, ornithine-citrulline exchange carrier; ASP, aspartate; Dicarb., dicarboxylate.

Urea synthesis from ammonia, which is achieved by five enzyme-catalyzed steps, occurs partly in the matrix space of mitochondria and partly in the cytosol (Figure 11). The first step is the formation of carbamoyl phosphate from ammonia, catalyzed by the mitochondrial enzyme carbamoyl phosphate synthetase I (ammonia):



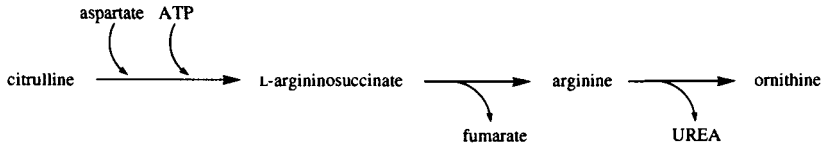
The consumption of two ATP molecules in this reaction makes the synthesis of carbamoyl phosphate (an energy-rich compound) essentially irreversible. The enzyme (M_r = 320 kDa) consists of two similar subunits. It requires MgATP (as substrate) and free Mg^{2+} for activity, and it is activated allosterically by N-acetylglutamate. The enzyme accounts for 15–25% of the total matrix-space protein in liver mitochondria. It differs from the cytosolic carbomoyl phosphate synthetase II (CPS II) functional in pyrimidine biosynthesis, which uses glutamine as a nitrogen source and is not regulated by N-acetylglutamate. Bicarbonate required for urea synthesis is produced by the dissociation of carbonic acid formed by the carbonic-anhydrase-dependent hydration of CO_2 released during citric-acid-cycle activity ($CO_2 + H_2O \rightarrow H_2CO_3 \rightarrow HCO_3^- + H^+$).

In the second step, carbamoyl phosphate reacts with ornithine to form citrulline in a reaction catalyzed by ornithine transcarbamoylase, which is also located in the matrix space:

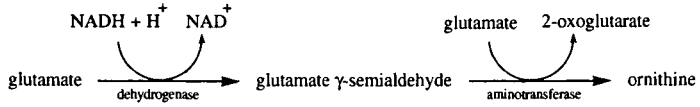


The equilibrium of the reaction lies strongly to the right, so that citrulline formation is greatly favored. The enzyme can also use lysine as substrate, with the formation of homocitrulline instead of citrulline. This activity explains the appearance of homocitrulline in the blood and urine of patients suffering from a lack of ornithine supply to the urea cycle, caused by a defective transport of ornithine from the cytosol to the matrix space. The name of this rare disorder (HHH syndrome) summarizes the features of the condition: hyperornithinemia, hyperammonemia, and homocitrullinuria.

Citrulline produced by the ornithine transcarbamoylase reaction passes to the cytosol in exchange for an ornithine molecule formed by urea cycle reactions taking place in the cytosol:



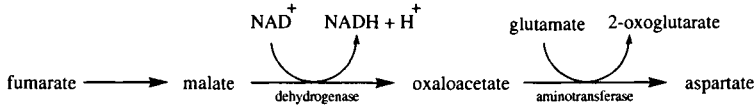
Provided that all the intermediates of the pathway are retained, they have a catalytic role, because one molecule of ornithine is regenerated by arginase for each molecule of ornithine consumed by the transcarbamoylase reaction. When arginine is removed from the cycle to satisfy the arginine requirements of extrahepatic tissues, the supply of ornithine for the urea cycle is replenished from glutamate by the activity of two additional mitochondrial enzymes:



The sum of the five reaction steps responsible for the synthesis of urea is as follows:

$$NH_3 + HCO_3^- + 3\ ATP^{4-} + H_2O + asparate^- \longrightarrow fumarate^{2-} + 2\ ADP^{3-} + AMP^{2-} + 4\ P_i^{2-} + 4\ H^+ + urea$$

An additional molecule of ATP is required to convert AMP into ADP, so that altogether four molecules of ATP are consumed during the synthesis of one urea molecule. The net energy cost is only one ATP, because three ATP molecules can be synthesized from NADH oxidation by the respiratory chain when fumarate is converted to aspartate by way of the malate dehydrogenase reaction:



Short-term regulation of urea cycle activity is achieved by the allosteric activation of carbamoyl phosphate synthetase by N-acetylglutamate, a compound synthesized in the matrix space in a reaction catalyzed by N-acetylglutamate synthase (acetyl-CoA + glutamate → N-acetylglutamate). In the longer term, regulation is achieved by reversible changes in the concentration of the five enzymes of urea synthesis, which increase within a few days when the protein content of the diet is increased.

A major or total deficiency of any enzyme in the urea cycle is lethal, because this condition abolishes the ability of the body to detoxify ammonia. Inborn errors of metabolism occur in each of the urea-cycle enzymes, including deficiencies in one or both of the mitochondrial enzymes of the cycle. The conditions caused by these deficiencies are severe, with a high incidence of mental retardation, seizures, coma, and early death. The most common disorder of the cycle is a deficiency in ornithine transcarbamoylase. A decrease in urea cycle activity often results in failure to remove sufficient ammonia when the protein content of the diet is high, so that the blood ammonia concentration increases (hyperammonemia) and symptoms of ammonia toxicity develop. In some patients, the condition can be alleviated by restriction of the protein intake and the administration of carbamoylglutamate, an analog of N-acetylglutamate, which (like N-acetylglutamate) activates carbamoyl phosphate synthetase I.

RESPIRATORY ENZYME SYSTEMS

The Respiratory-Chain System

Acetyl-CoA molecules derived from the carbon skeletons of foodstuffs are oxidized in the matrix space by the sequence of citric-acid-cycle reactions. During the oxidation of an acetyl-CoA molecule by the citric acid cycle, three pairs of hydrogen atoms are accepted by NAD molecules (from isocitrate dehydrogenase, 2-oxoglutarate dehydrogenase, and malate dehydrogenase, respectively), and the fourth pair is accepted by the FAD prosthetic group of succinate dehydrogenase present in Complex II. Hydrogen atoms are also transferred to NAD and FAD cofactors during the catalytic action of other dehydrogenases, especially those functional during the β-oxidation of fatty acids. In nearly all cases, these dehydrogenases are located in mitochondria, but in addition the operation of substrate "shuttle" systems enable mitochondria to oxidize cytosolic NADH. The oxidation process is completed by the action of the respiratory-chain system present in the inner mitochondrial membrane, which catalyzes the reoxidation of NADH and $FADH_2$ by molecular oxygen to regenerate their oxidized forms. The cofactors therefore act catalytically, becoming alternately reduced by dehydrogenase action and reoxidized by the respiratory chain. Oxidation is accompanied by a large change in free energy, and a substantial proportion of the energy liberated is linked or coupled to the synthesis of ATP. The respiratory-chain system shows many interesting adaptations to suit the specialized needs of individual cells and tissues. The simplest adaptation is seen in the cellular content of respiratory-chain components, which is broadly proportional to the cellular energy requirement. In some tissues mitochondrial respiration is adapted to support physiological functions operating in parallel with ATP synthesis, or even taking primacy over it. This is achieved by shifting the use of the proton electrochemical potential away from the ATP synthase to support other functions, including the generation of heat during facultative thermogenesis (especially in brown adipose tissue), or the translocation of monovalent ions (Na^+, K^+) and divalent ions (especially Ca^{2+}) between mitochondria and the surrounding cytoplasm.

The presence of tightly-bound prosthetic groups in the catalysts of the respiratory chain, and the high affinity between cytochrome *c* oxidase and oxygen, ensure that all the components are in a fully functional state and can catalyze rapid ATP synthesis even when the intracellular tissue oxygen tension is very low. The efficiency of the respiratory chain also depends on the physical organization of the redox catalysts within multisubunit complexes which span the inner membrane (Hatefi, 1985). The precise spatial arrangement of the functional components in Complexes I, III, and IV is crucial to their proton-translocating function. The prosthetic groups are important sites of action for specific and potent cell poisons, which inhibit or totally block respiratory-chain activity.

Redox Components

The respiratory chain system contains five main types of redox carrier, namely flavin, non-heme iron, heme iron, ubiquinone (Q or coenzyme Q), and NAD. In addition, copper functions as an essential catalyst in cytochrome *c* oxidase (Complex IV) (Capaldi, 1990). The synthesis of these carriers in the human body depends in most cases on the provision in the diet of small quantities of vitamins (riboflavin, niacin) or trace elements (iron, copper). The first three types of carrier function as the prosthetic groups of flavoproteins, iron-sulfur proteins, and cytochromes, respectively. Ubiquinone and cytochrome *c* function as small mobile redox carriers linking the large multisubunit respiratory-chain complexes, but part of the ubiquinone is associated with a fourth class of protein, the Q-binding proteins present in Complex I and Complex III. The presence of tightly-bound prosthetic groups in the catalysts of the respiratory chain ensures that all the components are in a fully functional state and can drive ATP synthesis even when the intracellular oxygen tension is very low. There is a large excess of NAD molecules in the matrix space, which have a "pool" function for a variety of NAD-linked dehydrogenases. This coenzyme is bound to dehydrogenases only transiently during oxidoreduction. Ubiquinone molecules are also in excess, and collect reducing equivalents from a variety of flavin-linked dehydrogenases as well as from Complex I and Complex II.

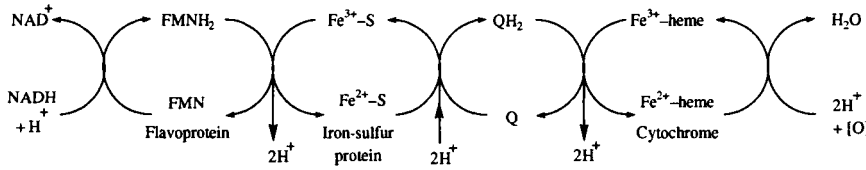
The redox potentials of the carriers span a wide range of values (Table 3), a property which enables them to function in a sequence of oxidoreactions, in which reducing equivalents removed from substrates by dehydrogenase action are transferred from one component to another and ultimately to molecular oxygen. Most of the redox centers function at an operating potential close to their mid-point potential, and this allows an adequate concentration of both oxidized and reduced

Table 3. Respiratory Chain Catalysts

Redox Carrier	*Reducing Equivalents Carried*	*Protein Component*	*Approximate Midpoint Potential, E_m Volts*
NAD*	2H ($H^- + H^+$)	Dehydrogenase	–0.32
Flavin	2H	Flavoproteins:	
		FMN form (Complex I)	–0.30
		FAD form (Complex II)	–0.06 to –0.09
Nonheme iron	e^-	Iron-sulfur protein	–0.27 to +0.28
Ubiquinone (Q)*	2H ($2H^+ + 2e^-$)	Q-binding proteins	+0.04
Heme iron	e^-	Cytochromes *b*	+0.07
		Cytochromes *c*	+0.25
		Cytochromes *a*	+0.29
Copper	e^-	Cytochrome *c* oxidase	+0.25
Oxygen			+0.82

Note: *Coenzymes, functioning as mobile carriers of reducing equivalents.

forms of each carrier to be available for reaction. In simplified form, the reaction sequence of the respiratory chain can be shown as follows:



Respiratory-Chain Complexes

The redox carriers are organized physically into four respiratory chain complexes and functionally into four equipotential redox groups (Table 4). The relatively large differences in redox potential between Complexes I, III, and IV provide the source of free energy that is conserved by formation of a proton electrochemical potential (the proton gradient), and is used ultimately to drive ATP synthesis by the ATP synthase (Complex V) (Harold, 1986). This arrangement enables the cell to exploit large free-energy changes occurring when NADH or flavin-linked substrates are oxidized by oxygen, so that more than one molecule of ATP can be synthesized per molecule of substrate oxidized. The coupling mechanism enables ATP to be synthesized by a common process at the expense of free energy released

Table 4. The Respiratory Chain Complexes and ATP Synthase

			Subunits			
Complex	*Function*	*Molecular Weight (kDa)*	*Total No.*	*No. of mt-DNA Encoded Subunits*	*Redox Groups*	*Inhibitors*
I	NADH-Q reductase	700	26	7	FMN; 4 to 6 FeS centers	Rotenone, amytal, piericidin
II	Succinate-Q reductase	126	4	0	FAD; 3 FeS centers	
III	QH_2-cytochrome *c* reductase	250	11	1	2 *b*-type hemes; 1 *c*-type heme (cyt. c_1); 1 FeS center; 2 Q-binding centers	Antimycin
IV	Cytochrome *c* oxidase	204	13	3	2 *a*-type hemes (cyt *a* and cyt a_3); 2 copper atoms	Cyanide, azide, sulfide, carbon monoxide
V	ATP Synthase	500			None	Oligomycin, DCCD, organotins
	F_O		8	2		
	F_1-ATPase		5	0		

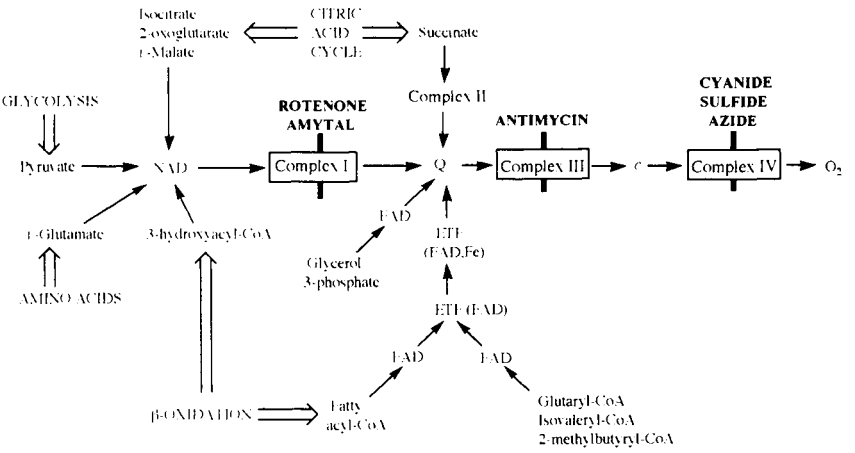


Figure 12. Pathways for the oxidation of substrates by the respiratory-chain system in mitochondria. Arrows represent the transfer of reducing equivalents. In many cells, the main sources of oxidizable substrates are the citric acid cycle and the fatty acid β-oxidation cycle. Specific inhibitors of the respiratory-chain complexes are shown in bold type.

by different kinds of oxidoreduction reaction, irrespective of the position of the reactants in the redox scale. Complex II is also located in the inner membrane, but there is only a small free-energy change taking place during the reaction catalyzed by the complex (succinate → Q) and it does not translocate protons.

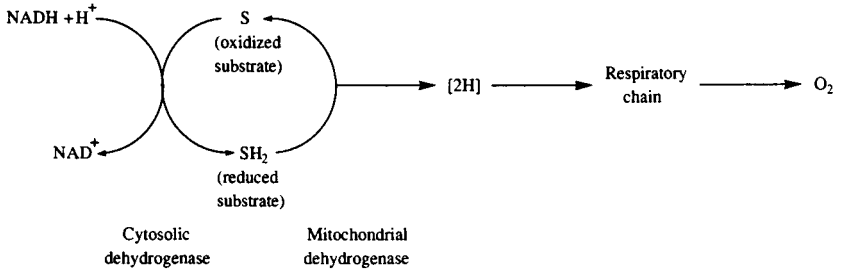
Pathways for Substrate Oxidation

Figure 12 summarizes the respiratory chain reactions responsible for the oxidation of intermediates in the citric acid cycle, the β-oxidation system, and other compounds. In many cases, reducing equivalents removed from substrates are transferred to NAD^+ by the action of NAD-linked dehydrogenases, or directly to Q by way of flavin-linked dehydrogenases. The intermediate transfer of reducing equivalents to NAD, Q, and ETF diminishes the crowding of individual dehydrogenase molecules close to the M-face of the inner membrane.

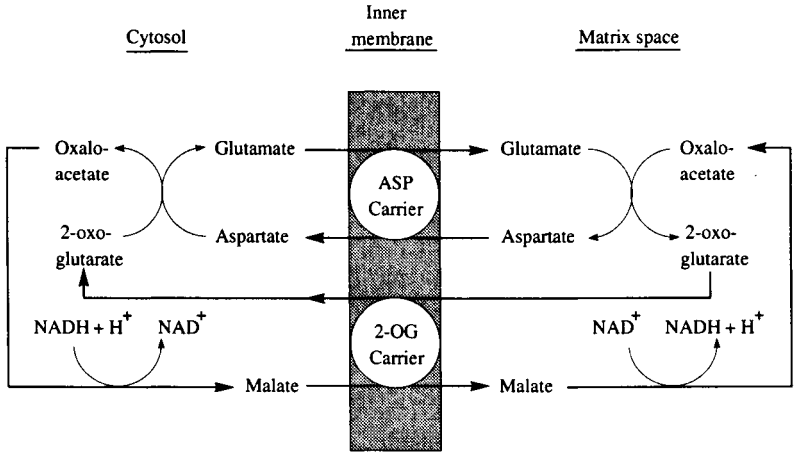
Reoxidation of Cytosolic NADH

NADH formed in the cytosol by the action of glyceraldehyde-3-phosphate dehydrogenase must be reoxidized rapidly to regenerate NAD^+ and thereby maintain glycolysis. Under anaerobic conditions this is achieved by the lactate dehydrogenase reaction, which converts the glycolytic products pyruvate and NADH into lactate and NAD^+. When oxygen is available, pyruvate formed by glycolysis in the

cytosol is transported into the mitochondrial matrix space for oxidation to CO_2 and H_2O, and cytosolic NADH is reoxidized by the transfer of reducing equivalents to the respiratory chain. In muscle, this is achieved by the action of an additional external respiratory-chain-linked NADH dehydrogenase located at the C-face of the inner membrane. In many other tissues, which lack this enzyme, the direct oxidation of NADH by mitochondria is prevented by the fact that the inner membrane is impermeable to NADH, and the reduced coenzyme is therefore unable to reach the catalytic center of Complex I located at the M-face of the inner membrane. Cytosolic NADH oxidation by mitochondria is achieved by the operation of two main substrate shuttle systems, in which NADH reoxidation by a cytosolic dehydrogenase is coupled to the formation of a substrate capable of oxidation by mitochondria, malate, or glycerol 3-phosphate. Each molecule of substrate acts catalytically by carrying reducing equivalents between two substrate dehydrogenases, one located in the cytosol and the other in the mitochondria:



The first shuttle system employs malate as the carrier of reducing equivalents from the cytosol to the mitochondria, and the presence of two isoenzyme forms of malate dehydrogenase in the cytosol and in the matrix space. It also requires the activity of both cytosolic and matrix-space aminotransferases, to circumvent the impermeability of the mitochondrial membrane to oxaloacetate:



The wide distribution and high activity of the 2-oxoglutarate and aspartate carriers in the mitochondrial inner membrane suggest an important role for this system in many tissues, including liver, kidney, heart, and brain. The second system, the glycerol 3-phosphate shuttle, functions in various tissues, including brain, skeletal muscle, and brown fat. It could also function in liver, especially in the hyperthyroid state, which induces additional synthesis of the mitochondrial enzyme. Other postulated systems include the malate-citrate shuttle, which could function under lipogenic conditions (see the section on Role of the Carriers in Gluconeogenesis and Lipogenesis).

Respiratory-Chain Inhibitors

In many cases, the action of a single inhibitor on the main energy-generating pathways of the mitochondrion *in vivo* renders the whole system inoperative, and therefore has a marked or even fatal effect upon the organism. The best-known example is the toxic action of cyanide on cytochrome *c* oxidase, which results in a powerful inhibition of the enzyme and a rapid cessation of aerobic metabolism. The extreme toxicity of cyanide illustrates the vital importance of the enzyme to the incessant activity of intracellular metabolism. Poisoning by cyanide usually blocks all but 1–2% of the respiration in animal tissues. The residual oxygen usage is due to the activity of peroxisomal oxidases and monooxygenase systems. About 2,000 species of cyanogenic plants contain cyanide precursor compounds which liberate cyanide when the plant is macerated. Many of these plants grow in the tropical and subtropical regions of the world and in these areas chronic cyanide toxicity is a common medical problem. Cyanide is also encountered by humans in manufacturing and industrial processes. The most vital site of cyanide action is the central nervous system, resulting in a cessation of respiration. When the electrical activity of the brain has ceased, the heart is often still functioning. Cyanide, azide, and sulfide each react with the oxidized enzyme and form stable Fe^{3+}—inhibitor complexes with the a_3 heme prosthetic group of the enzyme, whereas carbon monoxide combines only with the ferrous state (Fe^{2+}–CO). The toxicity of carbon monoxide in man is centered not in mitochondria but in the blood, where it forms carboxyhemoglobin and prevents the oxygen-transporting function of the protein.

The antimycins are potent fungal antibiotics produced by a species of *Streptomyces*, and rotenone is a powerful insecticide and fish poison extracted from the roots of the Derris plant. Both these compounds bind very tightly and specifically to their sites of action in the respiratory chain, and they are therefore potentially very toxic to all organisms dependent upon mitochondria for energy. They are often very much more toxic when given by injection than as an oral dose. The estimated lethal oral dose of rotenone in adult man is 200 mg, or about four times the value for cyanide. The precise action of antimycin, of rotenone, and of the barbiturate amytal is unknown. Like most inhibitors acting on complexes I, II, and IV, they share some

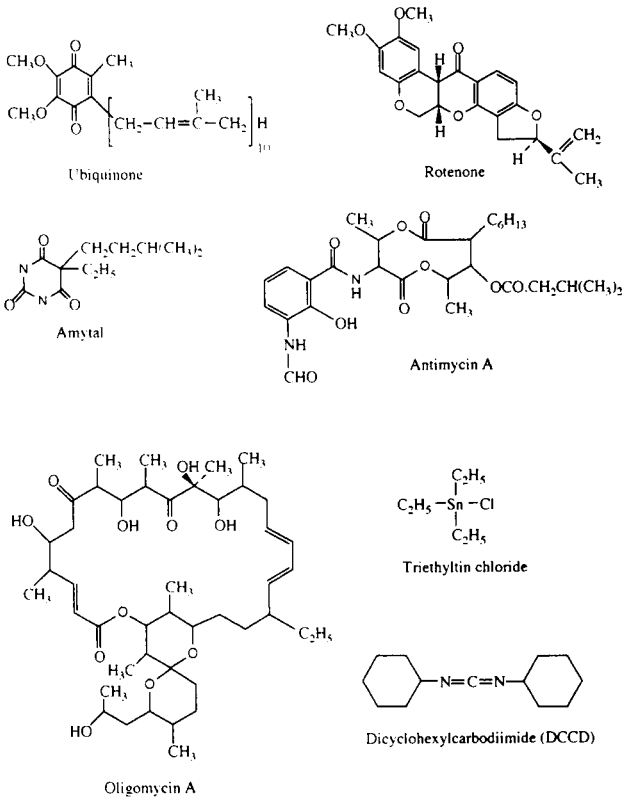


Figure 13. Structures of ubiquinone and some important inhibitors of oxidative phosphorylation reactions in mammalian mitochondria.

structural similarities with ubiquinone (Figure 13), and they probably prevent oxidoreduction reactions by blocking normal ubiquinone function.

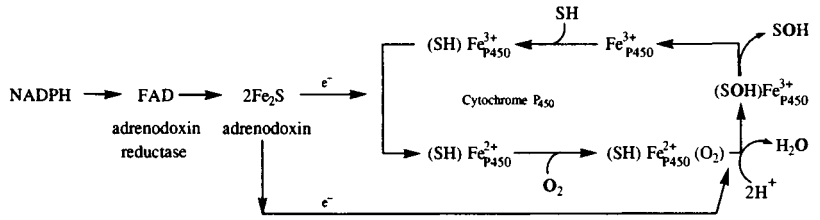
Monoxygenase Systems

These systems catalyze essential hydroxylation reactions during the conversion of cholesterol to steroid hormones and bile salts (Guenerich, 1992). The reactions can be summarized by the general equation:

$$SH + O_2 + NADPH + H^+ \longrightarrow SOH + H_2O + NADP^+$$

where SH represents the substance undergoing hydroxylation. One atom of the oxygen molecule is incorporated into the substrate, while the other atom is reduced to water. The reaction consists of a catalytic cycle, during which both substrate and oxygen are bound to the cytochrome P_{450} component of the monooxygenase

system, and the heme prosthetic group of the cytochrome undergoes oxidoreduction:



Cytochrome P_{450} consists of a family of closely related hemoproteins, each P_{450} species functioning as the terminal oxidase of a specific monooxygenase system. Mammalian cytochrome P_{450} is present in two distinct types of systems, one found in mitochondria (which resembles the system present in bacteria), and the other in the endoplasmic reticulum. In bovine adrenal cortex mitochondria, cytochrome P_{450} is present at a concentration approximately six times greater than that of the respiratory-chain cytochromes.

All three components of the system are located in the inner membrane, with their catalytic centers facing the matrix space. Cytochrome P_{450} is an integral protein. Adrenodoxin functions as a mobile shuttle carrying reducing equivalents between the reductase and oxidase components. One important source of NADPH is from the activity of an NADP-linked malic enzyme present in the matrix space:

$$\text{L-malate}^{2-} + \text{NADP}^{+} \longrightarrow \text{pyruvate}^{-} + \text{NADPH} + \text{CO}_2$$

Other possible sources include the reaction catalyzed by NADP-linked isocitrate dehydrogenase, and the conversion of NADH to NADPH by the energy-dependent transhydrogenase reaction. The mitochondrial monooxygenase system differs from that present in the endoplasmic reticulum, in which no iron-sulfur protein is required and the reductase transfers reducing equivalents directly to the P_{450} enzyme.

In the major steroidogenic tissues, including the adrenal cortex, testis, ovary, and the placenta during pregnancy, mitochondrial reactions play a very important role during the conversion of cholesterol into steroid hormones (Hall, 1984). In each tissue, the enzymes of the steroidogenic pathway are distributed between the mitochondria and the endoplasmic reticulum (Figure 14). Hormone synthesis begins in the mitochondria with a reaction in which cholesterol is converted into pregnenolone. This remarkable reaction, catalyzed by a distinct species of P_{450} (denoted by P_{450scc}) can be represented as:

$$\underset{\text{cholesterol}}{C_{27}H_{46}O} + 3\,O_2 + 3\,\text{NADPH} + 3\,H^{+} \longrightarrow \underset{\text{pregnenolone}}{C_{21}H_{32}O_2} + \underset{\substack{\text{4-methylpentanal}\\ \text{(isocaproic aldehyde)}}}{C_6H_{12}O} + 3\,\text{NADP}^{+} + H_2O$$

The reaction consists of three successive monooxygenase cycles, all catalyzed by the single reaction center in each molecule of P_{450scc}. The first two cycles are

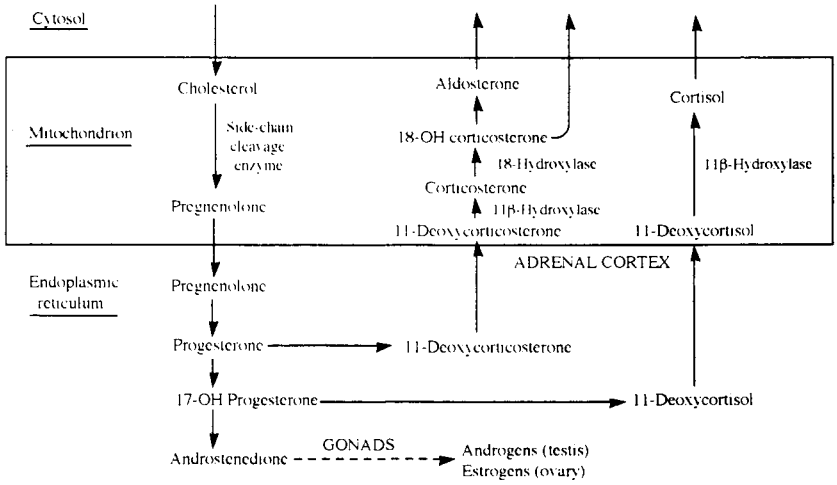
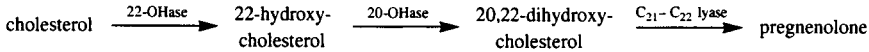


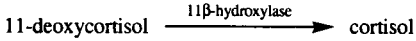
Figure 14. Role of mitochondrial monooxygenase reactions during the biosynthesis of steroid hormones. The side-chain cleavage reaction is a fundamental step in all steroidogenic tissues. In the adrenal cortex, intermediates formed by enzymes of the endoplasmic reticulum re-enter the mitochondria, where steroid hormone synthesis is completed.

hydroxylation reactions (OHase) which yield 20,22-dihydroxycholesterol, and the third cycle causes cleavage between these two carbon atoms (C_{20}–C_{22} lyase):

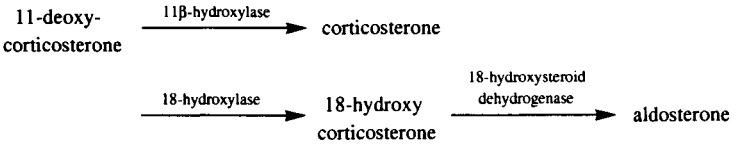


The side-chain cleavage of cholesterol to form pregnenolone, the common precursor molecule for steroid hormone synthesis, is the rate-limiting and regulated step in steroidogenesis. The stimulation of the reaction in cells of the adrenal cortex by adrencorticotrophic hormone (ACTH) is brought about in the short term by an accelerated transport of cholesterol to the inner-membrane sites of the monooxygenase system. In the longer term, ACTH induces the coordinated synthesis of the three monooxygenase enzymes. In the Leydig cells of the testis, which synthesize and secrete androgenic steroids, cholesterol transport into mitochondria is stimulated by luteinizing hormone (LH).

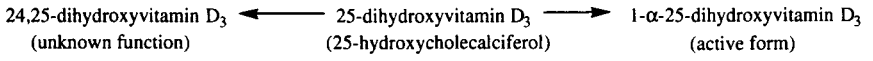
In the adrenal cortex, other monooxygenase systems are responsible for the final stages in the synthesis of glucocorticoid (cortisol, corticosterone) and mineralocorticoid (aldosterone) hormones (Figure 14). In both cases, 11β-hydroxylation is required, and is catalyzed by a P_{450} enzyme ($P_{450 11\beta}$) distinct from the side-chain cleavage enzyme:



In a similar reaction, the corresponding intermediate formed in the endoplasmic reticulum, 11-deoxycorticosterone, undergoes 11β-hydroxylation in mitochondria to yield corticosterone. Also in mitochondria, the 11β-hydroxylase is thought to catalyze 18-hydroxylase activity, required in an extension of this pathway leading to the synthesis of aldosterone:



Kidney and liver mitochondria also contain cytochrome P_{450}, although their content of it is only about 5% of the cytochrome *c* content. The kidney enzyme has a very important function during the metabolism of vitamin D_3 (cholecalciferol) because it catalyzes the 1- and 24-hydroxylation of 25-hydroxyvitamin D_3:



The 1α-25-dihydroxyvitamin D_3 formed by the kidney is the most potent known effector of increased intestinal absorption of calcium and mobilization of bone mineral. The liver mitochondrial monooxygenase system catalyzes a 26-hydroxylation step during bile acid synthesis. It is possible that a P_{450}-containing system is a common component of mitochondria, although present in very small amounts compared with adrenal-cortex mitochondria. Oxygen uptake during the 11β-hydroxylation of deoxycorticosterone by isolated bovine adrenal cortex mitochondria is at least as fast as that occurring during ATP synthesis, so that oxygen usage during hydroxylation reactions can be responsible for a substantial part of the total oxygen consumption by this tissue. By contrast, the contribution by the monoxygenases in kidney and liver mitochondria to tissue oxygen uptake is likely to be very small.

Role of Mitochondria in Thermogenesis

In warm-blooded adult animals exposed to cold, a fast initial thermogenic response is provided by shivering, during which respiratory energy is first conserved by ATP synthesis, and heat is then released by ATP hydrolysis during involuntary muscle contractions. Most newborn mammals, including man, respond to cold exposure with a large increase in oxygen consumption and heat production. This response is achieved largely without shivering, that is, by nonshivering thermogenesis, especially in brown adipose tissue (Trayhurn and Nicholls, 1986). The chemical basis for the intense heat production by this tissue is due to the specialized properties of their mitochondria. In brown adipose tissue, the linkage between the respiratory chain and the low ATP synthase activity is uncoupled in a regulated way by a unique inner-membrane protein called thermogenin or "uncou-

pler protein," so that much of the energy of oxidation can be released as heat when required. Thermogenin functions as a channel for the transport of anions across the membrane, including Cl^- and OH^-. The transport of OH^- from the matrix space is chemically equivalent to H^+ entry, and is equally effective in collapsing the proton electrochemical potential, resulting in the release of respiratory energy as heat (Figure 15). Thermogenin has been found only in brown fat mitochondria. The functional thermogenin molecule is a dimeric protein containing two identical 32-kDa subunits and a GDP-binding site, and for this reason it is also called the GDP-binding protein. Unlike the other metabolite carrier proteins in mitochondria, which seem to be always fully active and synthesized in constitutive amounts, both the content and activity of thermogenin are regulated to suit prevailing thermogenic demands. The thermogenin channel is thought to be regulated by two opposing influences: an inhibition caused by the binding to the protein of cytosolic purine

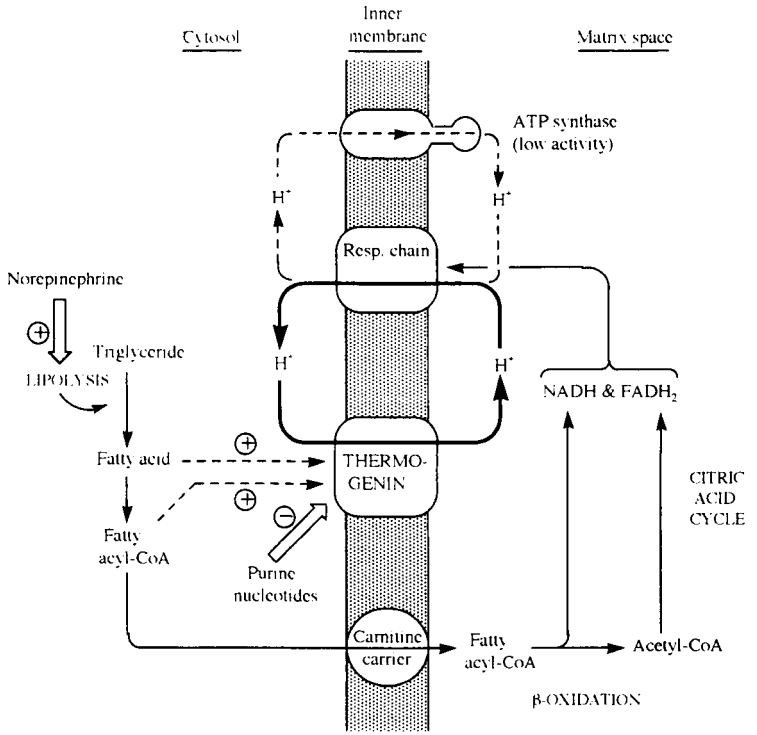
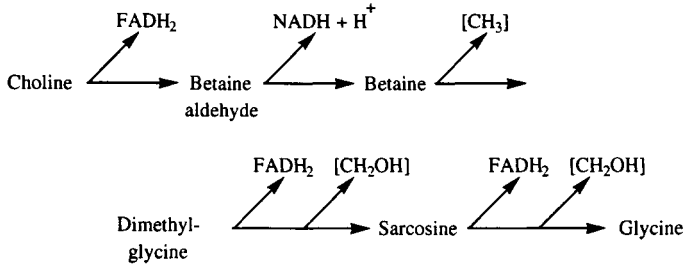


Figure 15. Regulation of thermogenesis in brown adipose tissue. The activation of lipolysis by norepinephrine increases the concentration of fatty acids and fatty acyl-CoA, which serve as oxidizable substrates and can also function as activators of proton transport through the thermogenin channel, resulting in the release of respiratory energy as heat.

nucleotides (GDP, GTP, ADP, ATP), and a reversal of this inhibition brought about by long-chain fatty acids and/or their acyl-CoA esters. Physiological activation of the tissue occurs by the release of the neurotransmitter norepinephrine. The binding of this hormone to the plasma membrane of brown adipocytes rapidly activates lipolysis and cell respiration. These effects are brought about by the stimulation of adenyl cyclase, resulting in the activation of cyclic-AMP-dependent triglyceride lipase. Long-chain fatty acids liberated by lipolysis serve a dual function, by providing both a substrate suitable for oxidation and an intracellular signal to switch metabolism to the thermogenic mode (Figure 15). When demand for additional body heating decreases, the steady-state concentration of fatty acid falls because of a decline in lipolysis and the removal of fatty acid by oxidation. There is some evidence for an uncoupling action of free fatty acids on muscle and liver mitochondria, which could also make a contribution to body thermogenesis. In adult rodents, brown fat thermogenesis appears to have a role in the regulation of calorie balance and body weight regulation. The adult human body seems to have very little brown fat and does not have a similar role. However, it remains possible that the content of active brown fat could be increased by suitable pharmacological treatment, and that this might then have a useful role in the control or treatment of obesity.

Ancillary Oxidation Reactions

In addition to the major oxidative pathways, mitochondria also contain a variety of other oxidizing enzymes, including the enzymes of choline oxidation (in liver and kidney), amine oxidase, and the glycine cleavage enzyme. During choline metabolism, choline dehydrogenase (an FAD-containing flavoprotein in the inner membrane) converts choline to betaine aldehyde, which is oxidized by NAD-linked aldehyde dehydrogenase to betaine. Betaine functions as a donor of 1-carbon methyl and hydroxylmethyl groups during a sequence of reactions in the matrix space leading to the formation of glycine:



Reducing equivalents from choline pass via FAD directly to ubiquinone, whereas those from the two later flavin-linked dehydrogenases (dimethyl glycine dehydrogenase and sarcosine dehydrogenase) are transferred from FAD to the ETF pathway shown in Figure 12.

Glycine is metabolized by the combined action of glycine cleavage enzyme (also called glycine synthase) and serine hydroxymethyltransferase, which together catalyze glycine decarboxylation:

$$\text{Glycine} + \text{NAD}^+ + \text{tetrahydrofolate} \longrightarrow \text{CO}_2 + \text{NH}_4^+ + \text{methylene tetrahydrofolate} + \text{NADH}$$

$$\text{Glycine} + \text{methylene tetrahydrofolate} \longrightarrow \text{Serine} + \text{tetrahydrofolate}$$

$$\text{Sum:} \quad 2\ \text{Glycine} + \text{NAD}^+ \longrightarrow \text{Serine} + \text{CO}_2 + \text{NH}_4^+ + \text{NADH}$$

The glycine cleavage enzyme in liver mitochondria consists of four protein components, which form a multienzyme system with several features in common with the 2-oxoacid dehydrogenases.

Amine oxidase (also called monoamine oxidase) is present in the mitochondrial outer membrane. The enzyme is a flavoprotein containing FAD as the prosthetic group. It catalyzes the oxidative deamination of amines ($R.CH_2.NH_2$) according to the general equation:

$$R\text{–}CH_2NH_2 + H_2O + O_2 \rightarrow R\text{–}CHO + NH_3 + H_2O_2$$

The physiological function of the enzyme is the catabolism of endogenous neurotransmitter amines and poisonous exogenous amines present in food or formed by bacterial action in the intestine. Two forms of the enzyme, types A and B, can be distinguished by their substrate specificity and their sensitivity to inhibitors. Amine oxidase inhibitors are used in the treatment of depression. Inhibition of the enzyme results in an increase in the epinephrine and serotonin content of the brain and the peripheral tissues. The increase in brain amines improves the availability of transmitter substances, and this has an antidepressant effect. Inhibitors of the type-B enzyme, including deprenyl, protect against the toxicity of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), a compound causing a condition resembling Parkinson's disease, in which the nerves of the nigrostriatal pathway are selectively destroyed. The active toxin is a pyridinium derivative (MPP^+) formed by the action of monoamine oxidase on MPTP.

Oxygen Toxicity

Oxygen present in air causes slow but progressive damaging effects on living organisms, including the process of aging (Cadenas, 1989). The reduction of oxygen by the addition of one, two, or four electrons leads to the formation of superoxide ($O_2^{-\bullet}$), peroxide (O_2^{2-}), and H_2O respectively:

$$O_2 \xrightarrow{e^-} O_2^{-\bullet} \xrightarrow[2H^+]{e^-} H_2O_2 \xrightarrow[H^+]{e^-} H_2O + OH^- \xrightarrow[H^+]{e^-} 2H_2O$$

Oxygen toxicity is thought to be due mainly to the conversion of superoxide to the hydroxyl radical ($OH^\bullet$), in a reaction with hydrogen peroxide which is catalyzed by traces of transition metal ions, especially iron and copper:

$$O_2^{-\bullet} + H_2O_2 \longrightarrow O_2 + OH^- + OH^\bullet$$

Superoxide radicals are generated in mitochondria, probably by the reaction of oxygen directly with the iron-sulfur centers in Complex I and the quinone centers in Complex III. These reactions consume only very small quantities of oxygen compared with the cytochrome *c* oxidase reaction. Oxygen radicals cause cell damage in a variety of ways: they are potent initiators of lipid peroxidation in intracellular membranes; they modify amino acid residues in proteins, leading to the inactivation of protein function by cross-linking reactions; and they cause modification to bases and strand breaks in DNA (Cross et al., 1987). Mitochondria contain catalase, glutathione peroxidase, and the manganoprotein superoxide dismutase, which together protect against oxygen toxicity by converting superoxide anions first to hydrogen peroxide and then to water:

$$2\ O_2^{-\bullet} + 2\ H^+ \xrightarrow{\text{superoxide dismutase}} O_2 + H_2O_2 \xrightarrow[\text{catalase}]{\text{glutathione peroxidase;}} O_2 + H_2O$$

ATP SYNTHESIS IN MITOCHONDRIA

The approximate redox potentials of oxidoreduction carriers functional in the various regions of the respiratory chain allow estimates to be made of the free-energy changes occurring during the three main stages of NADH oxidation by oxygen (Figure 16). These estimates show that the free energy change ΔG required to drive the synthesis of ATP (approximately 51.6 kJ/mol) is nearly equal to the free-energy changes occurring during oxidoreduction at coupling sites 1 and 2, and this result suggests that the efficiency of energy conservation at these sites is close to 100%. The free energy change at site 3 is sufficient to drive the synthesis of 2 ATP molecules per pair of electrons transferred, but in fact only 1 molecule of ATP is synthesized. In consequence there is a substantial loss of free energy at this site, ensuring that the cytochrome *c* oxidase reaction is irreversible and that the overall process of substrate oxidation by the respiratory chain is able to drive the incessant synthesis of ATP. The stoichiometry of ATP synthesis during substrate oxidation is expressed by the P:O ratio, that is, the ratio of the number of ATP molecules synthesized per atom of oxygen consumed. The value of the ratio is close to three during the oxidation of NAD-linked substrates and two during the oxidation of flavin-linked substrates. Table 5 shows the estimated ATP yield during the complete oxidation of glucose to CO_2 and H_2O. The values obtained illustrate the large difference in net ATP yield during aerobic metabolism (38 mole ATP per mole of glucose) compared with that obtained under anaerobic conditions (2 mole ATP per mole of glucose). The contribution of oxidative phosphorylation in mitochondria to the total aerobic synthesis of ATP is therefore $(36/38) \times 100 = 94.7\%$.

In living cells, a dynamic balance exists between ATP synthesis (largely in mitochondria) and ATP hydrolysis, which occurs during energy-requiring reactions taking place mainly in the surrounding cytoplasm:

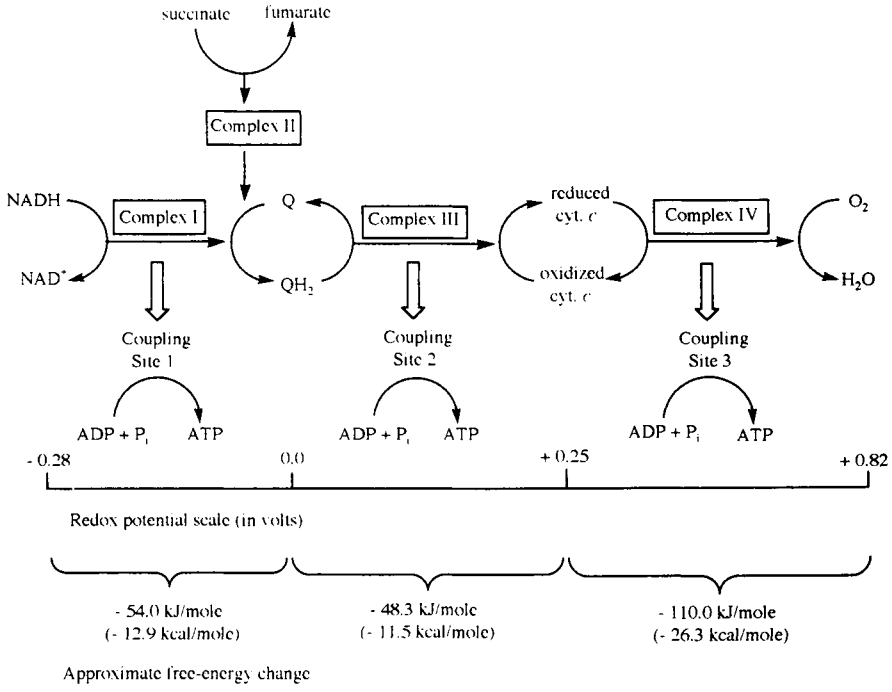


Figure 16. Diagram of the respiratory chain, showing the location of coupling sites 1, 2, and 3, where ATP synthesis is coupled to oxidoreduction steps. The diagram shows the estimated values for the free-energy changes occurring during oxidoreduction reactions, which are proportional to the differences in redox potential between the various electron carriers in the chain. Assuming that the free energy required for ATP synthesis *in vivo* is 51.6 kJ/mole, the values show that the free energy liberated at sites 1 and 2 is approximately sufficient to drive the synthesis of an ATP molecule, and energy liberated at site 3 is greatly in excess of that required (see text).

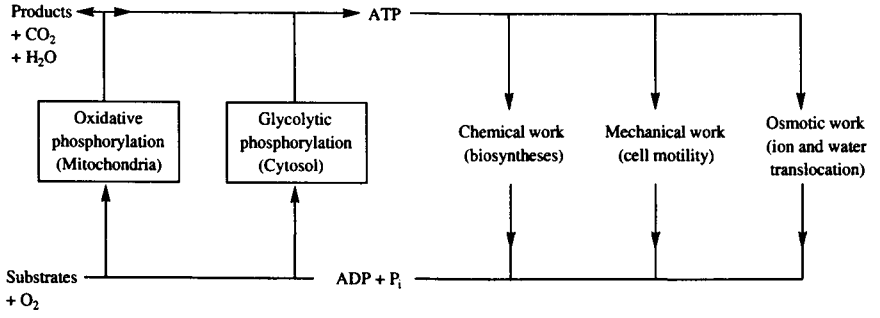


Table 5. Free-Energy Balance Sheet During Glucose Oxidation

Reaction	*Catalyst*	*Net ATP Formed mole/mole*		Free Energy Relased, ΔG^{O}	
		Anaerobic	*Aerobic*	*kJ/mole*	*kcal/mole*
Glucose → 2 lactate	Glycolysis	2		–197	–47
Glucose → 2 pyruvate	Glycolysis		2	–146	–35
$2(NADH + H^+) + O_2 \rightarrow 2NAD^+ + 2H_2O$	Malate-aspartate shuttle		6	–440	–105
$2\ Pyruvate + 5\ O_2 \rightarrow 6\ CO_2 + 4H_2O$	Citric acid cycle; respiratory-chain phosphorylation		30	–2284	–546
Total ATP synthesis					
Anaerobic		2			
Aerobic			38	(–2870)	(–686)

Notes: Free energy conserved as ATP, assuming ($ADP + P_i \rightarrow ATP + H_2O$ + 12.3 kcal/mole):
$38\ (ADP + P_i) \rightarrow 38\ (ATP + H_2O + \{38 \times 12.3 = 467.4\}$ kcal/mole
Efficiency of energy conservation = (467.4/686) × 100% = 68.1%

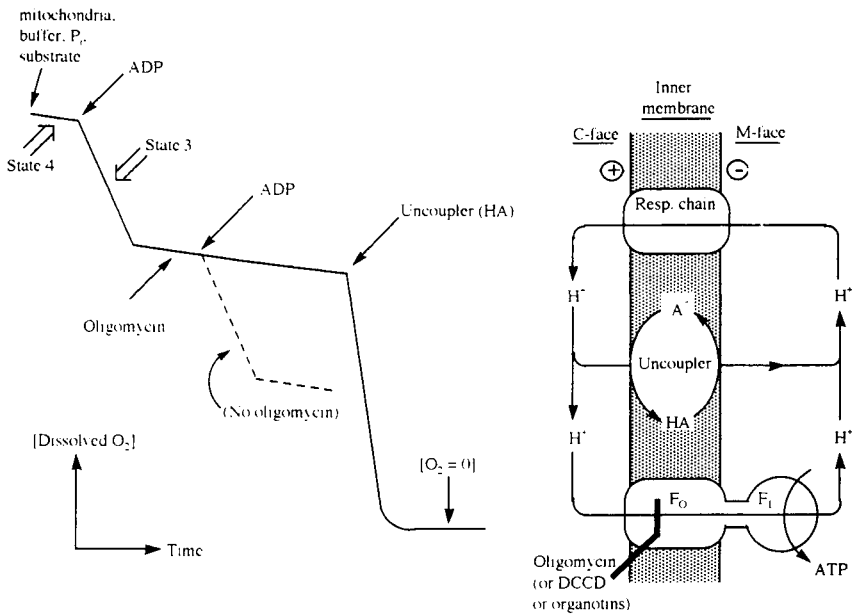


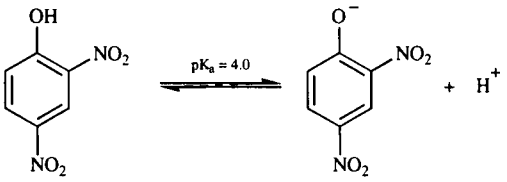
Figure 17. Assay of ATP synthesis in mitochondria and the action of inhibitors. (**a**) Oxygen electrode tracing which shows the influence of a typical F_O inhibitor (oligomycin) on the respiration of isolated mitochondria. Oligomycin inhibits the flow of protons through F_O during ATP synthesis, and the respiratory chain is inhibited by the back-pressure of the proton electrochemical potential. Rapid and sustained respiration is restored in the presence of an uncoupling agent (HA), which collapses the protonic potential and by-passes the blocked ATP synthase complex. In the absence of oligomycin, a second addition of ADP induces a second cycle of respiratory activity until the ADP added has been converted into ATP (shown by the dashed lines). (**b**) Diagram showing the sites of action of two classes of ATP synthesis inhibitors, those acting on the ATP synthase complex [oligomycin, dicyclohexylcarbodiimide (DCCD), organotins] and those acting as uncoupling agents (HA).

The oxidation of cell fuels by the respiratory chain and the synthesis of ATP by the ATP synthase are tightly-coupled, a property which ensures that fuel consumption is regulated by energy demand. When mitochondria are isolated from cell homogenates by the technique of differential centrifugation, they become largely separated from the ATP-hydrolyzing enzymes of the cell. The respiration rate of isolated mitochondria, incubated with sufficient oxidizable substrate, oxygen, and P_i, is stimulated strongly by the addition of ADP to initiate ATP synthesis (Figure 17). This experiment illustrates the two most important metabolic states of mitochondria, the resting or State 4 rate, and the phosphorylating or State 3 rate. When available ADP has been removed by conversion to ATP, the respiration rate reverts

to the State 4 condition. This phenomenon, called the respiratory control or acceptor control of mitochondrial respiration, is due to an inhibition of respiratory chain activity by the back-pressure of the proton gradient established under State 4 conditions. A similar respiratory control behavior of mitochondria in living cells and tissues is responsible for the increase in oxygen usage which occurs when their physiological activity is stimulated. For example, the oxygen uptake of secreting cells increases severalfold when rapid secretion is induced, and can increase by 50-fold or more in muscle when the resting tissue is stimulated into contractions. The increased energy usage promotes ATP hydrolysis and a fall in the ATP/ADP ratio, so that the concentration of ADP available to mitochondria increases and tissue respiration is accelerated. Changes in the ATP/ADP ratio have a controlling influence on other steps in intermediary metabolism, for example, the pyruvate dehydrogenase complex (see the section on Pyruvate Oxidation). In muscle and other tissues, the intracellular ATP/ADP ratio probably also provides an important metabolic link between mitochondrial respiration and tissue blood flow. When the steady-state ratio decreases, as occurs, for example, during increased energy demand or during hypoxia, the rise in ADP concentration leads to an increase in AMP concentration through adenylate kinase activity (2 ADP $\rightarrow$ ATP + AMP). The enzymic hydrolysis of AMP by cytosolic 5′-nucleotidase leads to the release of adenosine, a potent vasodilator, into the interstitial space (AMP + H_2O $\rightarrow$ adenosine + P_i). Adenosine binds to receptors in the smooth muscle of the arterioles, causing relaxation of the muscle and an increase in the blood supply of oxygen and fuels to the tissue. The intracellular adenine nucleotides lost in this way are replenished by biosynthesis.

Inhibitors of ATP Synthesis

In addition to the inhibition of ATP synthesis caused by the action of respiratory-chain inhibitors, two other types of inhibition occur due to a direct action on the ATP synthase complex and to the effect of uncoupling agents (uncouplers). Inhibitors of the ATP synthase complex include the antibiotic oligomycin, dicyclohexylcarbodiimide (DCCD), and organotin compounds. Each of these compounds abolishes ATP synthesis by blocking proton conduction through the F_O segment of the ATP synthase complex (Figure 17). Organotin compounds have been used widely in the plastics industry, as pesticides, and as antifouling compounds for ships' hulls. The estimated lethal oral dose of triethyltin in the human adult, about 70mg, is similar to that of cyanide. In the 1950s, the use of preparations containing diethyltin diiodide for the treatment of skin infections resulted in 100 deaths in France. The so-called classical uncoupling agents, including 2,4-dinitrophenol, are lipid soluble weak acids. They have a protonophore action, that is, they catalyze the transport of protons across the lipid bilayer of the inner mitochondrial membrane and therefore collapse the protonic potential formed by the respiratory chain.



2,4-dinitrophenol 2,4-dinitrophenolate anion

$$(HA \rightleftharpoons A^- + H^+)$$

This short-circuiting action results in a sustained and uncontrolled respiratory chain activity (with the release of respiratory energy as heat), a failure of ATP synthesis, and a simultaneous increase in ATPase activity (Figure 17). A similar uncoupling action on tissue mitochondria during poisoning by uncoupling agents results in a variety of symptoms, including deepened and accelerated respiration, listlessness and muscular weakness, difficulty in breathing, and death with extremely high fever (terminal hyperpyrexia). The approximate oral LD_{50} of 2,4-dinitrophenol is 25–50 mg/kg body weight. The extensive list of known uncoupling agents includes a variety of pesticides whose potent biological activity can be attributed to their high uncoupling activity. On a concentration basis, the uncoupling action of the pesticide benzimidazole compounds on isolated mitochondria is up to 1,000 times more potent compared with 2,4-dinitrophenol. An entirely different type of uncoupling action is brought about by arsenate, AsO_4^{3-}, which is chemically similar to phosphate, PO_4^{3-}, and competes with it at the catalytic centers of the ATP synthase. The product, ADP~arsenate, is much more unstable than the physiological product of the enzyme (ATP or ADP~phosphate) and it undergoes spontaneous hydrolysis with a release of energy as heat. Arsenate also has an uncoupling action on the substrate-level phosphorylation reactions in glycolysis and the succinyl-CoA synthetase reaction of the citric acid cycle. The uncoupling action of arsenate therefore differs from the protonophoric uncoupling agents such as 2,4-dinitrophenol. Arsenate, containing pentavalent arsenic, is much more readily excreted and far less toxic than trivalent arsenite, an inhibitor of oxoacid dehydrogenase complexes (see the section on Metabolism of 2-Oxoglutarate).

Mechanism of ATP Synthesis

The ATP synthase (Complex V) consists of two main sectors, a membrane complex F_O (where O stands for oligomycin and indicates the site of action of the inhibitor), and a roughly spherical F_1 complex (9 nm in diameter) which projects into the matrix space. These two sectors are connected by a stalk-like structure (Figure 18). The F_O sector contains several different subunits, including a group of DCCD-binding proteolipid subunits which are essential to its proton-conducting function. The F_1 complex, containing five different subunits ($\alpha\beta\gamma\delta\varepsilon$) has the

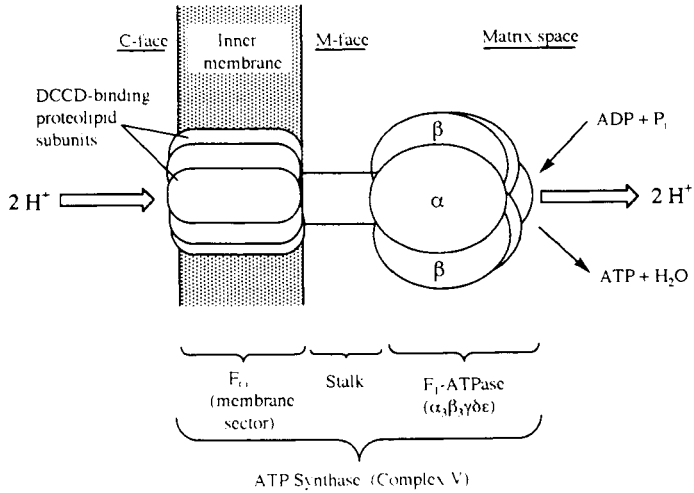


Figure 18. Simplified diagram to show the structure of the ATP synthase (Complex V). Subunits of the F_O sector form a structure spanning the lipid bilayer. Multiple copies of the DCCD-binding proteolipid subunit are organized to form a proton-conducting channel, which together with the stalk links the proton electrochemical potential across the membrane to the αβ catalytic centers of the F_1 sector. Other F_O subunits (not shown) probably organize the proteolipid subunits in the lipid bilayer and contribute to the stalk structure connecting the F_O and F_1 sectors. The F_1 sector also contains γδ and ε subunits (not shown). The ATP synthase is also called the F_O.F_1-ATPase, or the mitochondrial H^+-ATPase.

subunit composition $\alpha_3\beta_3\gamma\delta\varepsilon$, and it contains three catalytic centers formed at the three interfaces of the αβ subunit pairs. The stalk structure, about 4 nm in length, has at least two functions. First, it transmits the effect of the proton electrochemical potential, established by the proton-translocating respiratory chain, to the catalytic centers of the enzyme to drive ATP synthesis. Secondly, by separating the catalytic centers from the lipid bilayer it prevents the membrane potential from influencing the movements of substrates and products to and from the enzyme during ATP synthesis.

The most satisfactory explanation for the mechanism of respiratory-chain phosphorylation is provided by the chemiosmotic theory of Peter Mitchell (Senior, 1988). According to his theory, the respiratory chain and ATP synthase complexes are plugged through an ion-impermeable and proton-impermeable inner membrane, the osmotic barrier. During respiratory chain activity, complexes I, III, and IV function as proton pumps which translocate protons from the M-face to the C-face of the membrane. This process establishes a proton electrochemical potential across the membrane, and this potential is the driving force for ATP synthesis by the ATP synthase. The driving force is actually due to the combined effect of

two components, an electrical potential difference ($\Delta\Psi$) and a proton concentration difference (ΔpH) established across the membrane by the proton pumps. These two components constitute a protonic potential difference (Δp in millivolts, $\Delta\mu_{H^+}$ in kilojoules), defined by the equation:

$$\Delta p = \frac{\Delta\mu_{H^+}}{F} = \Delta\Psi - Z\,\Delta pH$$

where $\Delta\Psi$ and ΔpH represent differences in the electrical potential and pH, respectively, between the outside and the inside of the mitochondrial inner membrane. (Because ΔpH has a negative value, Δp is actually the numerical sum of the electrical and chemical potentials.)

When proton translocation is introduced into the ATP synthase mechanism, the reversible equation

$$ADP + P_i \rightleftharpoons ATP + H_2O$$

can be replaced by the equation

$$ADP + P_i + n\,H^+_{out} \rightleftharpoons ATP + H_2O + n\,H^+_{in}$$

where H^+_{out} and H^+_{in} represent protons in the aqueous media at the C-face and M-face of the inner membrane, respectively, and *n* protons are translocated per molecule of ATP synthesized. (The value of *n* has not been established definitively.) When the ratio H^+_{out}/H^+_{in} is increased by proton translocation during respiratory chain activity, the equilibrium of the latter reaction will be shifted to the right, in favor of ATP synthesis. How the flow of protons through the ATP synthase achieves this shift in equilibrium is unknown, but, broadly speaking, the chemiosmotic theory postulates that the process promotes separation and removal of the elements of the water molecule from the catalytic center of the enzyme, for example, in the form of $2H^+$ and O^{2-} (Figure 19).

The three catalytic centers of the ATP synthase show alternating-site cooperativity, that is, ATP formed at one catalytic center remains tightly-bound at a closed $\alpha\beta$ pair and it is not released until ADP and P_i bind to an adjacent open catalytic center. In an alternative and less precise coupling hypothesis, the conformational coupling mechanism, the alternating-site behavior is postulated to be part of a sequence of changes in subunit conformations initiated in the F_O sector by the protonic potential and transmitted to the catalytic centers via the DCCD-binding proteolipid subunits and the connecting stalk.

In Complex III, proton pumping occurs by a mechanism called the Q-cycle (Figure 20). The main feature of the cycle is the presence of two Q-reactive centers, $QH^{\bullet}_{in}$ and $QH^{\bullet}_{out}$, situated near the M-face and C-face of the inner membrane, respectively. The centers function as sites where ubisemiquinone is tightly bound and stabilized to establish $QH^{\bullet}$-Q and QH_2-$QH^{\bullet}$ couples at appropriate redox potentials, thereby preventing the dismutation of $QH^{\bullet}$ or its migration between the

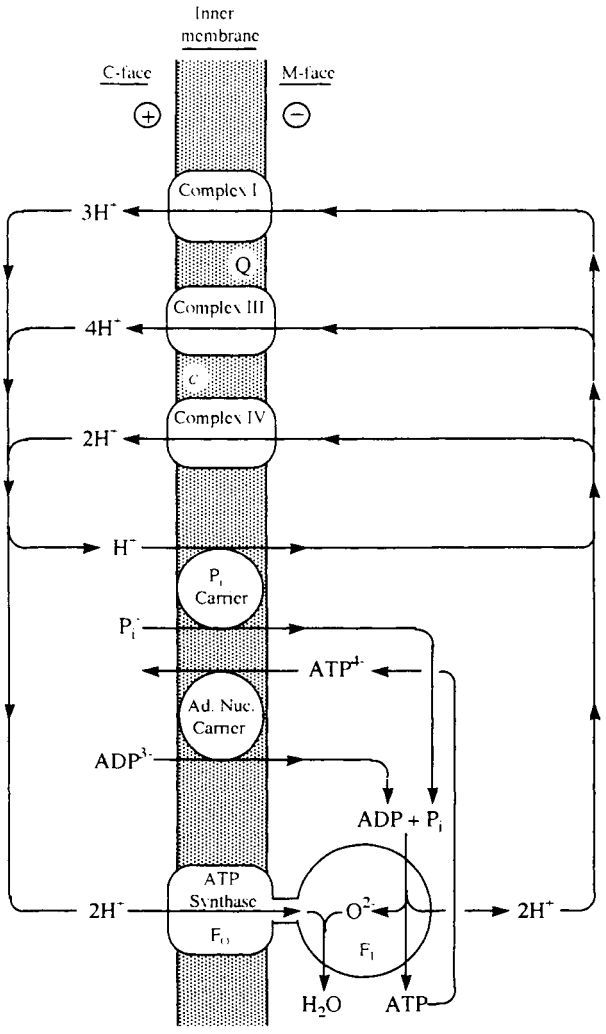


Figure 19. Chemiosmotic mechanism of ATP synthesis by the mitochondrial respiratory chain. The diagram shows the proton-translocating activity of complexes I, III, and IV to form a proton electrochemical potential across the inner membrane, and the return translocation of protons to the matrix space through the ATP synthase, to drive ATP synthesis. Metabolite carriers for inorganic phosphate (P_i) and adenine nucleotides (Ad. Nuc.) are also shown. The exact stoichiometry of proton movements is unknown, but is close to the values shown in the diagram, which suggest that nine protons are translocated during the oxidation of a molecule of NADH and six during succinate oxidation; and that three protons are used to drive the synthesis of ATP and the associated transport of P_i. These values indicate a P:O ratio of three during NADH oxidation and two during the oxidation of succinate to fumarate.

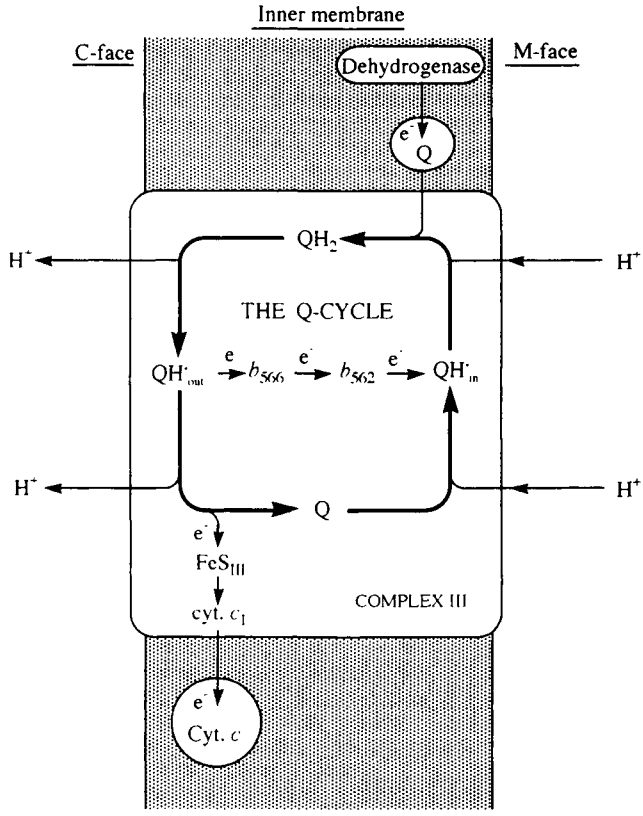


Figure 20. Reaction scheme for oxidation and proton translocation by Complex III, the Q-cycle. The scheme shows the location of two Q-reactive centers, $QH^{\bullet}_{in}$ and $QH^{\bullet}_{out}$, and two cytochrome *b* heme components, b_{566} and b_{562}. The complex translocates two protons per electron donated by a dehydrogenase, corresponding to four protons translocated per molecule of substrate oxidized by the complex.

centers, which would short-circuit the cycle. Hypothetical mechanisms have been proposed to explain proton-pumping by Complexes I and IV. They also include cyclical movements of hydrogen-carrying and electron-carrying compounds, for example, an H_2O_2/O_2 cycle or O-cycle in Complex IV.

Regulation of Energy Metabolism

The primary signal for the need to stimulate the rate of ATP synthesis is a fall in the cytosolic ATP/ADP ratio resulting from increased ATP hydrolysis mainly outside the mitochondria. An increased ADP concentration in the cytosol stimulates ADP entry into mitochondria in exchange for ATP exit on the adenine nucleotide

carrier. The rise in mitochondrial ADP concentration activates the respiratory control mechanism and leads to an increase in the $NAD^+/NADH$ ratio. This in turn stimulates isocitrate dehydrogenase and the dehydrogenase complexes for pyruvate and 2-oxoglutarate, whose activities are each critically dependent on the $NAD^+/NADH$ ratio. An additional form of regulation is provided by changes in cytoplasmic $[Ca^{2+}]$. Three exclusively mitochondrial enzymes, pyruvate dehydrogenase phosphatase, 2-oxoglutarate dehydrogenase, and NAD-linked isocitrate dehydrogenase, are activated by an increase in free $[Ca^{2+}]$ in the micromolar range. The fluctuation of matrix-space free $[Ca^{2+}]$ in response to changes in cytosolic free $[Ca^{2+}]$ provides precise control over the activity of these enzymes, and is another way in which oxidative metabolism in mitochondria keeps pace with the intensity of tissue energy demands (McCormack et al., 1990).

METABOLITE TRANSPORT ACROSS MITOCHONDRIAL MEMBRANES

Numerous metabolite carriers are present in the inner membrane of mitochondria in various tissues (Table 6). The study of the individual carriers is far from complete, but many of them constitute a homologous homodimer protein family with similar size and secondary structure, and they have probably evolved from a common ancestral protein (Kramer and Palmieri, 1989). A common feature is the likely presence of 12 transmembrane α-helices, with each monomer contributing six α-helices. Nearly all the carriers are inhibited by protein-modifying reagents, including sulfhydryl-binding reagents, and highly specific inhibitors are known for some of the carriers. During metabolite transport the carriers show saturation kinetics, high metabolite specificity (including isomer specificity) and a high energy of activation. (An exception is the calcium uniport carrier, which seems to function as a pore spanning the inner membrane.) All these features are characteristic of protein-catalyzed processes. The individual nature of the carriers is shown by their characteristic distribution, according to the physiological functions of the tissue (Azzi et al., 1989). The primary function of the carriers is to catalyze the uptake by mitochondria of metabolites required for the fundamental process of oxidative phosphorylation, without causing a collapse of the proton electrochemical potential generated across the inner membrane by the respiratory chain. These metabolites include ADP, P_i, and a variety of oxidizable substrates. The need for ADP and P_i transport from the cytosol arises because the catalytic centers of the ATP synthase complexes face the matrix space, whereas most of the ATP-consuming reactions of the cell which regenerate ADP and P_i occur outside mitochondria. This situation means that ATP must be exported from mitochondria, and this need is satisfied by linking ADP entry to ATP exit on the adenine nucleotide carrier. These essential transport requirements during ATP synthesis explain the presence of the adenine nucleotide and P_i carriers in all mitochondria. Arsenate, an uncoupler of oxidative phosphorylation reactions (see the section on Inhibitors of ATP

Table 6. Metabolite Carriers of Mitochondria

Carrier	*Transport Process Catalyzed*	*Function(s)*
Phosphate	Phosphate/H^+ symport	Oxidative phosphoryation; cotransport of P_i during cation uptake; linkage of P_i movements on other carriers
Adenine nucleotide	ADP/ATP antiport	Oxidative phosphorylation
Pyruvate (monocarboxylate)	Pyruvate/H^+ symport	Provision of pyruvate for the pyruvate dehydrogenase complex and pyruvate carboxylase
Carnitine	Carnitine/acylcarnitine antiport	Long-chain fatty acid transport
2-oxoglutarate	2-oxoglutarate/L-malate antiport	Malate-aspartate shuttle; gluconeogenesis from lactate; urea synthesis
Aspartate	L-glutamate + H^+/aspartate antiport	Malate-aspartate shuttle; gluconeogenesis from lactate; urea synthesis
Glutamate	L-glutamate/H^+ symport	Nitrogen metabolism; urea synthesis
Dicarboxylate	L-malate/P_i antiport	Gluconeogenesis from lactate, pyruvate, amino acids; urea synthesis
Tricarboxylate	Citrate + H^+/L-malate antiport	Lipogenesis; gluconeogenesis from phosphoenolpyruvate (PEP)
Glutamine	L-glutamine uniport	Glutamine metabolism
Ornithine	Ornithine/citrulline + H^+ antiport	Urea synthesis
Calcium	Calcium uniport	Regulation of cytoplasmic free $[Ca^{2+}]$
Sodium-calcium	2 Na^+-Ca^{2+} antiport	Sodium-calcium cycle
Sodium	Na^+-H^+ antiport	Sodium-calcium cycle
Potassium	K^+-H^+ antiport	Regulation of mitochondrial volume

Synthesis), is also transported from the cytosol to the matrix space by the phosphate carrier. Pyruvate and long-chain fatty acids commonly serve as the main fuels for mitochondrial oxidation, and carriers for these compounds also have a wide tissue distribution. Dicarboxylate carrier activity is high in liver and kidney, where it functions in gluconeogenesis, and tricarboxylate carrier activity is high in liver and adipose tissue, where transport of citrate from the matrix space to the cytosol is required during lipogenesis (see below). Liver mitochondria also contain carriers catalyzing ornithine import and citrulline export during urea synthesis (see the section on the Relationship of the Cycle to Urea Synthesis). Inhibitors of the metabolite carriers have been discovered during experimental studies on isolated mitochondria. The most specific inhibitors are atractyloside and ruthenium red. Atractyloside is a toxic glycoside compound present in the rhizomes of *Atractylis gumnifera*, a thistle growing in the southern Mediterranean region. Poisoning by

atractyloside blocks the activity of the adenine nucleotide carrier, causes a decrease in body oxygen consumption, and results in death during convulsive hypoglycemia. Ruthenium red, an inorganic hexavalent ruthenium complex, is a very specific inhibitor of the calcium uniport carrier.

The carrier-mediated transport of a metabolite molecule occurs in one of three ways: the transfer across the membrane is coupled either to the countertransfer of another metabolite molecule (antiport) or to the transfer of another molecule in the same direction (symport); or there is a unidirectional transfer of a metabolite molecule (uniport). All three kinds of transport occur across the inner membrane (Table 6). Each carrier protein is thought to carry out its transport function by alternating between two different conformational states during a translocation event. This allows carrier-bound substrate molecules to exchange with like molecules in solution, first on one side of the membrane and then on the other side, but never simultaneously on both sides. The channel carrying bound substrate is probably formed by closely-packed polypeptide structures which part and roll over the bound substrate during the transition from one conformational state to the other state of the carrier (alternating-access model). Charge separation occurs across the membrane both when tetravalent ATP is exchanged for trivalent ADP, and when one molecule of aspartate anion leaves the matrix space in exchange for one molecule of undissociated glutamate. In coupled mitochondria, the membrane potential generated across the membrane during respiration (negative inside) stimulates the ejection of ATP and aspartate to the cytosol, with the effect that both carriers catalyze unidirectional transport.

Mitochondria also contain carriers for metal-ion transport across the inner membrane, among which the activity of the calcium uniporter is very much higher than that of the other carriers. Mitochondria contain separate calcium uptake and release pathways involved in the regulation of free $[Ca^{2+}]$ in the matrix space and cytosol. The transport of Na^+ is linked to the movement of both Ca^{2+} and H^+ across the inner membrane. Fluctuations in the intracellular free $[Ca^{2+}]$ within the submicromolar concentration range are very important for the regulation of many cell functions. The cytosolic free $[Ca^{2+}]$ is determined by the coordinated activity of various Ca^{2+}-transporting systems located in the boundary membranes of the cell (plasma membrane and endoplasmic reticulum) and in the mitochondria. Under normal physiological conditions, the rate of cyclical calcium movements across the inner membrane is low and avoids an unnecessary dissipation of respiratory energy, because the calcium uniporter has a high apparent K_m for Ca^{2+} compared with the cytosolic free $[Ca^{2+}]$, and the V_{max} of the calcium release pathway is low. A rise in the cytosolic free $[Ca^{2+}]$ stimulates the calcium uniporter to transport Ca^{2+} into the matrix space, without causing a corresponding increase in calcium release, thus returning the cytosolic free $[Ca^{2+}]$ to its normal value. In this way, the high speed and capacity of calcium uptake enables mitochondria to provide a vital protection for cells when injurious influences cause a temporary influx of large amounts of

additional calcium into the cytoplasm. Conversely, a fall in cytosolic calcium decreases the uniporter activity without a corresponding decrease in the release rate, and this restores the cytosolic free [Ca^{2+}] by depleting mitochondrial calcium. The calcium uniporter probably also functions as a carrier for Fe^{2+} import into mitochondria for use in heme biosynthesis.

Calcium ions have an important role in cell death when mitochondrial activity is impaired. One early event after the loss of mitochondrial function due to anoxia or to poisoning is an uncontrolled and sustained increase in cytosolic calcium. This occurs for two reasons. First, there is a depletion of cytosolic ATP and hence a failure of cellular Ca^{2+}-efflux mechanisms in the boundary membranes of the cell, which depend on Ca^{2+}-ATPase activity. Secondly, the fall in the mitochondrial membrane potential prevents calcium uptake into the matrix space, and there is a loss of mitochondrial calcium to the cytosol. An increase in cytosolic [Ca^{2+}] leads to the rapid activation of Ca^{2+}-dependent degradative enzymes, including phospholipases, proteases, and endonucleases, which in turn causes damage to vital cell components and results in cell death.

Calcium ions have an uncoupling action on isolated mitochondria, probably brought about indirectly by their activation of phospholipase A_2. The action of this enzyme on bilayer phospholipids causes an increase in proton permeability of the inner membrane, and the release of fatty acids and lysophospholipids which also have an uncoupling activity. Uncoupling by Ca^{2+} is thought to occur in the condition of malignant hyperthermia (MH; also called malignant hyperpyrexia), an autosomal dominant hereditary condition induced in predisposed individuals by pharmacologic agents in common use, including volatile anesthetics (e.g., halothane) and muscle relaxants (e.g., succinyl-choline). The syndrome is characterized by a rapid increase in body temperature, tachycardia, arrhythmias, and muscle rigidity. It progresses to death unless it is recognized quickly and treated effectively (Foster, 1990). Muscle mitochondria accumulate large amounts of Ca^{2+} in the matrix space during a malignant hyperthermia crisis, and this results in swelling and structural damage to the organelles. The biochemical basis of the intense hyperthermia is thought to be an uncoupling action of Ca^{2+} on muscle mitochondria, following a sudden increase in cytoplasmic [Ca^{2+}] triggered by the pharmacological agent. Recent studies of the syndrome have implicated abnormalities in the phosphoinositide cycle and in the sarcoplasmic reticulum calcium-release channel of muscle cells.

Role of the Carriers in Gluconeogenesis and Lipogenesis

The role of several carriers during oxidative phosphorylation reactions and during urea synthesis is described elsewhere in the text. This section provides a brief account of carrier functions during the synthesis of glucose (gluconeogenesis) and fatty acids (lipogenesis). These functions are summarized in the diagrams

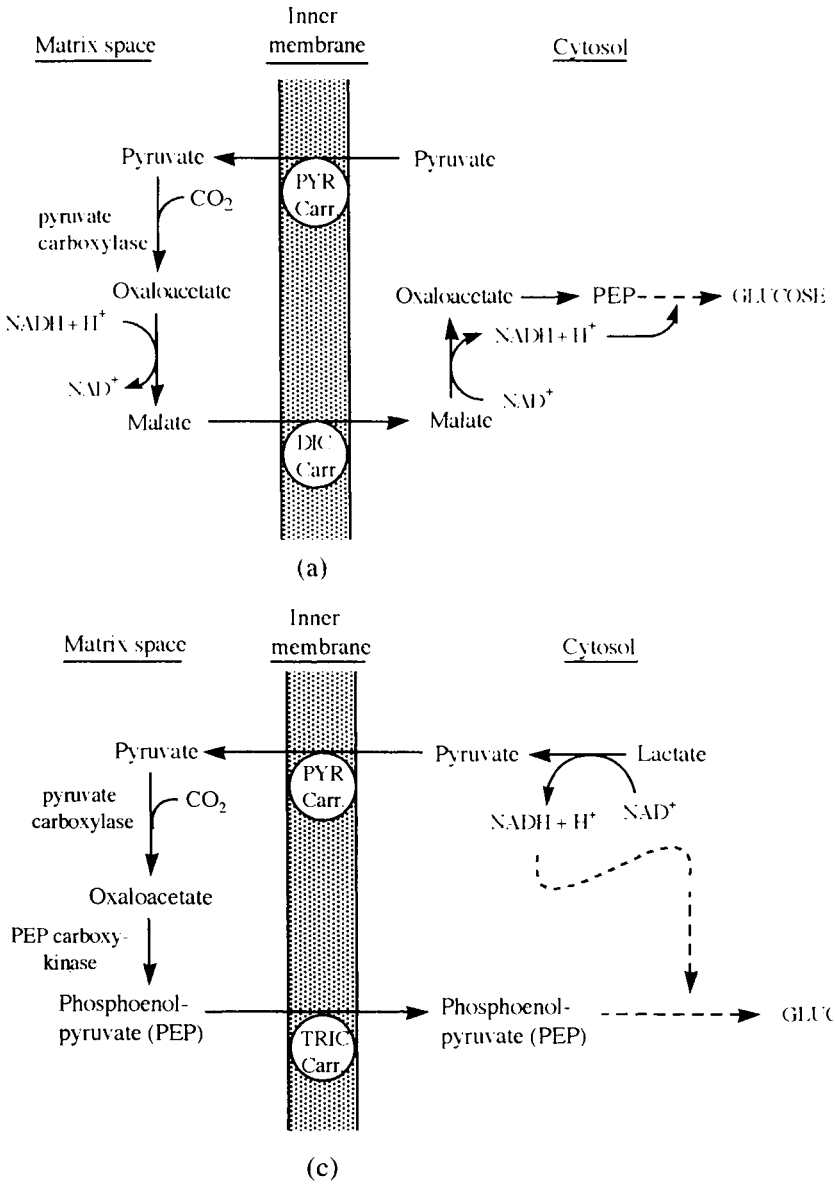


Figure 21. Role of mitochondrial metabolite carriers during gluconeogenesis and lipogenesis. (**a**) Gluconeogenesis from pyruvate; (**b**) gluconeogenesis from lactate, using cytosolic phosphoenolpyruvate carboxykinase;*(continued)*

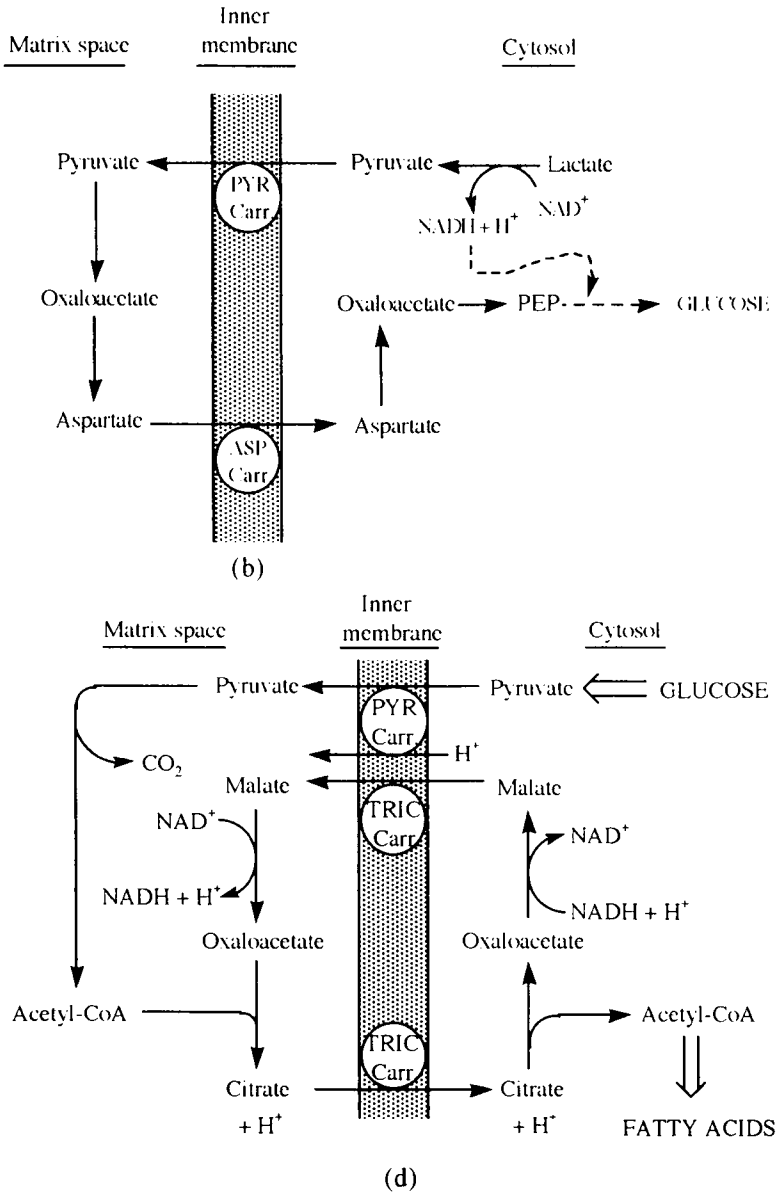


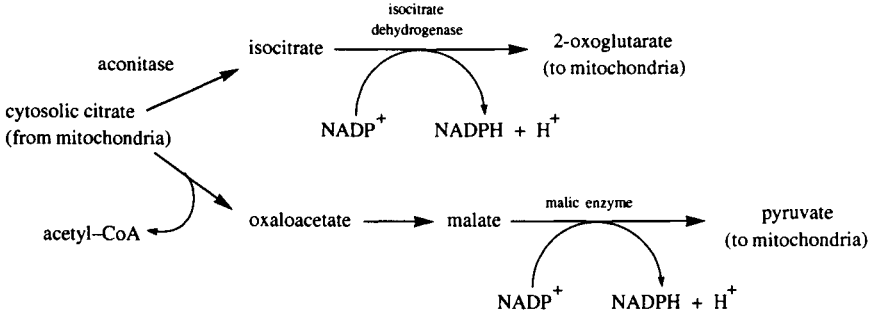
Figure 21. (continued) Role of mitochondrial metabolite carriers during gluconeogenesis and lipogenesis. (**c**) gluconeogenesis from lactate, using mitochondrial phosphoenolpyruvate carboxykinase; and (**d**) lipogenesis from glucose.

shown in Figure 21. The main precursors for gluconeogenesis (mainly in liver and kidney) are lactate, pyruvate, amino acids derived from skeletal muscle, and glycerol released during the hydrolysis of neutral lipid stores in adipose tissue. During gluconeogenesis, the carbon skeletons of the precursors undergo conversion to oxaloacetate and then to phosphoenolpyruvate (PEP). When pyruvate serves as precursor, the carbon skeleton for oxaloacetate formation is transported from the matrix space in the form of malate, which also carries reducing equivalents to the cytosol. When lactate is precursor no additional reducing equivalents are required, and the carbon skeleton is exported from mitochondria in the form of aspartate. In human liver, almost 50% of the total PEP CK activity is located in the matrix space. Some oxaloacetate is converted to PEP in mitochondria, and PEP is then transported to the cytosol on the tricarboxylate carrier.

Lipogenesis takes place in many tissues, including liver, kidney, brain, mammary gland, and adipose tissue. When fatty acids are synthesized from glucose, acetyl-CoA formed in the matrix space by the pyruvate dehydrogenase complex cannot be transferred directly to the cytosolic fatty acid synthase complex, because the inner membrane is impermeable to acetyl-CoA. Acetyl units are first converted to citrate, which is then transported from the matrix space to the cytosol on the tricarboxylate carrier (Figure 21). In the cytosol, acetyl-CoA is regenerated from citrate by the action of ATP-dependent citrate lyase:

$$\text{citrate}^{3-} + \text{CoASH} + \text{ATP}^{4-} \rightarrow \text{acetyl-CoA} + \text{oxaloacetate}^{2-} + \text{ADP}^{3-} + \text{P}_i^{2-}$$

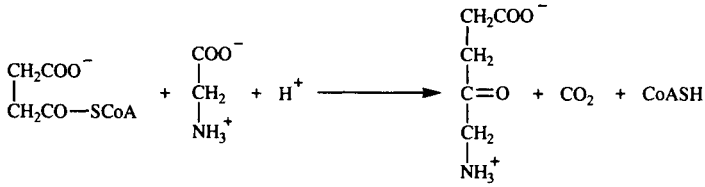
The oxaloacetate formed can be converted to malate, which then returns to the matrix space as counteranion on the tricarboxylate carrier. The regeneration of oxaloacetate in the matrix space completes a cycle of reactions, the malate-citrate shuttle, during which one molecule of ATP is consumed by citrate lyase for each acetyl unit exported from mitochondria, and a pair of reducing equivalents are transferred from NADH in the cytosol to the matrix space. More complex patterns of metabolism during lipogenesis not only facilitate acetyl-group transport from mitochondria, but also enable NADPH required for lipogenesis to be generated in the cytosol by the malic enzyme or by isocitrate dehydrogenase:



OTHER FUNCTIONS OF MITOCHONDRIA

Heme Biosynthesis

Mitochondrial enzymes fulfill an important role during the biosynthesis of heme, an essential component of cytochromes, hemoglobin, myoglobin, and catalase. The synthetic pathway, which depends on at least eight enzyme-catalyzed steps, begins and ends in the mitochondria (Figure 22). The intermediates of the pathway are relatively hydrophobic, and their transport between the matrix space and the cytosol does not seem to require specific metabolite carrier proteins. The first and rate-determining step of the pathway is catalyzed by the matrix-space enzyme 5-aminolevulinate synthase (ALA synthase), and provides the primary control point:



The synthase uses pyridoxal phosphate as cofactor. The turnover of ALA synthase is very rapid and synthesis of the enzyme is regulated by heme, the end product of the pathway. Heme probably regulates in one or more of the following ways: by repression of enzyme synthesis; by direct end-product inhibition; by inhibition of the transport and processing of the precursor form of the synthase. The three final

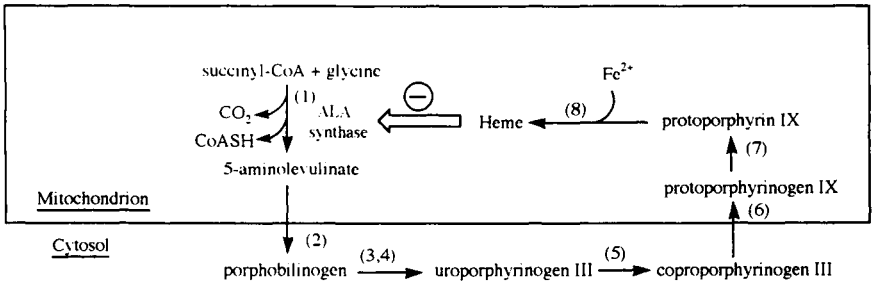
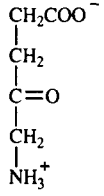


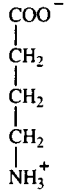
Figure 22. Role of mitochondria in heme biosynthesis. Numbers refer to enzyme-catalyzed steps, showing that steps 1,6,7, and 8 take place in mitochondria. The enzymes catalyzing these reactions are as follows: (1) 5-aminolevulinate synthase; (6) coproporphyrinogen III oxidase; (7) protoporphyrinogen oxidase; and (8) ferrochelatase. The diagram also shows the regulation of the pathway by heme, which acts upon the first and rate-determining step (see text).

steps of the pathway are catalyzed by mitochondrial enzymes associated with the inner membrane. Heme synthesized in mitochondria is retained for cytochrome synthesis or distributed to other regions of the cell for hemoprotein formation, for example: cytochrome P_{450} in the endoplasmic reticulum; catalase in peroxisomes; and hemoglobin in the cytosol of erythroblasts, the precursor cells of erythrocytes.

In the porphyrias, a group of inherited diseases, porphyrins and their precursors (including 5-aminolevulinate) accumulate and are excreted in excessive amounts. The condition is caused by defects in the individual enzymes of the pathway, including the mitochondrial enzymes. The main clinical symptoms are intermittent attacks of nervous system disorders (which are poorly understood) and/or sensitivity of the skin to sunlight. It is possible that 5-aminolevulinate acts as a neurotoxin. It bears a structural resemblance to 4-aminobutyrate, the major inhibitory transmitter substance in the vertebrate central nervous system:



5-aminolevulinate

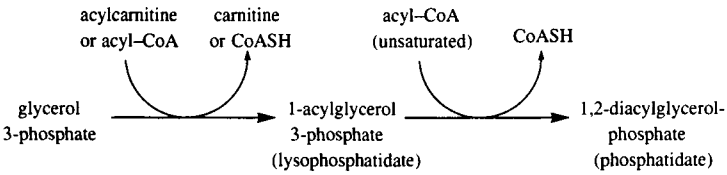


4-aminobutyrate

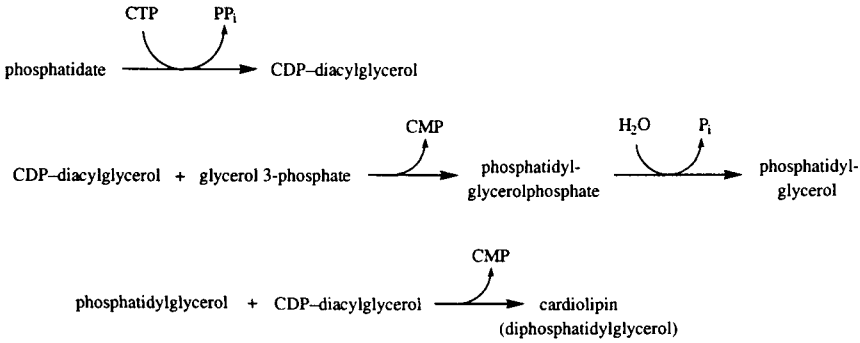
Porphyrins accumulated in the skin or circulating blood vessels of the skin absorb light in the ultraviolet region, which causes a transition of the porphyrin molecules to excited electronic states. The excited porphyrins can react directly with biological structures or with molecular oxygen, generating highly-reactive and damaging oxygen species, including singlet oxygen. These events explain the photocutaneous symptoms in the porphyrias.

Phospholipid Biosynthesis

Mitochondria have a limited capacity to synthesize lipids, apart from the synthesis of cardiolipin and the decarboxylation of phosphatidylserine to form phosphatidylethanolamine. (The latter reaction is catalyzed by a decarboxylase enzyme present in the inner membrane.) Other lipids required during mitochondrial biogenesis are synthesized elsewhere in the cell, especially in the endoplasmic reticulum, and are then transferred to the growing mitochondrion. Phosphatidic acid (phosphatidate) required for cardiolipin synthesis is formed by the acylation of glycerol 3-phosphate, in two reaction steps catalyzed by enzymes present in the mitochondrial outer membrane:



Phosphatidic acid is converted to CDP-diglycerol, an intermediate which functions as the phosphatidyl-group donor in two later steps leading to the formation of cardiolipin:



REFERENCES

Angelini, C., Trevisan, C., Isaya, G., Pegolo, G., & Vergani, L. (1987). Clinical varieties of carnitine and carnitine palmitoyltransferase deficiency. Clin. Biochem. 20, 1–7.

Azzi, A., Nalecz, K. A., Nalecz, M. J., & Wojtczak, L. (eds.) (1989). The Anion Carriers of the Mitochondrial Membranes. Springer-Verlag, Heidelberg.

Bereiter-Hahn, J. (1990). Behavior of mitochondria in the living cell. Int. Rev. Cytol. 122, 1–64.

Cadenas, E. (1989). Biochemistry of oxygen toxicity. Ann. Rev. Biochem. 58, 79–110.

Capaldi, R. A. (1990). Structure and function of cytochrome c oxidase. Ann. Rev. Biochem. 59, 569–596.

Cross, C. E., Halliwell, B., Borish, E. T., Pryor, W. A., Ames, B. N., Saul, R. L., McCord, J. M., & Harman, D. (1987). Oxygen radicals and human disease. Ann. Int. Med. 107, 526–545.

Devlin, T. M. (ed.) (1992). Textbook of Biochemistry with Clinical Correlations. Wiley-Liss, New York.

Foster, P. S. (1990). Malignant hyperpyrexia. Int. J. Biochem. 22, 1217–1222.

Guenerich, F. P. (1992). Cytochrome P450: advances and prospects. FASEB J. 6 (2), Thematic issue.

Hall, P. F. (1984). Cellular organization for steroidogenesis. Int. Rev. Cytol. 86, 53–95.

Harold, F. M. (1986). The Vital Force: A Study of Bioenergetics. Freeman, New York.

Hatefi, Y. (1985). The mitochondrial electron transport and oxidative phosphorylation system. Ann. Rev. Biochem. 54, 1015–1069.

Jenkins, T. M. & Weitzman, P. D. J. (1988). Physiological roles of animal succinate thiokinases. FEBS Lett. 230, 6–8.

Kramer, R. & Palmieri, F. (1989). Molecular aspects of isolated and reconstituted carrier proteins from animal mitochondria. Biochim. Biophys. Acta 974, 1–23.

McCormack, J. G., Halestrap, A. P., & Denton, R. M. (1990). Role of calcium ions in regulation of mammalian intramitochondrial metabolism. Physiol. Rev. 70, 391–426.

Meijer, A. J., Lamers, W. H., & Chamuleau R. A. F. M. (1990). Nitrogen metabolism and ornithine cycle function. Physiol. Rev. 70, 701–748.

Newsholme, E. A. & Leech, A. R. (1983). Biochemistry for the Medical Sciences. Wiley, Chichester.

Nicholls, D. G. (1982). Bioenergetics: An Introduction to the Chemiosmotic Theory. Academic Press, New York.

Pfanner, N. & Neupert, W. (1990). The mitochondrial protein import apparatus. Ann. Rev. Biochem. 59, 331–354.

Scholte, H. R. (1988). The biochemical basis of mitochondrial diseases. J. Bioenerg. Biomembr. 20, 161–191.

Schulz, H. & Kunau, W. H. (1987). Beta-oxidation of unsaturated fatty acids: a revised pathway. Trends Biochem. Sci. 12, 403–406.

Schulz, H. (1991). Beta-oxidation of fatty acids. Biochim. Biophys. Acta 1081, 109–121.

Senior, A. E. (1988). ATP synthesis by oxidative phosphorylation. Physiol. Rev. 68, 177–231.

Skulachev, V. P. (1988). Membrane Bioenergetics. Springer-Verlag, Berlin.

Trayhurn, P. & Nicholls, D. G. (eds.) (1986). Brown Adipose Tissue. Arnold, London.

Tyler, D. D. (1992). The Mitochondrion in Health and Disease. VCH Publishers, New York.

Yeaman, S. J. (1989). The 2-oxo acid dehydrogenases; a complex family. Trends Biochem. Sci. 257, 625–632.

Zammit, V. A. (1984). Mechanisms of regulation of the partition of fatty acids between oxidation and esterification in the liver. Prog. Lipid Res. 23, 39–67.

Chapter 5

Lysosomal Storage Diseases

GRAZIA M. S. MANCINI and FRANS W. VERHEIJEN

INTRODUCTION AND HISTORY

In the early 1960s, De Duve discovered a new organelle in eukaryotic cells containing a vast array of hydrolases active at acid pH and surrounded by a single membrane. He called these organelles lysosomes from the Greek λυω (dissolve) and σομα (body). At that time storage disorders (thesaurismoses) were known, but

Principles of Medical Biology, Volume 3
Cellular Organelles and the Extracellular Matrix, pages 133–154

ISBN: 1-55938-804-8

Table 1. Criteria for the Identification of an Inborn Lysosomal Disease*

1. The disease be a storage disease (thesaurismosis)
2. The depot be segregated in vacuoles related to lysosomes
3. The substance that accumulates be susceptible to degradation by an extract of normal lysosomes
4. One of the enzymes involved in this process be inactive in the patients

Note: *From Hers, 1965.

it was debatable whether the storage was the effect of hyperproduction or faulty degradation of substances.

The link between lysosomes and storage disorders was established in 1963 when Hers discovered that patients with Pompe disease, an inherited generalized glycogenosis, lacked the enzyme acid maltase. In these patients undegraded glycogen accumulates in the lysosomes which, consequently, increase in number and size and cause cell damage. This discovery was soon followed by the identification of lysosomal enzyme defects in Gaucher disease by Patrick in 1965 and in Niemann-Pick disease by Brady et al. in 1966.

These findings led to the concept of "lysosomal disease" which as described by Hers is a thesaurismosis caused by a genetic deficiency of a lysosomal hydrolase, leading to lysosomal storage of the specific undegraded substrate (see Table 1). For the first time, a group of human diseases could be attributed to dysfunction of a subcellular organelle. Thus, medical interest was redirected from organ to organelle disease. Nowadays, it is quite common to refer to mitochondrial or peroxisomal disorders as abnormalities of metabolic functions which are specifically regulated in each of these subcellular organelles.

By far, most lysosomal diseases known to date are caused by enzyme defects. Lysosomal enzymes are exohydrolases that work sequentially, usually starting at the nonreducing end for polysaccharidic substrates (from glycosaminoglycans or glycoproteins), and at the hydrophilic end for lipid substrates. The better characterized enzymes are glycohydrolases. To date, most of the known human lysosomal diseases have been ascribed to a deficiency of a glycohydrolase. Since these glycohydrolases work in a sequential fashion, the degradation product of the first enzyme becomes the substrate for the subsequent reaction. When one of the enzymes is deficient the whole degradative sequence is blocked and complete macromolecules remain undigested in the organelle. Unlike glycohydrolases, lysosomal proteases are either endo- or exo-peptidases, and probably do not need to act sequentially on the substrate. Thus far, no protease deficiency leading to peptide chain accumulation has been described. While no human lysosomal nucleosidase deficiency is known, lipase deficiencies cause lysosomal storage diseases such as Wolman disease, and cholesteryl ester storage disease (acid lipase defects), and Farber disease (ceramidase deficiency). After the description of the first patients, it soon became clear that lysosomal enzyme deficiencies can have

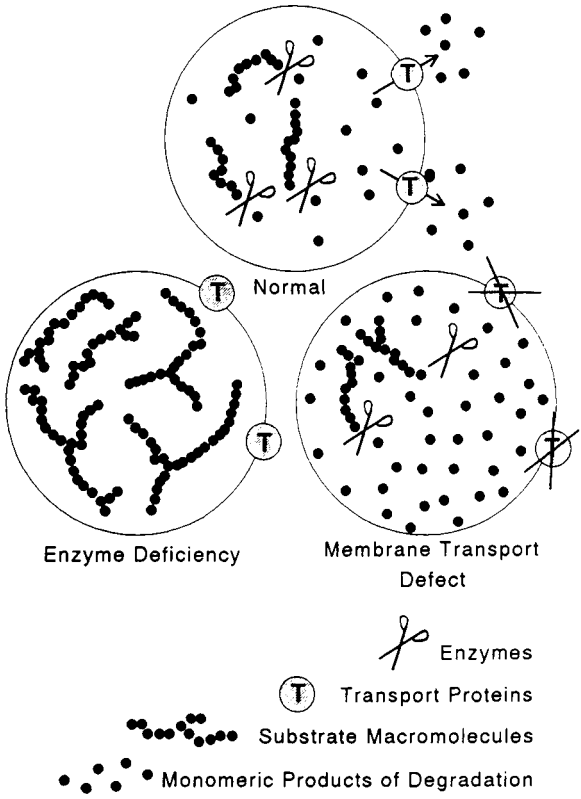


Figure 1. A schematic representation of the two main genetic causes of lysosomal storage disease. In the absence of one of the lysosomal hydrolases, undigested macromolecules accumulate, while in the case of a transport defect small molecular weight compounds fill the lysosomal compartment.

polymorphic clinical presentations, and such clinical heterogeneity is also observed in patients with the same enzyme deficiency.

The definition of lysosomal storage disease was not expected to remain confined to that of enzyme deficiency (see Figure 1). Cystinosis, for example, is a lysosomal storage disease where a single amino acid, cystine, accumulates in the lysosomal compartment, yet no enzyme defect is present. The discovery that in cystinosis a specific membrane transport system is defective provided the impetus for studies of the function of the lysosomal membrane in the control of metabolite flux. As a result of these efforts about twenty transport mechanisms have been characterized so far, and other diseases have been associated with a membrane transport defect. As it often happens in scientific research, the information obtained from the observation of naturally occurring mutants has largely contributed to a better understanding of lysosomal biogenesis and function.

CLASSIFICATION

About thirty five different lysosomal storage disorders (LSD) have been described and on the basis of the substrate involved in the degradative defect they are classified as follows:

- Mucopolysaccharidoses (MPS): defects in the degradation of different sorts of ubiquitous glycosaminoglycans.
- Lipidoses: degradation defects of glycolipids, e.g., gangliosides, sphingomyelin, cerebrosides, sulfatides, or other lipids including ceramide, triglycerides, and cholesteryl esters.
- Glycoproteinoses: degradation defects of the oligosaccharide chains of several glycoproteins, where the terminal undegraded sugar often provides the name of the disease. Prime examples are fucosidosis, mannosidosis, sialidosis, and aspartylglucosaminuria.
- Glycogenosis type II: the only lysosomal defect of glycogen digestion, which is caused by acid maltase (α-glucosidase) deficiency.
- Cystinosis: transport defect for the disulfide amino acid cystine.
- Sialic acid storage diseases: transport defect for acid monosaccharides, e.g., sialic acid and glucuronic acid.
- Vitamin B_{12} transport defect: (also classified as cobalamin F mutation) is one of the congenital disorders of vitamin B_{12} metabolism.
- Lysosomal storage diseases without a known etiology: some of them are worth mentioning. In Niemann-Pick disease type C, there is a defect in the recycling of cholesterol from the lysosome to the endoplasmic reticulum. However, the nature of the transport mechanism out of the lysosome is not yet known. Neuronal ceroid lipofuscinoses are other lysosomal diseases the causes of which remain unknown. In the late infantile and juvenile form, lysosomal accumulation of subunit C of the mitochondrial ATP synthase has been documented. Mucolipidosis IV is a storage disorder in which different sorts of gangliosides accumulate in the lysosome. The early suspicion of a defect in a ganglioside-specific neuraminidase has not been substantiated.

The defect of a lysosomal protein can be expressed in every cell type except the erythrocyte. However, the clinical manifestation of the disease is determined by the tissue specificity of the substrate involved in the degradation defect. For example, a degradation defect of keratansulfate, a mucopolysaccharide present in connective tissues, causes a mucopolysaccharidosis with bone dysplasia and degenerative changes of joint cartilage and other connective tissues without the involvement of other systems. On the other hand, defective degradation of ganglioside G_{M2}, a component of neuronal plasma membrane essential for proper signal transmission, produces signs of neuroaxonal degeneration in the central and peripheral nervous systems without the involvement of joints or other visceral organs. The expression

of the genetic defect in cell types not involved in the disease has been a useful contrivance for the diagnosis of disorders of poorly accessible systems such as the nervous system. In most lysosomal diseases, the study of enzyme activities in blood cells or cultured fibroblasts appears to be both reliable and sufficient for accurate diagnosis. In Table 2, lysosomal enzyme deficiencies are classified according to the defective enzyme with mention of the type of substrates which accumulate in the cells. Table 2 indicates how some of the glycohydrolases can catalyze hydrolysis of the same sugar attached to different compounds. This results in the storage of various macromolecules all ending with the same monosaccharide. An example is G_{M1} gangliosidosis where acid β-galactosidase deficiency causes accumulation of G_{M1} ganglioside, glycoprotein- and keratansulfate-derived oligosaccharide chains, all containing a terminal β-galactose at the nonreducing end.

Irrespective of the type of biochemical defect, all the disorders in which a sufficient number of patients has been described show a peculiar variability in the time of onset and severity of the symptoms.

LYSOSOME BIOGENESIS

The lysosome is one of the vesicular organelles involved in the endocytic pathway. To carry out its digestive function, the lysosome must communicate with several other intracellular compartments.

A chapter about lysosomal diseases would not be complete without briefly mentioning the biosynthetic steps of the most abundant lysosomal proteins: lysosomal enzymes. Most of the enzymes are soluble acid hydrolases. The genetic information for their synthesis is contained in genes expressed in almost every cell type (therefore called "house-keeping" genes). The enzymes are synthesized in the lumen of the rough endoplasmic reticulum as a polypeptide chain which is later glycosylated; this is similar to what happens to proteins destined for extracellular secretion. Glycosylation takes place on specific asparagine residues provided the residue is in the conformation -Asn-X-Thr/Ser-, where X can be any amino acid except proline. An event specific for lysosomal enzymes takes place in the Golgi complex, where the glycosidic chains acquire a terminal mannose-6-phosphate marker. This marker enables recognition of these proteins by a specific receptor (present in two isoforms: Ca^{2+}-dependent and -independent) present in the membrane of the *trans*-Golgi compartment. Binding to this receptor allows sorting to a prelysosomal acidic compartment (pH 6.0), where release from the receptor is caused by the low pH. The enzyme is then transferred to its final destination, which is a lysosomal compartment at lower pH (about 4.8), while the receptor is recycled to the Golgi system. This process is not 100% leakage-free, since a small amount of protein is secreted along with other secretory proteins. The Ca^{2+}-independent receptor is also present on the plasma membrane where it serves as the insulinlike growth factor II receptor and participates in the endocytosis of extracellular enzymes containing the mannose-6-phosphate recognition marker via clathrin-

Table 2. Lysosomal Storage Disorders: Classification

Disease	*Enzyme/Protein Defect*	*Gene Locus*	*Storage Product*
(Sphingo) Lipidoses			
G_{M2} gangliosidosis (Tay-Sachs disease)	hexosaminidase A (α chain)	α chain: 15q22-q25.1	G_{M2} ganglioside, asialoG_{M2}
Tay-Sachs disease, $A^{M}B$ variant	G_{M2} activator (for α chain)	chrom. 5	G_{M2} ganglioside
G_{M2} gangliosidosis (Sandhoff disease)	hexosaminidase A and B (β chain)	β chain: 5q13	G_{M2} ganglioside, globoside
G_{M1} gangliosidosis (Landing disease)	β-galactosidase	3p21-3pter	G_{M1} ganglioside, galactosyloligosaccharides, keratansulfate
Krabbe disease	β-galactocerebrosidase	14q21-q31	galactosylsphingosine
Gaucher disease	β-glucocerebrosidase	1q21	glucosylceramide
Gaucher disease variant	saposin C activator	10q21	glucosylceramide
Atypical Gaucher disease	saposin precursor	10q21	ceramide, glucosylceramide
Niemann-Pick disease type A and B	sphingomyelinase	11p15	sphingomyelin, cholesterol
Farber disease	ceramidase	?	ceramide
Wolman disease, cholesteryl ester storage disease	acid lipase	10q	cholesteryl esters, triglycerides
Metachromatic leukodystrophy	arylsulfatase A	22q13	sulfatides
Metachromatic leukodystrophy variant	saposin B	10q21	sulfatides, globotrialosylceramide, digalactosylceramide
Fabry disease	α-galactosidase A	Xq22	globosylceramide, galabiosylceramide, blood group B trisaccharide
Schindler disease	α-N-acetylgalactosaminidase	22q13-qter	glycoproteins, glycosaminoglycans, glycosphingolipids
Mucopolysaccharidoses (MPS)			
Hurler/Scheie disease (MPS I)	α-L-iduronidase	22pter-q11	heparan-, dermatan-sulfate
Hunter disease (MPS II)	iduronate-sulfate sulfatase	Xq26-q28	heparan-, dermatan-sulfate
Sanfilippo disease type A (MPS III A)	heparansulfamidase	?	heparansulfate
Sanfilippo disease type B (MPS III B)	N-acetyl-α-D-glucosaminidase	?	heparansulfate
Sanfilippo disease type C (MPS III C)	acetylCoA:α-glucosaminide acetyltransferase	?	heparansulfate
Sanfilippo disease type D (MPS III D)	N-acetylglucosamine-6-sulfate sulfatase	12q14	heparansulfate

Morquio disease type A (MPS IV A)	N-acetylgalactosamine-6-sulfate sulfatase	16q24.3	keratansulfate
Morquio disease type B (MPS IV B)	β-galactosidase	3p21-3pter	keratansulfate
Maroteaux-Lamy disease (MPS VI)	arylsulfatase B	5p11-q13	dermatansulfate
Sly disease (MPS VII)	β-glucuronidase	7q11.2-q22	dermatan-, heparan-sulfate
Glycogenoses			
Pompe disease (Glycogenosis type II)	α-glucosidase	17q23	glycogen
Glycoproteinoses			
Fucosidosis	α-fucosidase	1p34	fucosyl-oligosaccharides, fucosyl-glycolipid
Mannosidosis	α-mannosidase	19p13-q13	mannosyl-oligosaccharides
Aspartylglucosaminuria	aspartylglucosaminidase	4q	glycoasparagines
Sialidosis (Mucolipidosis I)	neuraminidase	10pter-q23?	
Galactosialidosis	protective protein/carboxypeptidase, secondary β-galactosidase and neuraminidase deficiency	20q	sialyloligosaccharides, gangliosides
I Cell disease (Mucolipidosis II)	phosphotransferase in the Golgi complex (see text)	4q21-q23	multiple oligosaccharides, mucopolysaccharides, and lipids
Transport disorders			
Cystinosis	cystine transport system	?	cystine
Sialic acid storage disease (Salla disease)	acid sugars transport system	6q14-q15	sialic acid, glucuronic acid
Vitamin B_{12} transport defect (Cbl F mutation)	vitamin B_{12} transport system	?	cobalamin
Lysosomal diseases with unknown defects			
Niemann-Pick disease type C	deficient cholesterol recycling	18	cholesterol
Mucolipidosis IV	abnormal ganglioside recycling?	?	gangliosides
Batten disease (Neuronal ceroidlipofuscinosis)	?	16q22	subunit C of mitochondrial ATP synthase
Mucosulfatidosis	9 different aryl- and steroid-sulfatases	?	sulfatides, mucopolysaccharides, sulfate-containing glycolipids

Table 3. Metabolites for Which the Mammalian Lysosomal Membrane Contains Specific Transport Systems

Amino acids
neutral small amino acids
proline and 3,4-deoxyproline
neutral bulk amino acids
branched and aromatic amino acids
anionic amino acids
cationic amino acids
cystine*
cysteine
small dipeptides
taurine
Sugars
neutral monosaccharides
N-acetylated hexoses
acid monosaccharides*
Nucleosides
purine nucleosides (adenosine, inosine)
pyrimidine nucleosides (cytidine, thymidine, uridine)
Ions
sulfate ions
phosphate ions
chloride ions
calcium ions
protons
folyl-polyglutamates
Vitamins
cobalamin (vitamin B_{12})*

Note: *Genetic deficiencies of these transport systems cause human lysosomal diseases.

coated pits and vesicles. Once the enzyme has reached the lysosome, further trimming (maturation) of the glycosidic chain takes place.

The biosynthetic steps for lysosomal enzymes are not the same as those associated with the lysosomal membrane, such as glucocerebrosidase (which is defective in Gaucher disease), acid phosphatase, and possibly the membrane transporters. In the case of glucocerebrosidase and acid phosphatase, it is known that they do not acquire the mannose-6-phosphate marker and are delivered to the lysosome via another unclear mechanism.

Until a decade ago, the general assumption was that products of enzymatic degradation would freely diffuse across the lysosomal membrane, and be recycled by the cell without a specific control mechanism. Thanks to improved technology in the investigation of membrane proteins, more than twenty lysosomal transport systems have thus far been found to control the traffic of various compounds,

including monosaccharides, amino acids, organic and inorganic ions, vitamins, and nucleosides (Table 3). But next to nothing is known about the synthesis of membrane transporters.

GENETICS

Although lysosomal storage diseases represent only 12% of the inborn errors of metabolism, and only 1.2% of all Mendelian disorders, the fact that most of them involve a mutation at the single gene level makes them scientifically very relevant.

Lysosomal diseases are individually very rare; the incidence of a single defect is about 1:100,000 to 1:25,000 in the general population. Some diseases, however, are much more frequent in certain ethnic groups, e.g., Ashkenazi Jews and French-Canadians (Gaucher disease, Tay-Sachs disease) or Finns (aspartylglucosaminuria, Salla disease) as a result of cultural and geographical or religious isolation. In these groups, the incidence reaches 1:1,000 or higher, with parallel high frequency of healthy carriers. The prevalence of one or a few major mutations in these communities (the so-called "ancestor effect") is evidence that isolation and inbreeding have perpetuated an original mutation.

The majority of lysosomal storage diseases are monogenetic disorders inherited according to the laws of Mendel for autosomal recessive traits. This means that the disease becomes manifest when both alleles bear a mutated gene. Two healthy heterozygotes have a 25% chance of getting an affected child. Only two disorders, namely Fabry disease and Hunter disease, are linked to mutations of genes in the sex chromosome X. Full symptomatology is therefore observed in affected males (one X chromosome). A healthy father, with a normal X chromosome, will not transmit the disease. A carrier mother, with one normal and one mutated X chromosome, has a 50% chance to give birth to an affected son, and a 50% chance to give birth to an heterozygote daughter. In some instances, because of the random and irreversible inactivation of the X chromosome in females (lyonization), female heterozygotes can show clinical symptoms of an X-linked disorder with variable severity. For unclear reasons this often occurs in Fabry disease, but only sporadically in Hunter disease. In lysosomal storage diseases, there is usually a correlation between the severity of the disease and the level of residual activity of the mutated protein. This also holds for female heterozygotes of Fabry disease.

The effort of several scientists and clinicians in the last decade has led to the sequencing of the genes encoding for most of the lysosomal enzymes, and thereby made it possible to trace the mutations underlying the different enzyme defects. Wide variability of symptoms and signs is a hallmark of lysosomal disorders. The age of onset, severity, and type of organ involvement can also vary, even though the patients have the same enzyme deficiency. This peculiar clinical picture has led to the suspicion of wide genetic heterogeneity long before the advent of gene cloning. After cloning of the genes, it has indeed become clear that several genetic mutations can lead to the same disease (meant as the same enzyme deficiency). In

fact, the far most common situation for a lysosomal disease is that a patient bears two different mutant alleles; thus, he or she is in reality a compound heterozygote. Exceptions are patients coming from highly consanguineous families or isolated communities, who can be homozygote for the same mutation.

The study of mutations at the gene level has led to a reliable correlation between the type of mutation(s) and the clinical picture. An example is provided by metachromatic leukodystrophy (MLD), a defect in the degradation of cerebroside sulfate caused by a deficiency in the enzyme arylsulfatase A. Two main mutant arylsulfatase A alleles, I and A, have been discovered in most of the patients with MLD. When both alleles are associated in the same individual (genotype A/I), the juvenile form of the disease is encountered. In general, the presence of allele I in combination with another mutant allele X (genotype I/X) determines a late infantile disease, and the same holds in the case of homozygosity for the I gene (genotype I/I). Heterozygosity for the allele A with another mutant gene (A/X genotype) and A/A homozygosity are both associated with a juvenile onset of the disease. A milder phenotype (later onset and less disabling symptoms) correlated well with the synthesis of an arylsulfatase enzyme protein with higher residual activity, while in severe clinical forms the residual activity is at the lower limit of detection. Incidentally, MLD can also be caused by a deficiency of the activator protein for the cerebroside sulfate degradation, saposin B. Thus, the genetic mutation in the latter case affects a genetic locus different from that of arylsulfatase A, showing how for the degradation of cerebroside sulfate at least two different genes are needed.

Extensive mutation analyses have shown a similar correlation in Gaucher disease, in Pompe disease, and in Tay-Sachs disease. This, however, has not been always possible. For example, homozygosity for a proline substitution in the place of leucine at position 444 (^{444}Leu→Pro) of the glucocerebrosidase gene is associated with the non-neuronopathic, as well as the neuronopathic form of Gaucher disease and homozygosity for arylsulfatase; A ^{426}Pro→Leu mutation has been found in juvenile and adult forms of MLD. These discrepancies emphasize the role of other genetic loci (or epigenetic), and the possible effect of environmental factors on the evolution of a monogenetic disorder. Apparently, caution must be exercised in predicting the clinical outcome from the genotype.

MOLECULAR PATHOGENESIS

Lysosomal storage disease represents the clinical expression of two types of defects in lysosomal function. The first is a lysosomal enzyme deficiency, resulting in the accumulation of undigested macromolecules, while the second is a defect in the transmembrane transport of metabolites that results in the intralysosomal accumulation of products of enzymatic digestion in the presence of a normal set of enzymes (see Figure 1).

A catalytic defect in the degradation of a substrate can be the result of many different situations. A defect in almost any step in enzyme biosynthesis, processing and maturation can generate a lysosomal storage disease. A mutation in the gene encoding for a lysosomal enzyme can cause abnormal transcription with total lack of enzyme synthesis (so-called CRM–, Cross Reacting Material-negative patients where antibodies specific to lysosomal enzymes fail to react with any synthesized protein). Alternatively, an abnormal polypeptide is synthesized which will undergo faulty biosynthetic steps, leading to early degradation, wrong compartmentalization, altered subunit association, and impaired catalytic function. When an immunoreactive enzyme is present, one speaks of CRM+ patients. Such patients have been described as having variants of gangliosidoses, e.g., the Tay-Sachs variant of G_{M2} gangliosidosis, G_{M1} gangliosidosis, glycogenosis type II, and the mucopolysaccharidoses.

A defective function of a lysosomal hydrolase, however, can also be secondary to a defect in another protein involved in the complex biosynthetic pathway of the enzyme. In the case of mucolipidosis II, also called I-cell (inclusion cell) disease, a phosphotransferase in the Golgi system fails to attach the mannose-6-phosphate recognition marker to lysosomal enzyme precursors, thereby preventing lysosomal delivery of catalytically normal enzymes. These are therefore secreted by the cell and can be measured in high quantities in body fluids, but not in cells.

Another mutation in a nonenzymic protein causing a secondary enzyme deficiency is the lack of an activator protein for the degradation of sphingolipids. Activator proteins have a detergentlike function and facilitate catalysis between a hydrophobic substrate and a hydrophilic enzyme. So far, five different activators have been identified: one activates degradation of ganglioside G_{M2} by hexosaminidase A, and the other four are called saposins A, B, C, and D. Saposins are all derived from processing of the same precursor, encoded by the same gene, and cooperate with different enzymes. Deficiency of an activator protein causes a disease similar to that observed in the corresponding primary enzyme deficiency. Known examples include metachromatic leukodystrophy caused by saposin B deficiency, the AB variant of G_{M2} gangliosidosis, a variant of Gaucher disease caused by saposin C deficiency, and a recently discovered syndrome with deficiency of all the saposins caused by a defect of the saposin precursor gene (Schnabel et al., 1992).

Lysosomal enzymes are present in the lysosome in excess of the substrates to be degraded. This is demonstrated by the fact that heterozygotes, with only one mutant allele, often have a reduced enzyme activity (about 50%) but are asymptomatic. In general, the severity of the symptoms correlates in all lysosomal storage diseases with the residual enzyme activity found in the cells toward the natural substrate of the enzyme, independently of the type of genetic mutation. The onset of lysosomal storage seems inversely correlated with the residual functional capacity of the enzyme. According to a theoretical model proposed by Sandhoff and

coworkers (Leinkugel et al., 1992), the substrate influx rate into the lysosome is not relevant for an efficient degradation when normal (100%) enzyme activity is present. But when the activity is severely decreased (between 5 and 10% of normal), small variations in substrate influx can result in an intracellular substrate concentration that exceeds the critical threshold, whereby the substrate degradation rate is not proportional to substrate concentration any more and accumulation starts. This model has recently been confirmed in studies with cultured patient cells fed with natural substrates.

Although only three inborn errors of metabolism are classified as lysosomal transport defects among circa three dozens of lysosomal disorders, they already constitute a new class of inherited defects of lysosomal function. The fact that only three of them are known is an indication of our limited knowledge of the lysosomal membrane. Information at the molecular level about this class of inherited defects is not yet available.

CLINICAL DIAGNOSIS

After having introduced the reader to the various causes of a defect in lysosomal function at the molecular level, it will be easier to understand why and how lysosomal storage disorders (LSD) are clinically very diverse. Although it seems difficult, considering the different causes, to suggest general criteria for the clinical diagnosis of LSD, it is worthwhile to mention those laid down by Hers (1965). These criteria are:

1. Progression of symptoms and signs of the disease in spite of therapeutic efforts, with gradual failure of organ function. The clinical state is never stable, even after apparently clinically silent periods.
2. Regression of acquired skills and/or delay of normal somatic and psychomotor development.
3. Progressive signs of storage (enlargement and functional failure) in visceral organs, among which the most often and severely affected are brain, liver, spleen, heart, kidney, eye, skin, bone, and cartilage. Sometimes, this leads to remarkable somatic changes which justify use of the term gargoylism, once used for some mucopolysaccharidoses.
4. Microscopical evidence of lysosomal dysfunction, e.g., an increase in number and size of lysosomes in different tissue biopsies. The morphology of the storage product in certain cell types can provide a clue to the etiology.

The same disease, as defined by a common protein (enzyme) deficiency, can vary according to: 1) the time of onset of symptoms, which usually correlates with the severity of the disease; and 2) type of accumulated substrate, and consequently, the type of organ involved. The cause of this variability lies in the type of primary genetic mutation, and in other individual constitutional factors, which are much more difficult to define, as well as in nongenetic factors such as lifestyle, nutrition,

Table 4. Lysosomal Disorders Which Can Present at Birth with "Fetal Hydrops"

Morquio disease type B (MPS* IV B)
Sly disease (MPS VII)
Sialidosis
Galactosialidosis
Gaucher disease
Niemann-Pick disease type C
I Cell disease
GM_1 gangliosidosis
Sialic acid storage disease

Note: *MPS: Mucopolysaccharidosis.

medications, and psychosocial environment. All these factors contribute to the maintenance of an equilibrium between the influx of substrate in the lysosome and the degradative capacity of the cell.

For almost each lysosomal enzyme deficiency, the age at which symptoms occur can vary. In general, the earlier the symptoms occur, the more severe is the clinical picture. Not infrequently, the neonatal disease presents the dramatic picture called "fetal hydrops" (see Table 4). Diseases of late onset are usually milder and often show only neurological signs. This is because the nervous system has the lowest regenerative capacity which makes it susceptible to permanent damage. In the case of G_{M1} gangliosidosis caused by a deficiency of β-galactosidase, three forms are known: infantile, juvenile and adult. While the infantile form can be evident at or immediately after birth, the juvenile and adult forms begin with regression of acquired neuromotor skills and mental performances after a period of normal development, and show slow progression. Infants with G_{M1} gangliosidosis show not only psychomotor delay but also hepatosplenomegaly, skeletal dysplasia, recurrent respiratory infections, cardiomyopathy, and visual impairment. Adults with G_{M1} gangliosidosis can show the first symptoms at the age of 15 or 20 years; this is characterized by an unsteady gait and uncoordinated movements, poor psychic performance, and dementia. However, hepatomegaly or bone deformities are rare. The mutations in the β-galactosidase gene underlying the different phenotypes result in a different residual enzyme activity (about 10% in the adult form and less than 1% in the infantile form).

For further description of LSD the reader is referred to two exhaustive treatises (Emery et al., 1990; Scriver et al., 1995).

An important factor which determines the clinical manifestations of a lysosomal disorder is the type and localization of stored material. Consider for example two different sphingolipidoses, such as Krabbe disease and Fabry disease. These are caused respectively by a deficiency of β-galactocerebrosidase and α-galactosidase A. In both cases, an enzyme cleaving galactose moiety is defective, but each is specific for a certain type of galactoside. β-Galactocerebroside is an essential

component of myelin sheaths. Instead, α-galactose-containing globo- and galabiocerebrosides are abundant in the plasma membranes of epithelial cells, smooth muscle, endothelial cells and erythrocytes. α-Galactosidase A recognizes also the blood group B circulating antigen. Consequently, Krabbe disease only involves the central and peripheral nervous system, while Fabry disease involves the vascular system, as well as the kidney, heart, and autonomic ganglia. All show progressive deterioration. In addition, the skin of patients with Fabry disease presents a diagnostic change called angiokeratoma corporis diffusum produced by subcutaneous telangiectasias and angiomas, combined with hyperkeratosis of the epidermis. A similar skin change is observed in other lysosomal glycoproteinoses.

The role of substrate in the pathogenesis of LSD is further demonstrated by the occurrence of a syndrome which can be caused by a deficiency of two distinct enzymes that act on the same natural substrate. Morquio disease type A is caused by a defect of α-N-acetyl-galactosamine-6-sulfate sulfatase, while Morquio disease type B arises from a β-galactosidase deficiency. Both enzymes are involved in keratansulfate catabolism, the disruption of which produces damage of the glycosaminoglycan-rich cornea, skeletal, and cardiovascular systems. The result is a clinically indistinguishable syndrome.

Another factor contributing to clinical heterogeneity of LSD is the type of protein (enzyme) defect. Some genetic mutations in fact can alter the catalytic activity of an enzyme towards one of its substrates only. This is the case for mutations in the β-galactosidase locus which cause Morquio disease type B (loss of activity toward keratansulfate) or G_{M1} gangliosidosis (loss of activity toward gangliosides, glycoproteins, and keratansulfate). Structural information about the β-galactosidase gene has identified a mutation in Morquio disease type B, leading to a ^{273}Trp→Leu substitution in the enzyme polypeptide. Several other mutations of the β-galactosidase gene lead to different forms of G_{M1} gangliosidosis.

Among the organs affected in LSDs, the eye deserves special mention, because a thorough ophthalmological examination can sometimes reveal specific lesions which suggest the etiology of the disorder. For example, cherry red spots in the fundus are observed in many gangliosidoses and glycoproteinoses; corneal opacities are found in some mucopolysaccharidoses, in Fabry disease, I cell disease, and mannosidosis. In some instances, ocular findings contribute to the differential diagnosis. This is the case in Hurler and Hunter disease, two mucopolysaccharidoses which can have similar clinical presentations. However, corneal clouding is the rule in the former, but rare in the latter (Table 5).

Among lysosomal disorders, the clinical diagnosis of a vitamin B_{12} transport defect merits particular consideration. While sialic acid and cystine, which are involved in other transport defects, derive from the turnover of structural substrates, vitamin B_{12} enters the cell as an essential dietary factor. The transport defect blocks the intracellular delivery of cobalamin which is endocytosed as a complex with transcobalamin II and should exit the lysosome to be used as a cofactor of several

Table 5. Ocular Findings in Lysosomal Storage Diseases

Corneal/Lens Opacities	*Macular "Cherry Red Spot"*	*Optic Nerve Atrophy*	*Strabismus/Nystagmus*	*Retinal Degeneration*
Hurler disease (MPS I)	GM_1 agangliosidosis	Metachromatic leukodystrophy	Sialic acid storage disease	Hurler disease (MPS I)
Fabry disease	Tay-Sachs disease	Krabbe disease	Tay-Sachs disease	Hunter disease (MPS II)
Morquio disease (MPS IV)	Sandhoff disease	Tay-Sachs disease	GM_1 gangliosidosis	Sanfilippo disease (MPS III)
I Cell disease	Sialidosis	G_{M1} gangliosidosis	Gaucher disease	Fucosidosis
Maroteaux-Lamy disease (MPS VI)	Fucosidosis			Neuronal ceroidlipofuscinosis
Mannosidosis	Niemann-Pick disease type A			
Aspartylglucosaminuria	Farber disease type V			
Mucolipidosis IV				

reactions. The clinical picture is not that of a lysosomal storage disease (no progressive signs of storage in viscera) but is characterized by the symptoms of cobalamin deprivation and acute metabolic decompensation. In genetic terms, this disease represents a cobalamin F mutation. The occurrence of this human disease emphasizes the role of lysosomes in the intracellular distribution of essential active factors which are delivered to the cell via endocytosis.

PATHOLOGY

Among the clinical aspects of LSD, pathological findings deserve special consideration. Microscopical evidence of lysosomal storage is a major criterion for the diagnosis of LSD, as laid down by Hers (1965). When the possibility of LSD is raised as the result of clinical examination and history taking, confirmation by microscopical examination of tissue biopsy specimens is usually required. Accurate observation of lymphocytes in a simple routine peripheral blood smear can reveal the presence of abnormal cytoplasmatic vacuoles, suggestive of lysosomal dysfunction. In other cases, a skin or rectal mucosal biopsy clinches the diagnosis, even in situations where the skin and mucosae show no overt signs. This avoids the need for more invasive procedures in less accessible systems such as the nervous system. Electron microscopy of biopsy specimens shows specific morphologic changes in the storage material in vacuoles from fibrocytes, nerve endings, Schwann cells, and endothelial cells (Figure 2). Rectal biopsies are particularly useful for providing a picture of peripheral nerve ganglia and macrophages in the lamina propria. In some cases the microscopical finding is pathognomonic, and so specific as to permit the diagnosis of the type of disorder. This is true of Pompe disease (membrane-bound glycogen aggregates), cystinosis (cystine crystals surrounded by a membrane), Krabbe disease (globoid cells in the brain white matter), Gaucher disease ("crumpled silk" hystiocytes), or metachromatic leukodystrophy (metachromatic granules in nerve endings). In other cases findings can be suggestive, but not diagnostic, such as osmiophilic bodies with onionlike lamellae seen in sphingolipidoses, or striped "zebra bodies" in neuronal lysosomes of many gangliosidoses, or empty vacuoles in reticuloendothelial cells seen in sialidosis, mannosidosis, and mucopolysaccharidoses. Electron microscopical observation also allows the diagnosis of certain lysosomal disorders in which the biochemical and genetic defect is unknown. In the case of neuronal ceroid lipofuscinosis, fingerprint and curvilinear bodies in skin and sural nerve biopsies are pathognomonic. Lamellar bodies observed in biopsies from mucolipidosis IV patients are pathognomonic as well. Fluorescent lysosomal staining with filipin, a compound with a high affinity for cholesterol, allows the diagnosis of Niemann-Pick disease type C, the primary defect of which is unknown.

In some exceptional situations, light microscopic observations fail to detect lysosomal abnormalities in the affected cells. For example, in Krabbe disease (also known as globoid cell leukodystrophy), the oligodendroglial cells which synthesize

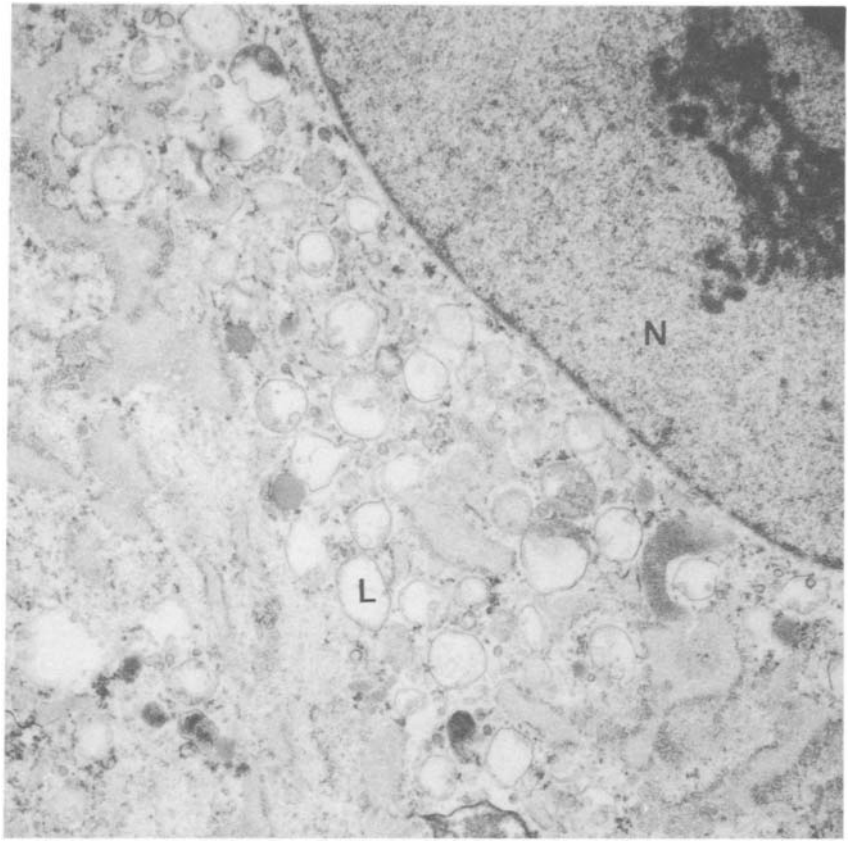


Figure 2. Electronmicroscopic micrograph of a skin fibroblast from a patient with sialic acid storage disease. The cytoplasm is filled with numerous swollen vacuoles containing electron-transparent or fine granular material. L: lysosomes, N: nucleus.

myelin are greatly decreased in number, but are morphologically normal. The abnormalities instead are restricted to a decrease in myelin and the presence of globoid cells, which are an active macrophage-like cells. It seems that myelin forming cells are unable to degrade myelin, hence are intoxicated by a product of alternative galactocerebroside degradation known as psychosine. Affected cells are thus incapable of storing because they undergo necrosis at an early stage.

BIOCHEMICAL DIAGNOSIS AND PREVENTION

The diagnosis of lysosomal storage disorders is completed by biochemical investigations which aim at the chemical characterization of storage products and the

definition of the primary protein (enzyme) defect. Biological fluids including urine, plasma, cerebrospinal and amniotic fluids are usually investigated because they can contain storage material either derived from cellular death or accumulated in the fluids as a physiological secretory component. In urine, for example, the finding of mucopolysaccharide fragments corroborates the diagnosis of mucopolysaccharidosis, while in the case of glycoproteinoses and some sort of gangliosidoses, specific types of oligosaccharide chains can be detected and distinguished by simple thin layer chromatography. In the case of Fabry disease, the undegraded sphingolipids are normal circulatory components, and can be measured in increased amounts in the plasma of patients. Plasma is also used to measure increased lysosomal enzymes in I cell disease. For routine diagnosis of transport defects such as cystinosis and sialic acid storage disease, an elevated content of cystine and sialic acid in cells and urine are of diagnostic value in the presence of typical symptomatology and pathology. There is no need for the assay of transport function.

For most lysosomal storage diseases, the definitive diagnosis is based on the finding of the specific defect of a lysosomal protein. This is usually accomplished by the determination of enzyme activity and only exceptionally by DNA analysis. Lysosomal hydrolases from peripheral blood leukocytes or fibroblasts, cultured from skin biopsies, are usually assayed by measuring degradative capacity toward artificial fluorogenic substrates. Fibroblasts are usually needed when the enzyme activity in leukocytes is not easily measurable with the available assay. Cultured cells are also needed when the metabolism of natural substrates must be studied in living cells, particularly in those cases where no enzyme assay is available or when an activator deficiency is suspected. For example, a patient with a Tay-Sachs variant of G_{M2} gangliosidosis can show a normal β-hexosaminidase activity against the fluorogenic substrate 4-methylumbellyferyl-β-hexosaminide. This assay determines the catalytic activity of the α-and β-chains of hexosaminidase but it does not test the function of the activator protein needed to digest the natural G_{M2} ganglioside substrate *in vivo*. In this case, the proper diagnosis is reached either by studying the catabolism of G_{M2} in living cells (containing their own set of lysosomal enzymes and activators) or by identifying immunologically the presence of the activator protein with specific antibodies. Besides activator deficiencies, fibroblasts are still needed for the enzymatic diagnosis of sialidosis, Niemann-Pick disease type C, and Sanfilippo disease type A.

Lysosomal enzymes are sometimes assayed in healthy individuals. Pseudodeficiencies is the name given to decreased enzyme activities (about 10-20% of normal) in clinically normal individuals. Pseudodeficiencies have been described for β-hexosaminidase, galactocerebrosidase, and arylsulfatase A in the absence of clinical symptoms of G_{M2} gangliosidosis, Krabbe disease, or metachromatic leukodystrophy, respectively. These cases are difficult to differentiate from healthy heterozygotes (for a pathogenetic mutation) simply on the basis of an enzyme assay, especially in cases where the pseudodeficiency allele frequently occurs in the

general population. An example is arylsulfatase A pseudodeficiency which occurs in 15% of normal individuals. DNA analysis has shown that pseudodeficient subjects have a non-pathogenetic mutation which only determines some decrease in enzyme activity but maintains normal functional capacity of the enzyme *in vivo*. DNA analysis, therefore, can be used to establish the correct diagnosis.

Assay of enzyme activities can also be required for identification of healthy heterozygotes, parents of known patients, or close relatives of patients who ask for genetic counseling fearing reproductive risk. In X-linked disorders, e.g., Hunter disease and Fabry disease, aunts or sisters in the reproductive age (of diagnosed patients) run the risk of having affected sons. In these women, enzymatic analysis of cells derived from the same clone (e.g., the hair root) can bypass pitfalls caused by random inactivation of the X-chromosome. Nonetheless, DNA analysis is the most reliable technique for heterozygote identification.

Successful heterozygote screening programs have been performed for autosomal recessive diseases, especially in certain ethnic groups. In Ashkenazi Jews and French-Canadians, the incidence of Tay-Sachs disease and Gaucher disease is dramatically reduced by more than 90% as a result of efficient screening programs which have involved more than one million people. In particular, after the identification of the major genetic mutations in the hexosaminidase A locus, DNA analysis of enzyme-defined carriers has been proposed to exclude pseudodeficiencies and also to provide accurate genetic counseling in the general population.

Prenatal analysis has proven to be a major consequence of the progress achieved in the biochemical characterization of lysosomal storage diseases. If an affected fetus is found positive, parents can properly prepare themselves for the assistance of the handicapped child or can request termination of the pregnancy. Prenatal diagnosis is indicated in families where one or more severely affected patients are diagnosed or after heterozygote screening in populations at high risk. Nowadays, chorionic biopsy during the first trimester, or placental biopsy and amniocentesis in the second trimester of pregnancy, provide sufficient chorionic tissue or amniotic fluid and cells for laboratory diagnosis. An essential prerequisite for prenatal diagnosis is the unequivocal demonstration of the biochemical defect in the previously diagnosed patient (so-called index patient). It is therefore important to have material of the index patient at one's disposal. The method of analysis is the determination of enzyme activity and/or storage products, and only rarely is identification of the gene mutation at the DNA level required. Prenatal diagnosis based on electron microscopical analysis of fetal tissue is an option for those cases where the protein defect is not definitively established; for example, ceroidlipofuscinoses and mucolipidosis IV.

THERAPY

Therapeutic trials of lysosomal storage diseases have been very frustrating. Pharmacological treatment is nowadays possible only for cystinosis. In cystinosis

patients, administration of cysteamine leads to depletion of lysosomal cystine via a mechanism of disulfide exchange. The mixed cysteine-cysteamine disulfide can leave the lysosome via another surviving transporter for cationic amino acids.

Replacement therapy for lysosomal enzyme deficiencies was encouraged by observations made shortly after the biochemical defect in this group of genetic disorders had been established. When fibroblasts of patients with two different lysosomal diseases were cocultivated, correction of the metabolic defect was achieved by exchange of normal enzymes (Fratantoni et al., 1968). These studies justified the first enzyme supplementation trials before establishing the nature of the uptake marker. Now we know that the mannose-6-phosphate marker allows uptake and lysosomal delivery of exogenous enzymes via receptor-mediated endocytosis. Administration of enzymes from various sources (human urine, plasma, leukocytes, platelets, placenta, bone marrow, and amniotic membranes) has been tried. "Bottle necks" of most trials have been: 1) the amount of material required for life-long supplementation; 2) the difficulty of proper targeting of the enzyme to the affected organs, and 3) difficult access to the brain because of the blood-brain barrier. The most successful results have been achieved where proper targeting has been exploited and in situations where nervous system involvement was not severe.

Bone marrow transplantation has been intensively pursued for the treatment of mucopolysaccharidoses and some sphingolipidoses, e.g., Gaucher disease. This technique has the theoretical advantage of producing a population of stem cells which permanently produce the lacking protein. Such cells colonize all compartments of the reticuloendothelial system (including brain microglia). Allogenic transplants have been used in about 150 patients, and autologous transplantation of genetically manipulated cells is forthcoming. The replaced enzyme seems to be expressed in the tissues of the patient with improvement of both biochemical and clinical parameters. However, several important problems, including discrepancies about clinical eligibility requirements, graft-versus-host disease, and immunosuppression, have slowed down the widespread use of this technique. A successful approach has been the infusion of an exogenous enzyme with a chemically modified uptake marker which targets it to the affected organ. Human placental glucerebrosidase has been enzymatically trimmed to expose terminal mannose moieties present in its oligosaccharide chains. Such a modified enzyme, commercially available with the name of Ceradase^{R}, is efficiently taken up by the liver in Gaucher patients via the liver mannose receptor. A limitation to this therapy is the forbidding cost, while the clinical results are satisfying in patients with nonneuronopathic Gaucher disease (Barton et al., 1991).

In this era of gene cloning, gene therapy promises to revolutionize traditional approaches to enzyme replacement. Animal models for lysosomal storage diseases are a useful tool in the development of gene therapy protocols. Two recent studies seem very promising for future trials in humans. A natural mutant mouse with β-glucuronidase deficiency lends itself for gene therapy of Sly disease (MPS VII).

When bone marrow cells from a MPS VII mouse are infected with a retrovirus containing the normal β-glucuronidase gene and later are implanted intravenously into irradiated recipient MPS VII mice, long term expression of low levels of β-glucuronidase occurs with partial correction of the lysosomal storage. Similar infection of skin fibroblasts from MPS VII mice, followed by intraperitoneal implantation has provided results analogous to those obtained with bone marrow cells, but without the disadvantages of irradiation and chemotherapy (Sly, 1993). Lastly, an exciting recent development in the therapy of lysosomal disorders involving the nervous system is that neuronal targeting of β-hexosaminidase A is achieved by coupling the enzyme via disulfide linkages to the atoxic fragment C of tetanus toxin (TTC). The latter is bound avidly by neuronal membranes. TTC-hexosaminidase A conjugation resulted in neuronal surface binding and enhanced endocytosis of enzyme, as observed in rat brain cultures. In cerebral cortex cultures from a feline model of G_{M2} gangliosidosis, both binding and uptake were similar to those found in rat brain cultures; that is, elimination of G_{M2} storage (Dobrenis et al., 1992).

SUMMARY

Lysosomal storage diseases are a group of about thirty inherited disorders of lysosomal biogenesis and function. Most of them are caused by primary or secondary enzyme deficiency with consequent accumulation of undigested substrate. A few are the result of defective transport across the lysosomal membrane with consequent accumulation of degradation products. Since the definition of the defect in Pompe disease in 1963, progress has been made in the diagnosis and prevention of lysosomal disorders. Easy biochemical diagnosis is available in specialized centers all over the world, with the possibility of secondary prevention by heterozygote identification and prenatal diagnosis in families at risk. Structural analysis of the gene encoding for the defective protein has been achieved for almost every lysosomal enzyme deficiency. In some cases, it is now possible to predict the type of clinical manifestations based on analysis of the type of genetic mutation. Therapeutic attempts have been made but promising results are confined to a small number of patients. Research in the field of gene therapy provides a more hopeful prospect.

ACKNOWLEDGMENTS

We thank Mr. Pim Visser for providing Figure 2.

REFERENCES

Barton, N. W., Brady, R. O., Dambrosia, R. M., Di Bisceglie, A. M., Doppelt, S. H., Hill, S. C., Mankin, H. J., Murray, G. J., Parker, R. I., Argoff, C. E., Grewal, R. P., Yu, K. T. & Collaborators. (1991).

Replacement therapy for inherited enzyme deficiency—Macrophage-targeted glucocerebrosidase for Gaucher's disease. New Eng. J. Med. 324, 1464–1470.

Brady, R. O., Kanfer, J. N., Mock, M. B., & Fredrickson, D. S. (1966). The metabolism of sphingomyelin II. Evidence of an enzymatic deficiency in Niemann-Pick disease. Proc. Natl. Acad. Sci. USA 55, 367.

Dobrenis, K., Joseph, A., & Rattazzi, M. C. (1992). Neuronal lysosomal enzyme replacement using fragment C of tetanus toxin. Proc. Natl. Acad. Sci. USA 89, 2297–2301.

Fratantoni, J. C., Hall, C. W., & Neufeld, E. F. (1968). Hurler and Hunter syndromes: mutual correction of the defect in cultured fibroblasts. Science 162, 570–572.

Hers, H. G. (1963). α-glucosidase deficiency in generalized glycogen storage disease (Pompe disease). Biochem. J. 86, 11.

Hers, H. G. (1965). Inborn lysosomal diseases. Gastroenterology 48, 625–633.

Leinkugel, P., Michel, S., Conzelmann, E., & Sandhoff, K. (1992). Quantitative correlation between the residual activity of β-hexosaminidase A and arylsulfatase A and the severity of the resulting lysosomal storage disease. Human Genetics 88, 513–523.

Patrick, A. D. (1965). A deficiency of glucocerebrosidase in Gaucher disease. Biochem. J. 97, 17 C.

Schnabel, D., Schröder, M., Fürst, W., Klein, A., Hurwitz, R., Zenk, T., Weber, J., Harzer, K., Paton, B.C., Poulos, A., Suzuki, K., & Sandhoff, K. (1992). Simultaneous deficiency of sphingolipid activator proteins 1 and 2 is caused by a mutation in the initiation codon of their common gene. J. Biol. Chem. 267, 3312–3315.

Sly, W. S. (1993). Gene therapy on the Sly. Nature Genetics 4, 105–106.

RECOMMENDED READINGS

Emery, A. E. H. & Rimoin, D. L. (eds.) (1990). Principles and Practice of Medical Genetics. 2nd edition. Sections Q101-103 and 110, Churchill Livingstone, Edinburgh.

Kleijer, W. J. (1990). Prenatal diagnosis. In: Inborn Metabolic Diseases. Diagnosis and Treatment (Fernandes J., Saudubray, J.-M., & Kada, K., eds.). Part XVII, pp. 683–695. Springer-Verlag, New York.

Neufeld, E. F. (1991). Lysosomal storage diseases. Ann. Rev. Biochem. 60, 257–280.

Scriver, C. R., Beaudet, A. L., Sly, W. S., & Valle, D. (eds.) (1995). The Metabolic and Molecular Bases of Inherited Disease. Part 12: Lysosomal enzymes; Part 16, Chap. 126: Lysosomal transport disorders: cystinosis and sialic acid storage disorders. McGraw-Hill Information Services Company, New York.

Wenger, D. A. & Williams, C. (1991). Screening for lysosomal disorders. In: Techniques in Diagnostic Human Biochemical Genetics: A Laboratory Manual (Hommes F., ed.). Chap.33. Wiley-Liss, Inc., New York.

Chapter 6

Peroxisomal Disorders

RONALD J. A. WANDERS and JOSEPH M. TAGER

Principles of Medical Biology, Volume 3
Cellular Organelles and the Extracellular Matrix, pages 155–172

ISBN: 1-55938-804-8

INTRODUCTION: THE ZELLWEGER SYNDROME

The peroxisomal disorders represent a newly recognized group of inborn errors of metabolism in man in which there is an impairment of one or more peroxisomal functions. The cerebro-hepato-renal (Zellweger) syndrome is generally considered to be the prototype of this group of disorders which now comprises at least 15 different diseases. The cerebro-hepato-renal (Zellweger) syndrome was first described in 1964 by Bowen, Lee, Zellweger, and Lindenberg. These investigators reported on two unrelated pairs of siblings who presented at birth with an apparently similar pattern of multiple malformations. Major abnormalities included severe hypotonia, midface hypoplasia, wide fontanels, open metopic and lambdoid sutures, high forehead, low-set ears, glaucoma with corneal opacities and cataracts, multiple joint contractures, enlarged liver, chondrodysplasia punctata, and multiple subcapsular renal cysts (Bowen et al., 1964). A few years later Passarge and McAdams (1967) described five similarly affected sisters and introduced the name cerebro-hepato-renal syndrome for this condition. In honor of Hans Zellweger's seminal role in identifying this syndrome the name cerebro-hepato-renal syndrome of Zellweger was proposed, which is now the generally accepted designation but is usually abbreviated to Zellweger syndrome.

A major discovery was made in 1973 by Sydney Goldfischer and his colleagues. In a historic paper, they reported that morphologically distinguishable peroxisomes are absent in liver and kidney from patients with the Zellweger syndrome (Goldfischer et al., 1973). The significance of this finding was not immediately appreciated since the functional importance of peroxisomes, at least in mammals (including man), was obscure at that time. In fact, much more attention was originally paid to the mitochondrial abnormalities that accompanied the absence of peroxisomes in some patients (see e.g., Goldfischer et al., 1973).

FUNCTIONS OF PEROXISOMES IN RELATION TO THE METABOLIC ABNORMALITIES IN THE ZELLWEGER SYNDROME

Peroxisomes were first identified by Rhodin (1954), a Swedish investigator, who coined the term microbody to describe these apparently structureless cellular inclusions present in the kidney.

Subsequent studies, notably by De Duve and coworkers (see De Duve and Baudhuin, 1966 for a review), showed that peroxisomes contain a group of oxidases that use molecular oxygen to generate hydrogen peroxide (H_2O_2). Furthermore, De Duve and coworkers found that peroxisomes also contain catalase, which is able to decompose H_2O_2 to molecular oxygen and water. Oxidases are found not only in peroxisomes but also elsewhere in the cell, such as in the cytosol (e.g., xanthine oxidase) and in mitochondria (e.g., monoamine oxidase and sulphite oxidase). Furthermore, catalase is not present exclusively in peroxisomes, some being found

in the cytosol. Thus peroxisomes cannot be considered the exclusive site of H_2O_2 metabolism. Furthermore, the relatively innocuous character of catalase deficiency in man raises the question of the functional significance of human catalase in H_2O_2 metabolism (Eaton, 1989).

Early biochemical studies in patients with the Zellweger syndrome had revealed only nonspecific abnormalities, including elevated serum transaminases, (conjugated) hyperbilirubinemia, hypothrombinemia and elevated serum iron levels (see Wanders et al., 1988a for a review). In 1975, however, Danks and coworkers reported that L-pipecolic acid, a metabolite derived from L-lysine, accumulates in the serum of patients with the Zellweger syndrome (Danks et al., 1975). In retrospect, this must be taken as the first peroxisomal abnormality detected in patients with the Zellweger syndrome. In the absence of any indication of a functionally significant role of peroxisomes in L-pipecolic acid metabolism, it is not perhaps surprising that the accumulation of L-pipecolic acid was thought to result from functionally deficient mitochondria. Subsequently, when Hanson et al., (1979) discovered that abnormal bile acids are present in serum and urine from Zellweger patients they, too, attributed this abnormality to mitochondrial dysfunction. Indeed, several authors, including Goldfischer et al., 1973 in their seminal paper, reported on the presence of ultrastructurally distorted and functionally deficient mitochondria in Zellweger patients. It is now generally agreed, however, that any mitochondrial abnormalities in Zellweger patients are secondary effects arising from the absence of peroxisomes. Although mitochondrial dysfunction may certainly contribute to the pathogenesis of the disease, it is the general experience that mitochondrial abnormalities are usually mild in the Zellweger syndrome, as may be gauged from the absence of significant lactic acidemia in these patients (Lazarow and Moser, 1989).

The final recognition that the Zellweger syndrome is, indeed, a peroxisomal disorder came in the early 1980s as a result of two key observations. First, Brown et al. (1982) showed that there are elevated levels of saturated very-long-chain fatty acids, notably hexacosanoic acid (C26:0) and lignoceric acid (C24:0), in the plasma of patients with the Zellweger syndrome. Second, Heymans et al. (1983, 1984) found that the levels of plasmalogens are greatly decreased in erythrocytes and tissues of the patients. These findings not only established the essential role of peroxisomes in human metabolism but also formed the basis for the identification of a number of inherited diseases in man as being due to peroxisomal dysfunction.

PEROXISOMAL DISORDERS: A CLASSIFICATION

We now recognize at least 15 different genetic diseases in man in which peroxisomal functions are impaired. These diseases are divided into three distinct groups depending on whether there is a generalized loss of peroxisomal functions (group A), a loss of multiple peroxisomal functions (group B), or a loss of a single

peroxisomal function (group C). Before discussing these diseases in some detail it is necessary to give a brief overview of the metabolic functions of peroxisomes.

PEROXISOMES IN MAN: FUNCTIONAL ASPECTS OF DIRECT RELEVANCE TO HUMAN DISEASE

In considering whether peroxisomes play an indispensable role in a metabolic pathway one can make use of one or more of the following criteria (see Tager et al., 1990 for a discussion):

1. The localization of one or more enzymes of the pathway is restricted to peroxisomes.
2. Balance studies involving the use of specific inhibitors and/or analysis of products show that peroxisomal enzymes participate in a metabolic pathway.
3. In patients with the Zellweger syndrome there is an accumulation of substrates or intermediates in a metabolic pathway, or a deficiency of the products of the pathway.

Studies carried out during the last 15 years in a large number of laboratories have shown that peroxisomes are highly versatile organelles involved in a great number of different metabolic pathways, both anabolic and catabolic. For a detailed discussion of these pathways, the reader is referred to the chapter by Masters and Crane.

β-Oxidation of Fatty Acids and Fatty Acid Derivatives

One of the most important functions of peroxisomes in man is the β-oxidative chain-shortening of fatty acids and fatty acid derivatives. Just as in mitochondria, this occurs by means of four consecutive reactions involving dehydrogenation, hydration, dehydrogenation, and thiolytic cleavage, respectively. Before fatty acids and fatty acid derivatives can undergo β-oxidation activation to coenzyme A must occur. This is brought about by a variety of fatty acid activating enzymes present in different compartments within the cell. In recent years much has been learned about the substrate specificity and functional organization of the peroxisomal β-oxidation system (see Mannaerts and Van Veldhoven, 1992 for review). It has become clear, for instance, that the first step in peroxisomal β-oxidation, i.e., the dehydrogenation of acyl-CoA esters to their *trans*-2-enoyl-CoA esters, is not brought about by a single oxidase. Instead, multiple oxidases are involved in the β-oxidation of the various fatty acids and fatty acid derivatives. Furthermore, there are major differences between rat and man in this respect. In the rat three different oxidases have been identified as palmitoyl-CoA oxidase, pristanoyl-CoA oxidase,

and di/trihydroxycholestanoyl-CoA oxidase, respectively, with profound differences in substrate specificity, as the names indicate (Van Veldhoven et al., 1992; Wanders et al., 1993a).

Unexpectedly, the situation is different in man, where pristanoyl-CoA and trihydroxycholestanoyl-CoA are substrates for a single enzyme (Van Hove et al., 1993). The next two steps, hydration and dehydrogenation, are carried out by a single protein harboring enoyl-CoA hydratase and L-3-hydroxyacyl-CoA dehydrogenase activities. This protein is multifunctional, at least in the rat; it also contains enoyl-CoA isomerase activity (Palosaari and Hiltunen, 1990). The final step in the β-oxidation spiral is catalyzed by a specific peroxisomal thiolase. The presence of enoyl-CoA isomerase and 2,4-dienoyl-CoA reductase activities in peroxisomes facilitates removal of the double bond in unsaturated fatty acids so that they can be β-oxidized in this organelle (Hiltunen et al., 1993). With respect to substrate specificity it is now clear that peroxisomes are responsible for the β-oxidative chain-shortening of a distinct set of compounds which cannot be handled by mitochondria. These compounds include:

1. Very-long-chain fatty acids (C24:0 and C26:0).
2. Di- and trihydroxycholestanoic acids.
3. Pristanic acid.
4. Long chain dicarboxylic acids.
5. Certain prostaglandins.
6. Certain leukotrienes.
7. 12- and 15-hydroxyeicosatetraenoic acids.
8. Certain mono- and polyunsaturated fatty acids.

For the purpose of describing the peroxisomal diseases that have been identified so far, compounds 1–3 are the most relevant. Although the oxidation of compounds 4–8 has been found to be impaired in diseases in which there is a generalized loss of peroxisomal functions, no disease has as yet been described in which the defect is restricted to an inability to oxidize only one of these compounds (see Table 1).

The activation of compounds 1–3 to their coenzyme A esters is brought about by three distinct synthetases with different subcellular localizations (Wanders et al., 1992a; Singh et al., 1987; Schepers et al., 1989). The CoA esters are subsequently handled by one of two acyl-CoA oxidases as described above (Van Hove et al., 1993). Pristanoyl-CoA and di/trihydroxycholestanoyl-CoA are dehydrogenated by the recently identified branched-chain acyl-CoA oxidase, whereas hexacosanoyl-CoA (C26:0-CoA) is primarily handled by the straight-chain acyl-CoA oxidase. The pathways of β-oxidation of the three substrates converge at the level of multifunctional protein and thiolase.

Table 1. Classification of the Peroxisomal Disorders

Group A: Generalized Loss of Peroxisomal Functions
Cerebro-hepato-renal (Zellweger) Syndrome (ZS)
Neonatal adrenoleukodystrophy (NALD)
Infantile Refsum disease (IRD)
Hyperpipecolic acidemia (HPA)
Group B: Multiple Loss of Peroxisomal Functions
Rhizomelic chondrodysplasia punctata (RCDP)
Zellweger-like syndrome
Group C: Loss of a Single Peroxisomal Function
X-linked adrenoleukodystrophy (XALD) and its phenotypic variants
Acyl-CoA oxidase deficiency (pseudo-NALD)
Bifunctional protein deficiency
Peroxisomal thiolase deficiency (pseudo-ZS)
Acyl-CoA: dihydroxyacetonephosphate acyltransferase (DHAPT) deficiency (pseudo-RCDP)
Glutaryl-CoA oxidase deficiency
Phytanic acid storage disease (Refsum disease)*
Alanine glyoxylate aminotransferase deficiency (hyperoxaluria type I)
Acatalasemia

Note: * Whether Refsum disease is a peroxisomal disease or not is disputed.

Ether-Phospholipid Biosynthesis

Peroxisomes play an essential role in ether-phospholipid biosynthesis since the two enzyme activities responsible for the introduction of the ether bond in ether-phospholipids, i.e., dihydroxyacetonephosphate acyltransferase (DHAPAT) and alkyldihydroxyacetonephosphate synthase (alkylDHAP synthase), are localized in peroxisomes (see Van den Bosch et al., 1992 for a review). The main end products of this biosynthetic pathway in man are plasmalogens, widely distributed in mammalian cell membranes and making up as much as 80–90% of ethanolamine phospholipids in brain white matter. The physiological role of ether-lipids has not yet been clarified except for platelet-activating factor, which plays a role in anaphylaxis, inflammation, and other reactions. It has been suggested that plasmalogens have an important function as scavengers of reactive oxygen species (Morand et al., 1988).

Glyoxylate Catabolism

In man, the enzyme alanine:glyoxylate aminotransferase is present almost exclusively in peroxisomes (Noguchi and Takada, 1979; Noguchi, 1987). An inherited deficiency of the enzyme gives rise to the disease known as Hyperoxaluria type I. This disease has severe clinical consequences resulting from accumulation of glyoxylate and, especially, oxalate, which is formed from glyoxylate in a reaction catalyzed by lactate dehydrogenase (see Danpure, 1989 for a review). It can be

concluded that peroxisomes are the main if not exclusive site of the catabolism of glyoxylate in man.

L-Pipecolic Acid Degradation

L-pipecolic acid is a metabolite formed during the degradation of L-lysine. It is likely that peroxisomes also play an indispensable role in the degradation of L-pipecolic acid since, in man, L-pipecolic acid oxidase is a peroxisomal enzyme (Wanders et al., 1988b; Mihalik et al., 1991). Indeed, in the Zellweger syndrome, which is characterized by a generalized impairment of peroxisomal functions, one of the abnormalities encountered is an accumulation of L-pipecolic acid (Danks et al., 1975). Although an isolated deficiency of L-pipecolic acid oxidase has not yet been demonstrated, several patients have been described with L-pipecolic acid accumulation in the absence of other abnormalities (see Roesel et al., 1991).

Biosynthesis of Certain Polyunsaturated Fatty Acids, e.g., Docosohexaenoic Acid

Based on the finding that docosahexaenoic acid levels are markedly low in erythrocytes and tissues from patients with Zellweger syndrome, Martinez (1989, 1990) has suggested that peroxisomes catalyze one of the intermediary reactions in the biosynthesis of docosahexaenoic acid, specifically the Δ 4-desaturation step.

Other Functions of Peroxisomes

Apart from the functions listed above, other metabolic pathways in which peroxisomes play a role are the following:

1. Hydrogen peroxide metabolism.
2. Cholesterol synthesis.
3. Dolichol synthesis.
4. Polyamine catabolism.

One or more enzymes in each of these pathways is localized in peroxisomes (see Van den Bosch et al., 1992 for a review). The question of whether or not the initial steps in the catabolism of phytanic acid are confined to peroxisomes is still a matter of discussion. This compound, which is derived from the phytol present in the diet, does accumulate in patients with the Zellweger syndrome (see below).

BIOGENESIS OF PEROXISOMES

It is now well-established that peroxisomes, like mitochondria and chloroplasts, arise by growth and division of pre-existing peroxisomes, rather than by budding

from the endoplasmic reticulum as originally thought. They may transiently or permanently be interconnected to form a peroxisomal reticulum containing elements resembling smooth endoplasmic reticulum. Unlike mitochondria and chloroplasts, however, peroxisomes contain no DNA, so that all peroxisomal proteins must be coded for by nuclear genes. All peroxisomal proteins investigated so far, including soluble matrix proteins, core proteins, and integral membrane proteins, are synthesized on free ribosomes and imported posttranslationally into peroxisomes. Almost all are synthesized in their final size. One exception is 3-oxo-acyl-CoA thiolase, which is synthesized as a 44 kDa precursor protein. After import into the peroxisomes it is converted to its 41 kDa mature form.

TOPOGENIC SIGNALS ON PEROXISOMAL PROTEINS

As with other proteins destined for particular subcellular sites, peroxisomal proteins must possess topogenic signals to direct them to peroxisomes. One such topogenic signal is the serine-lysine-leucine (SKL) motif (or one of its variants) present at the C-terminus of a number of peroxisomal proteins (see Subramani, 1993 for review). The fact that the SKL motif or one of its variants is not present in all peroxisomal proteins, indicates that such proteins must use some other targeting signal. Recent results from different laboratories (Swinkels et al., 1991; Osumi et al., 1991) have shown that a non-SKL type of topogenic signal is located at the amino terminal end of peroxisomal 3-oxoacyl-CoA thiolase. Whether or not there are any additional topogenic signals is not known at this time.

PEROXISOME ASSEMBLY FACTORS

The presence of topogenic signals that direct certain proteins specifically to peroxisomes implies that the peroxisomal membrane must contain receptors that recognize these signals. Such receptors will undoubtedly be integral membrane proteins. Several integral membrane proteins have been identified in peroxisomes (see Van den Bosch et al., 1992 for a review). However, although the structure of these proteins has been determined and much is known about the genes encoding them, their function is not yet known. One such protein, the 70 kDa peroxisome integral membrane protein, is a member of a superfamily of ATP-binding proteins known as the ABC transporters (Kamijo et al., 1990). These integral membrane proteins undoubtedly play an essential role in the assembly of peroxisomes so that mutations in the genes encoding these proteins can be expected to lead to defects in peroxisome assembly, i.e., to the Zellweger syndrome phenotype. Indeed, Shimozawa et al., 1992 have shown that transfection of fibroblasts from a Japanese patient with the cDNA for a 35 kDa peroxisomal membrane protein referred to as peroxisome assembly factor I (Tsukamoto et al., 1991) restores peroxisome biogenesis in these mutant cells. They have identified the defect in this patient as a point mutation leading to premature termination of peroxisome assembly factor I.

COMPLEMENTATION ANALYSIS OF CELLS FROM PATIENTS WITH DEFECTS IN PEROXISOME BIOGENESIS

Complementation analysis after somatic cell fusion of fibroblasts from patients with diseases of peroxisome biogenesis has indicated that the cell lines can be divided into at least nine complementation groups (Shimozawa et al., 1993). Thus at least nine genes must be involved in the biogenesis of peroxisomes. The products of these genes undoubtedly include not only integral membrane proteins such as receptors, as discussed above, but also soluble peroxisomal assembly factors (unfoldases; chaperonins). These complementation studies show that mutations in different genes can lead to the same clinical phenotype. On the other hand, different mutations in the same gene can lead to different clinical phenotypes; this, of course, is to be expected.

THE PEROXISOMAL DISORDERS: CLINICAL FEATURES AND BIOCHEMICAL CHARACTERISTICS

As discussed above, the peroxisomal disorders are usually subdivided into three groups depending upon the extent of peroxisomal dysfunction. It must be stressed that this classification is only of limited clinical use although it does provide some insight into the pathophysiology of the disorders. For instance, the clinical phenotype associated with the Zellweger syndrome may be found in patients with a loss of multiple peroxisomal functions (the Zellweger-like syndrome) and in patients with a deficiency of a single peroxisomal enzyme (e.g., the pseudo-Zellweger syndrome, which is caused by a deficiency of peroxisomal thiolase (see Wanders et al., 1988 for a review).

The Group A Peroxisomal Disorders

The cerebro-hepato-renal (Zellweger) syndrome, the prototype of the peroxisomal disorders, is a multi-system disorder involving both embryofetopathy and progressive changes which continue into postnatal life (Lazarow and Moser, 1989). The clinical presentation of the Zellweger syndrome is dominated in the first months of life by the typical craniofacial dysmorphia (high forehead, large anterior fontanels, hypoplastic supraorbital ridges, epicanthal folds and deformed ear lobes in >90% of cases), and profound neurological abnormalities. Impaired hearing and retinopathy are probably present in all cases and cataracts are frequently found. Furthermore, liver disease, calcific stippling of the epiphyses, and small renal cysts usually are found. Brain abnormalities include not only cortical dysplasia and neuronal heterotopia but also regressive changes. There is hypomyelination rather than demyelination. In the Zellweger syndrome peroxisomes are strongly deficient. A similar deficiency of peroxisomes is also found in other disorders of group A (Table 1).

In neonatal adrenoleukodystrophy degenerative changes take precedence over errors in morphogenesis. Patients may show some psychomotor development and achieve certain milestones before regression sets in. Kelley et al., (1986) have suggested the following criteria in order to discriminate between neonatal adrenoleukodystrophy and the Zellweger syndrome. Neonatal adrenoleukodystrophy patients show more prominent adrenal atrophy, cerebral demyelination, systemic infiltration of lipid-laden macrophages, and elevated levels of saturated very-long-chain fatty acids, whereas Zellweger syndrome patients have chondrodysplasia, glomerulocystic disease, central nervous system dysmyelination rather than demyelination, and accumulation of both saturated and unsaturated very-long-chain fatty acids.

The clinical presentation of the infantile form of Refsum disease is even milder than that of neonatal adrenoleukodystrophy. In infantile Refsum disease distinct abnormalities are absent during the neonatal period, there is only minor facial dysmorphism, and the patients may survive until the second decade of life (Poll-Thé et al., 1987). Hyperpipecolic acidemia shows a striking clinical resemblance to the Zellweger syndrome. In fact, there is still some discussion as to whether hyperpipecolic acidemia should be considered a distinct clinical entity or not (Wanders et al., 1988a).

Thus a deficiency of peroxisomes is associated with a wide spectrum of clinical abnormalities ranging in intensity from severe, as in the Zellweger syndrome, to relatively mild, as in infantile Refsum disease. The severity of the clinical abnormalities may reflect the extent of peroxisomal malfunctioning as suggested by studies of *de novo* plasmalogen biosynthesis in fibroblasts from patients (Schrakamp et al., 1988) and of the level of very-long-chain fatty acids in the patients' plasma (Lazarow and Moser, 1989). On the average the plasma C26:0 levels are lower in infantile Refsum disease patients than in the Zellweger syndrome or neonatal adrenoleukodystrophy patients. In both studies, however, the difference did not reach statistical significance due to a large spread making it doubtful whether the diagnosis Zellweger syndrome, neonatal adrenoleukodystrophy, or infantile Refsum disease can be made unequivocally on the basis of plasma very-long-chain fatty acid levels and/or the rate of plasmalogen biosynthesis in fibroblasts. Biochemically, patients affected by a group A disorder, i.e., a disorder of peroxisome biogenesis (also called peroxisomal deficiency disorder) show a great number of abnormalities (Table 2) which all follow logically from the functions of peroxisomes in man as described above.

The Group B Peroxisomal Disorders

At present, this group comprises two diseases, rhizomelic chondrodysplasia punctata and the Zellweger-like syndrome. Rhizomelic chondrodysplasia is clinically characterized by a disproportionally short stature primarily affecting the

Table 2. Biochemical Characteristics of the Peroxisomal Disorders

Parameter Measured	Type of Peroxisomal Disorder: Peroxisome Deficiency Disorders	Zellweger-like Syndrome	X-linked ALD	Acyl-CoA Oxidase Deficiency	Bifunctional Protein Deficiency	Thiolase Deficiency	Di/Trihydroxy Cholestanoic Acidemia	RCDP	Pseudo RCDP
Metabolites in body fluids									
• Very-long-chain fatty acids	Elevated	Elevated	Elevated[3]	Elevated	Elevated	Elevated	Normal	Normal	Normal
• Bile acid intermediates	Elevated	Elevated	Normal	Normal	Elevated	Elevated	Elevated	Normal	Normal
• Pipecolic acid	Elevated[1]	NA	Normal	Normal	Normal	Normal	Normal	Normal	Normal
• Phytanic acid	Elevated[2]	NA	Normal	Normal	Elevated[4]	Elevated[4]	Elevated[4]	Elevated[2]	Normal
• Pristanic acid	Elevated[2]	NA	Normal	Normal	Elevated[4]	Elevated[4]	Elevated[4]	Normal	Normal
Plasmalogen synthesis									
• DHAPAT	Deficient	Deficient	Normal	Normal	Normal	Normal	Normal	Deficient	Deficient
• Alkyl DHAP synthase	Deficient	NA	Normal	Normal	Normal	Normal	Normal	Deficient	Normal
• *De novo* synthesis	Impaired	Deficient	Normal	Normal	Normal	Normal	Normal	Deficient	Deficient
Peroxisomes									
• Hepatic peroxisomes	Deficient	Present	Present	Present	Present	Present	Present	Present	Present
• Particle-bound catalase	Deficient	NA	Present	Present	Present	Present	Present	Present	Present
Peroxisomal β -oxidation									
• Activity with C26:0	Deficient	Deficient	Deficient	Deficient	Deficient	Deficient	Deficient	Normal	Normal
• Enzyme proteins:									
- acyl-CoA oxidase	Deficient	Deficient	Normal	Deficient	Normal	Normal	Normal	Normal	Normal
- bifunctional protein	Deficient	Deficient	Normal	Normal	Deficient	Normal	Normal	Normal	Normal
- peroxisomal thiolase	Deficient	Deficient	Normal	Normal	Normal	Deficient	Normal	Abnormal[5]	Normal

Notes: Symbols used: [1] = age dependent; [2] = age- and diet-dependent; [3] = elevated except in some cases; [4] = may be normal depending on age and diet; [5] = abnormal molecular form, 44 kDa rather than 41 kDa. NA = not analyzed.

proximal parts of the extremities, typical facial appearance, congenital contractures, characteristic ocular involvement, severe growth deficiency, and mental retardation. Although death in early infancy has been described, most patients survive beyond the first year of life, sometimes well into their second decade. Four distinct abnormalities have been found in rhizomelic chondrodysplasia punctata: 1) deficient activity of acyl-CoA:dihydroxyacetonephosphate acyltransferase; 2) deficient activity of alkyldihydroxyacetonephosphate synthase; 3) deficient catabolism of phytanic acid; and 4) aberrant molecular mass of peroxisomal thiolase due to the fact that the precursor of the enzyme (molecular mass: 44 kDa) is not processed to the mature form (molecular mass: 41 kDa) (for a review see Lazarow and Moser, 1989). Recently we identified a patient with all the clinical signs and symptoms of rhizomelic chondrodysplasia punctata but with only one of the above biochemical abnormalities, namely a deficiency of acyl-CoA:dihydroxyacetonephosphate acyltransferase (Wanders et al., 1992b). The identification of this new peroxisomal disorder emphasizes the functional importance of ether-phospholipids even though it is as yet unclear exactly what all of these functions are. In the Zellweger-like syndrome, an entity clinically indistinguishable from the Zellweger syndrome, abnormalities in ether-phospholipid synthesis and peroxisomal β-oxidation are found, despite the fact that peroxisomes are abundantly present in the liver (Suzuki et al., 1988).

The Group C Peroxisomal Disorders

The group C peroxisomal disorders comprise diseases in which peroxisomes are present and only a single peroxisomal enzyme is deficient. This group includes X-linked adrenoleukodystrophy, the most frequently encountered of the peroxisomal disorders. The clinical presentation of adrenoleukodystrophy is highly variable (Moser and Moser, 1991). The childhood cerebral form of adrenoleukodystrophy is the most stereotypic form of the disease. In other patients the clinical presentation may begin only in adolescence or adult life as in adrenomyeloneuropathy. Childhood adrenoleukodystrophy patients typically display behavioral, visual and/or auditory disturbances as well as an abnormal gait. The disease usually culminates within a few years in dementia, blindness, quadriplegia, and death. Darkening of the skin, secondary to adrenal dysfunction, may develop. Adrenal insufficiency may precede the development of neurological symptoms and even remain the only clinical abnormality.

Biochemically, patients with adrenoleukodystrophy and adrenomyeloneuropathy show elevated very-long-chain fatty acids (notably C24:0 and C26:0) which is due to an impaired peroxisomal β-oxidation of these fatty acids (Moser and Moser, 1991). The impairment is due to a deficient activity of the peroxisomal very-long-chain acyl-CoA synthetase whereas the capacity of the microsomes to activate very-long-chain fatty acids is normal (Lazo et al., 1988;

Wanders et al., 1988c). This finding provided an explanation for the earlier observation that chemically synthesized very-long-chain acyl-CoA esters are oxidized normally in homogenates of fibroblasts from X-linked adrenoleukodystrophy patients (Hashmi et al. 1986).

Recently, Mosser et al. (1993) have identified the X-linked adrenoleukodystrophy gene. Surprisingly, the nucleotide sequence of the gene showed no homology to known acyl-CoA synthetases. Instead, it showed strong homology with a family of ATP-binding transport proteins known as the ABC transporters. Other members of this family are the cystic fibrosis chloride channel regulator, the multidrug resistance P-glycoprotein, and the 70 kDa peroxisomal integral membrane protein (see Higgins, 1992 for a review). At present it is not clear how a functional deficiency of very-long-chain acyl-CoA synthetase is related to a defect in the gene identified by Mosser et al. (1993). One possibility is that the gene encodes a protein involved in the correct insertion of very-long-chain acyl-CoA synthetase in the peroxisomal membrane.

Apart from X-linked adrenoleukodystrophy there are a number of other disorders in which peroxisomal β-oxidation is impaired. These disorders include acyl-CoA oxidase deficiency (referred to as pseudo-neonatal adrenoleukodystrophy), bifunctional protein deficiency and, as indicated above, peroxisomal thiolase deficiency (referred to as pseudo-Zellweger syndrome). The clinical phenotype of these latter disorders is quite different from those of the different forms of adrenoleukodystrophy in which symptoms do not appear before 3–4 years of age at the earliest. In contrast clinical abnormalities are already present at birth in patients with a deficiency of acyl-CoA oxidase (Poll-Thé et al., 1988), bifunctional protein (Watkins et al., 1989), or peroxisomal thiolase (Goldfischer et al., 1986; Schram et al., 1987). Other disorders belonging to group C are hyperoxaluria type I (see Danpure, 1989 for a review), glutaryl-CoA oxidase deficiency in a variant case of glutaric aciduria type I (Bennett et al., 1991), dihydroxyacetonephosphate acyltransferase deficiency in a variant case of rhizomelic chondrodysplasia punctata (Wanders et al., 1992b), and acatalasemia, a relatively innocuous disease. It is still disputed whether phytanic acid storage disease (Refsum disease) is a peroxisomal disorder or not, simply because the subcellular site of phytanic acid α-oxidation has not yet been settled.

Biochemical Identification of Peroxisomal Disorders

There is neurological involvement in all of the peroxisomal diseases identified so far with the exception of hyperoxaluria type I and acatalasemia. Furthermore, there is an accumulation of very-long-chain fatty acids in all peroxisomal disorders with neurological involvement except for rhizomelic chondrodysplasia punctata and isolated acyl-CoA:dihydroxyacetonephosphate acyltransferase deficiency (pseudo-rhizomelic chondrodysplasia punctata). Accordingly, analysis of very-

long-chain fatty acids in plasma or serum has generally been regarded as a good screening method for peroxisomal disorders. This is usually done by means of gas chromatography with or without mass-spectrometry. Until recently, most laboratories used the procedure developed by Moser and coworkers (see Moser and Moser, 1991). However, this method is rather laborious. Recently, Onkenhout et al. (1989) described a simple, one-step procedure for determination of plasma VLCFAs that seems to be superior to other methods used (see Moser and Moser, 1991 for a discussion).

Experience in our own laboratory over the last few years has shown that analysis of very long-chain fatty acids is, indeed, a reliable method of screening for peroxisomal disorders with an impairment in peroxisomal β-oxidation. As indicated above, this includes all disorders listed in Table 1 except for rhizomelic chondrodysplasia punctata, pseudo-rhizomelic chondrodysplasia punctata, hyperoxaluria type I, glutaryl-CoA oxidase deficiency, acatalasemia, and the cases with di- and trihydroxycholestanoic acidemia recently described in the literature (see Wanders et al., 1993b for references).

A caveat is that plasma levels of very long-chain fatty acids may not be aberrant in a minority of patients with X-linked adrenoleukodystrophy. We have recently observed two patients with all of the clinical signs and symptoms of the disease whose plasma levels of very long-chain fatty acids were repeatedly normal (Wanders et al., 1992c). However, the levels of very long-chain fatty acids were clearly abnormal in fibroblasts from the patients. Thus, additional studies with fibroblasts are imperative in all cases with clinical signs and symptoms suggestive of childhood X-linked adrenoleukodystrophy (or one of its phenotypic variants) but with normal plasma levels of very long-chain fatty acids. The recent identification of patients with normal plasma very long-chain fatty acid levels but elevated levels of di- and trihydroxycholestanoic acid suggests that these metabolites must be analyzed in those cases in which a peroxisomal disorder is strongly suspected on clinical grounds but in which plasma very long-chain fatty acids are normal (Wanders et al., 1993b).

Biochemical diagnosis of rhizomelic chondrodysplasia punctata or its variant with an isolated deficiency of acyl-CoA:dihydroxyacetonephospate acyltransferase activity should be based on an analysis of erythrocyte plasmalogen levels together with measurements of acyl-CoA:dihydroxyacetonephosphate acyltransferase activity in leukocytes or platelets, followed by more detailed studies in fibroblasts. Additionally, methods for the prenatal diagnosis of all peroxisomal disorders listed in Table 1 have become available (Schutgens et al., 1989).

CONCLUSIONS AND PERSPECTIVES

In the two decades that have elapsed since the discovery by Goldfischer and coworkers that peroxisomes are deficient in the Zellweger syndrome much has been learned about the clinical and biochemical characteristics of peroxisomal disorders,

although it is clear that many patients remain unrecognized. Furthermore, new peroxisomal disorders remain to be discovered, particularly those caused by a deficiency of a single peroxisomal enzyme.

The last two decades have also been characterized by an expansion of our knowledge of the functions of peroxisomes. Clearly, these organelles are highly versatile. Considerable progress is being made in understanding the mechanisms involved in the biogenesis of peroxisomes and in defining the molecular genetics of the peroxisomal disorders. Recent discoveries in these two fields should lead to a more accurate diagnosis of the peroxisomal disorders and to improved carrier detection. In addition, the molecular genetic studies open up the possibility of curing peroxisomal diseases by means of gene therapy. The next decade should be an exciting one in terms of recognizing, detecting, diagnosing and perhaps curing inborn errors of metabolism involving this versatile organelle.

ACKNOWLEDGMENTS

The authors' studies referred to in this chapter were supported by grants from the Section for Medical Sciences of the Netherlands Organization for Scientific Research (NWO) and the Princess Beatrix Fund (The Hague). Mrs. Iet van der Gracht and Sjoukje Mooybroek are gratefully acknowledged for expert preparation of the manuscript.

REFERENCES

Bennett, M. J., Pollitt, R. J., Goodman, S. I., Hale, D. E., & Vamecq, J. (1991). Atypical riboflavin-responsive glutaric aciduria and deficient peroxisomal glutaryl-CoA oxidase activity: a new peroxisomal disorder. J. Inher. Metab. Dis. 14, 165–173.

Bowen, P., Lee, C. S. M., Zellweger, H., & Lindenberg, R. (1964). A familial syndrome of multiple congenital defects. Bull. Johns Hopkins Hosp. 114, 402–414.

Brown, F. R., McAdams, A. J., Cummins, J. W., Konkol, R., Singh, I., Moser A. E., & Moser, H. W. (1982). Cerebro-hepato-renal (Zellweger) syndrome and neonatal adrenoleukodystrophy: similarities in phenotype and accumulation of very-long-chain fatty acids. Johns Hopkins Med. J. 151, 344–361.

Danks, D. M. P., Tippett, P., Adams, C., & Campbell, P. (1975). Cerebro-hapato-renal syndrome of Zellweger. A report of eight cases with comments upon the incidence, the liver lesion and a fault in pipecolic acid metabolism. J. Pediatr. 86, 382–387.

Danpure, C. J. (1989). Recent advances in the understanding, diagnosis and treatment of primary hyperoxaluria type I. J. Inher. Metab. Dis. 12, 210–224.

De Duve, C. & Baudhuin, P. (1966). Peroxisomes (microbodies and related particles). Physiol. Rev. 46, 323–357.

Eaton, J. W. (1989). Acatalasemia. In: The Metabolic Basis of Inherited Disease (Scriver, C. R., Beaudet, A. L., Sly, W. S., & Valle, D., eds.), pp. 1551–1561, McGraw-Hill, New-York.

Goldfischer, S., Moore, C. L., Johnson, A. B., Spiro, A. J., Valsamis, M. P., Wisniewski, H. K., Ritch, R. H., Norton, W. T., Rapin, I., & Gartner, L. M. (1973). Peroxisomal and mitochondrial defects in the cerebro-hepato-renal syndrome. Science 182, 62–64.

Goldfischer, S., Collins, J., Rapin, I., Neumann, P., Neglia, W., Spiro, A. J., Ishii, T., Roels, F., Vamecq, J., & Van Hoof, F. (1986). Pseudo-Zellweger syndrome: deficiencies in several peroxisomal oxidative activities, J. Pediatr. 108, 25–32.

Hanson, R. F., Szcepanik-van Leeuwen, P., Williams, G. C., Grabowski, G., & Sharp, H. L. (1979). Defects of bile acid synthesis in Zellweger's syndrome. Science 203, 1107–1108.

Hashmi, M., Stanly, W., & Singh, I. (1986). Lignoceroyl-CoASH ligase: enzyme defect in fatty acid β-oxidation in X-linked adrenoleukodystrophy. FEBS Lett. 196, 247–250.

Heymans, H. S. A., Schutgens, R. B. H., Tan, R., Van den Bosch, H., & Borst, P. (1983). Severe plasmalogen deficiency in tissues of infants without peroxisomes (Zellweger syndrome). Nature (London) 306, 69–70.

Heymans, H. S. A., Van den Bosch, H., Schutgens, R. B. H., Tegelaers, W. H. H., Walther, J. U., Mueller-Hoecker, J., & Borst, P. (1984). Deficiency of plasmalogens in the cerebro-hepato-renal (Zellweger) syndrome. Eur. J. Pediatr. 142, 10–15.

Higgins, C. F. (1992). ABC transporters: from microorganisms to man. Ann. Rev. Cell Biol. 8, 67–113.

Hiltunen, J. K., Filppula, S. A., Häyrinen, H. M., Koivuranta, K. T., & Hakkola, E. H. (1993). Peroxisomal β-oxidation of polyunsaturated fatty acids. Biochimie 75, 175–182.

Kamijo, K., Taketani, S., Yokota, S., Osumi, T., & Hashimoto, T. (1990). The 70-kDa peroxisomal membrane protein is a member of the Mdr (P-glycoprotein)-related ATP-binding protein superfamily. J. Biol. Chem. 265, 4534–4540.

Kelley, R. I., Datta, N. S., Dobijns, W. S., Hajra, A. K., Moser, A. B., Noetzel, M. J., Zackai, E. H., & Moser, H. W. (1986). Neonatal adrenoleukodystrophy: new cases, biochemical studies and differentiation from Zellweger and related peroxisomal polydystrophy syndromes. Am. J. Med. Genet. 23, 869–901.

Lazarow, P. B. & Moser, H. W. (1989). Disorders of peroxisome biogenesis. In: The Metabolic Basis of Inherited Disease (Scriver, C. R., Beaudet, A. L., Sly, W. S., & Valle, D., eds.), pp. 1479–1509, McGraw-Hill, New York.

Lazo, O., Contreras, M., Hashmi, M., Stanley, W., Irazu, C., & Singh, I. (1988). Peroxisomal lignoceroyl-CoA ligase deficiency in childhood adrenoleukodystrophy. Proc. Natl. Acad. Sci. USA 85, 7647–7651.

Mannaerts, G. P. & Van Veldhoven, P. P. (1992). Role of peroxisomes in mammalian metabolism. Cell. Biochem. Funct. 10, 141–151.

Martinez, M. (1989). Polyunsaturated fatty acid changes suggesting a new enzymatic defect in Zellweger syndrome. Lipids 24, 261–265.

Martinez, M. (1990). Severe deficiency of docosahexaenoic acid in peroxisomal disorders: a defect of Δ4-desaturation? Neurology 40, 1292–1298.

Mihalik, S.J., McGuinness, M., & Watkins, P.A. (1991). Purification and characterization of peroxisomal L-pipecolic acid oxidase from monkey liver. J. Biol. Chem. 266, 4822–4830.

Morand, O.H., Zoeller, R.A., & Raetz, C.R.H. (1988). Disappearance of plasmalogens from membranes of animal cells subjected to photosensitized oxidation. J. Biol. Chem. 263, 11597–11606.

Moser, H.W. & Moser, A.B. (1991). Measurement of saturated very-long-chain fatty acids in plasma. In: Techniques in Diagnostic Human Biochemical Genetics (Hommes, F.A., ed.), pp. 177–191, Wiley-Liss, New York.

Mosser, J., Douar, A. M., Sarde, C. O., Kioschis, P., Feil, R., Moser, H., Pouska, A. M., Mandel, J. L., & Aubourg, P. (1993). The putative X-linked adrenoleukodystrophy gene shows unexpected homology to "ABC" superfamily of transporters. Nature (London) 361, 726–730.

Noguchi, T. & Takada, Y. (1979). Peroxisomal localization of alanine glyoxylate aminotransferase in human liver. Arch. Biochem. Biophys. 196, 645–647.

Noguchi, T. (1987). Amino acid metabolism in animal peroxisomes. In: Peroxisomes in Biology and Medicine (Fahimi, H. D. & Sies, H., eds.), pp. 234–243, Springer-Verlag, Berlin/Heidelberg.

Onkenhout, W., Van der Poel, P. F. H., & Van den Heuvel, M. P. M. (1989). Improved determination of very-long-chain fatty acids in plasma and cultured skin fibroblasts: application to the diagnosis of peroxisomal disorders. J. Chromatogr. 494, 31–41.

Osumi, T., Tsukamoto, T., Hata, S., Yokota, S., Miura, S., Fujiki, Y., Hijikata, M., Miyazawa, S., & Hashimoto, T. (1991). Amino-terminal presequence of the precursor of peroxisomal 3-ketoacyl-

CoA thiolase is a cleavable signal peptide for peroxisomal targeting. Biochem. Biophys. Res. Commun. 181, 947–954.

Palosaari, P. M. & Hiltunen, J. K. (1990). Peroxisomal bifunctional protein from rat liver is a trifunctional enzyme possessing 2-enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase, and Δ^3,Δ^2-enoyl-CoA isomerase activities. J. Biol. Chem. 265, 2446–2449.

Passarge, E. & McAdams, A. J. (1967). Cerebro-hepato-renal syndrome. A newly recognized hereditary disorder of multiple congenital defects including sudanophilic leukodystrophy, cirrhosis of the liver and polycystic kidneys. J. Pediatr. 71, 691–702.

Poll-Thé, B. T., Saudubray, J. M., Ogier, H. A. M., Odievre, M., Scotto, J. M., Monnens, L., Govaerts, L. C. P., Roels, F., Cornelis, A., Schutgens, R. B. H., Wanders, R. J. A., Schram, A. W., & Tager, J. M. (1987). Infantile Refsum disease: an inherited peroxisomal disorder; comparison with Zellweger syndrome and neonatal adrenoleukodystrophy. Eur. J. Pediatr. 146, 477–483.

Poll-Thé, B. T., Roels, F., Ogier, H., Scotto, H., Vamecq, J., Schutgens, R. B. H., Van Roermund, C. W. T., Van Wijland, M. J. A., Schram, A. W., Tager, J. M., & Saudubray, J. M. (1988). A new peroxisomal disorder with enlarged peroxisomes and a specific deficiency of acyl-CoA oxidase (pseudo neonatal adrenoleukodystrophy). Am. J. Hum. Genet. 42, 422–434.

Rhodin, J. (1954). Correlation of ultrastructural organization and function in normal and experimentally charged proximal tubule cells of the mouse kidney. Ph. D. thesis, Karolinska Institutet, Stockholm, Akitbolaget, Godvil.

Roesel, R. A., Carroll, J. E., Rizzo, W. B., Van der Zalm, T., & Hahn, D. A. (1991). Dyggve-Melchior-Clausen syndrome with increased pipecolic acid in plasma and urine. J. Inher. Metab. Dis. 14, 876–880.

Schepers, L., Casteels, M., Verheyden, K., Parmentier, G., Asselbergs, S., Eyssen, H. J., & Mannaerts, G. P. (1989). Subcellular distribution and characteristics of trihydroxycoprostanoyl-CoA synthetase in rat liver. Biochem. J. 257, 221–229.

Schrakamp, G., Schalkwijk, C. G., Schutgens, R. B. H., Wanders, R. J. A., Tager, J. M., & Van den Bosch, H. (1988). Plasmalogen biosynthesis in peroxisomal disorders. J. Lipid. Res. 29, 325–334.

Schram, A. W., Goldfischer, S., Van Roermund, C. T. W., Brouwer-Kelder, E. M., Collins, J., Hashimoto, T., Heymans, H. S. A., Van den Bosch, H., Schutgens, R. B. H., Tager, J. M., & Wanders, R. J. A. (1987). Human peroxisomal 3-oxoacyl-coenzyme A thiolase deficiency. Proc. Natl. Acad. Sci. USA 84, 2494–2496.

Schutgens, R. B. H., Schrakamp, G., Wanders, R. J. A., Heymans, H. S. A., Tager, J. M., & Van den Bosch, H. (1989). Pre- and perinatal diagnosis of peroxisomal disorders. J. Inher. Metab. Dis. 12, 118–134.

Shimozawa, N., Tsukamoto, T., Suzuki, Y., Orii, T., Shirayochi, Y., Mori, T., & Fujuki, Y. (1992). A human gene responsible for Zellweger syndrome that affects peroxisome assembly. Science 155, 1132–1134.

Shimozawa, N., Suzuki, Y., Orii, T., Moser, A., Moser, H.W., & Wanders, R. J. A. (1993). Standardization of complementation grouping of peroxisome-deficient disorders and the second Zellweger patient with peroxisomal assembly factor-I (PAF-I) defect. Am. J. Hum. Genet. 52, 843–844.

Singh, H., Derwas, N., & Poulos, A. (1987). Very long chain fatty acid oxidation by rat liver mitochondria and peroxisomes. Arch. Biochem. Biophys. 259, 382–390.

Subramani, S. (1993). Protein import into peroxisomes and biogenesis of the organelle. Ann. Rev. Cell Biol. 9, 445–478.

Suzuki, Y., Shimozawa, N., Orii, T., Igarashi, N., Kono, N., Matsui, A., Inoue, Y., Yokota, S., & Hashimoto, T. (1988). Zellweger-like syndrome with detectable hepatic peroxisomes: a variant form of peroxisomal disorder J. Pediatr. 113, 841–845.

Swinkels, B. M., Gould, S. J., Bodnar, A. G., Rachubinski, R. A., & Subramani, S. (1991). A novel, cleavable peroxisomal targeting signal at the amino-terminus of rat 3-ketoacyl-CoA thiolase. EMBO J. 10, 3255–3262.

Tager, J. M., Brul, S., Wiemer, E. A. C., Heikoop, J. C., Middelkoop, E., Bout, B., Westerveld, A., & Wanders, R. J. A. (1990). Peroxisomal disorders: an updating. In: Adrenoleukodystrophy and Other Peroxisomal Disorders (Uziel, G., Wanders, R.J.A., & Cappa, M., eds.), pp. 17–36, Elsevier, Amsterdam.

Tsukamoto, T., Miura, S., & Fujiki, Y. (1991). Restoration by a 35kDa membrane protein of peroxisome assembly in a peroxisome deficient mammalian cell mutant. Nature (London) 350, 77–81.

Van den Bosch, H., Schutgens, R. B. H., Wanders, R. J. A., & Tager, J. M. (1992). Biochemistry of peroxisomes. Ann. Rev. Biochem. 263, 11590–11596.

Van Hove, G. F., Van Veldhoven, P. P., Fransen, M., Denis, S., Eyssen, H. J., Wanders, R. J. A., & Mannaerts, G. P. (1993). The CoA-esters of 2-methyl-branched chain fatty acids and of the bile acid intermediates di- and trihydroxycoprostanic acids are oxidized by one single peroxisomal branched chain acyl-CoA oxidase in human liver and kidney. J. Biol. Chem. 268, 10335–10344.

Van Veldhoven, P. P., Van Hove, G., Asselbergh, S., Eyssen, H. J., & Mannaerts, G. P. (1992). Substrate specificities of rat liver acyl-CoA oxidases: palmitoyl-CoA oxidase (inducible acyl-CoA oxidase), pristanoyl-CoA (non-inducible) acyl-CoA oxidase and trihydroxycholestanoyl-CoA oxidase. J. Biol. Chem. 267, 20065–20074.

Wanders, R. J. A., Heymans, H. S. A., Schutgens, R. B. H., Barth, P. G., Van den Bosch, H., & Tager, J. M. (1988a). Peroxisomal disorders in neurology. J. Neurol. Sci. 88, 1–39.

Wanders, R.J.A., Romeyn, G.J., Van Roermund, C.W.T., Schutgens, R.B.H., Van den Bosch, H., & Tager, J.M. (1988b). Identification of L-pipecolic oxidase in human liver and its deficiency in the Zellweger syndrome. Biochem. Biophys. Res. Commun. 154, 33–38.

Wanders, R.J.A., Van Roermund, C.T.W., Van Wijland, M.J.A., Schutgens, R.B.H., Van den Bosch, H., & Tager, J.M. (1988c). Direct demonstration that the deficient oxidation of very long chain fatty acids in X-linked adrenoleukodystrophy is due to an impaired ability of peroxisomes to activate very long chain fatty acids. Biochem. Biophys. Res. Commun. 153, 618–624.

Wanders, R.J.A., Denis, S., Van Roermund, C.W., Jakobs, C., & Ten Brink, H.J. (1992a). Characteristics and subcellular localization of pristanoyl-CoA synthetase in rat liver. Biochim. Biophys. Acta. 1125, 274–279.

Wanders, R.J.A., Schumacher, H., Heikoop, J., Schutgens, R.B.H., & Tager, J.M. (1992b). Human dihydroxyacetonephosphate acyltransferase deficiency: a new peroxisomal disorder. J. Inher. Metab. Dis. 15, 389–391.

Wanders, R.J.A., Van Roermund, C.W.T., Lageweg, W., Jakobs, B.S., Schutgens, R.B.H., Nijenhuis, A.A., & Tager, J.M. (1992c). X-linked adrenoleukodystrophy: biochemical diagnosis and enzyme defect. J. Inher. Metab. Dis. 15, 634–644.

Wanders, R.J.A., Denis, S.W., & Dacremont, G. (1993a). Studies on the substrate specificity of the inducible and non-inducible acyl-CoA oxidases from rat kidney peroxisomes. J. Biochem. 113, 577–582.

Wanders, R.J.A., Schutgens, R.B.H., Barth, P.G., Tager, J.M., & Van den Bosch, H. (1993b). Postnatal diagnosis of peroxisomal disorders: a practical approach. Biochimie. 75, 269–279.

Watkins, P.A., Chen, W.W., Harris, C.J., Hoefler, G., Hoefler, S., Blake, D.C., Balfe, A., Kelley, R.I., Moser, A.B., Beard, M.E., & Moser, H.W. (1989). Peroxisomal bifunctional enzyme deficiency J. Clin. Invest. 83, 771–777.

Chapter 7

Molecules in Living Cells

DAVID S. GOODSELL

INTRODUCTION

An experimentally invisible size range exists between the world of cells and the world of molecules. At one extreme, cell biologists use the techniques of microscopy to examine cells. With light microscopy, features on the order of microns can be discerned. Cell motility, cytoplasmic streaming, and mitosis yield their secrets. With electron microscopy, even finer features, close to molecular resolution, can be studied. With the scanning electron microscope, shadowed surfaces of cells may be seen. In thin sections, we see the partitioning of eukaryotic cells into functional compartments—nucleus, mitochondria, endomembrane structures—and study the complex details of transport and communication within the cell. At the highest resolution, the shapes of large molecular complexes may even be determined. Often, however, the rigors of sample preparation and radiation damage destroy molecular features.

Principles of Medical Biology, Volume 3
Cellular Organelles and the Extracellular Matrix, pages 173–180

ISBN: 1-55938-804-8

Biochemists and molecular biologists, on the other hand, see the atomic details of molecules. X-ray crystallography, NMR spectroscopy, and a variety of physical methods may be used to determine the exact shape and size of individual molecules, often revealing the location of each of the thousands of atoms in a biologically interesting macromolecule. Hundreds of proteins and dozens of nucleic acids have been studied to atomic detail and are available to researchers through the Brookhaven Protein Data Bank* (Bernstein et al., 1977). The biochemist's view is intrinsically myopic, however, as these techniques typically give detailed information only on a single, purified subject. All information on context and interaction is lost.

The pictures in this chapter attempt to bridge these two complementary, but non-overlapping, experimental views of cells. Data from microscopy is used to define the gross arrangement of membranes and location of large macromolecules like ribosomes and fibers of DNA. Data from chemical composition studies are used to determine the number of individual molecules which make up the grey areas and dark bands in electron micrographs. And finally, structural results are used to determine the shape and size of each of these molecules. Together, this data is sufficient to create a convincing simulation of the arrangement of molecules inside a living cell (Goodsell, 1991, 1992).

ESCHERICHIA COLI

The intestinal bacterium *Escherichia coli* is an ideal subject for creating a simulated view of a living organism. It is a relatively simple cell—a bag of enzymes—and has been characterized at many levels by numerous experimental methods (Neidhardt et al., 1990). Figure 1 shows a cross section through a typical *E. coli* cell magnified by one million times. At this magnification, the entire sausage-shaped cell would extend the length of your room and would come up to about your knees in height. The handful of flagella would extend 5 to 10 meters across the floor in graceful curves. Note that these illustrations include only the macromolecular components of the cell. Thousands of small molecules, ions, and water are omitted and must be imagined filling the interstices. Extending from top to bottom on the left of Figure 1a is the multi-layered cell wall. Outermost (furthest to the left) is the outer membrane, facing the surrounding water. Lipopolysaccharides make up most of the outer half of the lipid bilayer in this membrane, with long, gluey polysaccharide chains protecting the surface of the bacterium. Piercing the membrane are porin proteins forming stable holes through the membrane, allowing nutrients to enter and wastes to exit.

Small lipoproteins bridge the outer membrane to the next layer inside: the structural layer of peptidoglycan. This cross-linked net of protein and polysaccha-

*The Protein Data Bank may be reached at: Protein Data Bank, Chemistry Department, Building 555, Brookhaven National Laboratory, Upton, NY 11973, USA.

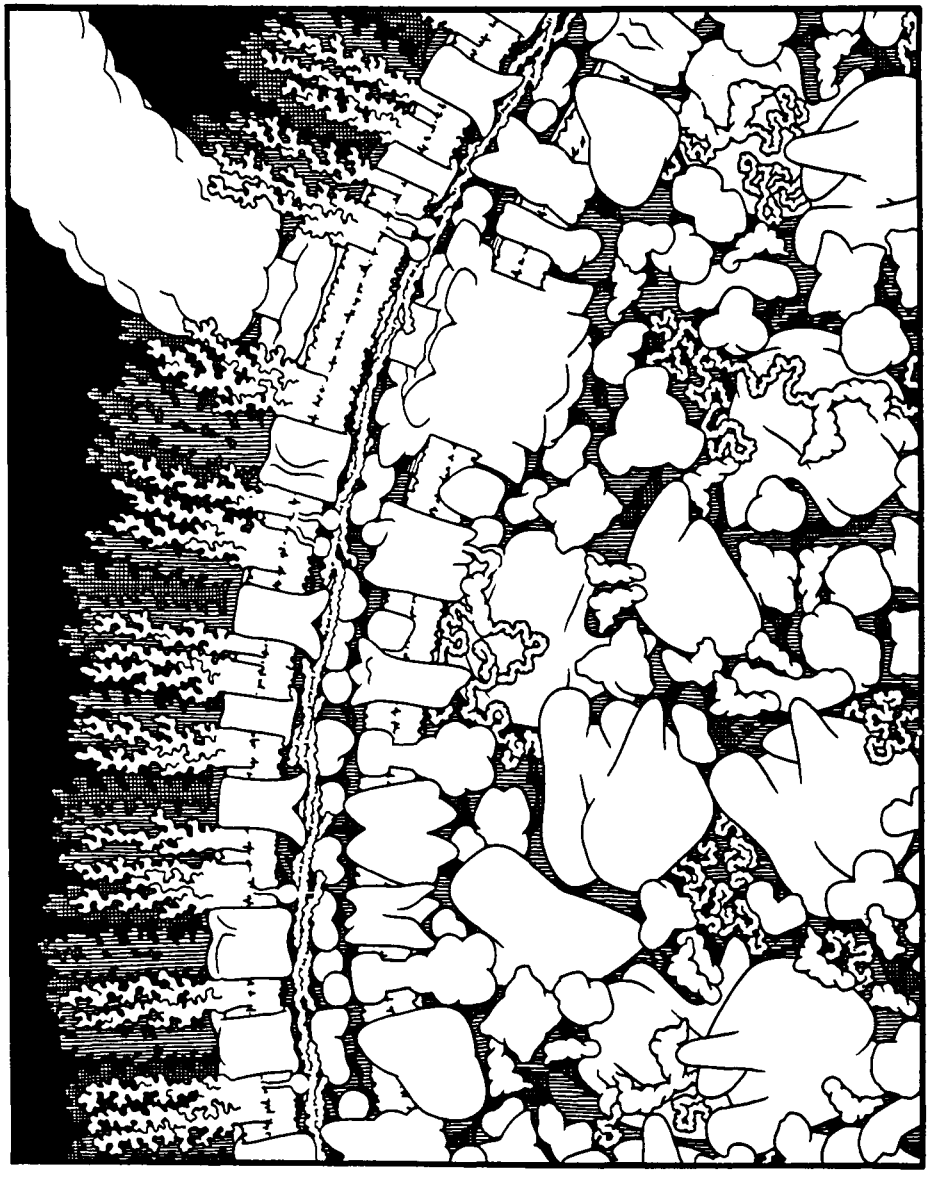


Figure 1. A section through *E. coli* at one million times magnification. a) The layered cell wall extends from top to bottom, with a flagellar motor and flagellum at top. *(continued)*

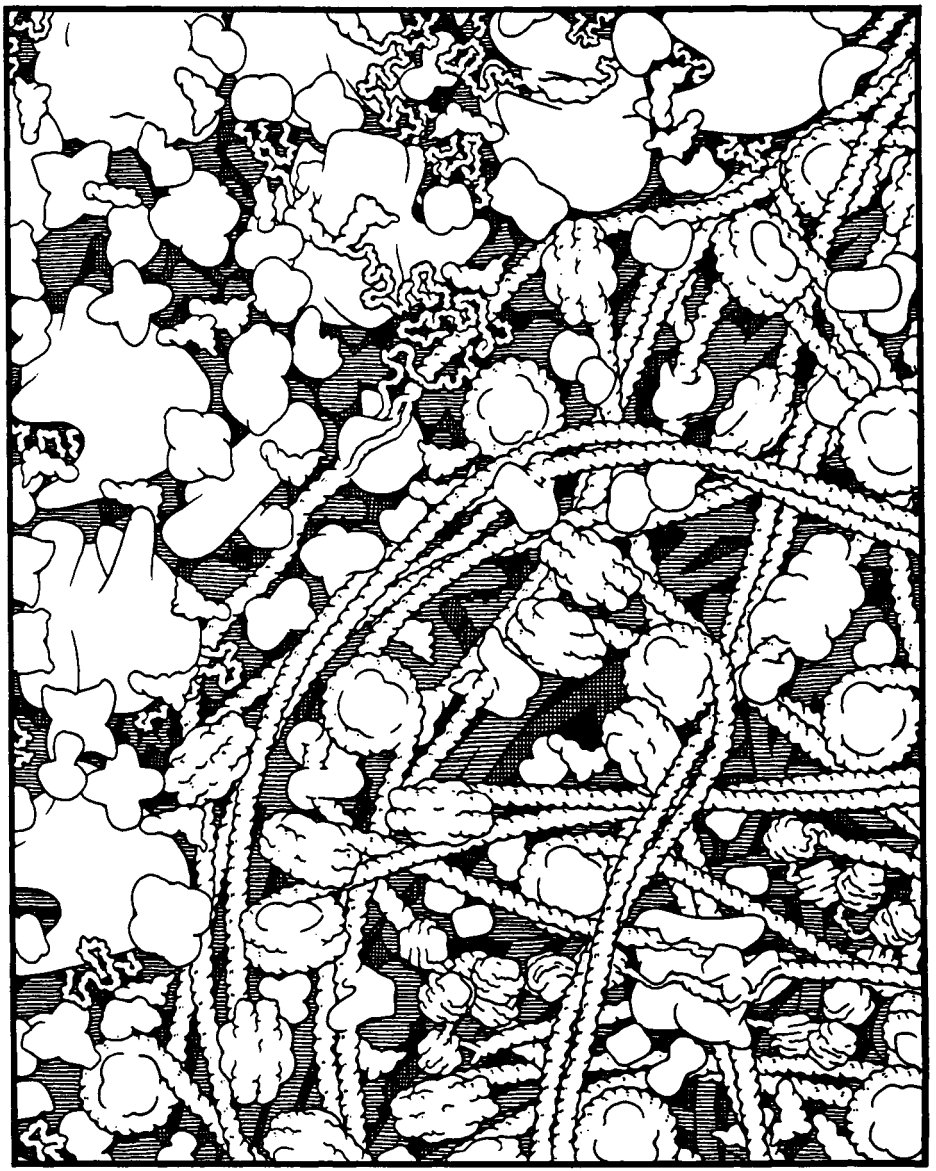


Figure 1. (continued) A section through *E. coli* at one million times magnification. b) The cytoplasm continues in the left half and a section of the nuclear region is at the right. See Figure 3 for the key identifying individual components.

ride forms a tough sheath to help hold the non-spherical shape of the bacterium. Moving inward (to the right), we cross the periplasmic space, filled with a collection of soluble carrier and electron transport proteins, to reach the complex inner membrane. This membrane is filled to capacity with proteins for specific transport, energy production, and signalling. In addition, spanning the entire cell wall of the bacterium are motor complexes which turn the flagella. Using energy from a proton gradient across the inner membrane, these motors turn the flagella in either direction to propel the bacterium through its environment.

Inside the cell wall is the cytoplasm (right side of Figure 1a and the left side of Figure 1b), where most of the synthetic tasks of the cell are performed. Dominating the view are many ribosomes, synthesizing new proteins from twisting RNA messages. Between ribosomes are about one thousand different types of enzymes, performing the individual chemical steps of biosynthesis and breakdown of foodstuffs. These enzymes come in many shapes and sizes (Goodsell and Olson, 1993). Most are oligomeric complexes of identical subunits: symmetrical dimers, trimers in a perfect triangle, octamers arranged in a cube, or even larger complexes.

Innermost (to the right in Figure 1b) is the nuclear material of the cell. The DNA of the cell—one large circle, one half millimeter in diameter if fully extended, and often several smaller plasmids—is coiled and supercoiled through the central regions of the cell. A diverse collection of proteins help package the DNA and regulate its transcription.

MITOCHONDRIA

Of more personal interest to us, as eukaryotes, are the details of the eukaryotic cell. Where the bacterial cell is relatively simple, eukaryotic cells are filled with complexity, with specialized compartments and an elaborate overhead of communication. Figure 2 shows one of these compartments, a mitochondrion, again at one million times magnification (Srere, 1982). At this magnification, the branching mitochondrion would be slightly larger in diameter than the *E. coli* cell. A typical eukaryotic cell at one million times magnification would fill an entire six-story building.

To the left of Figure 2a is the cytoplasm of the cell, surrounding the mitochondrion. Protein filaments, including several actin filaments seen here, permeate the cytoplasm, supporting the larger eukaryotic cell. Protein synthesis is performed here, by larger ribosomes than those seen in the bacterial cell. Myriad enzymes perform other chemical tasks.

The outer mitochondrial membrane, running from top to bottom in the center of Figure 2b, is pierced with proteins similar to the porin protein of *E. coli*, as well as a variety of other signalling and transport proteins. Continuing inward (to the right) we cross the intermembrane space, with various soluble electron transport proteins. The folded inner membrane is next, packed with the integral membrane proteins of oxidative phosphorylation and membrane transport. Innermost, to the right of

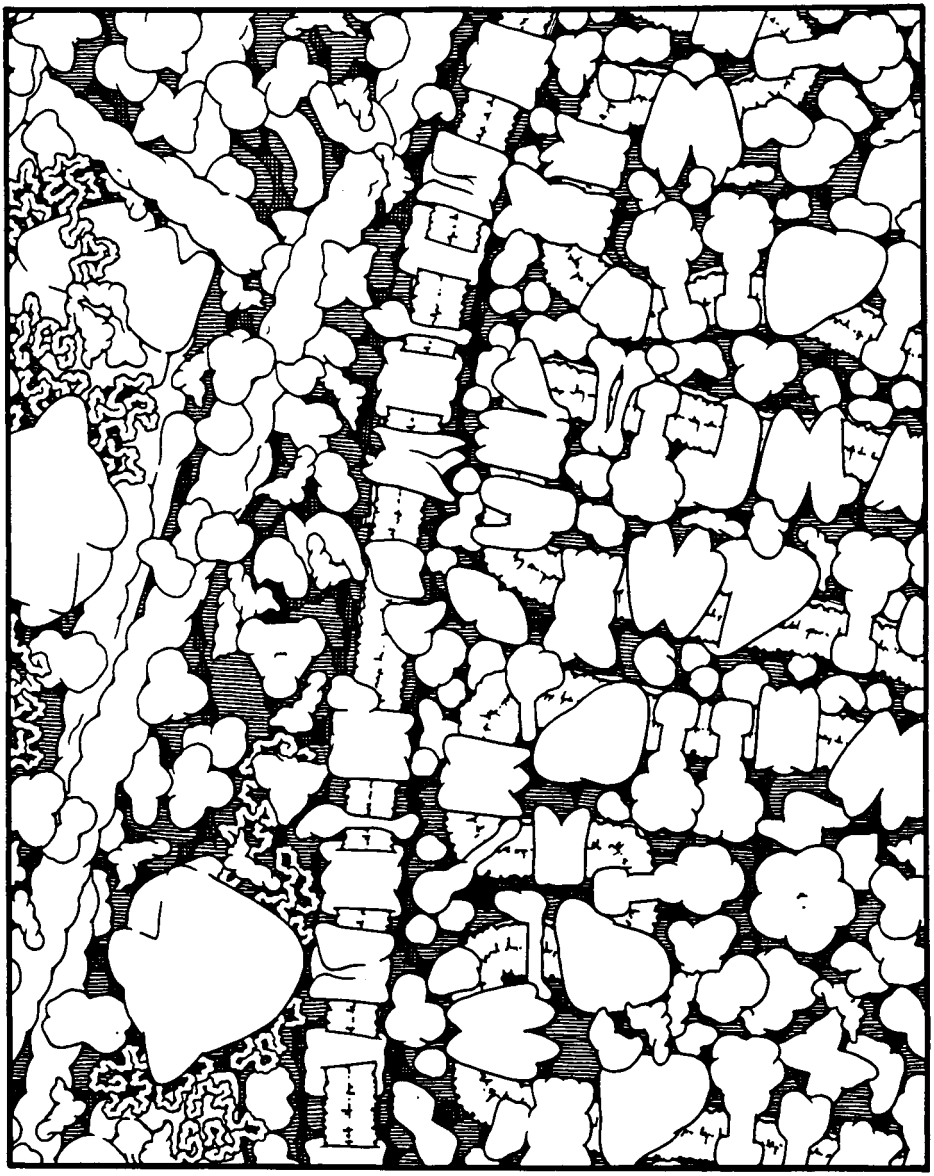


Figure 2. A section through a mitochondrion at one million times magnification. a) The two-layered cell wall of the mitochondrion extends from top to bottom, with the cytoplasm to the left and the mitochondrial matrix to right. *(continued)*

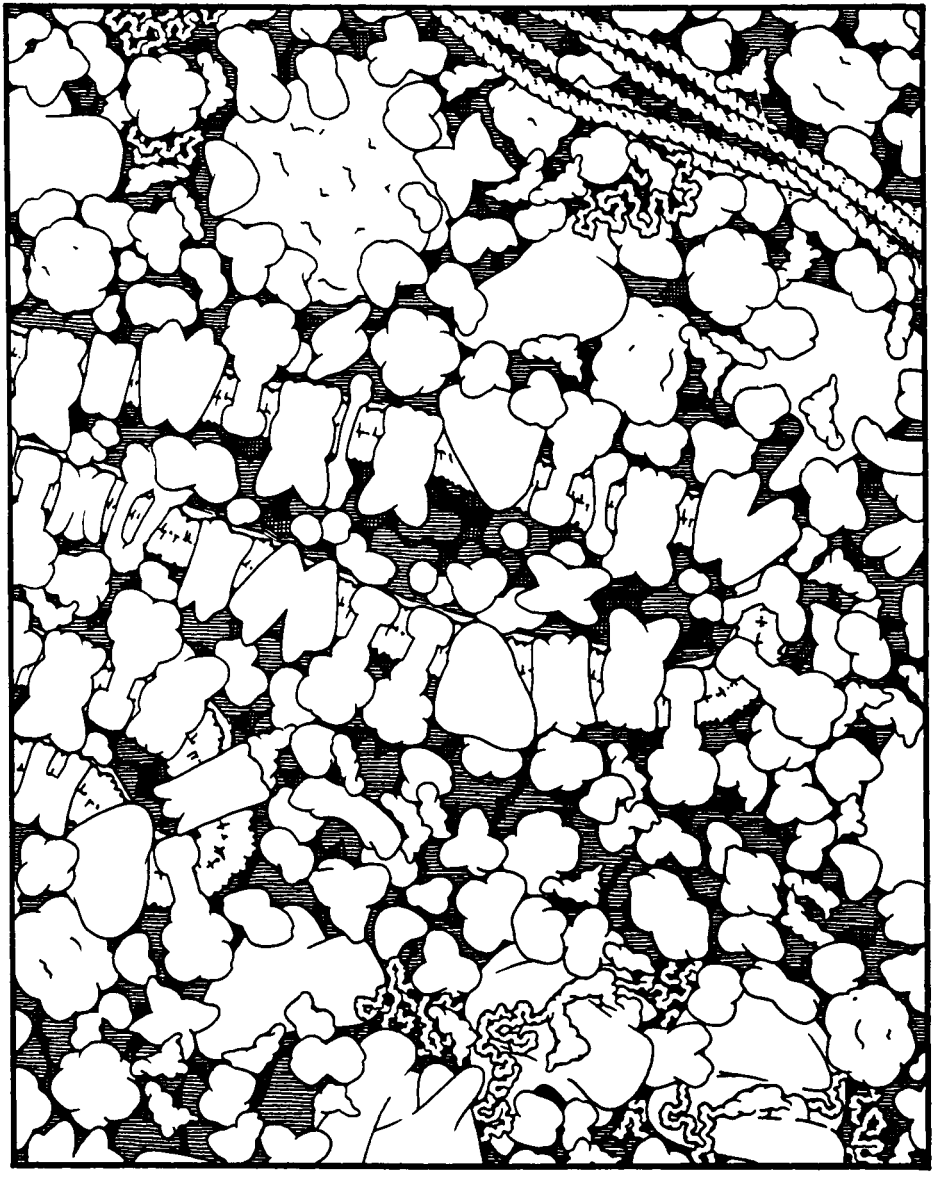


Figure 2. (continued) A section through a mitochondrion at one million times magnification. b) The mitochondrial matrix fills most of this illustration, with two invaginations of the inner membrane extending from the left side.

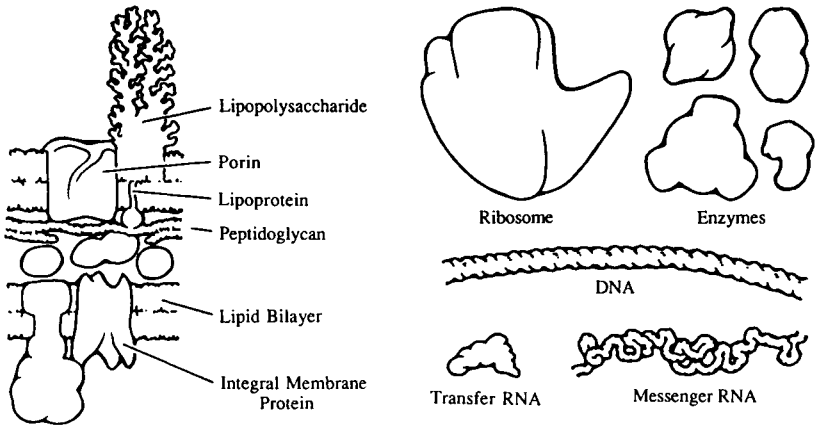


Figure 3. Key to Figures 1 and 2.

Figure 2a and comprising most of Figure 2b, is the mitochondrial matrix, containing the soluble proteins of the tricarboxylic acid cycle, other catabolic enzymes, and an entire set of protein synthesis machinery separate from that of the cytoplasm.

MOLECULAR MOTION

Of course, these illustrations only show a single instant of time. The densely packed molecules are in rapid motion, constantly colliding with neighbors, randomly diffusing in every direction. The enzymes diffuse the width of each illustration in about 2 milliseconds. The ribosomes move somewhat slower, the small molecules proportionately faster. The proteins embedded in the membranes are also in constant motion, diffusing laterally through the two-dimensional confines of the lipid bilayer.

REFERENCES

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr., Brice, M. D., Rodgers, J. R., & Kennard, O. (1977). The Protein Data Bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. 112, 535–542.

Goodsell, D. S. (1991). Inside a living cell. Trends Biochem. Sci. 16, 203–206.

Goodsell, D. S. (1992). A look inside the living cell. Am. Scientist. 80, 457–465.

Goodsell, D. S. & Olson, A. J. (1993). Soluble proteins: size, shape and function. Trends Biochem. Sci. 18, 65–68.

Neidhardt, F. C., Ingraham, J. L., & Schaechter, M. (1990). Physiology of the Bacterial Cell: A Molecular Approach. Sinauer Associates, Sunderland, Mass.

Srere, P. A. (1982). The structure of the mitochondrial inner membrane-matrix compartment. Trends Biochem. Sci. 7, 375–378.

Chapter 8

Extracellular Matrix

D. W. L. HUKINS, S. A. WESTON,
M. J. HUMPHRIES, and A. J. FREEMONT

Principles of Medical Biology, Volume 3
Cellular Organelles and the Extracellular Matrix, pages 181–232

ISBN: 1-55938-804-8

INTRODUCTION

The tissues and organs that make up the body are never composed solely of cells. Some tissues, such as brain, are densely packed with cells, but most, including skin, muscle and cartilage, contain large volumes of extracellular space filled by extracellular matrix (ECM). ECMs consist of hydrated assemblies of proteins and polysaccharides. They have mechanical functions and also influence the behavior of cells. The mechanical functions include packaging and protecting cells and tissues from external forces; in addition, some tissues, like tendons and bone, have important mechanical functions within the body. ECM can influence the behavior of cells in two ways. First, it provides a platform for cell movement and orientation; secondly, it signals to cells and thereby contributes to the regulation of tissue-specific gene expression. Understanding the structure of a tissue, the co-ordination of its functions, its creation during development, and its response to damage and disease, must therefore include an understanding of the function of its ECM.

Most ECMs contain similar classes of macromolecules. However, different tissues contain these classes of molecules in differing proportions and also contain different variants of them. Varying the composition of the ECM confers different physical properties on the tissue. At one extreme, it may be soft and highly deformable (e.g., dermal connective tissue): at the other extreme, hard and rigid (e.g., bone). In between these two extremes lie other kinds of ECM with intermediate properties; for example, cartilage is firm but more deformable than bone. At a gross level, these differences in physical structure reflect differences in the functions for which the matrices are intended. The primary functions of bone are to support the mass of the body, providing anchorage points for muscles and protection of soft tissues, such as the brain. Cartilage serves to cushion bone and promote movement of joints; it needs to be strong, deformable and to have a relatively smooth surface. Dermal connective tissue primarily serves a packing and support function for epithelial and endothelial cells and, therefore, must be extremely deformable yet elastic. The relatively open structure of matrices such as that found in the dermis also allows passage of both nutrients and cells within tissues. Finally, basement membranes, that underlie all epithelia and endothelia, are thin sheet-like structures. Their major functions are in delineating tissue boundaries and organizing the structures of tissues.

All cells produce their own ECM and so are responsible for its composition and structure. The cells which produce the greatest volume of ECM are those in the connective tissues, e.g., bone, cartilage, tendon, and so forth. They have a common precursor and so are closely related. This relationship is demonstrated by callus, formed during healing of a fracture, where bone, fibrous tissue and cartilage are all formed from the identical precursor lineage cells of the periosteum. Connective tissue cells have many similarities; they are usually physically separated from their neighbors and respond to mechanical damage to the tissue. In fibrous tissues (e.g.,

tendon) these cells are known as fibroblasts, in cartilage as chondrocytes and in bone as osteoblasts and osteocytes.

In very general terms, the ECM produced by these and other cells contains polymerized protein fibers which act as a scaffold. These fibers are decorated with accessory proteins, and together they are embedded in hydrated, polysaccharide-rich gels. The principal fibrous proteins are collagen and elastin; the ECM also contains proteoglycans (consisting of protein and polysaccharide) and glycoproteins. The hydrated gels are made up of proteoglycans and/or hyaluronic acid. The characteristic physical properties of different matrices are the result of variation in the composition and relative proportions of the particular molecules that are employed for each of these functions. In general, those tissues that are subjected to high tensile forces (such as tendon) tend to contain extensive depositions of fibrillar collagens, while tissues that have to withstand compression (such as cartilage) contain high concentrations of proteoglycans.

The propensity of ECM macromolecules to assemble is a characteristic feature of their structures. Both protein-protein and protein-carbohydrate interactions are important in the functioning of an ECM and many different kinds of binding are found, based on covalent, hydrophobic and electrostatic interactions. The association of one macromolecule with another, or with a cell surface, may not have a high affinity. However, when a complex matrix is being assembled, or when a cell interacts with an assembled matrix, the sum of the many individual interactions results in high affinity binding. Mature matrices are stabilized by covalent linkages, but initially a combination of hydrophobic and electrostatic interactions holds them together. In particular, most binding energy in protein-protein interactions involves hydrophobic bonding, while associations with polysaccharides often involve electrostatic interactions.

In addition to the ability of ECMs to provide a support or other mechanical function, they also form substrates for cell movement and provide instructional cues to help regulate tissue-specific gene expression. As described below, the molecular interactions employed by cells to bind to ECMs are reasonably well characterized. Some of these interactions appear to result in stable adhesion, others stimulate cells to move, while a few have been shown to enhance transcription of specific genes. Detailed investigations of the migratory and phenotypic consequences of individual receptor-ligand binding events are currently being carried out, but it remains to be seen how cell phenotype as a whole is regulated by a complex assembled matrix.

This chapter is divided into three major sections. First, we will consider in more detail the molecules that make up ECMs and how they mediate their functions. This section has been divided into two parts: (1) those molecules whose function is primarily structural (proteoglycans, collagens and elastin), and (2) those molecules (mainly cell adhesion molecules) that have accessory functions depending on the type of matrix. Secondly, we will consider the general functions of ECMs *in vivo*, progressing from roles in development, to changes in matrices that occur through

aging, to pathological alterations in matrix composition. Thirdly, we will describe in detail the structure, function, and contribution to health and disease of four model systems (articular cartilage, bone, skin, and the microvasculature). These have been selected because they each exemplify specific features of ECMs and permit a synthesis of many of the points described earlier in the chapter.

COMPOSITION

Water is the major component of ECM. The importance of water is illustrated by the composition of connective tissues which consist mostly of ECM because they contain few cells. For example, healthy articular cartilage consists of about 80% water, by mass, and even a fibrous tissue like tendon is 60% water. When these tissues are dehydrated, their mechanical properties change dramatically. Connective tissues then change from being flexible and resilient to become hard and brittle. It has been suggested that water acts as a "plasticizer" in the ECM (Jeronimidis and Vincent, 1984), i.e., that it separates the macromolecular components, reducing the strength of interactions between them and so lowering the energy required to deform the matrix.

Since water is an integral part of the structure of the ECM, it allows connective tissues to withstand applied pressure (Grodzinsky and Frank, 1990). Indeed, the glycosaminoglycans (see below) in ECM attract water so that it tends to swell and so exerts a "swelling pressure." Thus, the ECM possesses an internal fluid pressure or "turgor" which allows it to withstand a compressive load, in much the same way that air pressure in a tire supports the compressive load exerted by the weight of a vehicle.

When tissues are subjected to continuous compression, some water can be squeezed out of the ECM. This water loss can affect the size and mechanical properties of some connective tissues. Loss of water from the intervertebral disc during the course of a day leads an average person to shrink by about 2 cm. The water lost is replenished during the night's sleep when the compressive load on the spine is greatly reduced. Even slight water loss will have some effect on the mechanical properties of connective tissues. Since water loss takes a finite time, their mechanical properties (e.g., stiffness) depend on the rate at which they are loaded and will also be affected by repetitive loading. Fluid flow can also increase the rate at which large molecules (e.g., proteins) can diffuse through the ECM.

Water is attracted to ECM by its glycosaminoglycans (conventionally abbreviated to GAGs). All of the GAGs, except hyaluronic acid (often abbreviated to HA) are covalently bound to proteins in the ECM. Hyaluronic acid is also called "hyaluronan" and, since some of its carboxylic acid side chains will be dissociated at physiological pH values, is often referred to as "hyaluronate." Like the other GAGs it is an unbranched polymer chain consisting of repeating disaccharide units (Figure 1). There can be as many as 50,000 disaccharides in a single HA chain which means that it can be several micrometers long. When it is dissolved in water,

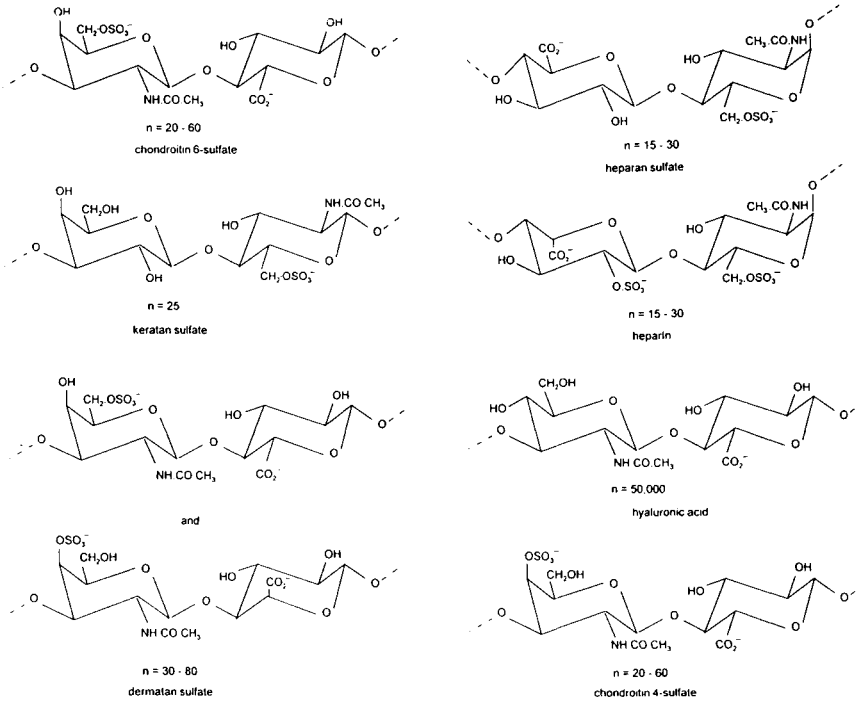


Figure 1. Glycosaminoglycan (GAG) structures. This figure shows the idealized structural formulae of the disaccharide units which repeat to form a GAG chain with the typical number, n, of disaccharides in each type of chain.

a single chain can fold up in three dimensions so that it fills a space which is 1,000–10,000 times its own volume.

The other GAGs are chondroitin sulfates, dermatan sulfate, keratan sulfate, heparin, and heparan sulfate; their idealized chemical structures are shown in Figure 1. This figure also shows the number of disaccharide residues in a typical chain for each type of GAG. In practice, a single chondroitin sulphate chain, for example, can have regions in which the repeating disaccharides correspond to chondroitin 4-sulfate and chondroitin 6-sulfate. The types of GAG which are present in the ECM depend on the tissue. Chondroitin sulfates occur in cartilage, cornea, bone, skin and arterial walls; dermatan sulfate occurs especially in skin, tendon, arterial walls and heart valves. Keratan sulfate is not so widely distributed but occurs in cornea, cartilage and the nucleus pulposus of the intervertebral disc.

Each GAG chain is covalently bonded at one end to the core protein of a proteoglycan (abbreviated to PG). At this end of the chain there is a short sequence of different carbohydrate residues; the final residue in this sequence binds to the protein through a serine residue. PGs are often likened to bottle brushes in which

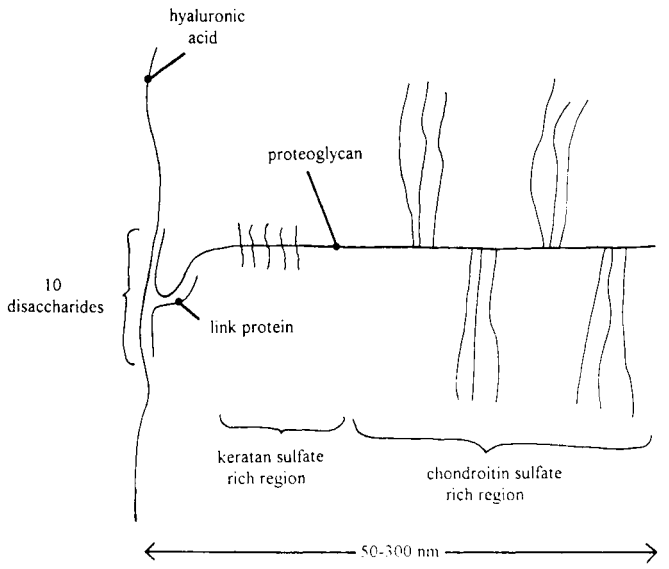


Figure 2. Cartilage proteoglycan (PG) structure and its interaction with hyaluronic acid (HA). This figure shows a schematic diagram of the PG monomer, which consists of a protein core to which chondroitin sulfate (CS) and keratan sulfate (KS) chains are attached. Monomers aggregate by interacting with 10 disaccharides in an HA chain. The interaction is stabilized by a link protein.

the GAG chains are the bristles. A core protein in a PG can contain about 2,000 amino acid residues. There can be many different core proteins in the ECM of a single tissue and each can have different numbers and types of GAG chains attached to it. Some PGs are attached to the cell membrane by their core proteins while others occur in the bulk ECM.

Cartilage PGs have been extensively studied and the larger ones can form "aggregates" in the presence of HA (Carney and Muir, 1988; Kjellan and Lindahl, 1991). In an aggregate, one end of a PG is attached (non-covalently) to ten disaccharide residues of an HA chain, as shown in Figure 2. This figure also shows the structure of an individual large PG molecule from cartilage. It has an HA-binding region at one end and a region which is rich in chondroitin sulfates at the other; there is also a region which is rich in keratan sulfate. Because of their ability to form aggregates, unaggregated PGs are sometimes called "PG monomers." A single cartilage PG can have a molecular weight as high as 2×10^6 with the result that the molecular weight of an aggregate can be 2×10^8. The volume of a hydrated PG aggregate can then be larger than that of a single bacterium. The interaction between the PG and the HA chain is stabilized by glycoproteins called "link

Table 1. Collagen Types

Type	*Structure formed*	*Examples of distribution*
I	Fibrils	Tendon, bone, skin, cornea etc.
II	Fibrils	Cartilage, nucleus pulposus, vitreous (eye), etc.
III	Fibrils	Connective tissues
IV	Mesh	Basement membranes
V	Fibrils	Tendon, bone, skin etc.
VI	Beaded fibrils periodicity 100nm	Tendon, bone, skin, cornea etc.
VII	Specialized fibrils	Skin, oral mucosa, cervix
VIII	Mesh	Descemet's membrane
IX	Attaches to type II fibrils	Cartilage, vitreous
X	Fine fibrils	Mineralizing cartilage
XI	May be incorporated into type II fibrils	Cartilage
XII	Surface of type I fibrils	Same as type I
XIII		Fetal epidermis, intestinal mucosa
XIV		
XV		
XVI		

proteins" which also bind to the two components non-covalently. Figure 2 shows the interactions which occur in a PG aggregate. Link proteins have been detected in the ECM from many different connective tissues.

Collagens are proteins which consist of three polypeptide chains; in each chain the amino acid sequence -glycine-X-Y- repeats itself along at least some of the length (Kielty et al., 1993). In this sequence, X is often proline and Y is hydroxyproline. When these repetitive "triplets" occur, the three chains wind around each other to form a coaxial triple helix. In the triple-helical regions, the three chains are packed closely together and so this part of the molecule is effectively rod-shaped. There are at least 16 different types of collagen; each type is the product of a different gene. They all appear to assemble into larger structures in the ECM. Table 1 lists the 16 types which have been identified at the time of writing, the structures which they form and the tissues in which they occur (Burgeson and Nimni, 1992).

Types I, II and III account for over 70% of the mass of collagen in the body and are sometimes called "interstitial collagens." They form fibrils which have a characteristic appearance in electron micrographs of sectioned tissues (Figure 3). These fibrils have a banded structure because of the way in which their rod-like molecules pack together (Figure 4). The only parts of the molecules which do not form a triple-helical structure are called "terminal peptides" or "telopeptides" and occur at the very ends. However, types I, II, III, V and XI, which all form fibrils, are synthesized as "procollagens" with much longer ends, which are not triple helices. The procollagens are secreted and their extended ends are cleaved by specific enzymes. The resulting molecules are sometimes called "tropocollagens" and, because of their rod-like shape, can pack together to form fibrils. There is no reason to suppose that a fibril will contain only one type of collagen; type V collagen

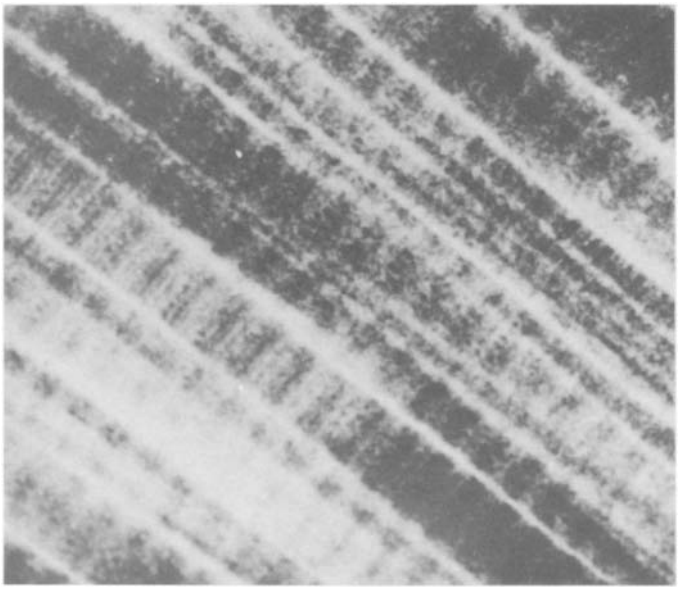


Figure 3. Collagen fibrils viewed in the electron microscope. This transmission electron micrograph was obtained from a section of stained tissue at an original magnification of about 30,000. The fibrils have a characteristic banded appearance (with a periodicity of 64 nm) and a diameter of about 30 nm.

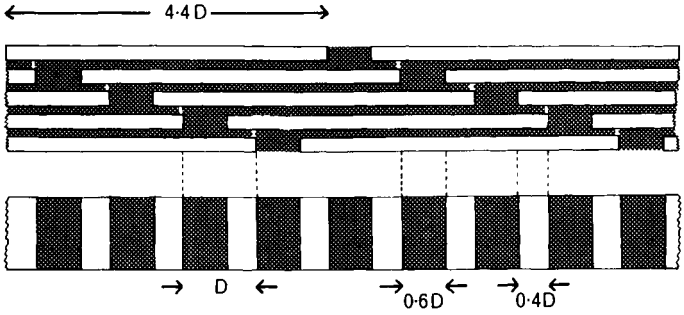


Figure 4. Model for the packing of collagen molecules within a fibril. The fibril has a periodicity of D = 67 nm; electron microscopy yields a value of D = 64 nm because dehydration shrinks the fibril. Each collagen molecule with a length of 4.4D and a diameter of about 1.2 nm is represented (top) by a rod. Each rod is staggered, with respect to its nearest neighbor, by a distance D. The head of one molecule is covalently linked to the tail of its neighbor. Five molecules then define the periodic structure shown, with a "gap" region (length 0.6D) and an "overlap" region (length 0.4D). The gap region can take up the stains used in electron microscopy but the overlap region tends to exclude it. Thus the schematic electron micrograph of a collagen fibril (below) has a periodicity of D and consists of alternating dark (stained) and light (unstained) bands. Note that this diagram indicates the axial relationship between molecules but does not show how they pack together in three dimensions.

occurs in type I fibrils where its function may be to limit fibril diameter (Birk et al., 1990). Fibrils are stabilized by covalent "cross-links" between the molecules. In types I and II collagen, these cross-links are chemically distinct from those which form in most other proteins in that they do not involve a bond between the sulphur atoms of neighboring cysteine residues ("disulphide bridges"); further details of the chemistry of these cross-links appear in the Aging Section. The cross-links join the head of one molecule to the tail of a neighbor, so that the whole fibril is held together by covalent bonds. In some connective tissues (e.g., tendon) the fibrils aggregate into larger fibrous bundles but in others (e.g., articular cartilage) individual fibrils are surrounded by a "ground substance" which consists largely of hydrated PGs.

The function of collagen fibrils is to provide tensile reinforcing for the ECM, i.e. the fibrils resist forces which tend to pull tissues apart (Hukins and Aspden, 1985). ECM needs to be reinforced by fibers because the hydrated PGs form a gel-like phase which would simply slide sideways when compressed, so that they could not then exploit their swelling pressure to support a compressive load. Furthermore, hydrated PGs on their own could not withstand the tensile stresses to which ECM is subjected in many tissues. Collagen fibrils can perform their function because their cross-linked structures are stiff and strong when they are pulled in the axial direction, i.e., when they are stretched. In a tissue like tendon, the fibrils are aligned roughly parallel. Tendon can then withstand the high axial tensions to which they are subjected because their collagen fibrils are aligned in this axial direction. If these fibrils are axially compressed, they simply buckle. Similarly, they can only reinforce tissues which bend or twist if the fibrils are oriented so that they tend to be stretched by these movements. Thus, collagen fibrils are oriented in the directions in which a tissue has to withstand tension. Interactions between collagen and other components of the ECM may be involved in the transfer of mechanical stress to the fibrils and, hence, influence the mechanical properties of the tissue (Hukins and Aspden, 1985).

Collagen fibrils also prevent damaged tissues from failing suddenly. For example, ligaments do not tear straight across when they are damaged. Instead small tears in the ECM are diverted when they encounter the strong collagen fibrils, as shown in Figure 5. The damaged ligament can then heal while, in the meantime, retaining the ability to withstand some load. Of course, the damaged ligament would fail completely if this load were excessive.

Not all collagen types form fibrils. Types IV and VII have triple-helical regions which are interrupted as a result of irregularities in their otherwise regularly repetitive sequences of amino acid triplets. Type IV collagen occurs only in basement membranes where its molecules aggregate into a mesh-like structure which is stabilized by cross-links involving disulphide bridges, as well as by what are presumed to be more the more usual collagen cross-links. Type VII collagen forms specialized fibrils which are thought to be stabilized by disulfide bridges. Its

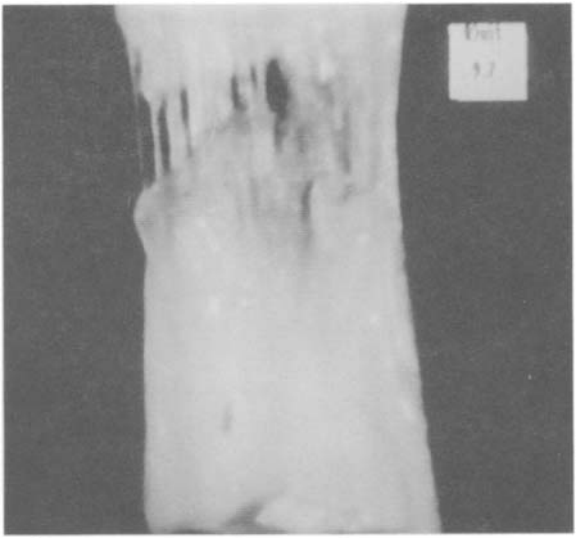


Figure 5. Torn ligament. The collagen fibrils divert the tear so that it does not propagate directly across the ligament. As a result, the damaged ligament does not fail completely.

function is to anchor basement membranes and epithelial cells to stroma in a variety of tissues. Its absence is associated with some forms of epidermolysis bullosa in which the skin becomes blistered as a result of separation of the dermal stroma from the epithelium (see below). Molecules of types VI, VIII, IX, X, and XIV collagens contain shorter stretches of triple helix separated by stretches of less regular amino acid sequence. The similarity between types IX, XII, XIV, and XVI collagens leads to their being collectively termed FACIT (fibril associated collagens with interrupted triple-helix) collagens. Type X collagen occurs in the growth plate of long bones and is believed to be involved in matrix mineralization. Types IX and XI occur in the predominantly type II collagen fibrils of hyaline cartilage (Brewton and Mayne, 1993).

Elastic fibers also occur in ECM (Winlove and Parker, 1990). They are much less stiff than collagen fibrils so that tensile stress can be transferred from elastic fibers to the stronger collagen as a tissue is stretched; this allows the rubber-like properties of elastic fibers to be exploited but enables collagen to protect them at high strains. Elastic fibers consist of elastin, which is a highly cross-linked protein, and glycoprotein fibers.

There are many other proteins in the ECM (Heinegard and Oldberg, 1989). Some of them are restricted to specific tissues and those of cartilage and bone have been extensively investigated. Osteocalcin is the most extensively studied bone protein in which some of the glutamic acid residues are modified to γ-carboxyglutamic acid

(GLA); it is sometimes called "GLA protein." The glycoproteins which are involved in cell-matrix interactions (e.g., fibronectin) are considered in the next section. Phosphoproteins are believed to control the deposition of calcium phosphates in calcified tissues and at sites of ectopic calcification. For example, the phosphoprotein osteonectin (also called SPARC or BM-40) occurs in bone. However, much more information about the ways in which phosphoproteins can influence mineral deposition and calcium transport can be obtained from the caseins (from milk), phosvitin (from eggs) and salivary phosphoproteins which can be readily extracted in much greater quantities (Holt and van Kemenade, 1989).

Molecules which affect the functions of cells must also diffuse through the ECM and may lead to cellular activity which changes its composition (Darnell et al., 1990). These mobile molecules include hormones, cytokines and growth factors. Interleukins provide an example of cytokines; there are several different types which are often denoted by IL-1, IL-6, and so forth. Growth factors have been discovered relatively recently in ECM; they include insulin-like growth factors (e.g., IGF-1) and transforming growth factors (e.g., TGF-β). All of these molecules can influence the cells to change the composition of the matrix in normal physiological activity or in pathology.

Although there are several enzymes which are able to degrade collagen, no single enzyme acting alone is able to degrade collagen in cross-linked fibrils (Weiss, 1984). Elastase is able to cleave the terminal peptides, which are the cross-linking sites, from type I collagen and so free the triple-helical molecule from the fibril. Collagenases have been identified which can then degrade the free triple-helical molecules of types I, II and III collagen. These collagenases are secreted in an inactive form and have to be activated by other enzymes.

Finally, the physical properties of ECM can be modified by mineral deposition. The minerals which are most commonly deposited are calcium phosphates (Hukins, 1989). For example, deposition of a poorly crystalline and impure form of the mineral hydroxyapatite ($Ca_5[PO_4]_3OH$) is responsible for the hardness of bone. In general, calcified matrix is a composite material whose properties are intermediate between those of soft tissues (which are soft and flexible, like most polymeric materials) and the mineral phase (which is hard and brittle, like most ceramics).

CELL-MATRIX INTERACTIONS

As already described, ECM has many functions which include provision of structural support and mediation of cell adhesion. Most eukaryotic cells have the ability to recognize and react functionally with adhesive ECMs (Akiyama et al., 1990; Hay, 1991; Mosher et al., 1992). This situation is true not only for actively migrating cells that use adhesive contact for traction and guidance, but also for stationary cells that require a platform for support and orientation. Cell-matrix adhesion has a homeostatic function in promoting tissue regeneration during wound healing, while aberrant adhesion contributes to the aetiology and pathogenesis of a number of

major human diseases including arthritis, cardiovascular disease and cancer (Humphries, 1990). Some of these are described in detail in subsequent sections. The structural complexity of ECM has hampered progress towards an understanding of adhesive recognition events at a cellular level. However, there has been progress in identifying the macromolecular interactions involved.

Cell Binding Sites in Adhesion Molecules

Cell-matrix interactions involve cell surface proteins and the components of the ECM. These interactions have been studied by disassembling the system into its components to determine the function of each. The most abundant components of ECM have now been identified, purified and characterized; those which are specifically concerned with cell adhesion are listed in Table 2; functions of fibronectin, integrins and laminin will be described in this chapter. Detailed biochemical dissection of these components has been aimed at elucidating the molecular basis of adhesive activity. Initially, proteolytic and chemical cleavage were used to identify major cell adhesive domains in fibronectin and fibrinogen. In

Table 2. Integrins and the Molecules to which They Bind

Subunits		*Molecules bound*
β1	α1	CO, LN
	α2	CO, LN
	α3	CO, FN, KN, LN
	α4	FN, VCAM
	α5	FN
	α6	FN
	α7	LN
	α8	LN
	αV	CO, FG, FN, VWF, VN
β2	αL	ICAM-1, ICAM-2
	αM	C3bi, Factor X, FG, ICAM-1
	αX	C3bi
β3	αIIb	BSP, FG, FN, LN, TSP, VN, VWF
	αV	FG, FN, TSP, VN, VWF
β4	α6	LN?
β5	αV	FG, FN, VN
β6	αV	FN
β7 (=βp?)	α4	MC
	α1EL	?
β8	αV	?

Note: *Key*: Abbreviations are: BSP—bone sialoprotein; C3bi—complement component; CO—collagens; FG—fibrinogen; FN—fibronectin; ICAM—intercellular adhesion molecule; KN—kalinen; LN—laminin; MC—mucosal addressin cellular adhesion molecule-1; TSP—thrombospondin; VCAM-1—vascular cell adhesion molecule-1; VN—vitronectin; VWF—von Willebrand factor.

fibronectin there appears to be two adhesive domains, while in fibrinogen the adhesive domain is located in the carboxy-terminal region of one of its protein chains. In both cases, the active sites within these domains have been localized further through the use of smaller protein fragments and synthetic peptides. For example, in the central cell-binding domain of fibronectin, adhesive activity has been assigned to a tetrapeptide with the sequence arginine-glycine-aspartate-serine (RGDS) (Pierschbacher and Ruoslahti, 1984a; Yamada and Kennedy, 1984). Synthetic peptides containing this sequence exhibit anti-adhesive activity, as seen by their ability to inhibit the attachment and spreading of cells on fibronectin *in vitro*, while analogs containing minor changes in the first three residues, such as an aspartate to glutamate substitution at position three, are inactive (Pierschbacher and Ruoslahti, 1984b; Yamada and Kennedy, 1985; Ruoslahti and Pierschbacher, 1987). The presence of this key RGD tripeptide is not restricted to fibronectin, indeed it has been found in almost all adhesive extracellular macromolecules.

Adhesive peptides may have potential applications, not only for understanding the molecular basis of cell-ECM interactions but also as potential anti-adhesive drugs. However, synthetic mimics of these sequences have a relatively low biological activity when compared with the molecules from which they are derived. This suggests that the current generation of anti-adhesive peptides bind with reduced affinity to cellular receptors, because of their relatively unrestricted structures in solution. There are a number of possible explanations which could account for this phenomenon. For example, binding could depend on three-dimensional molecular structure. Alternatively, adhesive proteins may possess additional regions that contribute to the affinity and specificity of receptor binding (Humphries, 1990).

Receptor Identification

The recognition of extracellular molecules by cell surface receptors is the principal mechanism used by cells to sense their environment. Cell surface receptors act as a bridge which links extracellular with intracellular environments. Whether the receptor interacts with soluble factors (e.g., growth factors and cytokines), or cell surface adhesion molecules, or immobilized ECM, its role is that of a bridge. Some receptors also act as signal transducers, thus permitting cells to react appropriately to changes in their extracellular environment. These signals make a major contribution to the regulation of cellular phenotype and are discussed in greater detail below.

Once cell adhesion sites were identified within matrix molecules, the search began for the cell surface receptors with which they interacted. For the purposes of this chapter, discussion will essentially be limited to a family of glycoproteins known as the integrins. This choice is made because integrins appear to be the major receptors used by cells to bind to ECM, although other types of molecules are involved in cell adhesion. Examples include the cadherins, selectins, PGs and

immunoglobulins. The integrins are distinguished not only by their ability to bind to both ECM and cell surface ligands but also by the diversity of their ligand binding (Humphries, 1990; Tuckwell et al., 1992).

Two different experimental approaches have identified integrins as receptors. First, binding molecules have been isolated from cell extracts by affinity chromatography on ligands or their fragments. Secondly, integrins have been identified as antigens recognized by monoclonal antibodies that perturb ligand-mediated cell adhesion (Humphries et al., 1993). As a result of these two approaches, a complex profile of integrin-ligand interactions has been established. Rather than each ligand having its own specific integrin receptor, current data indicate that a ligand may have multiple receptors. The converse is also true in that an integrin may bind to many different ligands.

A single integrin molecule consists of an α subunit and a β subunit. Fourteen different α subunits (α1–8, αM, αL, αX, αV, αIIb, αIEL) and eight β subunits (β1–8) have been identified in vertebrates, although only twenty different dimer combinations are known to exist. (For reviews, see Akiyawa et al., 1990; Hemler, 1990; Humphries, 1990; Hynes, 1992). Association of the α and β subunits requires divalent cations (e.g. Ca^{2+} or Mg^{2+}). Electron microscopy of isolated integrins (using the technique of rotary shadowing) shows that they have a globular head, which binds ligand, and two thin tail regions which form a stalk (Carrell et al., 1985; Nermut et al., 1988; Kelly et al., 1989; Weisel et al., 1992). A model of integrin structure is depicted in Figure 6.

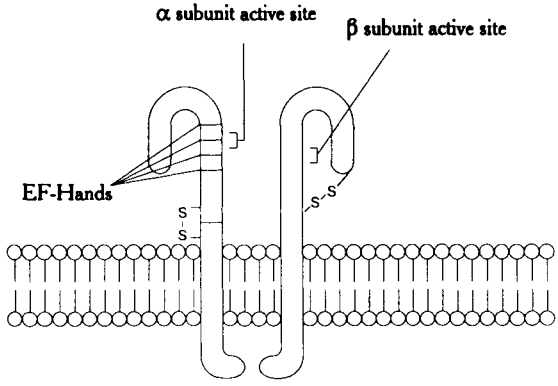


Figure 6. Schematic representation of the structure of an integrin receptor. Integrins appear to have a globular head and two stalk-like tails, with each subunit contributing one tail and part of the head to the overall structure. Both subunits have a large extracellular domain, a hydrophobic transmembrane region and a short cytoplasmic domain. The cation binding sites (EF-hands) and the intrachain disulfide bonds are depicted. The characterized active site regions are shown for both subunits, although their spatial relationship remains to be determined.

Although the most convenient method of grouping integrins is on the basis of a common β subunit, it is now clear that certain α subunits may complex with more than one β subunit, thereby complicating the classification. As a general rule, there are three main integrin classes, β1, β2 and β3 which combine with an α subunit so that the resulting dimers exhibit different ligand-binding specificities. The β1 and β3 subfamilies predominantly mediate cell-matrix adhesion, while the β2 class are cell–cell adhesion receptors. There is an additional loose distinction between β1 and β3 integrins in that β1 integrins are generally involved in adhesion to connective tissue macromolecules such as fibronectin, laminin and collagens, while the β3 receptors bind to vascular ligands such as fibrinogen, von Willebrand factor, thrombospondin and vitronectin. In terms of cellular distribution, β1 and β3 integrins have a widespread occurrence and are co-expressed on most cell types while, in contrast, β2 integrins are restricted to leukocytes.

As already noted, there are two general features of ligand-binding to integrins. First, certain integrins are able to bind different ligands that superficially have dissimilar structures. For example, α2β1 binds to collagen, fibronectin and laminin, and αVβ3 binds to bone sialoprotein, fibrinogen, fibronectin, laminin, thrombospondin and von Willebrand factor. While the molecular basis for these interactions is not yet fully understood, it is possible that each protein may have a shared structural feature, such as an adhesion recognition sequence, which represents a common binding signal. Second, the same ligand may have multiple integrin receptors. For example, fibronectin has α5β1 as the principal receptor for one of its cell-binding domains and α4β1 for the other, but there is also evidence implicating α2β1, α3β1, αVβ1, αVβ3, αIIbβ3 and αVβ5 as fibronectin receptors.

Regulation of Integrin Activity

Electron microscopy shows that the head of the integrin molecule lies on the cell surface while the stalk penetrates the cell membrane (Figure 6). Some modification of this model may be required to take account of possible contributions of the stalk region to ligand binding and also the role of additional components which may be necessary for integrin function. Integrin-ligand binding is also dependent on the presence of divalent cations, typically Ca^{2+} or Mg^{2+}.

Several mechanisms can be invoked to explain how integrins might transmit extracellular information across the plasma membrane. These could be independent of shape change and might include mechanisms that rely on clustering as a means of varying cytoplasmic presentation, or on changes in the binding and orientation of receptor-associated molecules. Alternatively, changes in receptor conformation could be used to affect signalling molecules or the cytoskeleton. Although there is evidence to suggest that each of these mechanisms makes a partial contribution to integrin signalling, the latter mechanism is currently considered most important. A simple prediction implied by each of these mechanisms is that integrins can exist

in a state in which they are transmitting signals, that is to say an active or open state, where they are able to bind ligand, and at least one alternative state in which signalling is either less active or is shut off completely. Signalling via integrins takes two forms. The first is regulation of the affinity and conformation of the receptor from inside the cell, so-called "inside-to-out signalling"; the second is triggering of intracellular events by ligand occupation of the receptors, so-called "outside-to-in" signalling (Humphries et al., 1993).

In the case of inside-to-out signalling, there are examples of intracellular events which increase or decrease affinity of various integrins for their ligands. These can result in either an active or inactive integrin-receptor conformation. For example, activation of T cells by antigen leads to activation of β1 integrins on lymphocytes and so to increased adhesion of these activated cells to collagen, fibronectin and laminin. In contrast, there are a number of cases where integrins lose their activity even though they persist on the cell surface. This loss of activity occurs during development in α5β1 of keratinocytes (Adams and Watt, 1990) and in α6β1 of retinal neurons (Neugebauer and Reichardt, 1991). In addition, phosphorylation events have been associated with apparent inactivation of certain β1 integrins (Tapley et al., 1989; Horvath et al., 1990).

There is evidence that integrins mediate information transfer in to cells (outside-to-in signalling). As already discussed, integrins are capable of interacting with their ligand when they are in their active state. This interaction can trigger a number of intracellular events. These include tyrosine phosphorylation in platelets, fibroblasts and carcinoma cells, cytoplasmic alkalinization in fibroblasts, endothelial cells and platelets, activation of lymphocytes and activation of secretion in synovial fibroblasts, monocytes and neutrophils and finally differentiation in myoblasts and keratinocytes (Hynes, 1992). From this discussion, it is evident that there is considerable overlap between the pathways which operate in inside-to-out and outside-to-in signalling.

Signal Transduction and Phenotypic Effects

The cytoplasmic domain of an integrin molecule (i.e., the part which penetrates into the cell) has the potential for transducing integrin-ligand binding to intracellular events, e.g. cell migration or integrin-mediated changes in gene expression.

The cytoplasmic domain of the β subunit is required for the localization of β1 integrins to focal contacts and for ligand binding (Hayashi et al., 1990). A region close to the transmembrane domain appears to be important in this respect, at least in the case of β1 integrin localization to focal contacts on fibronectin (Marcarrtonio et al., 1990; Reszka et al., 1992). These domains may also be important for other integrin subclasses. In addition, this domain also appears to be involved in integrin activation since expression of a mutant β3 subunit lacking almost all of its cytoplasmic domain generated a constitutively inactive form of αIIbβ3 (O'Toole

et al., 1991). It has been shown to interact with the actin cytoskeleton (Horwitz et al., 1986; Larjava et al., 1990; Otey et al., 1990). The role of the β subunit cytoplasmic domain in integrin-ligand binding suggests that it may be involved in the transduction of signals between intracellular pathways and extracellular domains of the integrin (inside-to-out signalling).

The role of the cytoplasmic domain of the α subunit is less clear, but it may be involved in the regulation of the association of the β subunit cytoplasmic domain with the cytoskeleton. In general, the cytoplasmic domains appear to be involved in controlling the state of activation of the integrin and transmitting the ligand-binding event to intracellular components.

The communication between integrin-ligand binding and the variety of cellular responses which ensue suggests that integrins may be associated with second-messenger signalling pathways. A number of second messengers used by receptors to transmit signals to intracellular effectors have now been characterized. In general, receptor-ligand binding can lead to the activation of GTP-binding and -hydrolyzing proteins (G proteins) which can in turn activate, or inhibit, adenylyl cyclase. The cyclic AMP produced in the case of activation can interact with cyclic AMP-dependent protein kinase, leading to protein phosphorylation. G protein activation can also stimulate phospholipase C with the production of inositol phosphates and diacyl glycerol, potentially leading to an increase in intracellular calcium, and the production of active lipid metabolites or protein kinase C activation, respectively. Receptor activation can also lead to its own phosphorylation or to phosphorylation of other proteins by tyrosine kinases, thus potentially affecting receptor function. Integrins appear to be able to mobilize most, if not all, of these second messenger systems. In addition to these second messengers, integrins are also able to produce a specific intracellular pH rise of 0.1–0.2 pH units on ligand binding (Schwartz et al., 1991). This may be significant since intracellular pH can act as a regulator of cell growth and differentiation.

A consistent feature of cell contact sites is their high content of proteins containing tyrosine phosphate which suggests a prominent role for tyrosine kinases in cytoskeletal assembly. Recently, a direct link between integrins and tyrosine phosphorylation has been made. The phosphorylation of specific proteins including focal adhesion kinase (FAK) has been shown to occur in response to ligand binding (Guan et al., 1991; Kornberg et al., 1991; Guan and Shalloway, 1992). As yet, no direct association between integrins and FAK has been demonstrated, and it is possible that there is a tyrosine kinase cascade waiting to be elucidated. There is a strong correlation between tyrosine phosphorylation of FAK and activation of its kinase activity (Kornberg et al., 1991; Lipfert et al., 1992), suggesting that other molecules will become phosphorylated as a result. Other minor phosphotyrosine-containing proteins have also been detected following integrin cross-linking. Since integrin cytoplasmic domains are themselves phosphorylated on tyrosine and/or serine in response to various agonists, integrin phosphorylation could be a possible

mechanism for regulation of receptor function. Clearly, if the diversity of signalling-associated molecules already identified is any indication, establishing the components of a putative tyrosine cascade, defining the consequences of integrin occupancy on cytoskeletal architecture, and assigning functions to integrin-associated molecules will be a long and complicated process.

DEVELOPMENT

Tooth eruption illustrates the extent to which ECM is implicated in developmental changes (Gorski et al., 1988). The crown of a developing tooth has to push its way through bone to reach its functional position before further growth and root development occur. This process is called "eruption" and is a major event in the development of dentition. Recent observations have been made on dogs because they are expected to resemble humans in the way in which their teeth erupt. In the dog, eruption of the premolars begins in the sixteenth week after birth. Bone is absorbed above the crown of the tooth to form a pathway for it to erupt. Bone formation and root growth below the crown are complete after 23 weeks.

Eruption appears to be controlled by the dental follicle. The follicle is a thin layer of vascular connective tissue which covers the developing crown. Surgical removal of the follicle (but not making incisions in it) prevents the eruption process. Between the ages of 12–20 weeks, changes occur in the composition and structure of the ECM of the follicle which appear to correlate with specific physiological events. Although most of the collagen remains as type I during this period, the proportion of collagen (expressed as a fraction of the dry mass of tissue) increases by 250%. At the same time, the PG content increases by 45%. Electron micrographs indicate that the collagen fibrils have random orientations in the follicle before eruption; they also have very similar diameters with a mean value of 36 nm. By the age of 20 weeks the fibrils are aligned into bundles and their individual diameters have increased to a mean value of 51 nm. In addition, bundles of smaller diameter fibrils (not necessarily collagen) appear within the bundles of larger fibrils. These observations do not immediately suggest a mechanism for the eruption process. But they do show that changes within the ECM appear to be correlated with an important macroscopic developmental change.

In general, ECM is important in development for two reasons: (1) it forms the bulk of the connective tissues which form many of the obvious structures within the body, and (2) it influences how tissues develop. In many connective tissues the ECM forms well defined structures which enable the tissue to perform its function. These structures include ligaments which have relatively few cells and a high proportion of collagen which is organized into fibers. Another example is provided by the intervertebral disc which will be used to illustrate the development of these structures in the next two paragraphs.

Intervertebral discs are roughly cylindrical structures composed of connective tissue with few cells, i.e., they consist of a very high proportion of ECM (Hukins,

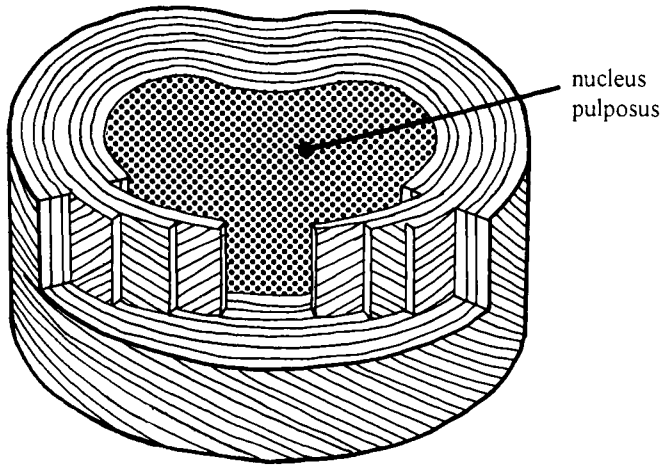


Figure 7. Structure of the intervertebral disc. The nucleus pulposus is surrounded by the lamellae of the annulus fibrosus. Collagen fibers in a single lamella are parallel and tilted with respect to the axis of the spine; the direction of tilt alternates in successive lamellae.

1988). Their superior and inferior faces, which are roughly circular, are coated with a thin layer of hyaline cartilage which separates the disc from its neighboring vertebrae in the spine. The structure of the disc is conventionally described as consisting of an inner gel-like region (the "nucleus pulposus" or "nucleus") surrounded by the tough, leathery lamellae of the outer region (the "annulus fibrosus" or "annulus"). In practice, it is not easy to distinguish a clear boundary between these regions—especially in older people. Collagen fibrils in the annulus are aligned to form coarse fibers which resemble tendons. In a single lamella, these fibers are tilted (by about 65°) with respect to the axis of the spine. The direction of tilt alternates in successive lamellae as shown in Figure 7.

The development of the disc has been described by Taylor and Twomey (1988). In published studies of human disc development, the age of the embryo or fetus is expressed in terms of its crown-to-rump length (CRL); at 3 weeks gestation the CRL is 2 mm, increasing to 10–12 mm at 6 weeks and 80 mm at 11 weeks. The precursor of the spine begins to appear when the CRL reaches 2 mm (3 weeks gestation). It consists of a mesenchymal column surrounding the notochord. When the CRL reaches 5 mm, this column becomes segmented into bands which will later develop into the discs and the vertebral bodies, as shown schematically in Figure 8. The bands which will become the disc (including the end-plates) are called the "perichordal discs." The other bands start to become cartilage at a CRL of 10–12 mm and more closely resemble bone at 40–50 mm. The division between nucleus and annulus starts to appear at a CRL of 20–40 mm. In the outer part of the

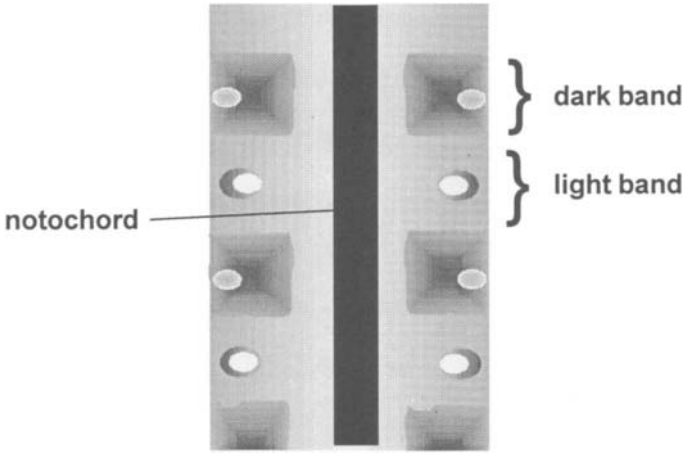


Figure 8. Schematic diagram showing segmentation of the mesenchymal column, during development of the spine, to form the alternating column of discs and vertebrae.

perichordal disc, the fibroblasts lie in distinct layers. Collagen fibers appear at about the same time. There are sufficient fibers to measure their orientations accurately (by x-ray diffraction) at a CRL of less than 80 mm; by then they form the same tilted pattern that is observed in the adult disc. It seems likely that they are laid down in this pattern, which is dictated by the arrangement of the fibroblasts, when they are first formed.

ECM influences how tissues develop in two ways: (1) it influences the behavior of cells, and (2) it is used by the cells to implement the changes which are required during development (Bard, 1990). It can influence the behavior of cells by providing a signal for cells to differentiate and by affecting their ability to migrate. There are several examples of ways in which ECM can initiate differentiation. The appearance of type II collagen in the developing chick head is correlated with the differentiation of neural crest cells into chondrocytes. However, this is not a general rule because neural crest cells become fibroblasts in the corneal stroma, which also contains type II collagen. Another example is provided by the ability of bone matrix to transform muscle cells into chondrocytes (cartilage cells). The composition of ECM may influence the ability of cells to migrate in differentiation. Cells can move more rapidly through a collagen gel if HA is added to it but more slowly if chondroitin sulfate is added. The orientations of collagen fibrils also appear to guide the migration of neural crest cells during the development of cornea. The adhesion of cells to ECM also influences their ability to migrate; this subject is covered in a later section.

One way in which cells use ECM to implement changes is to use it to generate space (Bard, 1990). Both HA and PGs are suited to this function because of the

large volume occupied by a single molecule (Wight et al., 1992). The development of the anterior chamber of the eye (between the cornea and the lens) provides an example of this function. Neural crest cells colonize the inner surface of the cornea. They then secrete PGs which lead the matrix between cornea and the lens to swell. The pressure exerted by this swelling matrix forces the cornea away from the lens and presumably helps it to adopt its curved shape.

Developmental changes can also occur in the components of the ECM and their relative proportions. Elastic fibers provide a clear example of such a change (Winlove and Parker, 1990). Fetal elastic fibers consist mainly of glycoprotein fibers but the proportion of the protein elastin increases during maturation. Some examples of changes in the proportion of the components are described above. Changes also occur in the PGs of human cartilage during development and maturation; a decrease occurs in both the total GAG content and the ratio of chondroitin 4-sulfate to chondroitin 6-sulfate (Carney and Muir, 1988). This result is explained by the cartilage containing two types of PG whose relative abundance changes.

AGING CHANGES

Changes occur in the composition and structure of ECM during aging. The consequence of these changes can be clearly seen in connective tissues where there is a high proportion of ECM. For example, skin becomes less compliant and wrinkles begin to form, i.e., gross changes are apparent in its structure and mechanical properties. It has been shown that the time taken for skin to return to its normal position, after being pinched out of shape, can increase from 1 second in a young adult to as long as 20 seconds in an elderly person. These changes are presumed to result from changes in the composition of the ECM. The time for skin to recover from pinching is correlated with the water content of its elastin. In some tissues pigments may be deposited in the ECM; for example, colored patches ("liver spots") appear in skin and hyaline cartilage may become yellowish during aging. More commonly, changes occur in the proportions or chemical properties of the normal matrix components. These changes influence the mechanical properties and, hence, many of the biological functions of connective tissues.

It is often difficult to distinguish aging changes from the consequences of degeneration. Indeed, one standard definition of "degeneration" as "a change to a functionally less active form" does not exclude the possibility that the functional change may be a natural consequence of increased age. Thus, many definitions do not really distinguish the two concepts. Normal aging changes might be expected to occur in all people, although not necessarily at the same rate. One set of criteria to distinguish aging changes is that they should be "universal, intrinsic, deleterious and progressive." Degenerative changes could then be considered as the consequences of wear and tear which were more common in some people than others. Such degenerative changes might include the appearance of microscopic fissures

on the surface of articular cartilage and formation of callus around healed trabeculae in cancellous bone. Of course, these changes will be more common with increasing age. Fissures in cartilage are also very common (perhaps almost "universal"), certainly "deleterious" and are "progressive". In practice, it may then be difficult to distinguish such degenerative changes from the effects of aging. The criterion that an aging change should be "universal" is not as reliable as it might appear. Changes become apparent at different ages in different individuals; indeed, the distinction is often drawn between "biological" and "chronological" age to indicate that some people age more rapidly than others. It is then possible for an individual to die before an aging change could be detected. The only remaining distinguishing feature is that aging changes are supposed to be "intrinsic," whereas degenerative changes are considered to be caused by wear and tear, i.e., they are extrinsic. However, a change can only be proved to be extrinsic if its mechanism is known. Since the cause of cartilage fibrillation is not always known, the presumption that it is a result of degeneration is not proven. In general, if a degenerative change is very common it may be difficult to distinguish it from aging. Conversely, an aging change which occurs late in life and which progresses slowly could easily be considered to be degenerative because many individuals will die before the change manifests itself.

Since aging changes do not confer any functional advantages, it may be pointless to attempt to distinguish them from degeneration. This is true even for aging changes which can be demonstrated to result from expression of a specific gene. Such a gene would be late-acting so that its consequences might not be disadvantageous during the reproductive years. It would then not be discriminated against by natural selection, despite the eventual disadvantage in later years. Similarly, it can be difficult to distinguish the consequences of aging from age-related pathology. Some pathological processes, like osteoarthritis, are more common with increased age. Sometimes the increased frequency of pathological changes in the elderly can be the result of a long time period between the cause and the effect; an example is provided by Paget's disease which becomes evident long after the initial viral infection.

The water content of the ECM appears to decrease in many tissues during aging. A good example of water loss is provided by the nucleus of the intervertebral disc (McDevitt, 1988). At birth about 88% of the mass of the nucleus consists of water. In young adults the water content has decreased to about 80% and it can be as low as 65% by the age of 77 years. This loss of water leads to a remarkable change in the properties of the nucleus. In children and very young adults it has the appearance of a gel whose texture is very different from that of the annulus. In older people the nucleus has a fibrous texture and the two regions of the disc are much more difficult to distinguish.

Differences in the water content of the nucleus and the annulus usually enable them to be distinguished by magnetic resonance imaging (MRI) (Hukins et al.,

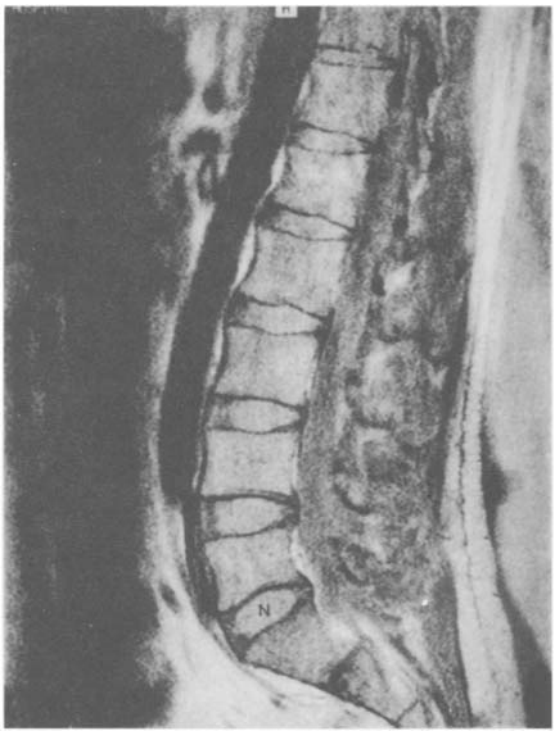


Figure 9. Magnetic resonance (MR) image of the lumbar spine. The nucleus pulposus (N) of each disc appears bright and is surrounded by the darker annulus fibrosus.

1988). For detecting degeneration, the image is obtained by a technique in which the nucleus pulposus of the normal disc appears bright and the annulus fibrosus appears black, as shown in Figure 9. However, this distinction is less clear in older people because of the loss of water from the nucleus associated with aging. Degenerate discs can be distinguished in MR scans by the complete loss of signal from the nucleus pulposus whose water content is comparable to that of the annulus. In the degenerate disc both regions then appear equally black. However, a degenerate disc detected by this technique is not necessarily a source of back pain; many asymptomatic people have degenerate intervertebral discs.

Many conclusions which are presented in the literature concerning aging changes, like loss of water, are subjective and unreliable because they are not supported by a proper statistical analysis. It is sometimes claimed that the annulus of the intervertebral disc also loses water but that the water loss is not so great as from the nucleus. Many of the data which have been published to support this claim do not appear to indicate a statistically significant loss, i.e., there is no evidence

that the annulus loses water during aging. In general, studies which purport to show changes in composition of the ECM during aging (or as a result of pathology) are often flawed because a small change in the proportion of a component has not been proved to be statistically significant.

Changes also occur in the proportions of the macromolecular components in the ECM of connective tissues. Aging changes in the PG content of cartilage are less marked than those which occur during maturation (Carney and Muir, 1988). However, there is a decrease in the ratio of keratan sulfate to chondroitin sulfate. This change is believed to be a result of cleavage of the core proteins by an enzyme so that some of the region of the PG which is rich in chondroitin sulfate is lost. It also appears that the aggregates formed with HA are less stable for the PGs of older cartilage. Changes in the ratio of keratan to chondroitin sulfates also occur in the intervertebral disc (McDevitt, 1988). Both maturation and aging of the nucleus are associated with an increase in the ratio of small PGs (which do not form aggregates) to larger PGs (which do form aggregates). The proportion of PGs which form aggregates in mouse nucleus decreases from 52% in the neonate to 28% in animals aged 6 months; similar changes appear to occur in human tissue. At the same time there is a decrease in the length of the HA molecules. In the neonatal mouse they are about 1,000 nm long, decreasing to about 300 nm at an age of 6 months.

Aging changes occur in the chemistry of the collagen cross-links (Kielty et al., 1993). Cross-links occur between specific amino acid residues in different collagen molecules. Initially they form between a lysine residue in one molecule and a hydroxylysine (modified lysine) residue in another. The formation of this cross-link is catalyzed by the enzyme lysyl oxidase which requires the collagen molecules to overlap as shown in Figure 4; cross-links are not formed between individual collagen molecules in solution, even in the presence of this enzyme. The resulting cross-links are said to be "reducible" because they can be broken, *in vitro*, by the reducing agent sodium borohydride ($NaBH_4$). During maturation and aging, these reducible cross-links are gradually replaced by "stable" or "non-reducible" cross-links. The physiological implications of this change in cross-link chemistry are not clear; the consequence for biochemists is that it is more difficult to extract the collagen from fibrils which are held together by stable cross-links. As a result, less is known about their chemistry. Formation of stable cross-links involves the compounds hydroxylysyl-pyridinoline and lysyl-pyridinoline. These compounds are especially common in the nucleus of the intervertebral disc and in articular cartilage but they can be detected in body fluids (including urine) because they are fluorescent. It has been proposed that the presence of these compounds in the urine can be used as a marker for collagen breakdown in the body, which might occur in joint disease (Robins et al., 1986).

Large diameter fibrils with a longer periodicity (about 100 nm) than normal collagen have been observed in electron micrographs of sections from several aging tissues (Timpl and Engl, 1987). These fibrils have been detected in Descemet's

membrane of the eye, the nucleus of the intervertebral disc, articular cartilage, skin, and nerve. They also appear in tumors in brain, bone and lymph nodes. Because their composition was originally unknown, they were given a variety of names—Luse bodies, fusiform bodies, spiny collagen, and zebra collagen. They are now believed to be type VI collagen molecules which have aggregated to form fibrils. Confusingly, these fibrils are sometimes referred to as "FLS collagen"; this is a case of mistaken identity. True FLS fibrils can be made *in vitro* by the addition of GAGs to collagen in solution but have never been identified with certainty *in vivo*. There are theoretical reasons to suppose that the formation of large diameter fibrils could influence the mechanical properties of ECM (Hukins and Aspden, 1986) but no direct evidence.

Amyloid deposition can be detected in aging tissues by histological techniques (Cottran et al., 1989). It may be associated with chronic inflammation, in which case the deposition process is called "secondary amyloidosis," but there may be no detectable underlying disease, in which case it is called "primary amyloidosis." Amyloid does not always have the same chemical composition. Evidence from a variety of techniques, including electron microscopy and x-ray diffraction, shows that it consists of fibrous deposits of denatured protein forming β-sheets.

PATHOLOGICAL CHANGES

Pathological changes in ECM can be considered to fall into two major categories: (1) an alteration in normal matrix, and (2) synthesis of abnormal matrix (Royce and Steinmann, 1993). The first usually takes the form of an increase or a decrease in the amount of matrix which retains its normal composition. The second may involve either the presence of a normal matrix component in an abnormal location or, alternatively, the synthesis of a totally alien matrix molecule. There are a variety of disease states that may be responsible for inducing these matrix changes. As matrix molecules are synthesized by cells, it follows that the matrix changes must be a reflection of altered cell function and that any stimulus, internal or external, that can lead to an alteration in cell function has the potential to affect the matrix. In very general terms, these influences include inflammation, trauma, neoplasia, ischemia, and metabolic disorders. The remainder of this section consists of a brief description of the pathogenic mechanisms, together with examples, by which each of these influences may induce matrix changes.

Inflammation

Inflammation is an exceptionally complex cellular and molecular response to tissue damage (Glynne et al., 1989; Whaley and Burt, 1992). Although there is a tendency to regard inflammation as being mediated by a defined small group of inflammatory cells, the processes involved recruit most of the local tissue cells into the response. Through the action of intercellular messengers, such as the cytokines

and locally synthesized growth factors, connective tissue cells, and in some cases non-connective tissue cells, are induced to synthesize either matrix proteins or enzymes capable of degrading matrix proteins (Duff, 1989).

A universal consequence of chronic inflammation is the formation of dense collagenous scar tissue. Scar tissue formation provides an example of increased matrix synthesis. It can be seen in the joints in rheumatoid arthritis, or in the bowel wall in Crohn's disease. The mechanisms linking inflammation and tissue fibrosis are complex but, at its most fundamental, the fibrotic process can be regarded as being due to the stimulation of local fibroblasts by cytokines synthesized by cells involved in the inflammatory process, notably cells of lymphoid and macrophage lineage. The nature of this cytokine response, despite a considerable research effort, is still unclear. Much of the research has been performed in *in-vitro* systems and has relied upon evidence of mitogenesis in the cultured fibroblasts as evidence of fibrogenic potential of a cytokine or growth factors. As cells are rarely able to synthesize matrix proteins whilst in the act of dividing, it would be more rational to examine synthesis of collagen directly, but in *in-vitro* systems this is not as straightforward as it may first seem.

Whatever the cause, one result of inflammation is the deposition of a new and often extensive matrix, usually of collagenous fibrous tissue. As collagen ages, increasing cross-linking causes the fibrils to stiffen with the result that post-inflammatory scarring tends to distort the affected tissue. In a tube, like the bowel, widespread inflammation of the type seen in Crohn's disease will then lead to luminal narrowing and the development of a stricture. In its turn, stricture may lead to gross dysfunction of the bowel necessitating local surgical resection.

Inflammation also involves altered matrix synthesis. In certain circumstances it is recognized that the scar tissue does not have the normal balance of molecular constituents. Thus, in repairing tendons the amount of type III collagen expressed as a ratio of type I collagen is much greater than in the uninjured state. Because type III collagen forms smaller diameter fibrils, the collagen fibrils can pack closer together in the repaired tissue and the surface area available for interaction with adhesion molecules is greater.

Increased matrix breakdown may accompany inflammation (Laurent, 1987). Certain inflammatory cells, notably macrophages, may synthesize enzymes which degrade ECM. In emphysema of the lung, for instance, the macrophages within the alveoli and dorsal bronchioles synthesize the enzyme elastase that will degrade the local pulmonary parenchyma leading to the typical bullae seen in this disease. The bulla is not directly caused by the inflammatory stimulus but is a reflection of the inflammatory response. Sometimes matrix degradation is not a direct consequence of the inflammation but is induced by stimulation of a specific matrix degrading cell by cytokines. In rheumatoid arthritis, osteoporosis surrounding the joint and marginal erosions are caused by stimulation of the bone degrading cells (osteoclasts), by IL-1 secreted by locally activated macrophages.

Trauma

Tissue trauma is invariably followed by attempted local tissue repair, except in severely compromised individuals such as those with severe protein malnutrition or receiving high-dose steroid therapy (Clark and Hewson, 1988).

One common type of repair mechanism illustrates very well how trauma may lead to both matrix degradation and increased matrix synthesis, through the same sequence of events. This is best illustrated using a relatively avascular connective tissue as an example. Professional athletes, particularly in high speed contact sports such as football, have a tendency to sustain twisting injuries to their knees causing trauma to the cruciate ligaments. Partial tears often do not heal well and when the knee is examined arthroscopically the surgeon describes the appearance of the ligaments as resembling spaghetti. The normal repair process requires the production of granulation tissue. This is a mixture of capillaries and fibroblasts which grow into the damaged ligament. The fibroblasts produce the repairing collagen and the capillaries bring nutrients essential for the normal function of the biosynthetically active fibroblasts. Much as a builder might remove bricks from the wall of an existing structure to bond in the wall of a new extension, so the granulation tissue burrows in to the normal tissue either side of the damaged area to produce a more effective bridging repair. Therefore, not only does the ingrowing granulation tissue synthesize new matrix but degradation is also necessary for effective repair to take place.

Neoplasia

Neoplastic cells may either synthesize an abnormal ECM themselves, or may indirectly stimulate changes in the local matrix by acting on normal matrix synthesizing cells. These effects are best exemplified by the changes seen in bones affected by neoplasms. Malignant tumors of bone-forming cells (osteogenic sarcomas) synthesize a variable matrix resembling osteoid (the ECM of normal bone) but with a composition and structure that is not typical of normal bone (Dahlin and Unni, 1991). Thus its disposition relative to bone cells, its ability to mineralize and its histological staining properties are different from those of normal bone.

Secondary malignant neoplasms that metastasize to bone are classically divided into two groups, particularly by radiologists, on the basis of their effects on bone matrix. The neoplasms may destroy bone, when they are described as lytic, or they may synthesize bone, when they are described as sclerotic. The mechanism underlying sclerosis is not known but it must be caused by a relative increase in osteoblastic activity over osteoclasis, since osteoblasts are responsible for building bone matrix which is destroyed by osteoclasts. Secondary neoplasms that classically give rise to sclerotic secondaries include carcinoma of the prostate and certain lymphomas.

Myeloma is a neoplasm that is commonly associated with lytic bone lesions. It is known that the malignant plasma cells of this neoplasm produce a substance, once known as osteoclast activating factor (OAF) and now recognized as IL-1, which stimulates osteoclasts, causing bone loss at the periphery of the tumor deposit. Because the tumor grows at an equal rate in all directions, these lesions tend to be spherical.

Ischemia

There is a hierarchy of cells within the body based on their degree of functional specialization. The more specialized the cell the more dependent it is upon an adequate oxygen supply for normal function. Specialized cells die when deprived of oxygen but, under identical conditions, non-specialized cells (such as fibroblasts) may survive, and so be stimulated to increased activity. In the case of fibroblasts, whose function is to produce the components of ECM, more matrix will be produced (Cowley and Trump, 1982). In patients with generalized narrowing of the coronary arteries, insufficient to cause infarction, cardiac muscle cells are then gradually replaced with fibrous tissue. The result is both a decrease in the effective contractility of the myocardium, due to loss of muscle cells, and increased stiffness, due to the presence of scar tissue.

Metabolic/Endocrine Disease

Metabolic disorders, such as the iron deposition disease, hemochromatosis, may stimulate local tissue cells, such as the macrophage derived tissue histiocyte, to produce fibrogenic cytokines. Iron deposition is greatest within the liver and, as a consequence, the liver in these patients may become grossly fibrotic and even cirrhotic.

Despite extensive study it is not clear which cells secrete the collagen in hepatic fibrosis. *In situ* hybridization studies suggest that the hepatocytes themselves are capable of synthesizing types I and III collagen but another candidate for this function is a cell in the wall of the hepatic sinus called the "cell of Ito."

In one of the commonest of all metabolic diseases, diabetes mellitus, abnormal ECM is deposited within the tissue, particularly in the walls of blood vessels (Monnier et al., 1986). Because of the high circulating levels of glucose, the formation and maturation of matrix macromolecules is affected by a process known as non-enzymic glycosylation. The resulting glycosylated macromolecules are relatively resistant to matrix degrading enzymes. These modified macromolecules then accumulate throughout the tissue and, when they accumulate in the walls of blood vessels, lead to abnormalities in the diffusion of metabolites and, hence, to tissue ischemia.

These are just a few examples of how ECM may be affected by disease processes. There are many others but these examples illustrate how virtually every disease process may induce these changes and that, if studied carefully, could be recognized in almost any pathological site. Many of the changes in the matrix themselves cause tissue dysfunction and are either permanent or take a long time to return towards normal. The implication is that therapeutic intervention would be very important to prevent long term tissue abnormalities but would need to be implemented early in the disease process.

SPECIFIC EXAMPLES

Articular Cartilage

Articular cartilage covers the ends of bones where they form synovial joints like the knee, hip, shoulder, and so on (Ghadially, 1983). In most human joints articular cartilage is a few millimeters thick, but there is considerable variation between joints as well as regional variations in the thickness of cartilage covering a given bone. Articular cartilage has two functions: (1) to act as a resilient coating to cushion the ends of the much stiffer bones at their points of contact, and (2) to act together with synovial fluid to lubricate the joint. These functions depend primarily on the properties of its ECM. Articular cartilage contains relatively few cells whose function is to produce the ECM components; these cells are a specialized type of fibroblast called "chondrocytes." It contains no nerves and very few blood vessels. (Although articular cartilage is usually described as "avascular," a few blood vessels can sometimes be seen in the tissue). The knee joint also contains a different type of cartilage known as "fibrocartilage" which forms the menisci or "semi-lunar cartilages." This fibrocartilage has an ECM which is very different from that of articular cartilage; its function is to ensure a good fit between the rounded ends (condyles) of the femur and the relatively flat surface of the tibia and is commonly damaged in sports injuries (Adams and Hukins, 1992).

The structure of articular cartilage is conventionally described as consisting of four zones, as shown in Figure 10. In the surface zone (sometimes called the superficial zone or zone 1) the cells are elongated with their long axes roughly parallel to the surface; the collagen fibrils tend to be oriented in the same direction. In the intermediate zone (also called the transition zone or zone 2) the cells have a more rounded shape and the collagen fibrils appear to be almost randomly oriented. Larger rounded cells occur in the deep zone (also called the radial zone or zone 3); here the collagen fibrils tend to be oriented perpendicular to the articular surface, as shown in Figure 10, and pass into the calcified zone (also called zone 4). The surface, intermediate and deep zones merge into each other. The relative thickness of these zones varies with the tissue site but the surface zone is usually the thinnest and the deep zone the thickest. However, there is a clear boundary between the deep

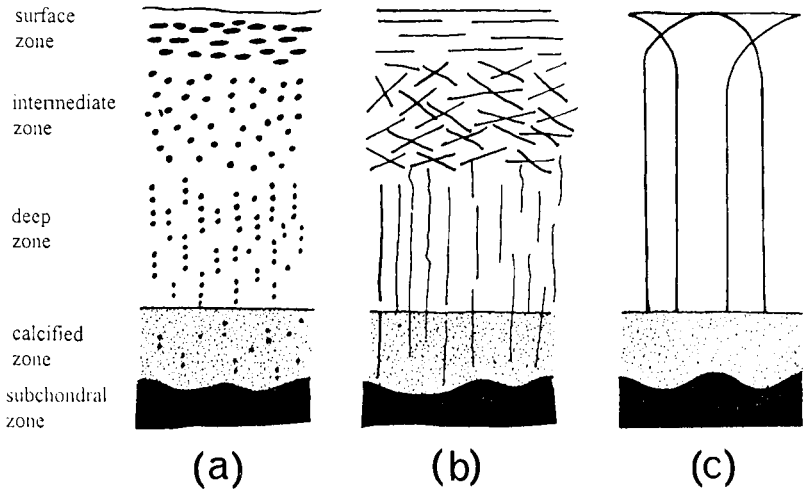


Figure 10. Zonal structure of articular cartilage. In (a) the shapes and orientations of the cells are shown, while (b) shows the orientations of the collagen fibrils. The arcades in (c) represent the paths of the collagen fibrils.

zone and the calcified zone, called the "tide mark." A boundary can also be seen between the calcified zone and the underlying bone (the "subchondral bone").

The ECM consists largely of type II collagen fibrils surrounded by hydrated PGs, many of which are aggregated by HA. Collagen fibrils in the surface zone tend, on average, to be thinner (diameter about 30 nm) than in the deep zones (diameter distribution 30–100 nm). Deposits of lipid can also be detected between the collagen fibrils by transmission electron microscopy. These extracellular lipids, together with the lipids in cell membranes, account for about 1% of the mass of fully hydrated articular cartilage. Aging cartilage accumulates a yellow pigment and it is blackened in a pathological condition called ochronosis. Most of the thickness of articular cartilage does not contain any mineral deposits but sodium urate crystals are formed in gout and calcium pyrophosphate in chondrocalcinosis ("pseudo-gout"). However, the calcified zone contains substantial deposits of impure, poorly crystalline hydroxyapatite which may also be deposited, much more generally, in osteoarthritic joints.

The ECM surrounding the chondrocytes has a specialized structure called the "pericellular matrix" (Poole et al., 1987). The term "territorial matrix" is sometimes used as a synonym but is also used to mean the region of ECM surrounding the pericellular matrix. "Interterritorial matrix" always means the bulk ECM. In the intermediate and deep zones, the chondrocyte has an obvious pericellular matrix which contains a high concentration of PGs and some type XI collagen. Whether or not these PGs interact directly with the cell surface is not known. This matrix is

surrounded by a fibrous capsule which contains types II, VI and IX collagens. This fibrous basket and its contents can be extracted intact from homogenized cartilage and is sometimes called a "chondron." There are pores in the basket which have a diameter of about 30 nm on its surface. It has been suggested that the high concentration of PGs in the chondron allows partition of water between it and the ECM. Fluid flow, in response to mechanical loads, could then occur backwards and forwards between the pericellular matrix and the ECM. It has also been proposed that this microcirculation could disperse matrix vesicles into the bulk ECM.

When examined by the naked eye, articular cartilage has a smooth glistening surface. When examined by reflected light microscopy or in the scanning electron microscope it appears to be covered with pits and hollows whose dimensions are of the order of a few micrometers. Lack of perfect smoothness at this level would not be surprising, but it has been suggested that at least some of the features which have been seen are caused by dehydration of the specimens. There are various reports of surface layers covering the surface of articular cartilage. Some involve lipid layers which can be present, absent or "heaped into mounds" on surfaces examined by transmission electron microscopy. They are likely to be ephemeral structures which could easily be wiped or washed off the surface by joint movement and are unlikely to have much functional significance. All surfaces are much more reactive than the bulk material and will tend to adsorb impurities; the lipid may be deposited on cartilage, while it is being processed for examination by electron microscopy, in much the same way as impurities adsorb on to clean metal surfaces (as occurs in the catalytic converter of a car exhaust system). An even more elusive structure rejoices in the name of the "lamina splendens." It has only been observed by light microscopy using techniques which rely on interference (phase contrast and, more recently, interference microscopy). There is no reason to believe that it is a structural feature of cartilage rather than being a result of interference at an optical discontinuity. It does, however, illustrate the danger of applying techniques without trying to understand how they work.

Articular cartilage is able to withstand applied pressure because of the swelling pressure exerted by its ECM (Grodzinsky and Frank, 1990) which is less rigid than the underlying bone because it is not calcified; it can also withstand the shear forces generated by movement of the joint because it is reinforced by collagen. It can then fulfill its first function of providing a strong cushion on the ends of bones. The ECM in articular cartilage exerts an internal pressure equivalent to 1–3 atmospheres (0.1–0.3 MPa) which is comparable to the air pressure in a car tire. However, it has been estimated that much higher compressive loads can be generated in joints, so that substantial deformation of the cartilage is expected to occur. Compressive loads can then initiate fluid flow within the ECM. This fluid flow has consequences for the mechanical properties and metabolism of the tissue (Grodzinsky and Frank, 1990). One mechanical consequence is that some of the energy supplied by compression is dissipated and so is not stored by the tissue where it might initiate

fractures. In the absence of an extensive vasculature, metabolites and waste products must be transported to and from the chondrocytes by diffusion. Small molecules and ions (e.g., urea and sulfates) can diffuse sufficiently rapidly through the interstices of cartilage ECM that fluid flow has little influence on their transport. However, the transport rates for larger molecules (e.g., proteins), which diffuse less easily through the ECM, are increased by fluid flow resulting from loading the cartilage.

Collagen fibrils maintain the mechanical stability of the cartilage when it is not subjected to an external load and enable it to withstand shear forces (Myers et al., 1984; Hukins and Aspden, 1989). If the cartilage is to remain in mechanical equilibrium, when it is not under load, its internal pressure must be balanced in the same way that a tire or a balloon retain the pressure of the air which they contain. Any such "pressure vessel" is in equilibrium because internal pressure stretches the wall of the container; the tension in the wall then balances the pressure. In articular cartilage the pressure will stretch the surface. Because the collagen fibrils of the surface zone are oriented parallel to the surface (i.e., in the direction in which this zone has to withstand tension) they are stretched and so provide tensile reinforcement. However, the internal pressure will also tend to lift the uncalcified tissue from the calcified tissue below. Since the collagen fibrils in the deep zone are perpendicular to the surface, they will be stretched in the cartilage; the tension in the fibrils balances the tendency to lift and tethers the cartilage in much the same way as ropes can tether a helium balloon to the ground. The almost random orientation of fibrils in the intermediate zone ensures that stresses can be transmitted throughout the thickness of the tissue so that there are no mechanical discontinuities. Any friction will tend to shear the articular surface away from the bulk tissue when a joint moves, as shown in Figure 11. It is clear that this frictional force will also tend to stretch those collagen fibrils in the surface zone which tend to be oriented parallel to the surface in the direction of the shear stress. Calculations also show that the tensile stress trajectories through the thickness of the tissue, which indicate the directions in which it will stretch, are closely similar to the directions in which the collagen fibrils are oriented (Myers et al., 1984).

How are these forces transferred to the collagen fibrils? It is possible that the fibrils are sufficiently long that both ends are embedded in the calcified zone and that they form arcades, as shown in Figure 10c. They then act like ropes which are anchored at both ends so that forces can be applied to them directly. This does not explain how they are stretched by the swelling pressure which requires transfer of stress from the hydrated PGs to the fibrils. Transfer of stress then requires interaction between the hydrated PGs and the collagen fibril surface. If such interactions exist, there is no need for the fibrils to be continuous (because stress can be transferred from one fibril to another through the PG gel) and the arcades shown in Figure 10c are best considered as a means of representing the loci of the discontinuous fibrils. Direct interactions between GAG chains of PGs and collagen

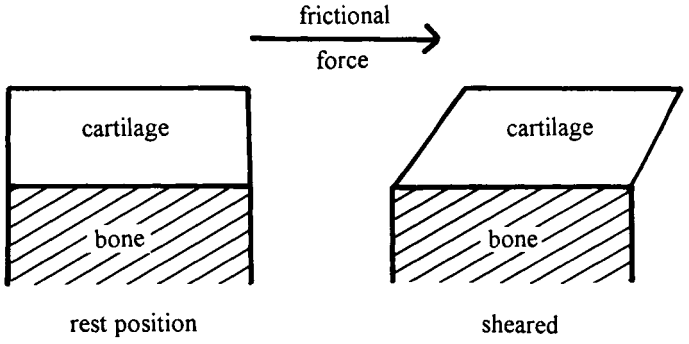


Figure 11. Friction tends to shear the articular cartilage in a joint. Since the cartilage resists this shear, the surface zone is placed in tension.

are possible as are interactions mediated by molecules which interact with both these components of the ECM, like fibronectin.

In a healthy joint there is very little friction when the two cartilage surfaces slide over each other in the presence of synovial fluid. The coefficient of friction for this movement is about 0.01 which is a third of the value for ice sliding over ice and about a tenth of the value for a lubricated engineering bearing. Lubricants form films which keep surfaces apart so that microscopic projections from the surface do not catch against each other and impede motion. The problem with fluid lubricants is that they can flow away from the surfaces. When a synovial joint is loaded, fluid is squeezed out of the cartilage and is spread over the surface by motion; presumably the fluid content of the cartilage is replenished when the load is removed. Therefore, synovial fluid is an effective lubricant only when one or preferably both of the surfaces are covered in healthy cartilage; it is a poor lubricant for metal articulating with metal or for metal on bone.

Osteoarthritis (also called "osteoarthrosis" and often abbreviated to OA) is characterized by erosion of articular cartilage which is not a direct result of inflammation (Scott, 1978). Some inflammation may accompany OA but cartilage loss as a result of inflammation is more predominant in rheumatoid arthritis. Cartilage loss (sometimes called "eburnation") in OA may lead to a narrowing of the joint space and be associated with subarticular bone sclerosis, and the formation of bone projections known as "osteophytes." These processes are illustrated in Figure 12. The changes which occur in the cartilage itself are often referred to as "fibrillation." They involve softening, splitting, and fragmentation within a focal area. Fibrillation then progresses and may lead to exposure of the underlying bone. Sometimes the cause can be identified—the disease is then classified as secondary OA. Often this cause may lead to inappropriate loading of the joint so that the cartilage is subjected to mechanical stresses which it cannot withstand. Such causes include congenital dysplasia leading to OA of the hip, meniscectomy (performed

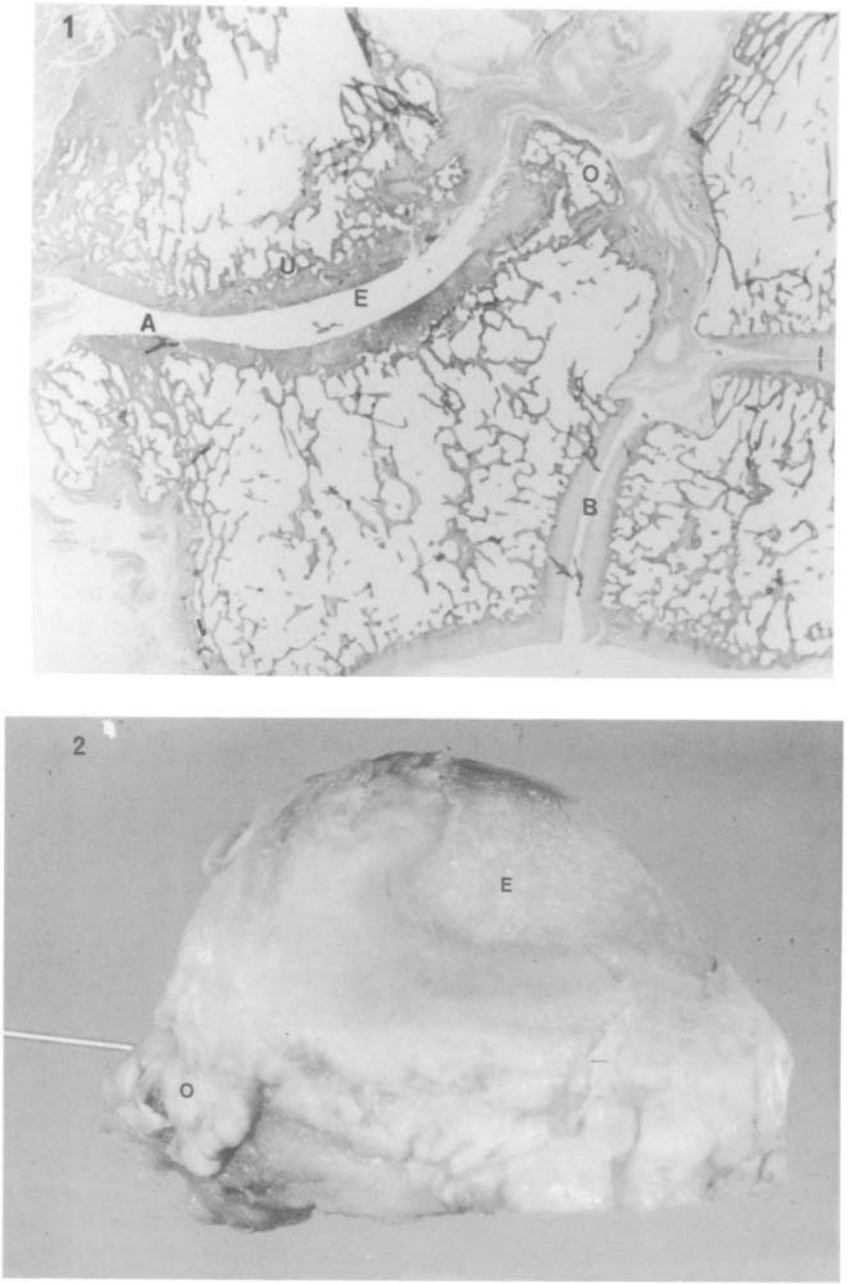


Figure 12. Osteoarthritis. Part 1 is a section through the bones of the wrist. One joint (A) shows loss of cartilage or "eburnation" (E), subarticular bone sclerosis (U) and osteophyte formation (O) when compared with the joint (B) which has normal cartilage and bone. Part 2 shows the macroscopic appearance of a femoral head from a patient with osteoarthritis showing eburnation (E) and osteophytes (O).

to prevent pain and loss of joint function resulting from a torn meniscus), OA of the knee and ankle in professional footballers, OA of the ankle in ballet dancers and so on. The problem can be especially serious when there is neuropathology which inhibits the sensation of pain from the joint. These joints are called "Charcot's joints" and the absence of pain means that they can be inadvertently abused. Several metabolic disorders also predispose to OA. When a cause cannot be identified, the disease is called primary or idiopathic or non-specific OA.

At present, there are no suitable techniques for monitoring the detailed changes in cartilage composition and properties accompanying OA in human patients. As a result, experiments have been performed on animals in an attempt to understand the processes occurring in fibrillation. Several animal models for OA have been developed. None completely mimics the human disease but inducing joint laxity (by cutting the cruciate ligaments of the knee or by removing the meniscus) do resemble human secondary OA. However, properly controlled experiments cannot be based on this type of model and most studies do not involve a statistical analysis of the data (Warskyj and Hukins, 1990). Nevertheless, some changes can be observed. These include loss of water and an increase in PG synthesis, with an increase in the ratio of chondroitin sulfate to keratan sulfate (Carney and Muir, 1988). A completely different kind of animal experiment has supported the suggestion that the initial lesions in OA are microfractures of the subchondral bone; when they heal, the bone is stiffer so that the cartilage is subjected to higher stress. Whatever the cause of cartilage fibrillation, it is not surprising that it is progressive. The mechanical stability of the tissue depends on tension in the organized collagen fibril network balancing the swelling pressure exerted by the hydrated PGs. Disruption of the structure of this network is expected to lead to mechanical instability in the internal structure of the tissue.

Both OA and rheumatoid arthritis are very painful and can lead to complete loss of joint function. Total replacement of the joint by a prosthetic device is commonly used to restore function and relieve pain (Charnley, 1979). Since synovial fluid requires a healthy articular cartilage surface to act as a lubricant, the materials used in the total joint prosthesis cannot mimic the tissues in the natural joint. Low friction is usually achieved by a metal component articulating with high density polyethylene. The main problems then encountered are loosening of the synthetic materials from the natural tissues and wear of the polyethylene component.

Bone

Bone, as opposed to a bone, has a matrix which contains water and three other major components: (1) collagen, (2) a mineral, which closely resembles hydroxyapatite, and (3) non-collagenous organic components (Freemont, 1989; 1993). The collagen is almost entirely type I and in normal bone is arranged in thin alternating rafts 2–3 μm thick. In each raft the collagen fibrils run at right angles to those in

the rafts on either side, producing a laminate of high tensile strength but which is deformable under compressive load. The mineral is responsible for the compressive stiffness of bone which also manifests itself as hardness. As in all calcified ECM, the combination of collagen fibrils and mineral enables bone to be strong without unworkable brittleness.

Non-collagenous organic components represent the most numerous, but not the most abundant of the organic molecules in bone. They fall into three major categories: (1) those molecules responsible for bone structure; (2) those molecules responsible for calcium binding, such as osteonectin, osteopontin and osteocalcin; and (3) those molecules responsible for the control of bone cell activity, the cytokines and growth factors. Both cytokines and growth factors are present in a latent form in bone matrix. When they are released by osteoclastic activity, they initiate cellular events in the region of the bone surface; TGF-β is the most studied of these molecules.

All the matrix molecules are synthesized by osteoblasts which also control their spatial orientations. The breakdown of matrix is controlled by the multinucleated osteoclasts. Through the activity of the latent incorporated messengers and the communicating network of intraosseous osteocytes, the activity of osteoclasts and osteoblasts is matched, or coupled, so that the total amount of bone remains constant.

The pathological changes that can occur to bone matrix are similar to those described in general terms earlier in this chapter (Revel, 1986). They can be summarized as too much or too little or abnormal component synthesis. In the Western world, particularly in postmenopausal women, too little bone matrix is becoming an important cause of morbidity. The general activity of womens' osteoblasts is under the control of the female sex hormone estrogen. After the menopause the cycle of hormonal activity is lost with the effect that osteoblastic bone synthesis is reduced. Osteoclasis is either unaffected or sometimes increased. The bone cell activity is, therefore, uncoupled and the bone mass falls (Figure 13). This reduction in bone mass is called osteopenia. A reduction in the amount of bone increases the risk of fracture. The fracturing osteopenia caused by removal of osteoblastic sex hormone drive in women is known as postmenopausal osteoporosis and is classically represented by crush fractures of the vertebral bodies (Figure 14).

Too little bone matrix is also seen when osteoclasis increases in the absence of a similar increase in bone formation. This has already been described in the context of lytic neoplastic bone deposits and is occasionally seen in the endocrine disorder hyperparathyroidism, where excessive secretion of parathyroid hormone leads to excessive osteoclasis and bone resorption.

An excess of bone matrix is much less common. It is seen in the virally induced bone disorder, Paget's disease. In this disorder, viral infection leads to a generalized increase in bone cell activity but the relatively ineffectual osteoclasis means that there is a slow acquisition of excess bone. Interestingly, this excessive bone is weak.

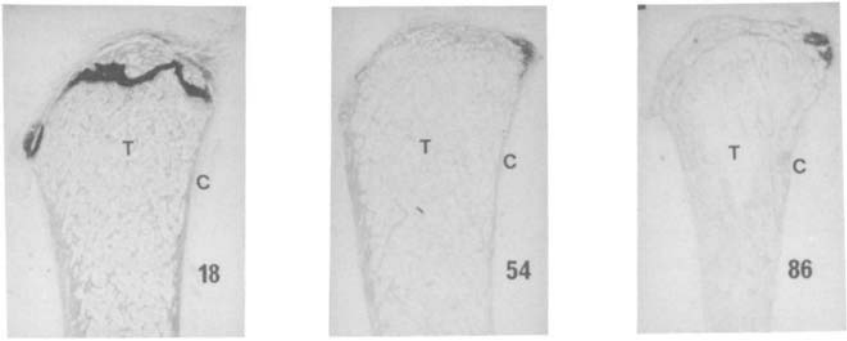


Figure 13. Sections of iliac crest biopsies stained to show cortical (C) and trabecular (T) bone. The specimens were obtained from three women whose ages (in years) are marked on the figure. The bone from the 18-year-old is normal. The 54-year-old has lost trabecular bone and the 86-year-old has little of either cortical or trabecular bone.

Although the osteoclastic activity does not match the osteoblastic activity, it is still grossly in excess of normal. Newly formed bone is eroded, new bone deposited and further erosion occurs. The resulting small packets of bone lack the strength of the bulk ECM, in much the same way as a pillar of pebbles is weaker than an equivalent pillar made from a single piece of stone. As a result, this bone is more likely to fracture under load than normal bone. The problem is compounded by the maximally stimulated osteoblasts that are typical of Paget's disease, producing an abnormal laminate structure. In this structure the collagen fibrils are randomly oriented. This random orientation weakens the bone in two ways. First, tensile reinforcement is no longer concentrated in the required directions. Secondly, the randomly oriented fibrils are unable to pack together as closely as in normal bone so that the collagen content in a given volume of ECM is decreased.

Another consequence of the increased volume of bone in Paget's disease is that holes or canals passing through the bone tend to be narrowed and their contents compressed. Most commonly this causes compression of the cranial nerves as they penetrate the skull, particularly the VIIIth nerve whose compression leads to deafness.

It is not possible to divorce the bone from the marrow adjacent to it. Osteoclasts are formed from hemopoietic precursors and in certain congenital disorders the stimulus to osteoclast formation is lacking. Infants with this disorder are able to deposit bone but not resorb it. The result is very dense bone and the disorder, which is named after the appearance of the bone, is called marble bone disease or, more properly, osteopetrosis.

Osteogenesis imperfecta (OI), also known as brittle bone disease, refers to a group of genetic disorders of collagen synthesis leading to changes in the skeleton and other sites of collagen deposition such as the sclerae, ligaments, skin, and so

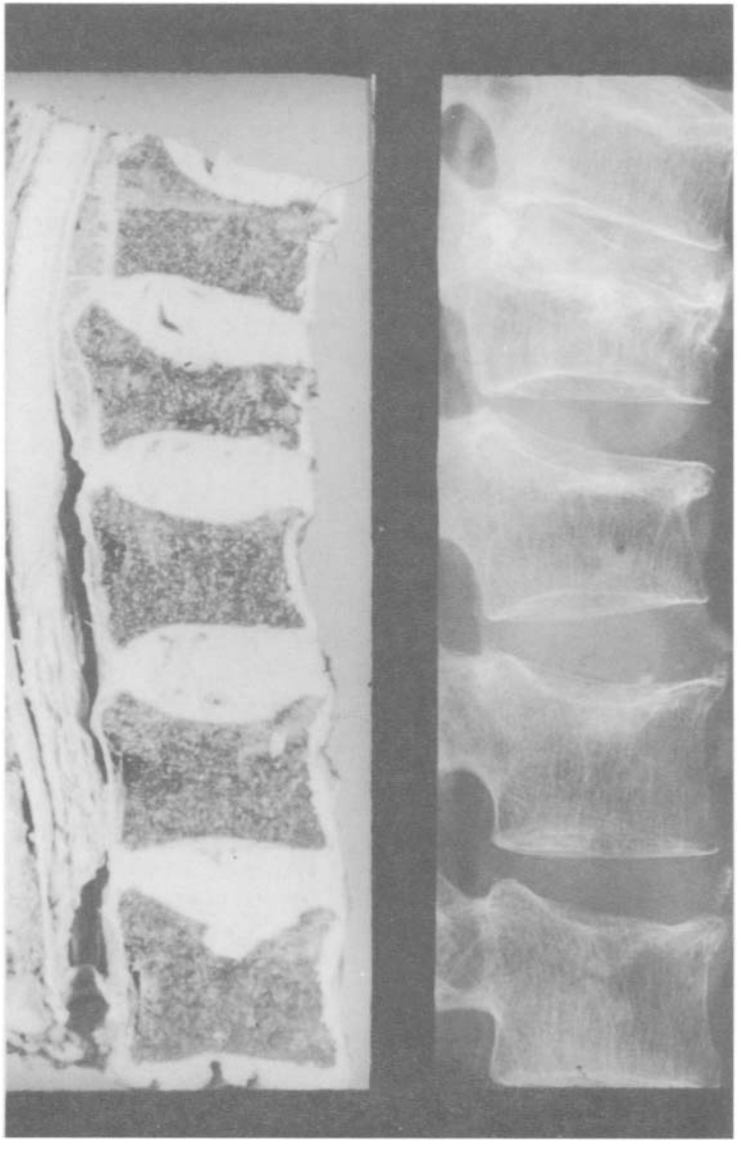


Figure 14. High resolution radiograph and macroscopic specimen of the lower thoracic and upper lumbar regions of an osteoporotic spine. There are varying degrees of vertebral crush fracture.

forth (Sykes and Smith, 1985). There are several forms of the disease but all share an abnormality of type I collagen synthesis. Each variant is characterized by a mutation of the genes encoding the synthesis of procollagen chains resulting in abnormalities in the molecular structure of collagen. The effect of this abnormal collagen depends on the tissue and the organism. In certain forms of the disease, the collagen is so deranged that the infant dies *in utero* or shortly after birth. In other forms the defect is so minor that the only abnormality is a form of mild premature osteoporosis.

Skin

The epidermis forms the outer layer of the skin; its structure is shown in Figure 15. It consists of cells, called keratinocytes, surrounded by ECM. As the keratinocytes age, they become filled with increasing amounts of the fibrous protein keratin. Surface layers of epidermis are constantly being shed so that the epidermis has to continuously renew itself. Cell division occurs in a "basal layer" in which keratinocytes are anchored to basement membrane; the basal layer contains type IV collagen, laminin and PGs containing heparan sulfate (Furthmayr, 1988). Cells that have left the basal layer undergo "terminal" differentiation as they move towards the tissue surface (Adams and Watt, 1990).

Integrins play an important role in determining the spatial organization of cells within the epidermis. Immunofluorescence studies have shown that integrins are expressed by the basal layer of proliferating keratinocytes but are largely absent from the layers of terminally differentiating cells (Adams and Watt, 1991). The major integrins expressed by keratinocytes in culture are $\alpha 2\beta 1$ and $\alpha 3\beta 1$ (variously identified as receptors for collagen and laminin), $\alpha 5\beta 1$ (the fibronectin receptor) and $\alpha 6\beta 4$ (Hertle et al., 1991). Expression of each integrin, whether in culture or in the epidermis, is confined to the basal layer. During terminal differentiation, keratinocytes lose their adhesiveness to fibronectin, laminin, and collagen types I

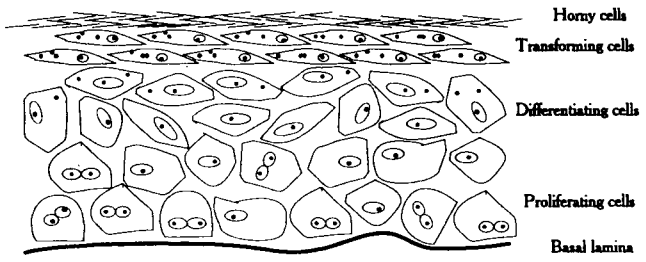


Figure 15. Schematic representation of the epidermis showing the vertical transition of keratinocytes from the basal layer to the skin surface. Keratinocytes proliferate in the basal layer, proximal to the basal lamina, differentiate in the spinous and granular layers and are finally transformed into tough horny cells which slough off.

and IV. These changes in adhesiveness are followed by the loss of β1 integrins from the cell surface. In the case of α5β1, there is direct evidence for a reduction of fibronectin-binding ability occurring several hours before the integrin is lost from the cell surface. This sequence of events may ensure that cells which are committed to terminal differentiation are selectively expelled from the basal layer (Adams and Watt, 1990, 1991; Hertle et al., 1991).

Keratinocytes migrate when the skin is wounded (Grinnell, 1992). In general terms, epidermal wounding involves the following processes, some of which occur concurrently. Initially, disruption of blood vessels causes loss of plasma fibronectin and fibrinogen into the perivascular space. These proteins covalently cross-link to themselves and to each other, providing a temporary thrombogenic substrate which supports platelet aggregation and blood coagulation. In addition, they provide a matrix which cells migrate into. Leukocytes are responsible for phagocytosis of bacteria and debris in this matrix, endothelial cells revascularize the wound space, and fibroblasts convert the matrix into granulation tissue (Humphries and Mould, 1993). Keratinocytes migrate over this matrix to reseal the wound. Since the matrix is composed mainly of fibronectin and fibrin, these keratinocytes need to use a different set of adhesion receptors to those involved in anchorage to basement membrane (Ding and Burstein, 1988). As a result, expression of the α5β1 fibronectin receptor occurs (Kurpakus et al., 1991). Normal adhesion is established once the epithelium is repaired (Humphries and Mould, 1993). The ability of fibronectin to support keratinocyte migration and adhesion during wound repair has suggested its potential use as a therapeutic agent for epithelial defects which are not healing well. However, addition of fibronectin sometimes has no visible effect on the rate of wound healing and further investigation is required to test its possible therapeutic applications (Humphries and Mould, 1993).

There are currently relatively few human diseases that are known to be caused by defects in ECM. More often, ECM is implicated in the etiology of a multifactorial disease, or contributes to pathogenesis. The family of inherited diseases collectively termed epidermolysis bullosa (EB) is one exception (Epstein, 1992; Vitto and Christiano, 1992). EB can be classified into at least three different categories, each of which is characterized by widespread blistering of skin in response to extremely mild trauma (Figure 16). The three sub-types of EB are termed "dystrophic," "junctional" and "simplex." While these are now thought to be the result of mutations in at least three different genes, the consequences are all manifested in the same way. In all cases, the propensity to blister appears early in life, but other aspects of the disease are highly variable. Inheritance can be recessive or dominant, skin involvement can be localized or widespread, the distress to the patient can vary immensely, and in some cases blisters heal normally, while in others there may be extensive scarring. In addition to the heritable forms of EB, EB acquisita is an acquired autoimmune disease which has a clinical presentation very similar to that of certain EB conditions (Cheng et al., 1992).

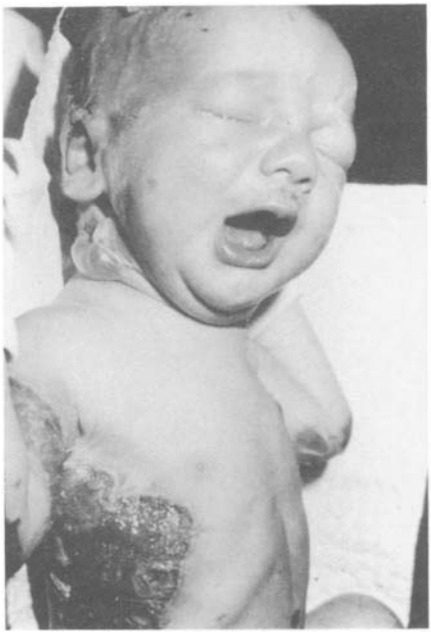


Figure 16. Skin blistering in a baby with epidermolysis bullosa (EB).

Classification of the sub-types of EB was originally carried out with the aid of electron microscopy. It was found that the point in the skin at which the split occurred varied from patient to patient: in EB simplex, the split was within the basal epidermal cells themselves (the roof of the blister contained the basal cell nuclei while the base of the blister contained the basal plasma membrane); in junctional EB, the split was in the basement membrane that underlies the basal cells; and in dystrophic EB, the split was in the upper regions of the dermis. None of these forms of blistering resembles classical friction blistering where the split generally occurs between layers of suprabasal keratinocytes. The differences in the site of fracture in the different EB conditions implied that each type of blistering was the result of a defect in a different molecule. EB simplex was almost certainly due to an intracellular molecule, while junctional and dystrophic EB could have been due to either malfunctioning of extracellular matrices or to defective keratinocyte interactions with a normal matrix. Largely through the efforts of a large number of medical geneticists, candidate molecules which might be mutated in each kind of EB have now been identified.

EB simplex is a relatively mild, non-scarring form of EB and usually is only irritating rather than life-threatening. In general, EB simplex is inherited in an autosomal dominant manner. A number of the underlying defects have now been

linked to mutations in components of the intermediate filament cytoskeleton (Bonifas et al., 1991; Coulombe et al., 1991a,b; Ryynanen et al., 1991; Lane et al., 1992). Intermediate filaments (IF) are linear structures whose main function is considered to be the stabilization of cytoplasmic architecture. The filaments have a diameter of 8–10 nm and are composed of polymerized dimeric proteins assembled into a coiled-coil arrangement. A number of different IF proteins have been identified, but in skin the principal constituents are the keratins, of which there are at least twenty characterized variants. Epidermal IFs are heterodimeric polymers containing one member of the type I keratin sub-family and one member of the type II sub-family. Type I keratin genes are clustered on human chromosome 17q (including the epidermal specific keratin 14), while type II genes are on chromosome 12q (including keratin 5).

The fragility of EB simplex basal keratinocytes suggested that the molecular defect could be related to the cytoskeleton and linkage analysis in large affected families revealed mapping to both chromosome 12q and 17q. In the 17q family, a point mutation which caused a leucine to proline alteration in a type I keratin was identified. This mutation would be predicted to disrupt the helical structure of the intermediate filament. Subsequently, a large number of mutations in a variety of keratin molecules have been identified as the causes of different EB simplex syndromes. It may be that the wide variation in severity of the disease could be a reflection of the role of the molecule that harbors the mutation and of the site within the molecule where the mutation is found. Confirmatory evidence for the involvement of intermediate filament keratins in EB simplex has come from the use of transgenic mice. Animals lacking a portion of keratin 14 have fragile skin and widespread cytolysis of their basal keratinocytes which leads to death shortly after birth.

In contrast to the relatively mild symptoms associated with EB simplex, dystrophic EB is an extremely serious disease in which the split in the epidermis is deeper and where scarring can be widespread. Frequently, there is recurrent blistering which hampers the ability of the skin to heal and this can even result in tooth and fluid loss. The site at which epidermal splits occur in dystrophic EB could be consistent with overproduction of a matrix proteinase such as collagenase, but to date there is no evidence for this. Indeed recent evidence has shown that the disease does not link to the collagenase gene. Instead, a breakthrough came with the observation that the number of anchoring fibrils, linking basal keratinocytes to the dermal connective tissue, was reduced in dominantly inherited dystrophic EB. In the more severe recessively inherited forms of the disease, the anchoring fibrils can be completely absent. This decrease might be expected to result in increased epidermal fragility. The anchoring fibril assemblies serve to link the major type I and III collagen fibers in the dermis to basement membrane components immediately beneath basal keratinocytes. The principal component of anchoring fibrils is type VII collagen, the gene for which (COL7A1) is localized to human chromosome

3p21. For both dominant and recessive forms of dystrophic EB, the defective gene maps exactly to this location (Hovnanian et al., 1992). Although specific mutations in type VII collagen have not yet been identified, it seems highly likely that they will be and that dystrophic EB will be the third known inherited disorder involving collagens (after osteogenesis imperfecta and certain classes of Ehlers Danlos syndrome).

The third form of EB, junctional, is also characterized by an extremely severe phenotype. As described above, the primary defect in this form of EB appears to be localized to the basement membrane underlying basal keratinocytes. As yet, there is no definitive evidence assigning the defect to a particular molecule. However, recent investigations of the molecular structures that link epidermal cells to the dermis have led to the identification of a series of epithelial-specific components which deserve further attention. In epidermis, groups of cells are bound together and to their underlying matrix by specialized anchoring junctions. Some of these structures (desmosomes and tight junctions) mediate cell-cell interactions, while others (adherens junctions and hemidesmosomes) link basal cells to the basement membrane. Hemidesmosomes are small, plaque-like structures that rivet the keratin intermediate filament skeleton to extracellular matrix and appear to have a key role in maintaining tissue architecture. At the microscopic level, hemidesmosomes are seen to link epidermal cells to type VII collagen-containing anchoring fibrils via thinner anchoring filaments (Figure 17). In junctional EB, the number and size of hemidesmosomes are greatly reduced and this might be predicted to weaken epidermal-dermal adhesion. The molecular organization of the hemidesmosome is largely unknown, although the cytoplasmic bullous pemphigoid antigen, BPAG1, the transmembrane collagenous bullous pemphigoid antigen, BPAG2, and the integrin receptor $\alpha 6\beta 4$ are known to be present (Jonkman et al., 1992). It is conceivable that either of these molecules might be targets for mutations causing junctional EB, as would a ligand for $\alpha 6\beta 4$.

A further molecule which might be involved in junctional EB has been identified through the use of monoclonal antibody staining of tissue sections. The GB3 antibody, originally raised against human amnion, was the first of a series of reagents shown to specifically recognize an epithelial basement membrane antigen present in normal anchoring filaments but that is absent in junctional EB (Verrando et al., 1988). Biochemical analyses of this antigen have shown it to be an isoform of laminin, termed kalinin, and since laminin is specifically recognized by the integrin $\alpha 6\beta 1$, kalinin may be a potential ligand for $\alpha 6\beta 4$. Studies are now underway to test the hypothesis that defects in the hemidesmosome/$\alpha 6\beta 4$/kalinin system may explain junctional EB and whether exogenous addition of the missing component might represent a possible therapeutic approach for correcting the disease.

At present, treatment of EB is exceedingly difficult, but in the future the inherited EB conditions may be well suited to therapies based on gene transfer. For those

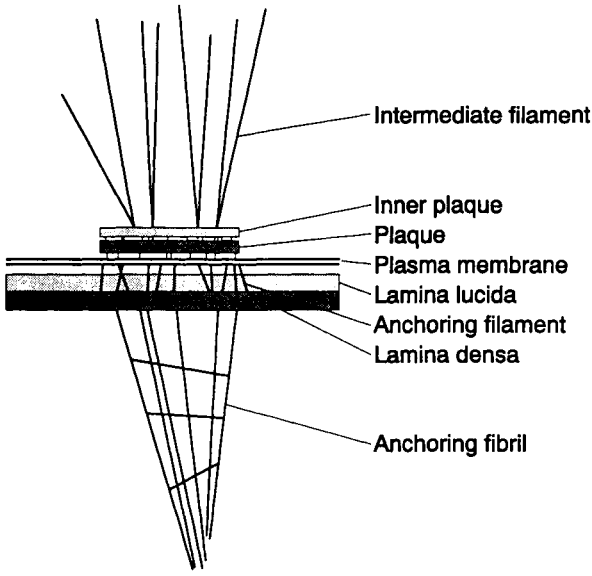


Figure 17. Diagrammatic representation of a hemidesmosome showing the structural features discernible by electron microscopy. The inner plaque region probably contains bullous pemphigoid antigen 1, while the plaque region contains bullous pemphigoid antigen 2 and the integrin $\alpha 6\beta 4$. Kalinin localizes to the lamina lucida and classical basement membrane molecules are found in the lamina densa (laminin, type IV collagen and heparan sulfate PG). The principal component of the anchoring fibril is type VII collagen.

mutations which lead to a loss of expression of a particular key protein, grafting with cultured keratinocytes expressing the transfected normal protein might offer the prospect of reversing the disease. For other mutations which produce dominant negative products which interfere with the functioning of the normal allelic product, other approaches such as antisense molecules targeted to the mutated allele would be preferred.

Microvasculature

At first sight the vessels of the microvasculature appear to be very unlikely candidates for the study of changes in the ECM. Most simply they consist of a lining of flattened endothelial cells surrounded by a thin basement membrane (Paulsson, 1992) containing rare pericytes. In this setting, the connective tissue matrix is the basement membrane. This is a highly organized structure consisting of a mixture of matrix molecules based upon a three-dimensional meshwork of type IV collagen fibers.

The importance of the integrity of this structure and the devastating effects that can occur if it becomes deranged is demonstrated by one of the more highly specialized basement membranes, that of the glomerular capillaries. The normal glomerular capillary basement membrane (GBM) is approximately 320 nm thick and consists of three distinct layers, at the ultrastructural level (Moore and Briggs, 1992). The central layer is thick and is densely stained in the preparative techniques involved in electron microscopy. It is called the "lamina densa." The peripheral thinner layers stain less densely and are known as the "lamina rara interna" and "externa." Type IV collagen is found in all three layers and accounts for 50% of the dry mass of the GBM. Laminin is also present in all three layers but is concentrated in the laminae rarae where it is responsible for the adhesion of the endothelial cells of the glomerular capillaries, on one side, and the glomerular epithelial cells, on the other. Polyanionic proteoglycans, particularly those containing heparan sulphate, are present in aggregates every 50 to 60 nm along both laminae rarae and are thought to be responsible for the charge-dependent filtration barrier effect of the basement membrane. In addition, there are glycoproteins including entactin and fibronectin.

The function of the GBM is to act as a selective semipermeable barrier to the transfer of plasma molecules from blood to urine. The structure of this basement membrane is somewhat different from that in other capillaries and this is reflected in its function as a filtration barrier. The GBM will exclude all proteins the size of albumin (3.6 nm in radius) but its permeability increases for progressively smaller molecules. Consequently, it is extremely permeable to water and small solutes. Permeability is also affected by charge, the polyanionic proteoglycans allowing neutral and cationic molecules to pass but inhibiting the passage of anions.

In disease the GBM becomes leaky permitting very large molecules, and even red blood cells, to pass through. Also, loss of proteoglycans leads to increased permeability of the GBM for anions (Tisher and Brenner, 1989).

Although there are rare disorders in which the GBM is thinned, by far the most common abnormality is basement membrane thickening which represents an abnormality of matrix molecules. This thickening is of two types: (1) diffuse thickening, as seen in diabetes mellitus, and (2) amyloidosis and the nodular thickening typical of a disorder known as membranous glomerulonephritis.

In membranous glomerulonephritis immune complexes, consisting of antigen and antibody are deposited on, or form on, the basement membrane. Here they elicit a response which causes deposition of large quantities of new basement membrane, first between the complexes and, eventually, surrounding them altogether. This process represents a proliferation of basement membrane material which, paradoxically, is associated with increased permeability. The basement membrane constituents are normal. Occasionally, patients with this disorder have a resolution of their disease which involves a loss of the excess matrix indicating that its molecules can be degraded.

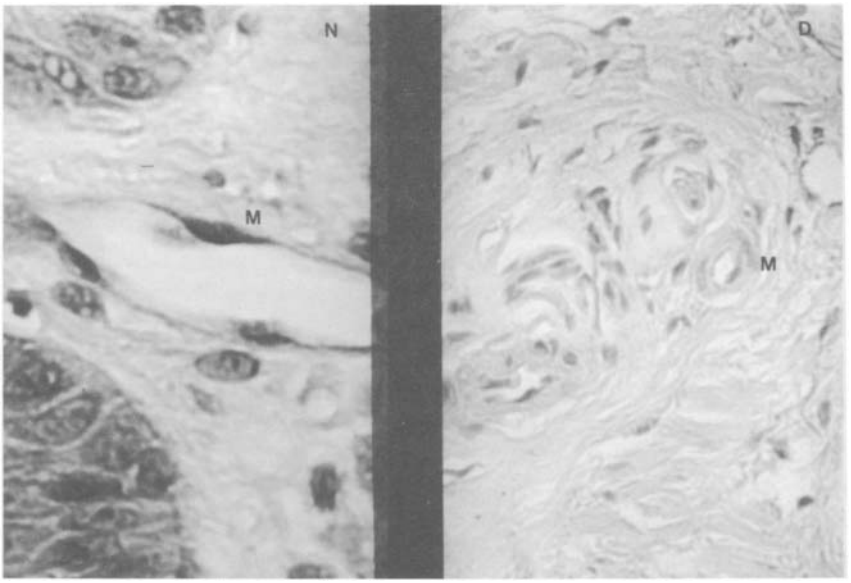


Figure 18. In normal skin (N) the microvascular basement membrane is so thin that it is not visible in this figure. In diabetic skin (D) the microvascular basement membrane (M) is thickened.

In diabetes there is a diffuse and often uniform thickening of the GBM (Figure 18). This is caused by an increase in type IV collagen and a decrease in other glycoproteins. These events are almost certainly the result of abnormal non-enzymic glycosylation of the collagen making it less susceptible to breakdown. As the rate of synthesis does not fall, accumulation occurs. This membrane is also more permeable because of the decrease in the concentration of polyanionic PGs.

Thickening of diabetic microvascular basement membrane is not restricted to the renal glomerular vessels. It is also seen in other vessels as diffuse thickening of vascular basement membranes. Another disease with very similar changes is systemic sclerosis. Recent studies have shown that, in systemic sclerosis, the microvascular basement membranes are thickened because they are reduplicated. By *in-situ* hybridization techniques, it has been shown that this is a consequence of an alteration in the synthesis of the type IV collagen by the endothelial cells. Two of the usual chains are replaced by another, as yet unidentified, chain that is less susceptible to degradation and therefore accumulates. The thickened basement membrane in systemic sclerosis has reduced permeability. As a result, the concentrations of oxygen and nutrients are reduced which stimulates fibroblasts to deposit type I collagen in the surrounding tissue (see Chapter 10).

This section illustrates how changes in a very small proportion of the ECM macromolecules can cause gross changes in organ function. In the case of mem-

branous glomerulonephritis and diabetic microvascular nephropathy, the outcome is most usually chronic renal failure which is universally fatal unless supportive regimes such as dialysis or transplantation are instituted. The vascular changes in systemic sclerosis lead to gross fibrosis of the skin forming a hard carapace that inhibits movement.

In general, this section also provides a useful conclusion to the chapter by emphasizing the relationship between chemical composition, structure, physical properties, and biological function of the ECM. It also illustrates how changes can be brought about by the production of increased amounts of normal material, externally induced changes in the matrix molecules and synthesis of abnormal matrix constituents by cells.

SUMMARY

Extracellular matrix (ECM) surrounds the cells in tissues. It contains glycosaminoglycans (GAGs) which attract the water that is its major component. Most GAGs are covalently linked to the protein core of a proteoglycan (PG). There are at least 16 genetically distinct types of collagen in ECM; over 70% of the collagen consists of types I, II and III which form fibrils. Collagen fibrils provide tensile reinforcing in ECM. ECM also contains elastic fibers, glycoproteins, cytokines, growth factors, and enzymes which are capable of degrading the matrix. Molecules like fibronectin are capable of binding cells to the macromolecular components of ECM. Receptors enable the cell surface to interact with these macromolecules. Integrins provide an important example of these receptor molecules which extend into the plasma membrane. The binding affinity to ECM can then be regulated from within the cell (inside-to-out signalling) and the matrix can influence events within the cell (outside-to-in signalling), including differentiation. Thus, ECM can influence development at the cellular level as well as forming the structures of the body. Changes in the ECM occur during aging but it may be difficult to distinguish aging from degeneration. Pathological changes involve an alteration in the normal matrix or synthesis of abnormal matrix; alteration in normal matrix usually involves an increase or a decrease in the amount of matrix which retains its normal composition.

REFERENCES

Adams, J. C. & Watt, F. M. (1990). Changes in keratinocyte adhesion during terminal differentiation: reduction in fibronectin binding precedes a5b1 integrin loss from the cell surface. Cell 63, 425–435.

Adams, J. C. & Watt, F. M. (1991). Expression of β1, β3, β4 and β5 integrins by human epidermal keratinocytes and non-differentiating keratinocytes. J. Cell. Biol. 115, 829–841.

Adams, M. E. & Hukins, D. W. L. (1992). Extracellular matrix of the meniscus. In: Knee Meniscus—Basic and Clinical Foundations (Mow, V. C., Arnoczky, S. P. & Jackson, D. W., eds.), pp 15–28, Raven Press, New York.

Birk, D. E., Fitch, J. M., Babiarz, J. P., Doane, K. J., & Linsenmayer, T. F. (1990). Collagen fibrillogenesis: interaction of types I and V collagen regulates fibril diameter. J. Cell. Sci. 95, 649–657.

Bonifas, J. M., Rothman, A. L., & Epstein, E. H. (1991). Epidermolysis bullosa simplex: evidence in two families for keratin gene abnormalities. Science 254, 1202–1205.

Brewton, R. G. & Mayne, R. M. (1993). Heterotypic type II, IX and XI fibrils: comparison of vitreous and cartilage forms. In: Extracellular Matrix Assembly and Function (Yurchenko, P. D., Birk, D. E., & Mecham, R. P., eds.), Academic Press, Orlando.

Carrell, N. A., Fitzgerald, L. A., Steiner, B., Erickson, H. P., & Phillips, D. R. (1985). Structure of human platelet membrane glycoproteins IIb and IIIa as determined by electron microscopy. J. Biol. Chem. 260, 1743–1749.

Cheng, J., Syder, A. J., Yu, Q. C., Letai, A., Paller, A. S., & Fuchs, E. (1992). The genetic basis of epidermolytic hyperkeratosis: a disorder of differentiation-specific epidermal keratin genes. Cell 70, 811–819.

Coulombe, P. A., Hutton, M. E., Letai, A., Hebert, A., Paller, A. S., & Fuchs, E. (1991a). Point mutations in human keratin 14 genes of epidermolysis bullosa simplex patients: Genetic and functional analyses. Cell 66, 1301–1311.

Coulombe, P. A., Hutton, M. E., Vassar, R., & Fuchs, E. (1991b). A function for keratins and a common thread among different types of epidermolysis bullosa simplex diseases. J. Cell Biol. 115, 1661–1674.

Cowley, R. A. & Trump, B. F. (1982). Pathophysiology of Shock, Anoxia and Ischaemia. Williams and Wilkins, Baltimore.

Dahlin, D. C. & Unni, K. K. (1991). Bone Tumors, 4th edn. Charles C. Thomas, Springfield.

Ding, M. & Burstein, N. L. (1988). Fibronectin in corneal wound healing. J. Oc. Pharm., 4, 75–91.

Duff, G. W. (1989). Peptide regulatory factors in non-malignant disease. Lancet 1, 1432–1434.

Gorski, J. P., Marks, S. C., Cahill, D. R., & Wise, G. E. (1988). Developmental changes in the extracellular matrix of the dental follicle during tooth eruption. Conn. Tissue Res. 18, 175–190.

Grinnell, F. (1992). Wound repair, keratinocyte activation and integrin modulation. J. Cell Sci. 101, 1–5.

Grodzinsky, A. J. & Frank, E. H. (1990). Electromechanical and physiochemical regulation of cartilage strength and metabolism. In: Connective Tissue Matrix, Part 2 (Hukins, D. W. L. ed.), London: Macmillan. pp. 91–126.

Guan, J. L., Trevithick, J. E., & Hynes, R. O. (1991). Fibronectin/integrin interaction induces tyrosine phosphorylation of a 120-kDa protein. Cell Reg. 2, 951–964.

Guan, J. L. & Shalloway, D. (1992). Regulation of focal adhesion-associated protein tyrosine kinase by both cellular adhesion and oncogene transformation. Nature 358, 690–692.

Hayashi, Y., Haimovich, B., Reszka, A., Boettiger, D., & Horowitz, A. (1990). Expression and function of chicken integrin β1 subunit and its cytoplasmic domain mutants in mouse NIH 3T3 cells. J. Cell Biol. 110, 175–184.

Hertle, M. D., Adams, J. C., & Watt, F. M. (1991). Integrin expression during human epidermal development *in vivo* and *in vitro*. Development 112, 193–206.

Holt, C. & van Kemenade, M. J. J. M. (1989). The interaction of phosphoproteins with calcium phosphate. In: Calcified Tissue, (Hukins, D. W. L., ed.), pp. 175–213 Macmillan, London.

Horowitz, A., Duggan, K., Buck, C., Beckerle, M. C., & Burridge, K. (1986). Interaction of plasma membrane fibronectin receptor with talin—a transmembrane linkage. Nature, 320, 531–533.

Horvath, A. R., Elmore, M. A., & Kellie, S. (1990). Differential tyrosine-specific phosphorylation of integrin in Rous sarcoma virus transformed cells with differing transformed phenotypes. Oncogene 5, 1349–1357.

Hovnanian, A., Duquesnoy, P., Blanchet-Bardon, C., Knowlton, R. G., Amselem, S., Lathrop, M., Dubertret, L., Uitto, J., & Goossens, M. (1992). Genetic linkage of recessive dystrophic epidermolysis bullosa to the type VII collagen gene. J. Clin. Inv. 90, 1032–1036.

Hukins, D. W. L. & Aspden, R. M. (1985). Composition and properties of connective tissues. Trends in Bio. Sci. 10, 260–264.

Hukins, D. W. L. & Aspden, R. M. (1989). Fibre reinforcing in articular cartilage. In: Material Properties and Stress Analysis In Biomechanics, (Yettram, A. L., ed.), pp. 44–59; Manchester University Press, Manchester.

Humphries, M. J. (1990). The molecular basis and specificity of integrin-ligand interactions. J. Cell Sci. 97, 585–592.

Humphries, M. J. & Mould, A. P. (1993). Integrin adhesion receptors in epidermal wound healing. Wounds 5, 24–31.

Humphries, M. J., Mould, A. P., & Tuckwell, D. S. (1993). Dynamic aspects of adhesion receptor function: integrins both twist and shout. BioEssays 15, 391–397.

Hynes, R. O. (1992). Integrins: versatility, modulation and signalling in cell adhesion. Cell, 69, 11–25.

Jeronimidis, G. & Vincent, J. F. V. (1984). Composite materials. In: Connective Tissue Matrix (Hukins, D. W. L., ed.), pp. 187–210, Macmillan, London.

Jonkman, M. F., De Jong, M. C. J. M., Heeres, K., & Sonnenberg, A. (1992). Expression of integrin α6β4 in junctional epidermolysis bullosa. J. Inv. Derm. 99, 489–96.

Kelly, T., Molony, L., & Burridge, K. (1987). Purification of two smooth muscle glycoproteins related to integrin. Distribution in cultured chick embryo fibroblasts. J. Biol. Chem. 262, 17189–17199.

Kornberg, L. J., Earp, H. S., Turner, C. E., Prockop, C. & Juliano, R. L. (1991). Signal transduction by integrins: increased phosphorylation caused by clustering of β1 integrins. Proc. Natl. Acad. Sci. USA, 88, 8392–8396.

Kurpakus, M. A., Quaranta, V., & Jones, J. C. R. (1991). Surface relocation of α6β4 integrins and assembly of hemidesmosomes in an *in vitro* model of wound healing. J. Cell Biol. 115, 1737–1750.

Lane, E. B., Rugg, E. L., Navsaria, H., Leigh, I. M., Heagerty, A. H. M., Ishida-Yamamoto, A., & Eady, R. A. J. (1992). A mutation in the conserved helix termination peptide of keratin 5 in hereditary skin blistering. Nature 356, 244–246.

Larjava, H., Peltonen, J., Akiyama, S. K., Yamada, S. S., Gralnick, H. R., Uitto, J., & Yamada, K. M. (1990). Novel function for β1 integrins in keratinocyte cell-cell interactions. J. Cell Biol. 110, 803–815.

Laurent, G. J. (1987). Dynamic state of collagen. Pathways of collagen degradation *in vivo* and their possible role in regulation of collagen mass. Am. J. Phys. 252, c1–c9.

Lipfert, L., Haimovich, B., Schaller, M. D., Cobb, B. S., Parsons, J. T., & Brugge, J. S. (1992). Integrin-independent phosphorylation and activation of the protein tyrosine kinase pp125^{FAK} in platelets. J. Cell Biol. 119, 905–912.

McDevitt, C. (1988). Proteoglycans of the intervertebral disc. In: The Biology of the Intervertebral Disc, Vol 1, (Ghosh, P., ed.), pp. 151–170, CRC Press, Boca Raton.

Marcantonio, E. E., Guan, J. L., Trevithick, J. E., & Hynes, R. O. (1990). Mapping of the functional determinants of the integrin β1 cytoplasmic domain by site-directed mutagenesis. Cell Reg. 1, 597–604.

Monnier, V. M., Vishwanath, V., Frank, K. A., Elmets, C. A., Douchat, P., & Kohn, R. R. (1986). Relation between complications of type I diabetes mellitus and collagen linked fluoresence. N. Eng. J. Med. 314, 405–408.

More, I. A. R. & Briggs, J. D. (1992). The kidneys and urinary tract. In: Muir's Textbook of Pathology (MacSween, R. N. M., ed.), pp. 891–945, Edward Arnold, London.

Myers, E. R., Armstrong, C. G., & Mow, V. C. (1984). Swelling pressure and collagen tension. In: Connective Tissue Matrix, (Hukins, D. W. L., ed.), pp. 161–186, Macmillan, London.

Nermut, M. V., Green, N. M., Eason, P., Yamada, S. S., & Yamada, K. M. (1988). Electron microscopy and structural model of human fibronectin receptor. EMBO J. 7, 4093–4099.

Neugebauer, K. M. & Reichardt, L. F. (1991). Cell-surface regulation of β1-integrin activity on developing neurons. Nature 350, 68–71.

O'Toole, T. E., Mandelman, D., Forsyth, J., Shattil, S. J., Plow, E. F., & Ginsberg, M. H. (1991). Modulation of the affinity of integrin αIIbβ3 (GPIIb-IIIa) by the cytoplasmic domain of αIIb. Science 254, 845–847.

Otey, C. A., Pavalko, F. M., & Burridge, K. (1990). An interaction between α-actinin and the β1 integrin subunit *in vitro*. J. Cell Biol. 111, 721–729.

Pierschbacher, M. D. & Ruoslahti, E. (1984a). The cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. Nature 309, 30–33.

Pierschbacher, M. D. & Ruoslahti, E. (1984b). Variants of the cell recognition site of fibronectin that retain attachment-promoting activity. Proc. Natl. Acad. Sci. USA 81, 5985–5988.

Poole, C. A., Flint, M. H., & Beaumont, B. W. (1987). Chondrons in cartilage. J. Orth. Res. 5, 509–521.

Rezska, A. A., Hayashi, Y., & Horwitz, A. F. (1992). Identification of amino acid sequences in the integrin bl cytoplasmic domain implicated in cytoskeletal association. J. Cell Biol. 117, 1321–1330.

Robins, S. P., Stewart, P., Astbury, C., & Bird, H. A. (1986). Measurement of the cross-linking compound, pyridinoline, in urine as an index of collagen degradation in joint disease. Ann. Rheu. Dis. 45, 969–973.

Ruoslahti, E. & Pierschbacher, M. D. (1987). New perspectives in cell adhesion: RGD and integrins. Science 238, 491–497.

Ryynanen, M., Knowlton, R. G., & Uitto, J. (1991). Mapping of epidermolysis bullosa simplex mutation to chromosome 12. Am. J. Hum. Gen. 49, 978–984.

Schwartz, M. A., Lechene, C., & Ingber, D. E. (1991). Insoluble fibronectin activates the sodium/hydrogen antiporter by clustering and immobilizing integrin α5β1, independent of cell shape. Proc. Natl. Acad. Sci. USA 88, 7849–7853.

Sykes, B. & Smith, R. (1985). Collagen and collagen gene disorders. Quar. J. Med. 56, 533–547.

Tapley, P., Horwitz, A. F., Buck, C. A., Burridge, K., Duggan, K., Hirst, R., & Rohrschneider, L. (1989). Analysis of the avian fibronectin receptor (integrin) as direct substrate for $pp60^{v\text{-}src}$. Oncogene 4, 325–333.

Taylor, J. R. & Twomey, L. T. (1988). The development of the human intervertebral disc. In: The Biology of the Intervertebral Disc, Vol. 1, (Ghosh, P., ed.), pp. 39–82, CRC Press, Boca Raton, FL.

Timpl, R. & Engl, J. (1987). Type VI collagen. In: Structure and Function of Collagen Types, (Mayne, R. & Burgeson, R. E., eds.), pp. 105–143, Academic Press, Orlando, FL.

Tisher, C. C. & Brenner, B. M. (1989). Renal Pathology with Clinical and Functional Correlations. J. B. Lippincott, Philadelphia.

Tuckwell, D. S., Weston, S. A., & Humphries, M. J. (1993). Integrins: a review of their structure and mechanisms of ligand binding. J. Cell Sci. Supplement. in press.

Verrando, P., Hsi, B. L., Yeh, C.-J., Pisani, A., Serieys, N., & Ortonne, J. P. (1988). Monoclonal antibody GB3, a new probe for the study of human basement membranes and hemidesmosomes. Exp. Cell Res. 170, 116–128.

Warskyj, M. & Hukins, D. W. L. (1990). Animal models for osteoarthritis—ensuring experimental validity. Br. J. Rheum. 29, 219–221.

Weisel, J. W., Nagaswami, C., Viliare, G., & Bennett, J. S. (1992). Examination of the platelet membrane glycoprotein IIb-IIa complex and its interaction with fibrinogen and other ligands by electron microscopy. J. Biol. Chem. 267, 16637–16643.

Yamada, K. M. & Kennedy, D. W. (1984). Dualistic nature of adhesive protein function: fibronectin and its biologically active peptide fragments can autoinhibit fibronectin function. J. Cell Biol. 99, 29–36.

Yamada, K. M. & Kennedy, D. W. (1985). Amino acid sequence specificities of an adhesion recognition signal. J. Cell Biochem. 28, 99–104.

FURTHER READINGS

Akiyama, S. K., Nagata, K., & Yamada, K. M. (1990). Cell surface receptors for extracellular components. Biochimica et Biophysica Acta, 1031, 91–110.

Bard, J. B. L. (1990). The role of extracellular matrix in development. In: Connective Tissue Matrix, Part 2, (Hukins, D. W. L., ed.), pp. 11–43, Macmillan, London.

Burgeson, R. E. & Nimni, R. E. (1992). Collagen types. Molecular structure and tissue distribution. Clin. Orth. and Rel. Res. 282, 250–272.

Carney, S. L. & Muir, H. (1988). The structure and function of cartilage proteoglycans. Phys. Rev. 68, 858–910.

Charnley, J. (1979). Low Friction Arthroplasty of the Hip: Theory and Practice, Springer-Verlag, Berlin.

Clark, R. A. F. & Hewson, P. R. (1988). The Molecular and Cellular Biology of Wound Repair. Plenum, New York.

Cottran, R. S., Kumar, V., & Robbins, S.L. (1989). Robbins Pathologic Basis of Disease, 4th ed., W.B. Saunders, Baltimore, MD.

Darnell, J. Lodish, H., & Baltimore, D. (1990). Molecular Cell Biology, 2d ed., W.H. Freeman, New York.

Epstein, E. H. (1992). Molecular genetics of epidermolysis bullosa. Science 256, 799–804.

Freemont, A. J. (1989). Histology of mineralized tissues. In: Calcified Tissue (Hukins, D. W. L., ed.), pp. 21–40, Macmillan, London.

Freemont, A. J. (1993). Basic bone biology. Intl. J. Exp Path. 74, 411–416.

Ghadially, F. N. (1983). Fine Structure of Synovial Joints. Butterworths, London.

Glynne, L. E., Houck, J. C., & Weissman, G. (1989). The Handbook of Inflammation, Vols. 1–5, Elsevier, Amsterdam.

Hay, E. D. (1991). Cell Biology of the Extracellular Matrix, 2d ed., Plenum, New York.

Heinegard, B. & Oldberg, A. (1989). Structure and biology of cartilage and bone matrix noncollagenous macromolecules. FASEB J. 3, 2042–2051.

Hukins, D. W. L. (1988). Disc structure and function. In: The Biology of the Intervertebral Disc, Vol. 1, pp. 161–197 (Ghosh, P., ed.), CRC Press, Boca Raton, FL.

Hukins, D. W. L. (1989). Mineral deposits in tissues. In: Calcified Tissue, pp. 1–19 (Hukins, D. W. L., ed.), Macmillan, London.

Kielty, C. M., Hopkinson, I., & Grant, M. E. (1993). The collagen family: Structure, assembly, and organization in the extracellular matrix. In: Connective Tissue and its Heritable Disorders, Molecular, Genetic and Medical Aspects, pp. 139–147 (Royce, P. M. & Steinmann, B. eds.), Wiley-Liss, New York.

Kjellen, L. & Lindahl, U. (1991). Proteoglycans: Structure and interactions. Ann. Rev. Biochem 60, 443–475.

Mosher, D. F., Sottile, J., Wu, C., & McDonald, J. A. (1992). Assembly of the extracellular matrix. Curr. Op. Cell Bio. 4, 810–818.

Paulsson, M. (1992). Basement membrane proteins: Structure, assembly and cellular interactions. Crit. Rev. in Biochem. and Mol. Bio. 27, 93–127.

Revel, P. A. (1986). Pathology of Bone. Springer-Verlag, Berlin.

Royce, P. M. & Steinmann, B. (1993). Connective Tissue and its Heritable Disorders, Molecular, Genetic, and Medical Aspects. Wiley-Liss, New York.

Scott, J. T. (1978). Copeman's Textbook of the Rheumatic Diseases, 5th ed., Churchill-Livingston, Edinburgh.

Uitto, J. & Christiano, A. M. (1992). Molecular genetics of the cutaneous basement membrane zone. Perspectives on epidermolysis bullosa and other blistering skin diseases. J. Clin. Inves. 90, 687–692.

Weiss, J. B. (1984). Collagens and collagenolytic enzymes. In: Connective Tissue Matrix, pp. 17–53 (Hukins, D. W. L., ed.), Macmillan, London.

Whaley, K. & Burt, A. D. (1992). Inflammation, healing, and repair. In: Muir's Textbook of Pathology, 13th ed., pp. 112–165 (MacSween, R. N. M., ed.), Edward Arnold, London.

Wight, T. N., Kinsella, M. G., & Quarnstrom, E. E. (1992). The role of protoglycans in cell adhesion, migration and proliferation. Curr. Op. Cell Bio. 4, 793–801.

Winlove, C. P. & Parker, K. H. (1990). Physicochemical properties of vascular elastin. In: Connective Tissue Matrix, Part 9, pp. 167–198 (Hukins, D. W. L., ed.), Macmillan, London.

Chapter 9

Posttranslational Processing of Collagens

KARI I. KIVIRIKKO

INTRODUCTION

Collagens are a family of extracellular matrix proteins that play a dominant role in maintaining the structural integrity of various tissues. Their biosynthesis is characterized by the presence of an unusually large number of cotranslational and posttranslational modifications, many of which are unique to collagens and a few other proteins with collagen-like amino acid sequences. The aim of this chapter is

Principles of Medical Biology, Volume 3
Cellular Organelles and the Extracellular Matrix, pages 233–254

ISBN: 1-55938-804-8

to review current information on the occurrence and functions of these modifications with special emphasis on the processing reactions and the enzymes involved. The literature on the topics covered in this chapter is now voluminous, and it is therefore not possible to cite even many of the key articles. For more detailed accounts and more complete references, the reader is referred to several reviews, as indicated in various parts of this article.

GENERAL FEATURES OF THE MODIFICATIONS

The molecule of type I collagen, the most abundant collagen, is a long, thin rod. It consists of three polypeptide chains called α chains, two identical α1(I) chains and one α2(I) chain, which are coiled around one other into a triple-helical conformation similar to a three-stranded rope. This molecule is synthesized in the form of a precursor which differs from the final molecule in that it has propeptide extensions at both its N- and C-terminal ends. This precursor is called the type I procollagen molecule, and its polypeptide chains are correspondingly proα chains (Figure 1). Every third amino acid in each of the α chains of the collagen domain is glycine, with the exception of short sequences at the ends of the α chains, and thus the molecular formula of an α chain can be approximated as $(Gly\text{-}X\text{-}Y)_n$, in which X and Y represent amino acids other than glycine.

Nineteen collagen types, containing altogether more than 30 genetically distinct α chains have now been identified. All collagen molecules are composed of three α chains in a triple-helical conformation. In some collagen types all three α chains of the molecule are identical while in others the molecule contains two or three different α chains. All collagen molecules form supramolecular aggregates that are stabilized in part by interactions between the triple-helical domains. Types I, II, III, V and XI form fibrils (Figure 2), and are hence called fibrillar collagens. Examples of other supramolecular aggregates are sheets constituting basement membranes in the case of type IV collagen, and laterally aggregated antiparallel dimers, which are the main components of anchoring fibrils in the case of type VII collagen (for recent

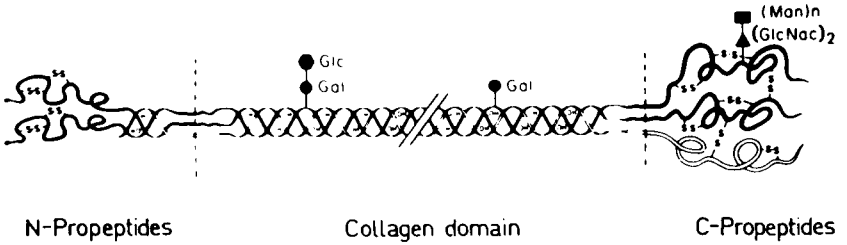


Figure 1. Schematic representation of the structure of the type I procollagen molecule. Glc, glucose; Gal, galactose; Man, mannose; and GlcNac, N-acetylglucosamine. Reproduced from Prockop and Kivirikko (1984) with permission.

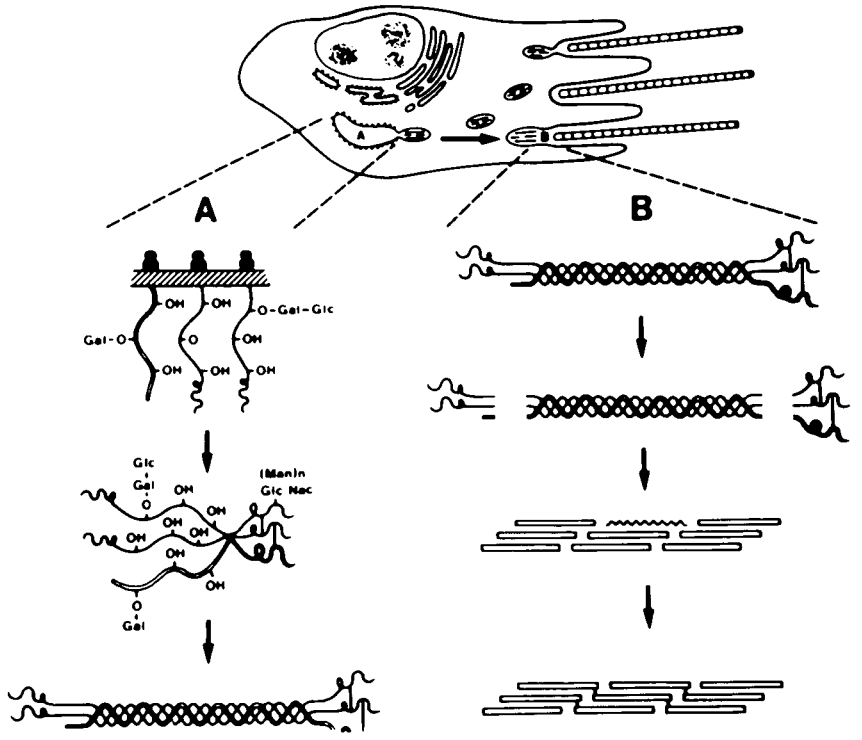


Figure 2. Intracellular (A) and extracellular (B) events in the biosynthesis of a fibrillar collagen. The intracellular modifications include hydroxylation of certain proline and lysine residues to 4-hydroxyproline, 3-hydroxyproline and hydroxylysine; glycosylation of some of the hydroxylysine residues to galactosylhydroxylysine and glucosylgalactosylhydroxylysine; glycosylation of certain asparagine residues in the C-propeptides; and zipper-like folding of the triple helix from the C-terminus towards the N-terminus. Cleavage of the signal peptide is not shown. The extracellular events include cleavage of the N- and C-propeptides from the procollagen molecule; self-assembly of the collagen molecules into fibrils by nucleated growth, and formation of covalent crosslinks. Reproduced from Prockop and Kivirikko (1984) with permission.

reviews on collagen structure and types, see Mayne and Burgeson, 1987; Fleischmajer et al., 1990; van der Rest and Garrone, 1991; Hulmes, 1992; Kielty et al., 1993; Kivirikko, 1993).

The posttranslational processing of a fibrillar collagen can be regarded as taking place in two stages (Figure 2). Intracellular cotranslational and posttranslational modifications, together with synthesis of the proα chains, result in the formation of triple-helical procollagen molecules, and extracellular processing then converts these molecules into collagens and incorporates the collagen molecules into stable,

Table 1. Posttranslational Events in the Biosynthesis of a Fibrillar Collagen and Their Main Functions

Events	*Biological Significance*
Intracellular	
Cleavage of signal peptide	Translocation across membrane
4-Hydroxylation of proline	Essential for triple helix formation at 37°
3-Hydroxylation of proline	Unknown
Hydroxylation of lysine	Essential for glycosylation of hydroxylysine Essential for stable crosslinks
Glycosylation of hydroxylysine	Influences fibril diameter May influence the rate of cleavage of N-propeptides
Glycosylation of propeptides	Unknown
Chain association and disulfide bonding	Essential for triple helix formation
Triple helix formation	Essential for a normal rate of secretion Essential for a functional protein
Extracellular	
Cleavage of N-propeptides	Essential for normal fibril morphology
Cleavage of C-propeptides	Essential for fibril formation
Fibril assembly	Formation of fibrils
Crosslink formation	Essential for stable fibrils

crosslinked fibrils. The posttranslational processing of a nonfibrillar collagen is basically similar but in many cases of the N- and/or C-terminal noncollagenous domains are not removed, and hence these domains are not called propeptides. Also, the supramolecular assemblies formed are not fibrils. The posttranslational events involved in the biosynthesis of a fibrillar collagen and their main biological functions are outlined in Table 1.

The posttranslational processing of collagens requires at least nine specific enzymes and several nonspecific ones (Table 2). One of the specific enzymes, procollagen N-proteinase, has two isoenzymes, one which cleaves type I and type II procollagens and possibly also others, but not type III, and another which appears to be specific to type III procollagen. In the case of all the other specific modifications a single enzyme appears to act on all the collagen types. The posttranslational enzymes also require several cofactors or cosubstrates (Table 2), as will be discussed in more detail in connection with the individual enzymes.

Almost all the specific posttranslational processing reactions show an unusual relationship to the conformation of the collagen being modified. The α chains must be nontriple-helical in order to be modified by any of the five specific intracellular enzymes, i.e., the three collagen hydroxylases and the two hydroxylysine glycosyltransferases (Figure 2). Once the α chains fold into the triple helix, processing by all these five enzymes ceases, and the time available before folding in fact limits the extent to which the specific intracellular modifications can proceed. Once folded, the protein is secreted from the cell, and the time required for folding to occur directly influences the time needed for secretion. Most of the extracellular

Table 2. Specific Posttranslational Enzymes of Collagen Synthesis

Enzyme	*Molecular weight* ($\bullet 10^{-3}$)	*Subunit composition*	*Cofactors and cosubstrates*
Intracellular			
Prolyl 4-hydroxylase	240	$\alpha_2\beta_2$	Fe^{2+}, 2-oxoglutarate, O_2, ascorbate[a]
Prolyl 3-hydroxylase	160	nd[b]	Fe^{2+}, 2-oxoglutarate, O_2, ascorbate[a]
Lysyl hydroxylase	180	α_2	Fe^{2+}, 2-oxoglutarate, O_2, ascorbate[a]
Hyl-galtransferase[c]	450, 200	nd	Mn^{2+}, UDP-galactose
Gal-hyl-glctransferase[c]	70	α	Mn^{2+}, UDP-glucose
Extracellular			
Procollagen I N-proteinase	500	$\alpha\beta\gamma\zeta$	Ca^{2+}
Procollagen III N-proteinase	70	α	Ca^{2+}
Procollagen C-proteinase	100	α	Ca^{2+}
Lysyl oxidase	30	α	Cu^{2+}, O_2, quinone

Notes: [a]Ascorbate is consumed only after uncoupled decarboxylation cycles (see text).
[b]Nd = not determined.
[c]Hyl-galactosyltransferase = hydroxylysyl galactosyltransferase; gal-hyl-glctransferase = galactosylhydroxylysyl glucosyltransferase.

events also demonstrate strict requirements concerning the conformation of the collagen. The protein must be triple-helical in order to be cleaved by a procollagen N-proteinase, and it must be triple-helical and incorporated into native-type fibrils or other supramolecular aggregates in order to serve as a substrate for lysyl oxidase (Figure 2). For detailed accounts of the general features of collagen posttranslational processing, the reader is referred to reviews by Kivirikko and Myllylä (1985), Kivirikko and Kuivaniemi (1987), Olsen (1991) and Kielty et al. (1993).

INTRACELLULAR MODIFICATIONS

All five specific intracellular enzymic reactions probably occur within the cisternae of the rough endoplasmic reticulum (Figure 2). Most of the hydroxylation of proline and lysine residues and the glycosylation of hydroxylysine residues takes place in the form of cotranslational modifications while the nascent polypeptide chains are being assembled on the ribosomes. All these reactions are continued, however, as posttranslational modifications after the release of the complete polypeptide chains from the ribosomes, until triple helix formation prevents any further intracellular processing. The signal sequences are already cleaved during or shortly after translocation across the membrane. Certain asparagine residues in the propeptides of fibrillar collagens and noncollagenous domains in the other collagens are probably likewise glycosylated as a cotranslational modification, whereas the interchain disulfide bonds are probably not formed until the polypeptide chains have been released from the ribosomes into the cisternae of the rough endoplasmic reticulum.

Hydroxylation of Proline and Lysine Residues

Occurrence and Functions of 4-Hydroxyproline, 3-Hydroxyproline and Hydroxylysine

The vast majority of the 4-hydroxyproline, 3-hydroxyproline and hydroxylysine contained in animal proteins is to be found in various collagens. 4-Hydroxyproline, and hydroxylysine in most but not all cases, are also found in the collagen-like domains of a few other proteins: the C1q subcomponent of complement, acetylcholinesterase, pulmonary surfactant apoproteins, mannose-binding proteins, types I and II macrophage scavenger receptors, and bovine conglutin. 4-Hydroxyproline, but not hydroxylysine, is also present in elastin, while 3-hydroxyproline has not been identified in any protein except collagens. A single 4-hydroxyproline residue is found in hydroxyproline-lysyl-bradykinin, and hydroxyproline luteinizing-releasing hormone, while a single hydroxylysine residue is found in anglerfish somatostatin (for reviews, see Kivirikko et al., 1989, 1992).

The $(Gly\text{-}X\text{-}Y)_n$ sequences of the $\alpha 1(I)$ and $\alpha 2(I)$ chains of type I collagen contain about 1000 amino acid residues. In these sequences about 100 of the X position amino acids are proline and about 100 of the Y position amino acids are 4-hydroxyproline. The 4-hydroxyproline content shows only relatively small, though distinct, variations between collagen types, whereas the 3-hydroxyproline and hydroxylysine contents vary markedly, ranging from 0 to over 10 per 1000 amino acids in the case of 3-hydroxyproline, and from about 5 to 70 in that of hydroxylysine. Additional variations in the content of the three amino acids are seen within the same collagen type from different tissues, and even in the same tissues in different physiological and pathological states.

4-Hydroxyproline has essentially been identified in vertebrate proteins exclusively in the Y positions of the repeating, or in some cases single, -X-Y-Gly- sequences (the repeating -Gly-X-Y- sequences are written here in the form -X-Y-Gly- due to the properties of the three hydroxylases, see below). Hydroxylysine is likewise found almost exclusively in the Y positions, while 3-hydroxyproline has been identified only in the sequence -3Hyp-4Hyp-Gly-. In just a few exceptional cases, however, 4-hydroxyproline has been identified in a sequence -X-4Hyp-Ala-, and hydroxylysine in a sequence -X-Hyl-Ala- or -X-Hyl-Ser- (see Kivirikko et al., 1992).

The function of the hydroxy group of the 4-hydroxyproline residues is to stabilize the collagen triple helix under physiological conditions. Nonhydroxylated collagen α chains can fold into the triple helix at low temperatures, but the midpoint of the thermal transition from helix to coil (T_m) in the case of type I procollagen molecules consisting of nonhydroxylated polypeptide chains is only 24°, about 15° lower than that for molecules consisting of fully hydroxylated polypeptide chains (i.e. chains in which all the Y position prolines have been converted to

4-hydroxyproline). Thus nonhydroxylated α chains cannot form triple-helical molecules *in vivo*, and an almost complete 4-hydroxylation of proline residues in the Y positions of the -Gly-X-Y- triplets is needed for the formation of a molecule that is stable at 37°. The means by which 4-hydroxyproline stabilizes the triple helix is not known, however. The most likely possibility is that the hydroxy group forms an additional bond, via a water molecule, either between two polypeptide chains or within the same chain in the triple helix (see Kivirikko et al., 1992).

The hydroxylysine residues of collagens have two important functions: their hydroxy groups act as sites of attachment for carbohydrate units (see glycosylation of hydroxylysine residues, below), and they are essential for the stability of the intermolecular collagen crosslinks (see below). The significance of hydroxylysine is clearly demonstrated by the profound changes in the mechanical properties of certain tissues that are seen in patients with a genetic deficiency in lysyl hydroxylase activity (see Prockop and Kivirikko, 1984; Kivirikko et al., 1992; Kivirikko, 1993).

The function of the 3-hydroxyproline residues is unknown at present.

Molecular Properties of the Three Collagen Hydroxylases

The hydroxylations of proline and lysine residues are catalyzed by three separate enzymes, prolyl 4-hydroxylase (EC 1.14.11.2), prolyl 3-hydroxylase (EC 1.14.11.7) and lysyl hydroxylase (EC 1.14.11.4). Prolyl 4-hydroxylase is an $\alpha_2\beta_2$ tetramer and lysyl hydroxylase an α_2 dimer, while the subunit structure of prolyl 3-hydroxylase is currently unknown (Table 2). Complete cDNA-derived amino acid sequences have now been reported for the α subunit (Helaakoski et al., 1989; Bassuk et al., 1989) and β subunit (Pihlajaniemi et al., 1987; Parkkonen et al., 1988; Kao et al., 1988) of human and chick prolyl 4-hydroxylase and for human (Hautala et al., 1992) and chick (Myllylä et al., 1991) lysyl hydroxylase. A highly surprising feature is that no significant homology is found between the primary structures of lysyl hydroxylase and the two types of subunit of prolyl 4-hydroxylase in spite of the marked similarities in the catalytic properties between these two enzymes.

Various binding studies indicate that the α subunits contribute most parts to the catalytic sites of the prolyl 4-hydroxylase tetramer, but some parts of the large catalytic sites may be cooperatively built up of both α and β subunits. The binding studies further indicate that the $\alpha_2\beta_2$ enzyme tetramer has two catalytic sites, i.e., one site per pair of dissimilar subunits (see Kivirikko et al., 1989, 1990, 1992).

Molecular cloning and nucleotide sequencing of the β subunit of prolyl 4-hydroxylase indicated, surprisingly, that this polypeptide is identical to the enzyme protein disulfide isomerase (PDI) (Pihlajaniemi et al., 1987; Parkkonen et al., 1988). Moreover, the β subunit was found to have PDI activity even when present in the native prolyl 4-hydroxylase tetramer (Koivu et al., 1987). PDI (EC 5.3.4.1) catalyzes rearrangements of disulfide bonds in various proteins *in vitro* and is likely to be the *in vivo* catalyst of disulfide bond formation in the biosynthesis of

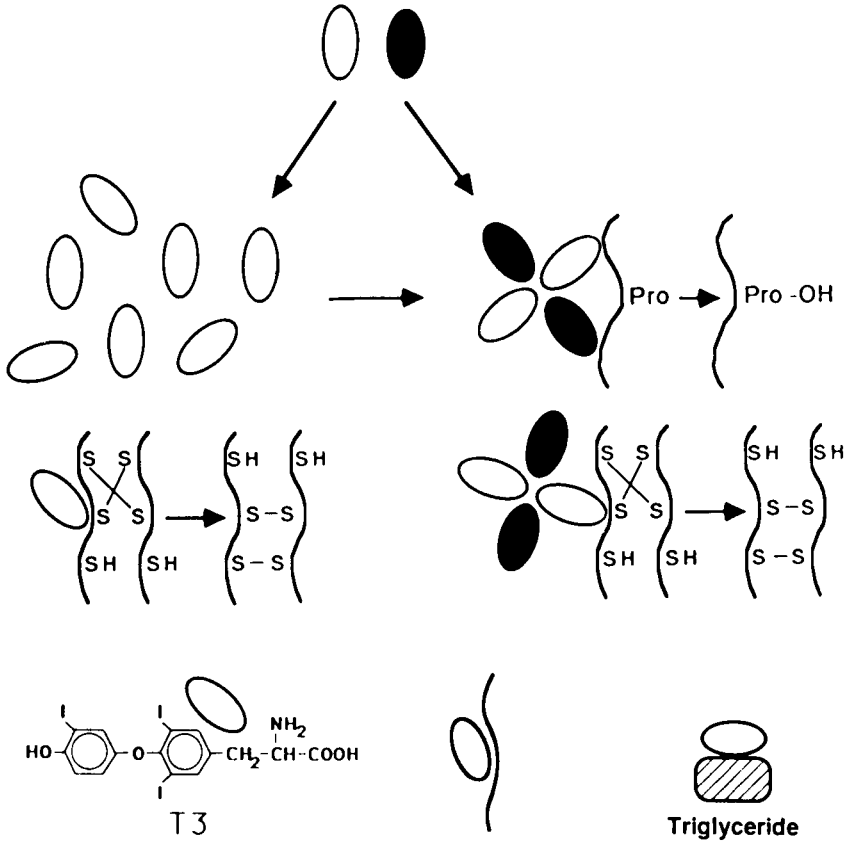


Figure 3. Functions of the multifunctional β subunit of prolyl 4-hydroxylase. The β subunit (open oval) is synthesized in an excess over the α subunit (black oval) and enters a β subunit pool before being incorporated into the prolyl 4-hydroxylase tetramer. The α subunit is incorporated into the tetramer immediately after synthesis. The β subunit has PDI activity both when present as a monomer and when present in the prolyl 4-hydroxylase tetramer. In the latter case it also participates in the catalysis of the 4-hydroxylation of proline residues (Pro → Pro-OH). The β subunit also acts as a major cellular thyroid hormone (T_3) binding protein; an endoplasmic reticulum luminal polypeptide uniquely binding various peptides; and as the smaller subunit of the microsomal triglyceride transfer protein dimer. Modified from Kivirikko et al. (1989) and (1990).

various secretory and cell surface proteins (for reviews, see Freedman, 1989; Noiva and Lennarz, 1992), including collagens (see other intracellular modifications, below). Work on both the cDNA and protein level thus demonstrates that a single polypeptide is involved in two entirely different enzyme reactions (Figure 3). The polypeptide has PDI activity both when present as a monomer and when incorpo-

rated into the prolyl 4-hydroxylase tetramer, and in the latter case it also participates in catalysis of the 4-hydroxylation of proline residues. Previous research had already demonstrated that a number of cell types contain, in addition to the prolyl 4-hydroxylase tetramer, large amounts of a protein that is identical to the β subunit monomer on a number of criteria (see Kivirikko et al., 1989, 1990, 1992). It is now evident that one function of this protein is to act as the enzyme PDI.

The PDI activity of the PDI/β subunit is not directly involved in the mechanism by which the prolyl 4-hydroxylase reaction is catalyzed. This finding is based on recent expression data for a recombinant human prolyl 4-hydroxylase tetramer in insect cells (Vuori et al., 1992 a,b). The PDI/β subunit polypeptide has two -Cys-Gly-His-Cys- sequences which represent two independently acting catalytic sites for the isomerase activity. When both these sequences were modified to -Ser-Gly-His-Cys-, the polypeptide had no PDI activity but still associated with the α subunit to form the $\alpha_2\beta_2$ tetramer, and this tetramer proved to be fully active prolyl 4-hydroxylase (Vuori et al., 1992b). These data are in agreement with the finding that lysyl hydroxylase, an α_2 dimer with a reaction mechanism similar to that of prolyl 4-hydroxylase, has no protein disulfide isomerase activity (see Kivirikko et al., 1989, 1992).

Recent reports indicate that the cellular PDI/β subunit polypeptide may have at least three additional functions (Figure 3). One of these is to serve as a major cellular thyroid hormone binding protein present in the endoplasmic reticulum, a second is to act as an endoplasmic reticulum luminal polypeptide uniquely binding various peptides, and a third is to serve as the smaller subunit of the microsomal triglyceride transfer protein dimer (see Freedman, 1989; Noiva et al., 1991; Wetterau et al., 1991; Noiva and Lennarz, 1992; Kivirikko et al., 1992). The PDI/β subunit thus appears to be a highly unusual multifunctional polypeptide.

Catalytic Properties of the Three Collagen Hydroxylases

Prolyl 4-hydroxylase, prolyl 3-hydroxylase and lysyl hydroxylase all hydroxylate proline and lysine residues only in peptide linkages (Figure 4). The minimum sequence requirement for hydroxylation by a vertebrate prolyl 4-hydroxylase is an -X-Pro-Gly- triplet, and that for interaction with lysyl hydroxylase an -X-Lys-Gly- triplet, although both enzymes may also hydroxylate certain other triplets in exceptional cases. These data agree with the occurrence of 4-hydroxyproline and hydroxylysine in collagens almost exclusively in the Y positions of -X-Y-Gly- sequences (see above). Prolyl 3-hydroxylase appears to require a -Pro-4Hyp-Gly- sequence, whereas a -Pro-Pro-Gly- sequence is probably not hydroxylated by this enzyme. The interaction with prolyl 4-hydroxylase and lysyl hydroxylase is further influenced by the amino acid in the X position of the -X-Y-Gly- triplet that is hydroxylated, and the interaction with all three enzymes is affected by other nearby amino acids, the peptide chain length and the peptide conformation. Longer

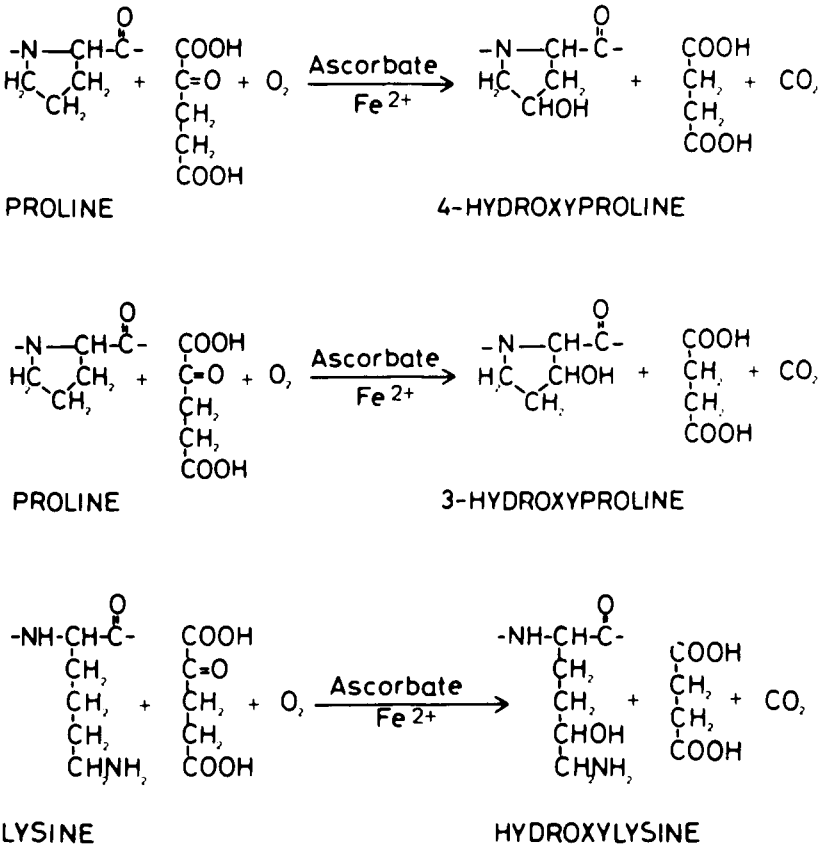


Figure 4. Reactions catalyzed by prolyl 4-hydroxylase, prolyl 3-hydroxylase and lysyl hydroxylase.

peptides are better substrates for all three enzymes than shorter ones, and the triple-helical conformation prevents all three hydroxylations, indicating that these must occur in the cell before triple helix formation (see Kivirikko et al., 1989, 1992).

All three enzymes require Fe^{2+}, 2-oxoglutarate, O_2 and ascorbate (Figure 4). The 2-oxoglutarate is stoichiometrically decarboxylated during hydroxylation, with one atom of the O_2 molecule being incorporated into the succinate and the other into the hydroxy group formed on the proline or lysine residue. Ascorbate is a highly specific requirement, but it is not consumed stoichiometrically, and the enzymes can catalyze their reactions for a number of catalytic cycles in the absence of ascorbate. The reaction requiring ascorbate is an uncoupled decarboxylation of 2-oxoglutarate, i.e. decarboxylation without subsequent hydroxylation. Collagen hydroxylases catalyze such uncoupled decarboxylation events at a low rate even in

the presence of an optimal concentration of their polypeptide substrates, and therefore the main biological function of ascorbate in the collagen hydroxylase reactions, both *in vitro* and *in vivo*, is probably that of serving as an alternative oxygen acceptor in the uncoupled reaction cycles.

Inhibitors of Prolyl 4-Hydroxylase, Implications for the Development of Antifibrotic Drugs

The central role of collagens in fibrosis of the liver, lungs and other organs has prompted attempts to develop drugs that inhibit their accumulation. Such drugs can naturally act by modulating the synthesis of collagen polypeptide chains, but so far no compound has been found that does this without also influencing the synthesis of a number of other proteins. It seems evident, however, that inhibition of some of the specific posttranslational events will also lead either to reduced formation of extracellular collagen fibers or to an accumulation of fibers with altered properties. The 4-hydroxylation of proline residues would seem a particularly suitable target for antifibrotic therapy, as inhibition of this reaction will prevent collagen triple helix formation and thus lead to a nonfunctional protein that is rapidly degraded.

A large number of compounds are now known that inhibit prolyl 4-hydroxylase competitively with respect to its polypeptide substrate or some of its cosubstrates. In addition, three groups of suicide inhibitors have been characterized. Particular interest has been focused in recent years on pyridine 2,4-dicarboxylate, pyridine 2,5-dicarboxylate (Figure 5) and their derivatives (Hanauske-Abel and Günzler, 1982; Majamaa et al., 1984). These compounds are highly effective competitive

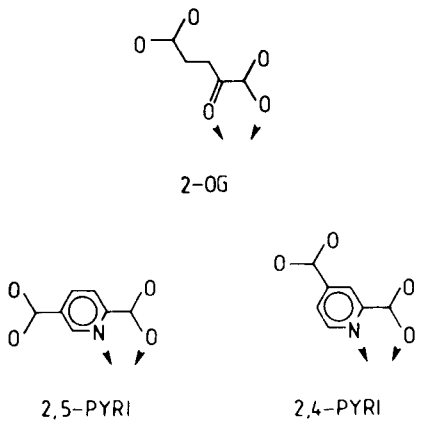


Figure 5. Structures of pyridine 2,5-dicarboxylate (2,5-pyri) and pyridine 2,4-dicarboxylate (2,4-pyri), compounds effectively inhibiting prolyl 4-hydroxylase competitively with respect to 2-oxoglutarate (2-Og). The compounds bind to the iron at a catalytic site of the enzyme in the manner indicated by the arrowheads.

inhibitors of prolyl 4-hydroxylase with respect to 2-oxoglutarate, but are only very weak inhibitors of 2-oxoglutarate dehydrogenase. Lipophilic proinhibitor derivatives of these compounds, such as pyridine 2,4-dicarboxylic acid di-(methoxyethyl)amide (HOE 077, Hoechst), do not in themselves inhibit the pure enzyme, but they readily pass through cell membranes and become converted to the active inhibitor inside the cell (Hanauske-Abel, 1991). HOE 077 has been shown to selectively inhibit collagen accumulation in the liver of rats with two models of hepatic fibrosis (see Kivirikko, 1993), and it now seems possible that some of the prolyl 4-hydroxylase inhibitors could be developed into effective drugs for the treatment of patients with fibrotic disorders.

It should be noted, however, that few detailed data are currently available on the structures of the large catalytic sites of prolyl 4-hydroxylase. Recent success in the expression of an active recombinant human enzyme tetramer in insect cells (Vuori et al., 1992a) has made it possible to begin experiments on site-directed mutagenesis of the enzyme (Vuori et al., 1992b) and attempts to mass produce the enzyme for crystallization. These should make it possible to obtain detailed information on the structures of its catalytic sites, which in turn should help the development of more effective inhibitors of the enzyme and thus the design of antifibrotic drugs.

Glycosylation of Hydroxylysine Residues

Collagens characteristically contain carbohydrate units linked to hydroxylysine residues. Since hydroxylysine is found only in collagens and a few other proteins with collagen-like amino acid sequences, the occurrence of such carbohydrate units is limited to these proteins. Some of the carbohydrate is present as the monosaccharide galactose and some as the disaccharide glucosylgalactose, these being the only carbohydrate prosthetic groups present in the collagen domains of various collagens. The extent of glycosylation of hydroxylysine residues varies markedly between collagen types, and even within the same collagen type in various physiological and pathological states. An additional variation is found in the ratio of the monosaccharide to the disaccharide, as almost all the carbohydrate is found in the form of the disaccharide in some collagen types while the predominant form is the monosaccharide in others (for a review, see Kivirikko and Myllylä, 1979).

The functions of the hydroxylysine-linked carbohydrate units have not been fully established. As these are the most extrusive groups on the surface of the collagen molecule, it has been suggested that they may regulate the packing of collagen molecules into supramolecular assemblies. In agreement with this suggestion, an inverse relationship has been found between carbohydrate content and fibril diameter. Furthermore, recent experiments on *de novo* generation of type I collagen fibrils from type I procollagen by enzymic cleavage of the propeptides have demonstrated that this collagen with a high degree of lysine hydroxylation and

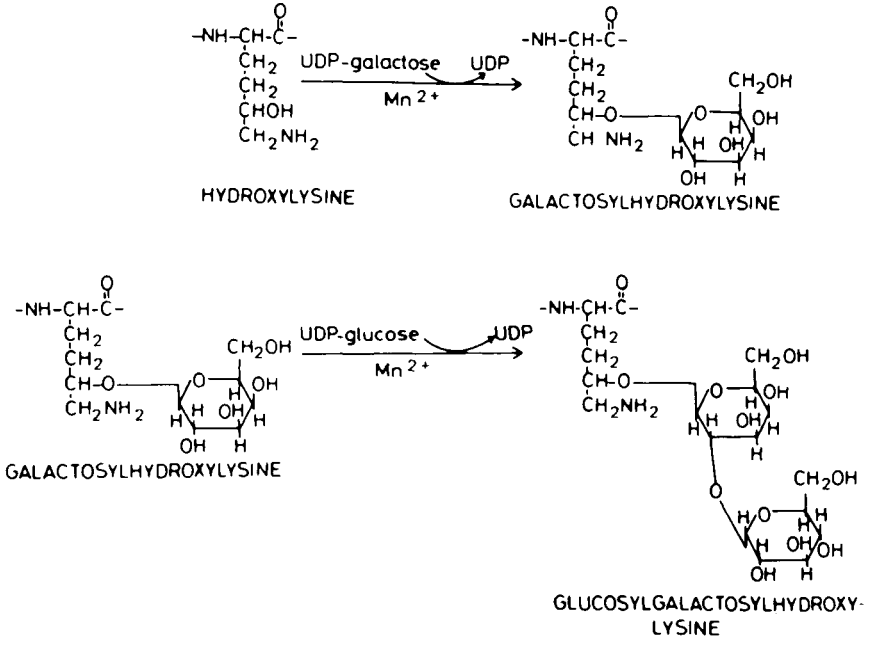


Figure 6. Reactions catalyzed by hydroxylysyl galactosyltransferase and galactosylhydroxylysyl glucosyltransferase.

hydroxylysine glycosylation forms thinner fibrils than the same protein with a normal degree of these modifications (Torre-Blanco et al., 1992). In addition, the overmodified type I procollagen used in these experiments was found to be cleaved by type I procollagen N-proteinase at a reduced rate, suggesting that the extent of collagen glycosylation may also influence interaction with this enzyme.

The formation of the hydroxylysine-linked carbohydrate units is catalyzed by two specific enzymes, first hydroxylysyl galactosyltransferase (EC 2.4.1.50), transferring galactose to some of the hydroxylysine residues, and then galactosylhydroxylysyl glucosyltransferase (EC 2.4.1.66), transferring glucose to some of the galactosylhydroxylysine residues (Figure 6). The glucosyltransferase is an enzyme monomer, whereas the subunit structure of the galactosyltransferase is currently unknown. These two enzymes and their reactions have been extensively characterized in the past (see Kivirikko and Myllylä, 1979), but little new information is available on their properties, and neither of the enzymes has been cloned so far.

Hydroxylysine is formed only as a cotranslational and posttranslational modification, and hence it is apparent that formation of the hydroxylysine-linked carbohydrate units *in vivo* occurs only on a peptide substrate. The free ε-amino group of

the hydroxylysine residue is an absolute requirement for both transferase reactions. Interaction with both enzymes is further influenced by the amino acid sequence of the peptide, the peptide chain length, and the peptide conformation. As in the cases of the three collagen hydroxylases, the triple-helical conformation prevents both glycosylations.

The preferential carbohydrate donor for both glycosyltransferases is the corresponding UDP glycoside, and both reactions require a divalent cation, this requirement being best fulfilled by Mn^{2+} (Figure 6). Several bivalent cations, the reaction product UDP, and certain other nucleotides inhibit both enzymes.

Other Intracellular Modifications

The nascent collagen polypeptide chains have signal peptides similar to those in most secreted proteins. The signal peptides are cleaved as in other proteins.

The propeptides of precursors of fibrillar collagens and the noncollagenous domains of some other collagen types contain asparagine-linked carbohydrate units. These are synthesized as in other proteins, the consensus sequence for asparagine glycosylation being an -Asn-X-Ser/Thr- triplet. Some collagens also contain carbohydrate units of other types in their noncollagenous domains, and some contain other posttranslationally formed units, such as tyrosine sulfate.

The propeptides of procollagens of the fibrillar collagen types and the noncollagenous domains of several other collagens contain both intrachain and interchain disulfide bonds, while interchain disulfide bonds are also found between the collagen domains in some collagen types. All these disulfide bonds are formed as in other secretory proteins, probably catalyzed by the enzyme protein disulfide isomerase (PDI, see above).

The function of intrachain and interchain disulfide bonds in proteins is to provide covalent linkages which stabilize the structures. A further important function of the interchain disulfide bonds with respect to the propeptides of procollagens is to facilitate rapid triple helix formation. Chain association in most collagens begins with hydrophobic and electrostatic interactions among the C-propeptides or C-terminal noncollagenous domains, these interactions being subsequently stabilized by the formation of interchain disulfide bonds. The triple helix is then formed from the C-terminus towards the N-terminus in a zipper-like manner, the rate being limited by the *cis-trans* isomerization of peptide bonds. The triple-helical molecules are then secreted from the rough endoplasmic reticulum through the Golgi complex before leaving the cell.

Triple helix formation is an absolute requirement for the normal rate of collagen secretion. If it is prevented, e.g. by the administration of inhibitors of prolyl 4-hydroxylase, random-coil polypeptide chains will first accumulate within the cisternae of the endoplasmic reticulum and will then be in part degraded and in part

secreted as a nonfunctional protein at a reduced rate (for references on these other intracellular events, see Kivirikko and Myllylä, 1985; Kivirikko and Kuivaniemi, 1987; Engel and Prockop, 1991; Olsen, 1991; Kielty et al., 1993).

EXTRACELLULAR MODIFICATIONS

Cleavage of the Propeptides

The first extracellular processing reactions in the biosynthesis of a fibrillar collagen involve cleavage of the propeptides. The function of these modifications is to permit formation of the native type collagen fibers. Removal of the C-propeptides appears to be an absolute requirement for fibril formation, whereas partially processed molecules with intact N-propeptides will form thin, irregular fibrils with diminished mechanical strength (see Prockop and Hulmes, 1994).

Cleavage of the propeptides from a procollagen molecule (Figure 7) requires two specific enzymes, a procollagen N-proteinase (EC 3.4.24.14) and a procollagen C-proteinase (EC 3.4.24). The actual number of enzymes is in fact larger, as the N-proteinase has collagen type-specific forms (Table 2). Procollagen type III N-proteinase and procollagen C-proteinase are both enzyme monomers, whereas the type I N-proteinase appears to have four different subunits which have not been characterized (see Hojima et al., 1989; Olsen, 1991). None of these enzymes has been cloned so far.

The procollagen proteinases are all endopeptidases that operate at a neutral pH and require a divalent cation such as Ca^{2+}. The type I procollagen N-proteinase cleaves type I and type II procollagens and possibly also others, but not type III, while the other N-proteinase appears to be specific to type III procollagen. The bond cleaved by the type I N-proteinase is either Pro-Gln or Ala-Gln in a number of collagens, the type III N-proteinase also cleaving a Pro-Gln bond. The C-proteinase cleaves at least an Ala-Asp, Gly-Asp and Arg-Asp bond. The procollagen N-proteinases have an absolute requirement of a triple-helical conformation in their substrates, whereas the C-proteinase has been reported to act on both native and denatured procollagens (for reviews, see Kivirikko and Myllylä, 1985; Kivirikko and Kuivaniemi, 1987; Olsen, 1991).

There appears to be no obligatory sequence for cleavage of the propeptides from a procollagen molecule (Figure 7). Nevertheless, the synthesis of type I procollagen in many systems involves removal of the N-propeptides before the C-propeptides. Some processing occurs in the opposite order in the biosynthesis of type II procollagen, while in the case of type III procollagen the N-propeptides are cleaved very slowly. This explains why collagen fibers in many tissues contain significant amounts of partially processed type III molecules with intact N-propeptides.

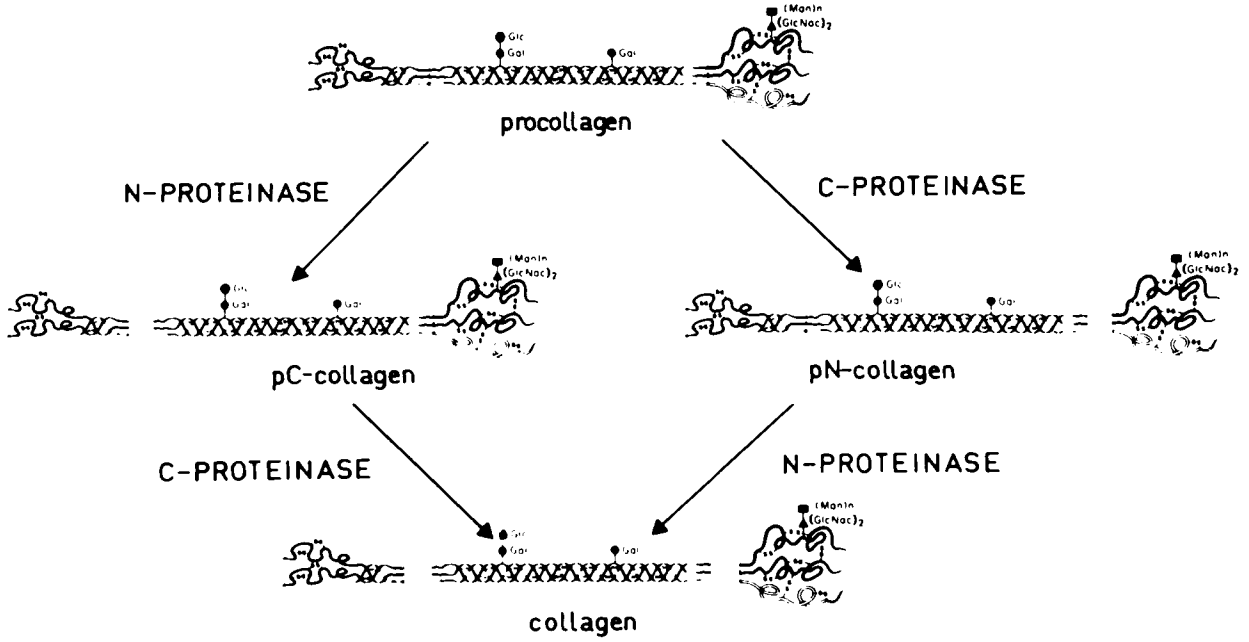


Figure 7. Cleavage of the propeptides in the biosynthesis of a fibrillar collagen. There is no obligatory sequence for the removal of the N- and C-propeptides from a procollagen molecule, but most of the processing of type I and II procollagens occurs as shown in the left-hand pathway, whereas type III procollagen is mainly processed as shown in the right-hand pathway.

Conversion to smaller forms also appears to take place in the biosynthesis of some nonfibrillar collagens but little is currently known about the enzymology of such processing reactions.

Ordered Aggregation

The molecules of the fibrillar collagens produced by removal of the propeptides have a marked tendency for self-assembly and spontaneous fibril formation. This process does not require any enzymes or other factors, and it will occur easily in collagen solutions *in vitro*. Such *in vitro* assembly takes place according to the principles of nucleation and growth, in which the rate-limiting step is the formation of a critical nucleus for a small number of molecules. Once this nucleus has been formed, growth proceeds rapidly by aggregation of collagen molecules (for reviews, see Hulmes, 1992; Prockop and Hulmes, 1994).

A system has recently been developed for studying the *de novo* formation of fibrils by initiating the process by enzymatic cleavage of procollagen using purified N- and C-proteinases. This system has made it possible to demonstrate that the fibrils grow from symmetrical pointed tips. Such experiments have further demonstrated that type I collagen copolymerizes with partially processed type I or type III procollagen containing the N-propeptide and that copolymerization can affect both the diameter of the fibrils and their shape (see Prockop and Holmes, 1994). In fact, it is now well documented that most collagen fibrils *in vivo* are composed of two or more different collagen types (see van der Rest and Garrone, 1991; Hulmes, 1992).

The nonfibrillar collagens assemble into other types of supramolecular structures in the extracellular matrix.

Crosslink Formation

Crosslink formation takes place in at least two stages. The first event is oxidative deamination of the ε-amino group in certain lysine and hydroxylysine residues into a reactive aldehyde catalyzed by lysyl oxidase (Figure 8). These aldehydes subsequently participate in the formation of the various crosslinks. None of the events beyond the lysyl oxidase reaction is known to require any enzyme.

The function of collagen crosslinks is to provide the collagen fibrils or other supramolecular assemblies with their mechanical strength. The significance of crosslinks is demonstrated especially clearly by the manifestations in experimental lathyrism, an acquired molecular disease characterized by a marked fragility of growing connective tissues. Lathyrism can be induced by the administration of β-aminopropionitrile or related compounds that act as irreversible inhibitors of lysyl oxidase and thus prevent the crosslinking of collagens and elastin. Deficient crosslinking is also found in animals with copper deficiency and mice and patients

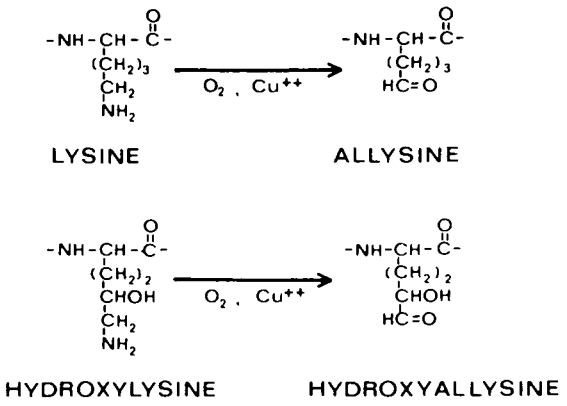


Figure 8. Reactions catalyzed by lysyl oxidase.

with heritable abnormalities of copper metabolism leading to deficient lysyl oxidase activity (for reviews on lysyl oxidase, see Kagan, 1986; Kagan and Trackman, 1991).

Lysyl oxidase (EC 1.4.3.13) is a highly insoluble enzyme that is synthesized in the form of a proenzyme with a N-terminal propeptide. This propeptide has been shown to be cleaved by an extracellular processing enzyme, but only initial data are available on the cleavage reaction (Trackman et al., 1992). Complete cDNA-derived amino acid sequences are now available for lysyl oxidase from several sources (e.g., Trackman et al., 1990; Hämäläinen et al., 1991).

Lysyl oxidase is a copper protein that contains a covalently bound quinone cofactor and requires O_2. The enzyme acts on both lysine and hydroxylysine residues present in the short nontriple-helical regions that remain at the ends of fibrillar collagen molecules after removal of the propeptides (Figure 1) and in the corresponding domains of nonfibrillar collagens. It utilizes both collagens and elastin as its substrates, the highest activity being seen with reconstituted collagen fibrils and certain ordered aggregates of elastin, whereas little or no activity is found with intact procollagen or monomeric collagen molecules or denatured collagen α chains.

The aldehydes formed in the lysyl oxidase reaction can then form two types of collagen crosslink, either by aldol condensation between two of the aldehydes or by condensation between one aldehyde and one ε-amino group of an unmodified lysine, hydroxylysine or glycosylated hydroxylysine residue. The crosslinks formed from a hydroxylysine-derived aldehyde are much more stable than those from a lysine-derived aldehyde, possibly due to migration of the double bond to a stable oxo form. With time, and depending on whether the initially oxidized residue was lysine or hydroxylysine, these difunctional crosslinks undergo further intra-

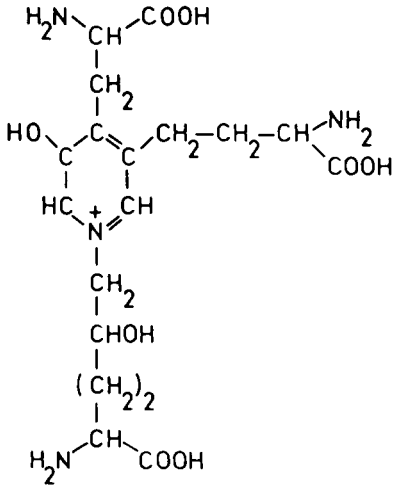


Figure 9. Structure of the hydroxylysyl-pyridinoline crosslink.

molecular and intermolecular reactions to yield several kinds of mature trifunctional and tetrafunctional crosslinks. Two such well-characterized trifunctional compounds are hydroxylysyl-pyridinoline (Figure 9), which is formed from two hydroxylysine-derived aldehydes and one hydroxylysine residue, and lysyl-pyridinoline, in which one of the three residues is derived from lysine instead of hydroxylysine (see Kivirikko and Kuivaniemi, 1987; Hulmes, 1992; Kielty et al., 1993).

In some collagens disulfide bonds also form covalent intermolecular crosslinks that join polypeptide chains in adjacent molecules. These disulfide bonds are formed in the extracellular matrix, and hence their formation does not appear to involve any catalysis by PDI.

SUMMARY

Collagen synthesis is characterized by the presence of an unusually large number of cotranslational and posttranslational modifications, many of which are unique to collagens and a few other proteins with collagen-like amino acid sequences. The posttranslational processing of a fibrillar collagen can be regarded as occurring in two stages: intracellular modifications, together with synthesis of the polypeptide chains, result in the formation of triple-helical procollagen molecules, and extracellular processing then converts these molecules into collagens and assembles the collagen molecules into stable, crosslinked fibrils. The processing of a nonfibrillar collagen is basically similar to that of a fibrillar collagen, except that it may not involve a procollagen intermediate and the extracellular supramolecular assemblies

are not fibrils. These processing reactions require at least nine specific enzymes and several nonspecific ones. The intracellular modifications require three specific hydroxylases that convert certain proline and lysine residues to 4-hydroxyproline, 3-hydroxyproline and hydroxylysine, and two specific glycosyltransferases that convert some of the hydroxylysine residues to galactosylhydroxylysine and some of the galactosylhydroxylysine residues to glucosylgalactosylhydroxylysine. The extracellular modifications require two specific procollagen proteinases, one to cleave the N-terminal propeptide and the other the C-terminal propeptide, and lysyl oxidase, which initiates crosslink formation by catalyzing oxidative deamination of certain lysine and hydroxylysine residues to reactive aldehyde derivatives. One of these enzymes, procollagen N-proteinase, has two collagen type-specific isoenzymes, whereas in all the other specific posttranslational modifications a single enzyme appears to act on all the collagen types. Most of these nine enzymes are now well characterized and three of them have been cloned. Detailed data are also available on the properties and functions of most of the posttranslational modifications.

REFERENCES

Bassuk, J. A., Kao, W. W.-Y., Herzer, P., Kedersha, N. L., Seyer, J. A., Demartino, J. A., Daugherty, B. L., Mark, G. E. III, & Berg, R. A. (1989). Prolyl 4-hydroxylase: Molecular cloning and the primary structure of the α-subunit from chicken embryo. Proc. Natl. Acad. Sci. USA 86, 7382–7386.

Engel, J. & Prockop, D. J. (1991). The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper. Ann. Rev. Biophys. Biophys. Chem. 20, 137–152.

Fleischmajer, R., Olsen, B., & Kühn, K., eds. (1990). Structure, Molecular Biology and Pathology of Collagen. Ann. N.Y. Acad. Sci. 580, 1–592.

Freedman, R. B. (1989). Protein disulfide isomerase: multiple roles in the modification of nascent secretory proteins. Cell 57, 1069–1072.

Hämäläinen, E.-R., Jones, T. A., Sheer, D., Taskinen, K., Pihlajaniemi, T., and Kivirikko, K. I. (1991). Molecular cloning of human lysyl oxidase and assignment of the gene to chromosome 5q23.3-31.2. Genomics 11, 508–516.

Hanuske-Abel, H. M. (1991). Prolyl 4-hydroxylase, a target enzyme for drug development. Design of suppressive agents and the in vitro effects of inhibitors and proinhibitors. J. Hepatol. 13, S8–S16.

Hanuske-Abel, H. M. & Günzler, V. A. (1982). A stereochemical concept for the catalytic mechanism of prolyl hydroxylase. Applicability to classification and design of inhibitors. J. Theor. Biol. 94, 421–455.

Hautala, T., Byers, M. G., Eddy, R. L., Shows, T. B., Kivirikko, K. I., & Myllylä, R. (1992). Cloning of human lysyl hydroxylase: complete cDNA-derived amino acid sequence and assignment of the gene (PLOD) to chromosome 1p36.2.-p36.3. Genomics 13, 62–69.

Helaakoski, T., Vuori, K., Myllylä, R., Kivirikko, K. I., & Pihlajaniemi, T. (1989). Molecular cloning of the α-subunit of human prolyl 4-hydroxylase: The complete cDNA-derived amino acid sequence and evidence for alternative splicing of RNA transcripts. Proc. Natl. Acad. Sci. USA 86, 4392–4396.

Hojima, Y., McKenzie, J., van der Rest, M., & Prockop, D. J. (1989). Type I N-proteinase from chick embryo tendons. Purification of a new 500-kDa form of the enzyme and identification of the catalytically active polypeptides. J. Biol. Chem. 264, 11336–11345.

Hulmes, D. J. S. (1992). The collagen superfamily—diverse structures and assemblies. Essays Biochem. 27, 49–67.

Kagan, H. M. (1986). Characterization and regulation of lysyl oxidase. In: Biology of Extracellular Matrix (Mecham, R. P., ed.), Vol. 1, pp. 321–389, Academic Press, Orlando.

Kagan, H. M. & Trackman, P. C. (1991). Properties and function of lysyl oxidase. Am. J. Respir. Cell Mol. Biol. 5, 206–210.

Kao, W. W.-Y., Nakazawa, M., Aida, T., Everson, W. V., Kao, C. W.-C., Seyer, J. M., & Hughes, S. H. (1988). Isolation of cDNA and genomic clones of β-subunit of chicken prolyl 4-hydroxylase. Connect. Tissue Res. 18, 157–174.

Kielty, C. M., Hopkinson, I., & Grant, M. E. (1993). Collagen: The collagen family: Structure, assembly and organization in the extracellular matrix. In: Connective Tissue and Its Heritable Disorders (Royce, P. M. & Steinmann, B., eds.), pp. 103–147, Wiley-Liss, New York.

Kivirikko, K. I. (1993). Collagens and their abnormalities in a wide spectrum of diseases. Ann. Med. 25, 113–126.

Kivirikko, K. I. & Kuivaniemi, H. (1987). Posttranslational modifications of collagen. In: Connective Tissue Disease. Molecular Pathology of the Extracellular Matrix (Uitto, J. & Peredja, A. J., eds.), pp. 263–292, Marcel Dekker, New York.

Kivirikko, K. I. & Myllylä, R. (1979). Collagen glycosyltransferases. Int. Rev. Connect. Tissue Res. 8, 23–72.

Kivirikko, K. I. & Myllylä, R. (1985). Post-translational processing of procollagens. Ann. N.Y. Acad. Sci. 460, 187–201.

Kivirikko, K. I., Myllylä, R., & Pihlajaniemi, T. (1989). Protein hydroxylation: prolyl 4-hydroxylase, an enzyme with four cosubstrates and a multifunctional subunit. FASEB J. 3, 1609–1617.

Kivirikko, K. I., Helaakoski, T., Tasanen, K., Vuori, K., Myllylä, R., Parkkonen, T., & Pihlajaniemi, T. (1990). Molecular biology of prolyl 4-hydroxylase. Ann. N.Y. Acad. Sci. 580, 132–142.

Kivirikko, K. I., Myllylä, R., & Pihlajaniemi, T. (1992). Hydroxylation of proline and lysine residues in collagens and other animal and plant proteins. In: Post-translational Modifications of Proteins (Harding, J. J. & Crabbe, M. J. C., eds.), pp. 1–51, CRC Press, Boca Raton.

Koivu, J., Myllylä, R., Helaakoski, T., Pihlajaniemi, T., Tasanen, K., & Kivirikko, K. I. (1987). A single polypeptide acts both as the β-subunit of prolyl 4-hydroxylase and as a protein disulfide-isomerase. J. Biol. Chem. 262, 6447–6449.

Majamaa, K., Hanauske-Abel, H. M., Günzler, V., & Kivirikko, K. I. (1984). The 2-oxoglutarate binding site of prolyl 4-hydroxylase. Identification of distinct subsites and evidence for 2-oxoglutarate decarboxylation in a ligand reaction at the enzyme-bound ferrous ion. Eur. J. Biochem. 138, 239–245.

Mayne, R. & Burgeson, R. E., (eds.) (1987). Structure and Function of Collagen Types. Academic Press, Orlando.

Myllylä, R., Pihlajaniemi, T., Pajunen, L., Turpeenniemi-Hujanen, T., & Kivirikko, K. I. (1991). Molecular cloning of chick lysyl hydroxylase. Little homology in primary structure to the two types of subunit of prolyl 4-hydroxylase. J. Biol. Chem. 266, 2805–2810.

Noiva, R. & Lennarz, W. J. (1992). Protein disulfide isomerase. A multifunctional protein resident in the lumen of the endoplasmic reticulum. J. Biol. Chem. 267, 3553–3556.

Noiva, R., Kimura, H., Roos, J., & Lennarz, W. J. (1991). Peptide binding by protein disulfide isomerase, a resident protein of the endoplasmic reticulum lumen. J. Biol. Chem. 266, 19645–19649.

Olsen, B. R. (1991). Collagen biosynthesis. In: Cell Biology of Extracellular Matrix, 2nd ed. (Hay, E. D., ed.), pp. 177–220, Plenum Press, New York.

Parkkonen, T., Kivirikko, K. I., & Pihlajaniemi, T. (1988). Molecular cloning of a multifunctional chicken protein acting as the prolyl 4-hydroxylase β-subunit, protein disulphide-isomerase and a cellular thyroid-hormone-binding protein. Comparison of cDNA-deduced amino acid sequences with those in other species. Biochem. J. 256, 1005–1011.

Pihlajaniemi, T., Helaakoski, T., Tasanen, K., Myllylä, R., Huhtala, M.-L., Koivu, J., & Kivirikko, K. I. (1987). Molecular cloning of the β-subunit of human prolyl 4-hydroxylase. This subunit and protein disulphide isomerase are products of the same gene. EMBO J. 6, 643–649.

Prockop, D. J. & Hulmes, D. J. S. (1994). Assembly of collagen fibrils *de novo* from soluble precursors. Polymerization and co-polymerization of procollagen, pNcollagen, and mutated collagens. In: Extracellular Matrix Assembly and Structure (Yurchenco, P. D. & Birk, D. E., eds.), pp. 47–90. Academic Press, Orlando.

Prockop, D. J. & Kivirikko, K. I. (1984). Heritable diseases of collagen. New Eng. J. Med. 311, 376–386.

Trackman, P. C., Pratt, A. M., Wolanski, A., Tang, S. S., Offner, G. D., Troxler, R. F., & Kagan, H. M. (1990). Cloning of rat aorta lysyl oxidase cDNA: Complete codons and predicted amino acid sequence. Biochemistry 29, 4863–4870.

Trackman, P. C., Bedell-Hogan, D., Tang, J., & Kagan, H. M. (1992). Post-translational glycosylation and proteolytic processing of a lysyl oxidase precursor. J. Biol. Chem. 267, 8666–8671.

Torre-Blanco, A., Adachi, E., Hojima, Y., Wooton, J. A. M., Minor, R. R., & Prockop, D. J. (1992). Temperature-induced post-translational over-modification of type I procollagen. Effects of over-modification of the protein on the rate of cleavage by procollagen N-proteinase and on self-assembly of collagen into fibrils. J. Biol. Chem. 267, 2650–2655.

van der Rest, M. & Garrone, R. (1991). Collagen family of proteins. FASEB J. 5, 2814–2823.

Vuori, K., Pihlajaniemi, T., Marttila, M., & Kivirikko, K. I. (1992a). Characterization of the human prolyl 4-hydroxylase tetramer and its multifunctional protein disulfide-isomerase subunit synthesized in a baculovirus expression system. Proc. Natl. Acad. Sci. USA 89, 7467–7470.

Vuori, K., Pihlajaniemi, T., Myllylä, R., & Kivirikko, K. I. (1992b). Site-directed mutagenesis of human protein disulphide-isomerase: effect on the assembly, activity and endoplasmic reticulum retention of human prolyl 4-hydroxylase in *Spodoptera frugiperda* insect cells. EMBO J. 11, 4213–4217.

Wetterau, J. R., Combs, K. A., McLean, L. R., Spinner, S. N., & Aggerbeck, L. P. (1991). Protein disulfide isomerase appears necessary to maintain the catalytically active structure of the microsomal triglyceride transfer protein. Biochemistry 30, 9728–9735.

Chapter 10

Cellular and Molecular Aspects of Selected Collagen Diseases

Rajendra Raghow

INTRODUCTION

Organs and tissues, the basic anatomical and functional units of the metazoan body, are organized and bound together by connective tissues, which include all membranes and septa, the supporting structures surrounding the vasculature and the contents of the synovial spaces. Although connective tissues are predominantly

Principles of Medical Biology, Volume 3
Cellular Organelles and the Extracellular Matrix, pages 255–273

ISBN: 1-55938-804-8

made up of mesenchymal cells and extracellular matrix (ECM), they are also highly enriched in water, salts and several common ingredients of circulating plasma (Hay, 1991; Ettinger and Doljanski, 1992; Turner, 1992). ECM is the fibrous component of the connective tissue, which is largely composed of collagen and elastin fibrils enmeshed within an array of glycoproteins and proteoglycans; many of the collagenous and noncollagenous constituents of ECM are well characterized at the molecular level. Biochemical analysis of ECM from various sources has uncovered extraordinary structural and functional heterogeneity; the composition of ECM may vary dramatically depending on the tissue of origin, the state of development of the organism and the cell-cell or cell-substratum interactions *in vivo* and *in vitro*. This remarkable heterogeneity of structure of ECM results from the intrinsic ability of its individual molecular components, which are made up of modular domains, to associate with each other in many permutations. Structures of ECM thus created are able to generate functional versatility to fulfill the specific needs of the unique physiological and pathological milieus of particular organs and tissues. For instance, the collagen-rich ECM is the primary determinant of the mechanical strength of the bony skeleton, the rugged elasticity of tendons and dermis, and the remarkable compressibility and resistance of cartilage (Vuorio and deCrombrugghe, 1990; Hay, 1991; Miller and Gay, 1992).

Although the mechanical support and structural roles of collagens have been appreciated for a long time, the dramatic phenotypes of *osteogenesis imperfecta* and Marfan syndrome in humans, resulting from specific mutations in genes encoding various collagens, have demonstrated their crucial importance in the early development (Prockop, 1991; Chu and Prockop, 1992). Likewise, the embryonic lethality of the MOV13 mice, homozygous for Type I collagen deficiency caused by insertion of a retrovirus in the mouse genome (Schnieke et al., 1983), underscores the role of precise regulation of collagen biosynthesis during early development. In different anatomic contexts, collagen-rich ECM functions as biomechanical scaffolding, as an adhesive promoting cell-ECM interaction, as a filtration barrier in the kidney glomerulus and, as a potential reservoir for cytokines and growth factors, as well as an integrator of signaling pathways governing cellular phenotype. Studies of cell-ECM interactions *in vivo* and *in vitro* have demonstrated clearly that cells of all embryological lineages possess specific cell surface receptor molecules, e.g., integrins, which are heterodimeric molecules containing α and β subunits. The individual constituents of ECM interact with the integrin receptors with high specificity and are able to transduce intracellular signals altering important cellular responses; both short term and long term changes occur following ECM-integrin interaction. These interactions guarantee that the dynamic changes occurring in ECM are continuously relayed intracellularly (Hynes, 1992; Damsky and Werb, 1992; Hynes and Lander, 1992; Juliano and Haskill, 1993).

PARADIGMS OF COLLAGEN DISORDERS

Coordinated spatial and temporal expression of various ECM macromolecules and cell-surface integrins during embryogenesis, as well as in tissues engaged in post-inflammatory repair and regeneration have been documented in exquisite detail (Hynes, 1992; Yang et al., 1993; George et al., 1993). Based on these studies, it may be concluded that cell-ECM interactions influence every one of the three basic attributes of all living cells-proliferation, migration and phenotypic differentiation. Studies of experimental models and clinical observations in humans have established numerous causative links between abnormal biosynthesis and accumulation of ECM and pathological manifestations commonly known as connective tissue diseases. The fibroses of vital organs, resulting from an excessive accumulation of ECM, represent some of the most common pathological conditions seen in the rheumatology clinic. I will attempt to outline the major cellular and molecular events thought to be associated with three common fibrotic diseases; I will restrict my discussion to scleroderma (systemic sclerosis) and fibroses of lung and liver, as paradigms of systemic and localized "collagen diseases," respectively. Rather than review the vast literature dealing with all aspects of these interesting sequelae, my aim here is to briefly consider a few selected phenomena which may shed mechanistic light on the genesis of these disorders. Needless to say, that a description of such a limited magnitude will contain apparent biases and speculations which may not be universally shared.

Scleroderma

Systemic sclerosis or scleroderma (Greek, *sclerose*, hard and *derma*, skin) is a generalized aberration of connective tissues characterized by prominent thickening and fibrosis of skin. As described in a number of excellent recent reviews, the term scleroderma encompasses a rather large spectrum of clinical disorders which may only be minimally related to each other (Le Roy, 1992; Seibold, 1993; Wigley, 1994; Clements, 1994). The hallmarks of scleroderma include subtle but significant inflammatory changes in the target tissues, vascular lesions and skin fibrosis. Depending on the extent of involvement of the internal organs, scleroderma may be divided into two main categories. The more common form, termed limited cutaneous systemic sclerosis (lSSc) may be characterized by thickening of skin limited to fingers, extremities or face; in this syndrome, there is only occasional involvement of the internal organs, and such involvement is commonly limited to the esophagus in the early stages of the disease. The symptoms of the second, less prevalent but more devastating form of the disease, termed the diffuse cutaneous systemic sclerosis (dSSc), are exhibited by both external and internal organs. Thus, in dSSc patients, skin covering extremities as well as trunk, undergo rapid thickening and the disease invariably targets kidneys (hypertension with renal failure), lungs (pulmonary alveolitis with interstitial fibrosis), gastrointestinal tract (severe

esophageal hypomotility), heart (myocarditis with myocardial fibrosis) and bony articulations (peripheral joint contractures). It should be stressed, however, that this somewhat arbitrary classification of scleroderma obscures the underlying heterogeneity of several pathological disorders, which fall somewhere between these two extremes. Additionally, there are a number of other diseases that present symptoms overlapping those seen in scleroderma; these include rheumatoid arthritis, systemic lupus erythematosus and a number of vasculitis syndromes. It is believed that, with the possible exception of the inherited diseases of the ECM like *osteogenesis imperfecta* or Marfan syndrome, the "collagen diseases" involve autoimmune mechanisms; the presence of circulating autoantibodies and accumulation of immune inflammatory cells in the involved tissues lend support to such a mechanism (White, 1994; Wigley, 1994).

Pulmonary Fibrosis

The primary function of the lungs is gas exchange and oxygenation of the blood. The basic symptoms of pulmonary fibrosis, breathlessness and non-productive cough, are frequently accompanied by systemic manifestations of weight loss, fever, lethargy, and arthralgia. These symptoms reflect changes in the alveolar architecture leading to impairment of the gas exchange apparatus. In the advanced stages, signs of pulmonary edema, hypertension and right-sided heart failure may be seen and the disease invariably progresses to respiratory failure. High resolution computer-assisted tomography scanning may reveal subpleural reticular opacities, cystic air spaces and peripheral fibrosis.

The functional and anatomic pulmonary unit, the alveolus, may be viewed as made of three compartments: the capillary endothelium, consisting of blood-lung interface, the epithelium representing the air-lung interface, and the interstitium, a thin barrier between the endothelium and epithelium. The thin interstitium of a healthy alveolus consists of cells and stroma and two basement membranes; the major interstitial cells are fibroblasts with an occasional inflammatory cell. In contrast, in pulmonary fibrosis, the interstitium is characterized by overt inflammation as evidenced by accumulation of immune inflammatory cells and excessive deposition of ECM (Raghow et al., 1985; 1989; Brockelmann et al., 1991; Khalil and Greenberg, 1991; Raghow, 1992; Hertz et al., 1992). A histologic examination of the lung biopsy specimens generally reveals infiltration of inflammatory cells with or without desquamation and thickened alveolar septa with fibrosis. Examination of the broncho-alveolar lavage fluid shows increased numbers of neutrophils, lymphocytes, and macrophages during the inflammatory phase of the acute pulmonary damage; the relative proportion of the various cell types may vary depending on the stage of the disease (Seibold, 1993).

Fibrotic Liver Diseases

The liver is a complex parenchymal organ functioning as a collection point for most of the noxious agents that enter the blood stream. Under some conditions, the elaborate detoxification mechanisms of the liver, become overwhelmed and hepatic injury and inflammation are inevitable. The process of post-inflammatory repair and regeneration of the damaged hepatic parenchyma, similar to other soft tissues, often culminate in fibrosis. A spectrum of fibrotic disorders involving liver include chronic liver disease, primary biliary cirrhosis, sclerosing cholangitis and Dupuytren's contracture associated with alcoholic liver. These diseases are characterized by dysmorphology of the hepatic parenchyma and associated fibrosis, resulting in compromised liver function. Chronological sequence of events leading to hepatic cirrhosis and fibrosis has been extensively studied in humans and experimental animals; histological, ultra-structural and biochemical assessment of various parameters of structure and function has been described in detail (Seyer and Raghow, 1992). The hepatic lobule, the basic anatomic unit of the liver, is made up of epithelial parenchyma and a complicated array of anastomosing vascular sinusoids. The microvasculature of the liver lobule is represented by the basic unit known as portal triad, consisting of the hepatic arteriole, the portal venule, and the bile duct. The portal triad is separated from the surrounding liver parenchyma by an elaborate ECM. The sinusoids, which receive a mixture of portal and arterial blood, are specialized sinuses containing a one-cell thick, porous lining of endothelial and Kupffer cells with a primitive basement membrane containing Type IV collagen but lacking laminin. The perisinusoidal space, called space of Disse, separates endothelial cells and the sinusoidal lumen from the hepatocytes. Fibers of Types I and III collagens are present in this region interspersed with sporadic lipocytes or Ito cells. The biliary passages begin with fine bile canaliculi between individual hepatocytes; through a network of canaliculi bile is drained into the bile duct near the portal triad. These ductile cells form a tubular arrangement with surrounding basement membranes and, at a further distance, a typical ECM containing Types I and III collagens. The cellular organization of the liver, dominated by the parenchymal hepatocytes, is optimized for its synthetic and detoxification functions. Hepatocytes account for 60–70% of the total cells; the remaining cellular population is made of Kupffer cells, epithelial cells and lipocytes. In contrast, in the fibrotic liver, there is substantial loss of parenchymal cells, which are replaced by connective tissues rich in Type I collagen.

CELLULAR AND MOLECULAR EVENTS INVOLVED IN FIBROSIS

Our concepts of initiation, progression and resolution of the tissue damage have been derived from two kinds of observations. First of these include the histological observations of the tissue biopsies or samples of human tissues obtained postmor-

tem. Furthermore, careful and systematic morphological, cytokinetic and biochemical analyses of tissues undergoing post inflammatory repair and regeneration in experimental animals have added valuable insights to data obtained from patients. Thus, models of acute or chronic injury of specific organs, like bleomycin-induced pulmonary fibrosis or carbon tetrachloride-induced hepatic fibrosis, have elucidated cellular and molecular mechanisms involved in tissue repair and fibrosis. Based on these observations, we can surmise that the events involved in soft tissue repair, subsequent to an inflammatory injury, follow a somewhat predictable sequence encompassing three distinct yet partially overlapping stages (Raghow, 1994). The first stage may be characterized by the attendant phenomena caused by damage to the integrity of the vasculature; hemorrhage, platelet degranulation, activation of the complement pathway leading to thrombus formation and hemostasis are readily evident at this stage. The second phase may be noted by massive chemotactic recruitment of leukocytes to the site of injury; generally, infiltration of neutrophils followed by macrophages and later, B and T lymphocytes, occurs during this phase of the repair. The elaboration of potent growth factors and cytokines by the newly-arrived immune inflammatory cells triggers the final phase of tissue repair. During this stage, autocrine and paracrine activation of different cells by cytokines/growth factors leads to brisk biosynthetic and morphological changes which include significant neovascularization, mitogenesis and ECM remodeling (Raghow, 1994). It should also be kept in mind that, in addition to the cytokines, various constituents of ECM are also involved as paramount modulators of many activities crucial to the successful completion of the repair process.

Although the involvement of the autoimmune mechanisms in the pathogenesis of connective tissue diseases is generally accepted, there is far less agreement as to the molecular identities of the agents which initiate the inflammatory process. The list of potential offending antigens (Table 1) is indeed large and consists of a number of well-known components of the apparatus involved in DNA replication and biogenesis of messenger RNA (Pisetsky, 1994; Rothfield, 1994; Reeves, 1994; White, 1994). Based on an extensive analysis of the many connective tissue disorders, it is apparent that in addition to the failure of the immune system to distinguish self from non-self, exogenous factors like bacterial and viral infections, and exposure to environmental or chemotherapeutic agents may significantly change the course of the inflammatory process and the impending regeneration and fibrosis, which follows it (Seibold, 1993). Experimental animal models have contributed tremendously to our understanding of these factors (Bowden, 1984; Rennard et al., 1984; Raghow et al., 1985; 1989, 1992; Armendariz-Borunda et al., 1990; 1993; 1994). For instance, intratracheal instillation of the chemotherapeutic agent, bleomycin, induces pulmonary injury, followed by repair and regeneration accompanying fibrosis, which essentially resembles the sequelae of adult respiratory distress syndrome (ARDS). Similarly, some of the basic cellular and molecular

Table 1. A Partial List of Common Autoantibodies Associated With Connective Tissue Diseases

Antigen	*Molecular Identity*	*Clinical Associations*
Centromere-binding proteins	CENP-A, 17 kD CENP-B, 80 kD CENP-C, 140 kD	Scleroderma
Topoisomerase I	70 kD DNA replication, Transcription	Scleroderma
RNA polymerase I	Multiple proteins, 12.5 kD to 210 kD Transcription of ribosomal RNA	Subset of scleroderma
Ribonucleoprotein particles	U1, U2 68, 70 kD U4/U6 snRNA Ro/SSA; 52,60 kD La/SSB; 48kD Transcription, splicing of mRNA	Myositis, Scleroderma, Sjögren's syndrome
Lamins A, B, C	Proteins of chromatin	Biliary cirrhosis, Systemic lupus erythematosus
Ku DNA binding proteins	70 kD and 80 kD heterodimer	Scleroderma, Systemic lupus erythematosus
PCNA	Protein associated with DNA polymerase delta	Scleroderma

manifestations of cirrhosis of the alcoholic liver are mimicked by postinflammatory hepatic regeneration and repair elicited in carbon tetrachloride- injured liver. Under normal conditions, healthy tissues are maintained in a quiescent state, albeit with the ability to constantly remove and replace effete cells, but without cellular overgrowth. Tissue injury and inflammation perturb this apparent equilibrium and elicit chemotaxis, mitogenesis, and phenotypic differentiation of cells (Figure 1) and excessive connective tissue formation. The process of postinflammatory tissue regeneration and repair recapitulate a number of well-characterized events commonly encountered during early embryogenesis; there is also close relationship of these events with the process of excisional dermal wound healing (Raghow, 1991; 1994). Thus, depending on the stage of tissue injury and repair, measurable changes in cell migration, mitogenesis, phenotypic differentiation as judged by altered gene expression and heightened biosynthetic activities can be demonstrated in all connective tissue diseases.

Cellular Components

To fully appreciate the process of tissue fibrosis, we have to consider the relative contribution of cell-cell and cell-ECM interactions. Regeneration and repair of the injured tissue is dependent upon the intercellular communication and, thus, disrup-

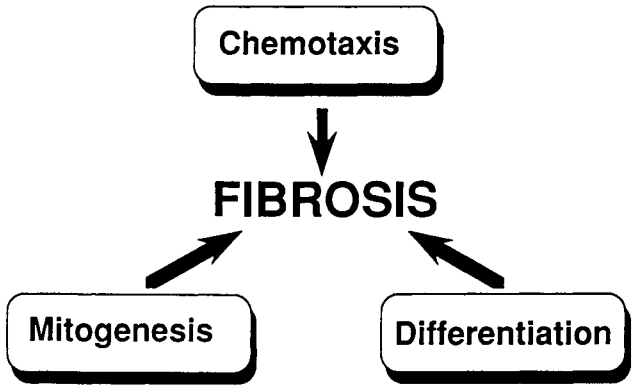


Figure 1. The trinity of fundamental cellular reactions required to drive tissue fibrogenesis. Interactions among the constituents of extracellular matrix, cytokines/growth modulatory factors and target cells regulate all three of these reactions. These basic reactions, involving directional migration of cells, proliferation and phenotypic differentiation as judged by preferential gene expression, are also central to the processes like organogenesis during early development, carcinogenesis, and postinflammatory wound repair and regeneration.

tion of signals between epithelial, endothelial, and mesenchymal cells and the cells of the immune system may lead to inappropriate repair and fibrosis. For example, in the normal lung, alveolar, epithelial, and endothelial cells are in close contact with the basement membranes. Following the disruption of basement membranes by inflammatory processes, unbridled proliferation of fibroblasts and concomitant changes in the quality and quantity of the constituents of ECM laid by these cells is often seen. However, fibroblasts do not function autonomously, but are subject to modulatory signals from a number of other cells, particularly the immune effector cells. The mutually regulatory relationships among the major cell types involved in the fibrogenesis commonly seen in connective tissue diseases are depicted in Figure 2 and are briefly considered below.

Neutrophils

As key phagocytic cells of the immune system, the neutrophils release a number of potent hydrolyzing enzymes that can degrade ECM. Furthermore, neutrophils also produce highly toxic free radicals and are therefore capable of modulating the initial toxicity, the nature and extent of which contributes directly to the process of repair and regeneration. However, neutrophils seem to exert a paradoxical effect on the process of fibrosis. For example, in bleomycin-treated beige mice, in which neutrophils failed to degranulate and release hydrolytic enzymes, but were capable of producing normal amounts of superoxide radicals, higher rates of collagen

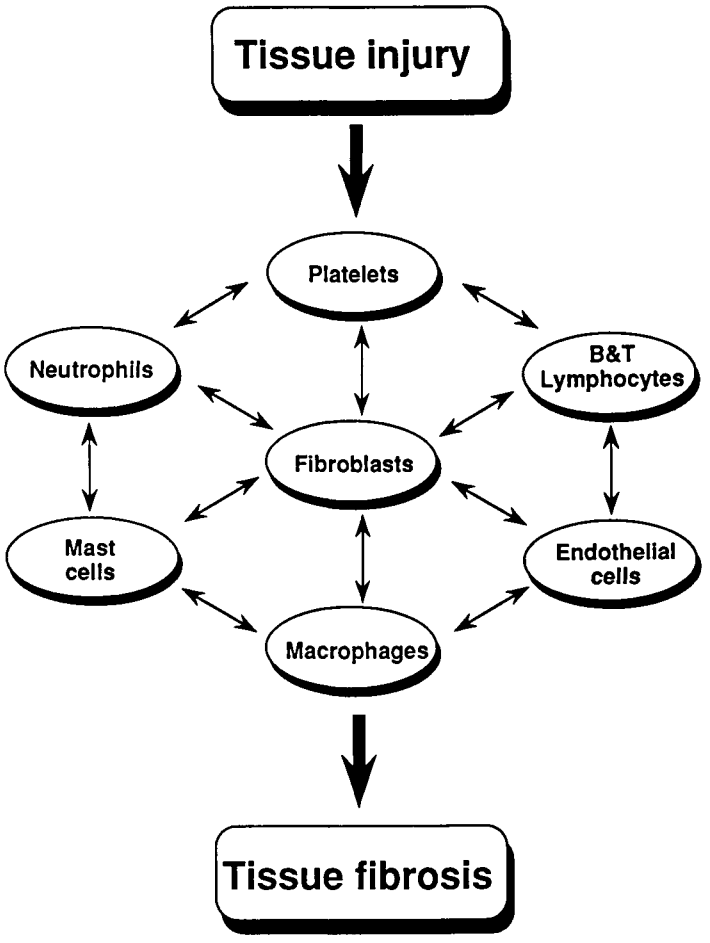


Figure 2. Complex circuitry of cell-cell, cell-cytokine, or cell-extracellular matrix interactions involved in the pathogenesis of fibrotic disease. Tissue injury and damage to the vasculature caused by spontaneous autoimmune mechanisms, induced by noxious chemicals or initiated by microbial pathogens begins a cascade of inflammatory reactions in which many cell types participate. These reactions involve either direct cell-cell interaction or are mediated indirectly through the production of cytokines; two-way arrows depict the potential feedback and mutual regulatory pathways. Increased biosynthesis and reduced degradation of the various constituents of ECM is the final outcome of these reactions.

synthesis and a more severe pulmonary fibrosis was seen (Thrall et al., 1981; Phan and Thrall, 1982; Clark and Kuhn, 1982). A direct role of the toxic oxygen radicals was also demonstrated by showing that bleomycin-induced injury was exacerbated by hyperoxic conditions in a dose-dependent manner (Rinaldo et el., 1982).

Macrophages

Regardless of the nomenclature, activated monocytes, tissue macrophages, or Kupffer cells, monocytes/macrophages are central modulators of the process of tissue fibrosis (Nathan, 1987; Kovacs and DiPietro, 1994). In addition to being active phagocytes, macrophages synthesize and secrete a host of potent biological modulators. Included in the ever-expanding armamentarium of cytokines produced by macrophages are interleukins 1 and 6, transforming growth factors (TGFα and β), platelet-derived growth factors (PDGF), tumor necrosis factor α (TNFα), interferon β and γ, (IFN β and γ)insulin-like growth factor (IGF), granulocyte/macrophage-colony stimulating factor (GM-CSF), and colony stimulating factor 1 (CSF-1). These potent cytokines can dramatically alter the phenotype of many cells including smooth muscle cells, endothelial cells and fibroblasts. Macrophages synthesize substantial amounts of fibronectin and thrombospondin, thus contributing directly to the composition of ECM. Macrophages also produce a number of proteases which degrade ECM components and promote its remodeling. In addition to these effects, macrophage-derived cytokines modulate the quality and quantity of ECM which dramatically alters the behavior of target cells. For instance, chemotaxis, cell proliferation and differentiation of endothelial cells, in response to remodeled ECM, are the key events in capillary formation and angiogenesis (Oppenheim et al., 1991; Koch et al., 1992; Kovacs and DiPietro, 1994; Sunderkotter et al., 1994).

Lymphocytes

The direct involvement of lymphocytes in the fibrotic process has been implicated in a number of studies of biopsied tissue samples and studies in experimental animals (Wahl et al., 1986). Thus, compared to their normal counterparts, the athymic nude mice lacking T-lymphocytes showed diminished inflammatory cell infiltration, fibroblast proliferation, and reduced collagen synthesis in response to bleomycin treatment (Raghow, 1992). Although the exact mechanism of how depletion of T-lymphocytes results in reduced collagen synthesis is currently unknown, a potential explanation lies in the fact that activated lymphocytes are capable of releasing a variety of mediators, which include interleukin-2 (IL-2) and IFN γ, which modulate a number of key macrophage functions (Kovacs and DiPietro, 1994). Furthermore, lymphocytes produce fibrogenic cytokines such as fibrosin, which may directly activate fibroblasts to produce excessive ECM (Prakash et al., 1995).

Eosinophils and Mast Cells

Eosinophils may be potentially important in the progression and resolution of tissue fibrosis in two different ways. Eosinophils produce collagenolytic enzymes

capable of degrading interstitial collagens, thus modifying the composition of the ECM; in addition, eosinophils appear to be toxic to the mesenchymal and parenchymal cells. The cytotoxic response of eosinophils towards lung fibroblasts and mesothelial cells apparently does not require prior activation to degranulate and the mechanism of direct cellular cytotoxicity remains speculative (Davis et al., 1984; Raghow, 1992).

Increased numbers of mast cells are associated with the development of tissue fibrosis in experimental models; thus, the genetic mouse model of scleroderma, the tight skin (Tsk/+), and the chronic graft *versus* host reaction-induced immune model of scleroderma show a large number of mast cells in the involved areas of the skin. Increased accumulation of mast cells has also been documented in the sclerodermatous skin lesions of patients (LeRoy, 1992; Seibold, 1993). Mast cells store histamine, heparan sulfate proteoglycan (HSPG) and a number of proteases. Histamine is able to induce fibroblast proliferation while HSPG may be indirectly involved as an ancillary factor(s) in the action of cytokines like fibroblast growth factor (FGF). In light of these observations, it is likely that other immune effector cells, capable of producing IL-3 and IL-4, the potent mast cell growth factors, contribute directly to the fibrosis seen in scleroderma.

Fibroblasts

Among the various cells intimately involved in the process of tissue fibrosis, fibroblast plays a unique and central role (Raghow, 1992; Mainardi, 1993; Postlethwaite, 1994). Fibroblasts are multifaceted cells of mesenchymal origin whose functions can be influenced by many stimuli; they accumulate *en masse* at the site of inflammatory damage through chemotactic migration and cell proliferation. Fibroblasts are the primary producers of the major constituents of ECM; they are also responsible for remodeling the connective tissue by releasing a variety of proteases, proteoglycanases and glycosaminodases. Heightened activity and increased numbers of fibroblasts are responsible for enhanced ECM. Although both promoters and inhibitors of fibroblast growth have been described, the effect of mitogenic cytokines like PDGF, IL-4 and TGFβ seem to predominate under conditions of fibrogenesis (Postlethwaite et al., 1988; 1992; Postlethwaite, 1994). TGFβ is of special interest since it is a growth inhibitor for hematopoietic, epithelial, endothelial cells but is an inducer of mitogenesis in fibroblasts. Excessive accumulation of ECM in response to TGFβ occurs since it regulates both the capacity of fibroblast to synthesize ECM and reduce its degradation. Finally, fibroblasts are not only the targets of cytokines but are themselves capable of releasing a variety of important cytokines; biosynthesis of a variety of interleukins, prostaglandins, and other modulatory substances by fibroblasts has been documented (Mainardi, 1993; Postlethwaite, 1994).

Endothelial Cells

Regulation of endothelial cell behavior is intimately involved in the pathogenesis of fibrotic diseases (LeRoy,1992; Seibold, 1993; Kovacs and DiPietro, 1994). Prominent intimal proliferation of arteries and disruption of arterioles and capillaries are the hallmark symptoms of systemic sclerosis. Increased angiogenic activity is an essential component of many inflammatory states. Cytokines in association with the constituents of the ECM profoundly influence proliferation, chemotactic migration, and phenotype of endothelial cells. The biosynthetic, chemotactic and phenotypic activities of endothelial cells are affected by numerous cytokines, which include TGFβ, PDGF, IL-1, IL-6 and TNFα. Although the identity of the cytokines, which modulate endothelial cell phenotype and trigger angiogenesis *in vitro* remain uncertain, both TNFα and IL-8 are known to induce angiogenesis in experimental systems. Contributing to this apparent confusion is the realization that the composition of ECM can modulate the activity of the same cytokines (see below); furthermore, ECM can influence endothelial cell shape and phenotype directly and thus determine the formation of the vascular elements. In light of these observations, it is generally acknowledged that the specific cytokines and the precise cell-cell and cell-ECM interactions underlying angiogenesis may differ depending on the site of the post inflammatory repair and regeneration.

FIBROGENIC MODULATORS

Cytokines

The initiation of the clotting cascade and other attendant biochemical reactions which lead to the formation of provisional matrix are the necessary substrates upon which the granulation tissue is assembled. The granulation tissue is made up of mesenchymal cells and blood capillaries embedded within the loosely assembled ECM of Types I and III collagens, fibronectin, and proteoglycans. The macrophages, lymphocytes, endothelial cells, and fibroblasts accumulated in vast numbers in the granulation tissue release many multipotent cytokines, which orchestrate the process of repair (Kelley, 1990; Whitby and Ferguson, 1991; Postelthwaite et al., 1992; Morgan and Pledger, 1992; Postelthwaite, 1994; Mainardi, 1994). Although, based on their predominant activities, these cytokines have been classified as pro-inflammatory, fibrogenic and antifibrogenic (Kovacs nd DiPietro, 1994), a careful examination of their properties (Table 2) would reveal that these cytokines possess overlapping activities, and an unequivocal classification into such categories is too simplistic.

In addition to promoting proliferation of fibroblasts, cytokines also directly modulate the production of ECM (Raghow et al., 1987; Raghow, 1991; 1992; 1994; Kovacs and DiPietro, 1994). Thus, enhanced rates of synthesis of macromolecules

Table 2. A Partial List of Cytokines/Growth Modulators Implicated in the Connective Tissue Diseases

Cytokine/Agent	*Likely Cell of Origin*	*Likely Target and Action*
Histamine	Mast Cell	Fibroblast, Proliferation
TNFα	Mast Cell, Macrophage Lymphocyte	Fibroblast, Inhibition of Growth, ECM Synthesis
PDGF	Platelet	Fibroblast, Smooth Muscle Cell, Proliferation
MCP-1	Endothelial Cell	Fibroblast, Inhibit Proliferation
Interleukin-1α,β	Monocyte	Fibroblast, Endothelial Cell, Proliferation, ECM
Interleukin-2	Lymphocyte	Macrophage, IL-1, TNFα, TGFβ, PDGF Synthesis
Interleukin-4	Lymphocyte	Fibroblast, ECM Synthesis, Endothelial Cell, MCP-1 synthesis
Interleukin-6	Lymphocyte	Endothelial Cell, Fibroblast Inhibit Mitogenesis
Interleukin-8	Macrophage	Lymphocyte, PMNS, Chemotaxtic Endothelial Cell, Angiogenesis
Fibrosin	Lymphocyte	Fibroblast Proliferation, ECM Synthesis
TGFβ	Platelet, Macrophage, Fibroblast	Many Cell Types, Chemotaxis, ECM Synthesis, Proteases, Mitogenesis, Differentiation

of ECM and their reduced degradation can be clearly demonstrated in response to TGFβ1 *in vitro* and *in vivo* using animal models of repair and fibrosis. Promotion of ECM synthesis by cytokines is mediated by transcriptional and posttranscriptional mechanisms; increased steady state levels of mRNA for Types I and III collagens, fibronectin and proteoglycans have been demonstrated in the involved areas of the sclerodermatous skin and in the experimental models of pulmonary and hepatic fibrosis (Armendariz-Borunda et al., 1990; 1993; 1994; LeRoy, 1992; Raghow, 1992; Seyer and Raghow, 1992; Siebold, 1993). Accumulation of ECM in response to cytokines may be further accelerated by inhibiting ECM degradation; thus, TGFβ inhibits the production of collagenase and stromelysin and promotes biosynthesis of plasminogen activator inhibitor, actions which further reduce proteolysis of ECM. In considering the causative roles of profibrogenic cytokines, it may be useful to remember that rather than contemplating the action of a given cytokine in isolation, we have to consider the complex mixtures of various cytokines which must indeed exist *in vivo*. Alas, our current models and the implicit reductionist assumptions underlying them are rather inadequate and will prompt further refinements as more accurate data become available.

Extracellular Matrix

A large fraction of the dry mass of all soft tissues is due to ECM, which comprises primarily collagen (60 to 70%) and elastin (25 to 35%), with contributions from glycosaminoglycans (<1%) and other glycoproteins (e.g., fibronectin, laminin, and minor collagens). Nearly two dozen different kinds of collagen α chains, encoded by unique genes, have been identified in different tissues of the vertebrate body (Vuorio and deCrombrugghe, 1990; Miller and Gay, 1992; Williams et al., 1993). Collagens may be divided into: (i) fibrillar (e.g., Type I); (ii) basement membrane-associated (e.g. Type IV); (iii) fibril-associated (e.g., Type IX) and (iv) short chain (e.g., Type X) collagen types (see Table 3). Types I and III collagen are the major ECM components of the vasculature, skin, and lungs; Type

Table 3. Common Types of Collagen Found in Vertebrate Tissues[a]

Type	*Molecular Composition*	*Distribution*	*Function*
I	$[\alpha 1(I)]_2\alpha 2(I)$	All connective tissues except hyaline cartilage and basement membranes	Formation of striated fibers of varying diameter and basement and high tensile strength
II	$[\alpha 1(II)]_3$	Hyaline cartilages and cartilage-like tissues, (e.g., vitreous humor)	Formation of thin fibrils smaller in diameter than Type I collagen fibers
III	$[\alpha 1(III)]_3$	The more distensible connective tissues (blood vessels)	Formation of thin fibrous elements
IV	$[a1(IV)]_2\alpha 2(IV)$	Basement membranes under epithelium and blood vessels	Formation of mesh-like scaffold
V	$[\alpha 1(V)]_2\alpha 2(V)$	Essentially all tissues	Similar to Type III collagen
VI	$[\alpha 1(VI),]\alpha 2(VI),\alpha 3(VI)$	Essentially all tissues	Formation of micro-fibrillar elements
VII	$\alpha 1(VII)]_3$	Dermal-epidermal junctions	Anchoring fibrils
VIII	Unknown	Descemet's membrane produced by endothelial cells	Unknown
IX	$[\alpha 1(IX),\alpha 2(IX),\alpha 3(IX)]$	Hyaline cartilage	Forms coaggregates with Type II collagen
X	$[\alpha 1(X)]_3$	Hypertrophic cartilage	Unknown
XI	$[\alpha 1(XI),\alpha 2(XI),\alpha 1(II)]$	Hyaline cartilage	Unknown
XII	$[\alpha 1(XII)]_3$	May be similar to Type I collagen	Unknown, but may form coaggregates with Type I collagen

Note: [a]Adapted from Miller & Gay, 1992.

III collagen is preferentially distributed in the alveolar interstitium and in the intima and media of the arteries. Type IV is the major collagen of the basement membranes, and is found in alveolar septa and subepithelial lining. In lungs and dermis, elastin not only is a critical determinant of their mechanical elasticity but provides a framework for cell growth and orientation. Elastin is the functional protein constituent of the elastic microfibrils; disruption of the microfibrillar network has devastating consequences as demonstrated by the emphysematous lesions created by exogenous administration of elastolytic enzymes in the lung. In addition to collagens and elastin, a number of glycoproteins and proteoglycans are also important constituents of ECM (Trelstad and Kemp, 1993).

Fibrosis is characterized by accelerated rates of synthesis and deposition of ECM; it has been shown that the rates of synthesis of interstitial collagens, and other components of ECM, are preferentially elevated in the scleroderma skin and in the target organs of various animal models of fibrosis. Based on numerous studies, it has become clear that ECM, rather than being an inactive bystander in the process of postinflammatory repair, actively modulates it in multiple ways. Some ECM constituents like laminin contain EGF-like domains which are capable of inducing fibroblast mitogenesis *in vitro*. It is conceivable that during inflammation, some of these domains may be uncovered as a result of proteolytic degradation of the parent molecule. However, whether the EGF-like mitogenic domains are actually involved *in vivo* is unknown (Raghow, 1994). Perhaps an even more important relationship between ECM and cytokines is their synergistic or antagonistic interaction with each other (Ruoslahti and Yamaguchi, 1991; Thiery and Boyer, 1992; Ruoslahti et al., 1992; Raghow, 1994). It has been documented that fibronectin-mediated cell adhesion synergizes the action of PDGF to trigger cytoplasmic alkalinization through PDGF receptor (Schwartz and Lechene, 1992). Similarly, bFGF requires obligatory binding to heparan sulfate proteoglycan (HSPG) for signal transduction mechanisms (Yayon et al., 1991). Since the original discovery of the bFGF-HSPG association, a number of additional cytokines (GM-CSF, CSF-1, PF-4 and TGFβ have been shown to specifically interact with proteoglycans (Raghow, 1994). The functional consequences of TGFβ binding to proteoglycans illustrates this paradoxical relationship between cytokines and ECM. In addition to binding to signaling receptors containing serine/threonine kinase domains, Type I and II receptors, TGFβ specifically binds to a proteoglycan, betaglycan, found on the surface of many cells; the latter, Type III receptor, is speculated to present TGFβ to the signaling receptors (Massague, 1992; Kingsley, 1994). In contrast, the binding of TGFβ to another proteoglycan, decorin, found in association with Type I collagen in ECM, can competitively inhibit TGFβ-receptor interactions (Border and Ruoslahti, 1991; Ruoslahti and Yamaguchi, 1991; Ruoslahti et al., 1992). These observations compel us to rethink the multiple roles of ECM as a modulator of cytokine action and also as a feedback modulator of its own synthesis.

SUMMARY AND CONCLUSIONS

Organs and tissues are formed by a complex interplay among the three fundamental activities of living cells: proliferation, directional migration, and phenotypic differentiation. These basic functional units are maintained in a rather quiescent state, notwithstanding the occasional replacement of damaged cells, without cellular overgrowth and structural disintegration. Dysfunctional immune responses, perhaps exacerbated by exogenous factors, can perturb the dynamic quiescence of the normal tissues and activate a series of events that lead to excessive ECM formation, either systemically or in a localized compartment of the body. Inflammatory mediators and cytokines, in close collaboration with the macromolecular constituents of ECM, mount a regenerative cascade, the successful outcome of which demands that many key activation reactions be terminated and even reversed in an orderly fashion. The central mechanisms by which complex and mutually modulatory interactions between cells and ECM, and ECM and cytokine/growth factors lead to orderly restoration of the injured tissue have only begun to be elucidated. Stochastic redundancy of these interactions is most likely responsible for their effectiveness in phenomena as wide ranging as embryogenesis, wound healing, neoplasia, and autoimmune fibrotic disorders. Although so far impressive progress toward understanding the molecular mechanisms of these processes has been made, extraordinary gaps in our knowledge still remain. For example, even educated guesses cannot be offered to explain how such a wide spectrum of symptoms are generated in different patients with an identical profile of antinuclear antibodies. How is the tissue damage targeted to restricted areas if the autoimmunity is elicited against a common antigen present in every cell of the body? What primary and secondary events modulate the precise make up of cytokines/growth factors, both spatially and temporally, during fibrogenesis? With the rapid emergence of novel and precise tools of molecular and cellular biology, many of these vexing questions are surely going to yield mechanistic insights in the coming years and will lead to more effective therapeutic interventions in collagen diseases.

ACKNOWLEDGMENTS

The research described from the author's laboratory has been generously supported from the National Institutes of Health (AR-39166) and Merit Award from the Department of Veterans Affairs (DVA). RR is a Research Career Scientist of the DVA. Cheerful and competent assistance of Donna Jackett and Lynne Blalock in the preparation of this chapter is acknowledged.

REFERENCES

Armendariz-Borunda, J., Seyer, J. M., Kang, A. H., & Raghow, R. (1990). Regulation of TGF-β gene expression in rat liver intoxicated with carbon tetrachloride. FASEB J. 4, 215–221.

Armendariz-Borunda, J., Katai, H., Jones, C. M., Seyer, J. M., Kang, A. H., & Raghow, R. (1993). Transforming growth factor β expression is transiently enhanced at a critical stage during liver regeneration following CCl_4 treatment. Lab. Invest. 69, 283–294.

Armendariz-Borunda, J., Simkevich, C. P., Roy, N., Raghow, R., Kang, A. H., & Seyer, J. M. (1994). Activation of Ito cells involves regulation of AP-1 binding proteins and induction of type I collagen gene expression. Biochem. J. 304, 817–824.

Border, W. A. & Ruoslahti, E. (1992). Transforming growth factor β in disease: the dark side of tissue repair. J. Clin. Invest. 90, 1–5.

Bowden, D. H. (1984). Unraveling pulmonary fibrosis: the bleomycin model. Lab. Invest. 50, 487–490.

Brockelmann, T. J., Limper, A. H., Colby, T. V., & McDonald, J. A. (1991). Transforming growth factor β is present in sites of extracellular matrix gene expression in human pulmonary fibrosis. Proc. Natl. Acad. Sci. USA 88, 6642–6646.

Chu, M. L. & Prockop, D. J. (1994). Collagen: Gene Structures. In: Extracellular Matrix and Inheritable Disorders of Connective Tissue. (Royce, P. M. and Steinmann, B., Eds.). Alan R. Liss, New York.

Clark, J. G. & Kuhn, C. (1982). Bleomycin-induced pulmonary fibrosis in hamsters: effect of neutrophil depletion on lung collagen synthesis. Am. Rev. Respir. Dis. 126, 737–742.

Clements, P. J. (1994). Raynaud's phenomenon, scleroderma, overlap syndromes and other fibrosing syndromes. Curr. Opin. Rheumatol. 6, 603–606.

Damsky, C. H. & Werb, Z. (1992). Signal transaction by integrin receptors for extracellular matrix: cooperative processing of extracellular information. Curr. Opin. Cell. Biol. 4, 772–781.

Davis, W. B., Fells, G. A., Sun, X., Gadek, J. E., Venet, A., & Crystal, R. G. (1984). Eosinophil-mediated injury to lung parenchymal cells and interstitial matrix: a possible role for eosinophils in chronic inflammatory disorders of the lower respiratory tract. J. Clin. Invest. 74, 269–274.

Ettinger, L. & Doljanski, F. (1992). On the generation of form by the continuous interactions between cells and their extracellular matrix. Biol. Rev. 67, 459–489.

George, E. L., Georges-Labouesse, E., Patel-King, R. S., Rayburn, H.; & Hynes, R. O. (1993). Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin. Development 119, 1079–1091.

Hay, E. D. (1991). Cell Biology of Extracellular Matrix, 2nd edn. Plenum, New York.

Hertz, M. I., Snyder, L. S., Hamon, K. R., & Bitterman, P. B. (1992). Acute Lung Injury. In: Wound Healing: Biochemical and Clinical Aspects (Cohen, I. K., Diegelmann, R. F., & Lindblad, W. J., Eds.), pp. 416–432, W. B. Saunders Company, Philadelphia, PA.

Hynes, R. O. (1992). Integrins: versatility, modulation, and signaling in cell adhesion. Cell 69, 11–25.

Hynes, R. O. & Lander, A. D. (1992). Contact and adhesive specificities in the associations, migrations, and targeting of cells and axons. Cell 68, 303–322.

Juliano, R. L. & Haskill, S. (1993). Signal transduction from the extracellular matrix. J. Cell. Biol. 120, 577–585.

Kelley, J. (1990). Cytokines of the lung. Am. Rev. Respir. Dis. 141, 765–788.

Keski-Oja, J., Raghow, R., Sawdey, M., Loskutoff, D. J., Postlethwaite, A. E., Kang, A. H., & Moses, H. L. (1988). Regulation of mRNAs for type 1 plasminogen activator inhibitor, fibronectin, and type I procollagen by transforming growth factor-β. J. Biol. Chem. 263, 3111–3115.

Khalil, N. & Greenberg, A. H. (1991). The role of TGF-β in pulmonary fibrosis. Ciba Found. Symp. 157, 194–211.

Kingsley, D. M. (1994). The TGF-β superfamily: new members, new receptors, and new genetic tests of function in different organisms. Genes & Dev. 8, 133–146.

Koch, A. E., Polverini, P. J., Kundel, S. L., Harlow, L. A., DiPietro, L. A., Elner, V. M., Elner, S. G., & Strieter, R. M. (1992). Interleukin-8 as a macrophage-derived mediator of angiogenesis. Science 258, 1798–1801.

Le Roy, E. C. (1992). Scleroderma (systemic sclerosis): Comparison with wound healing. In: Wound Healing: Biochemical and Clinical Aspects (Cohen, I. K., Diegelmann, R. F., & Lindblad, W. J., Eds.), pp. 510–522, W. B. Saunders, Philadelphia, PA.

Mainardi, C. L. (1993). Fibroblast function and fibrosis. In: Textbook of Rheumatology. 4th edn. (Kelley, W. N., Harris, E. D., Ruddy, S., Sledge, C. B., Eds), pp, 337–349, W. B. Saunders Company, Philadelphia, PA.

Massague, J. (1992). Receptors for the TGFβ family. Cell 69, 1067–1070.

Miller, E. J. & Gay, S. (1992). Collagen structure and function. In: Wound healing: Biochemical and clinical aspects (Cohen, I. K., Diggelman, R. F., & Lindblad, W. J., Eds.), pp. 130–151, W. B. Saunders, Philadelphia, PA.

Morgan, C. J. & Pledger, W. J. (1992) Fibroblast proliferation. In: Wound Healing: biochemical and clinical aspects (Cohen, I. K., Diggelmann, R. F., & Lindblad, W. J., Eds.), pp. 63–76, W. B. Saunders Company, PA.

Nathan, C. (1987). Secretory products of macrophages. J. Clin. Invest. 79, 319–326.

Oppenheim, J. J., Zachariae, C. O. C., Mukaida, N., & Matsushima, K. (1991). Properties of the novel pro inflammatory supergene "intercrine" cytokine family. Ann. Rev. Immunol. 9, 617–648.

Phan, S. H. & Thrall, R. S. (1982). Inhibition of bleomycin-induced pulmonary fibrosis by cobra venom factor. Am. J. Pathol. 107, 25.

Pisetsky, D.S. (1992). Anti-DNA antibodies in systemic lupus erythematosus. In: Rheumatic Disease Clinics of North America. W. B. Saunders Company, Philadelphia, PA.

Postlethwaite, A. E. (1994). Connective tissue metabolism including cytokines in scleroderma. Curr. Opin. Rheumatol. 6, 616–620.

Postlethwaite, A. E., Holness, M. A., Katai, H., & Raghow, R. (1992). Human fibroblasts synthesize elevated levels of extracellular matrix proteins in response to interleukin 4. J. Clin. Invest. 90, 1479–1485.

Postlethwaite, A. E., Raghow, R., Stricklin, G. P., Poppleton, H., Seyer, J. M., & Kang, A. H. (1988). Modulation of fibroblast functions by interleukin 1: increased steady-state accumulation of type I procollagen messenger RNAs and stimulation of other functions but not chemotaxis by human recombinant interleukin 1α and β. J. Cell. Biol. 106, 311–318.

Prakash, S., Robbins, P. W., & Wyler, D. J. (1995). Cloning and analysis of murine cDNA that encodes a fibrogenic lymphokine, fibrosin. Proc. Natl. Acad. Sci. USA 92, 2154–2158.

Prockop, D. J. (1990). Mutations that alter the primary structure of type I collagen: the perils of a system for generating large structures by the principle of nucleated growth. J. Biol. Chem. 265, 15349–15453.

Raghow, R. (1991). The role of transforming growth factor b in repair and fibrosis. Chest 99, 61S–65S.

Raghow, R. (1992). Heterogeneity and plasticity of extracellular matrix gene expression in fibroblasts: a pulmonary fibrosis perspective. In: Pulmonary Fibroblast Heterogeneity (Phipps, R. P., Ed.), pp 79–104. CRC Press, Boca Raton, FL.

Raghow, R. (1994). The role of extracellular matrix in post inflammatory wound healing and fibrosis. FASEB J. 8, 823–831.

Raghow, R., Lurie, S., Seyer, J. M., & Kang, A. H. (1985). Profiles of steady state levels of messenger RNAs coding for type I procollagen, elastin and fibronectin in hamster lungs undergoing bleomycin-induced interstitial pulmonary fibrosis. J. Clin. Invest. 76, 1733–1739.

Raghow, R., Postlethwaite, A. E., Keski-Oja, J., Moses, H. L., & Kang, A. H. (1987). Transforming growth factor β increases steady state levels of type I procollagen and fibronectin mRNAs posttranscriptionally in cultured human dermal fibroblasts. J. Clin. Invest. 19, 1285–1288.

Raghow, R., Irish, P., & Kang, A. H. (1989). Coordinate regulation of transforming gross factor beta-gene expression and cell proliferation in hamster lungs undergoing bleomycin-induced pulmonary fibrosis. J. Clin. Invest. 84, 1836–1842.

Reeves, W. H. (1992). Antibodies to the p70/p80 (Ku) antigens in systemic lupus erythematosus. In: Rheumatic Disease Clinics of North America. W. B. Saunders, Philadelphia, PA.

Rennard, S. I., Bitterman, P. B., & Crystal, R. G. (1984). Mechanisms of pulmonary fibrosis. Am. Rev. Respir. Dis. 130, 492.

Rinaldo, J., Goldstein, R. H., & Snider, G. L. (1982). Modification of oxygen toxicity after lung injury by bleomycin in hamsters. Am. Rev. Respir. Dis. 126, 1030.

Rothfield, N. M. (1992). Autoantibodies in scleroderma. In: Rheumatic Disease Clinics of North America. W. B. Saunders, Philadelphia, PA.

Ruoslahti, E. & Yamaguchi, Y. (1991). Proteoglycans as modulators of growth factor activities. Cell 64, 867–869.

Ruoslahti, E., Yamaguchi, Y., Hildebrand, A., & Border, W. A. (1992). Extracellular matrix/growth factor interactions. Cold Spring Harbor Symp. Quant. Biol. LVII, 309–315.

Schnieke, A., Harbers, K., & Jaenisch, R. (1983). Embryonic lethal mutation in mice introduced by retrovirus insertion into $\alpha 1$(I) collagen gene. Nature 332, 131–136.

Schwartz, M. A. & Lechene, C. (1992). Adhesion is required for protein kinase C dependent activation of the Na^+/H^+ antiporter by PDGF. Proc. Natl. Acad. Sci. USA 89, 6138–6141.

Seibold, J. R. (1993). Scleroderma. In: Textbook of Rheumatology, IV Edition (Kelley, W. N., Harris, E. D., Ruddy, S., & Sledge, C. B., Eds.), pp. 1113–1143, W. B. Saunders Company, Philadelphia, PA.

Seyer, J. M. & Raghow, R. (1992). Hepatic fibrosis. In: Wound Healing: Biochemical and Clinical Aspects (Cohen, I. K., Diegelmann, R. F., & Lindblad, W. J., Eds.), pp. 416–432, W. B. Saunders, Philadelphia, PA.

Sunderkotter, C., Steinbrink, K., Goebeler, M., Bhardwaj, R., & Sorg, C. (1994). Macrophages and angiogenesis. J. Leukocyte Biol. 55, 410–422.

Thiery, J. P. & Boyer, B. (1992). The junction between cytokines and cell adhesion. Curr. Opin. Cell Biol. 4, 782–792.

Thrall, R. S., Phan, S. H., McCormick, J. R., & Ward, P. A. (1981). The development of bleomycin-induced pulmonary fibrosis in neutrophil-depleted and complement-depleted rats. Am. J. Pathol. 105, 76.

Trelstad, R. L. & Kemp, B. D. (1993) Matrix glycoproteins and proteoglycans. In: Textbook of Rheumatology. 4th edn. (Kelley, W. N., Harris, E. D., Ruddy, S., & Sledge, C. B., Eds.), W. B. Saunders Company, Philadelphia, PA.

Turner, M. L. (1992). Cell adhesion molecules: a unifying approach to topographic biology. Biol. Rev. (Great Britain) 67, 359–377.

Vuorio, E. & deCrombrugghe, B. (1990). The family of collagen genes. Ann. Rev. Biochem. 59, 837–872.

Wahl, S. M., Hunt, D. A., Allen, J. B., Wilder, R. L., Paglia, L., & Hand, A. R. (1986). Bacterial cell wall induced hepatic granulomas: an in vivo model of T-cell-dependent fibrosis. J. Exp. Med. 163, 884–902.

Whitby, D. J. & Ferguson, M. W. J. (1991). Immunohistochemical localization of growth factors in fetal wound healing. Dev. Biol. 147, 207–215.

White, B. (1994). Immunologic aspects of scleroderma. Curr. Opin. Rheumatol. 6, 612–615.

Wigley, F. M. (1994). Clinical aspects of systemic and localized scleroderma. Curr. Opin. Rheumatol. 6, 628–636.

Williams, C. J., Vandenberg, P., & Prockop, D. J. (1993). Collagen and elastin. In: Textbook of Rheumatology (Kelley, W. N., Harris, E. D., Ruddy, S., & Sledge, C. B., Eds.) 4th edn., pp. 22–34, W. B. Saunders Company, Philadelphia, PA.

Yang, J. T., Rayburn, H., & Hynes, R. O. (1993). Embryonic mesodermal defects in α_5 integrin-deficient mice. Development 119, 1092–1105.

Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P., & Ornitz, D. M. (1991). Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to this high affinity receptor. Cell 64, 841–848.

This Page Intentionally Left Blank

INDEX