MOLECULES IN PHYSICS, CHEMISTRY, AND BIOLOGY

TOPICS IN MOLECULAR ORGANIZATION AND ENGINEERING

Molecules in Physics, Chemistry, and Biology

Volume 1

General Introduction to Molecular Sciences

Edited by

JEAN MARUANI
Centre de Mécanique Ondulatoire Appliquée,
Laboratoire de Chimie Physique,
CNRS and University of Paris, France.

Kluwer Academic Publishers
Dordrecht / Boston / London

Library of Congress Cataloging in Publication Data

CIP

Molecules in physics, chemistry, and biology.

(Topics in molecular organization and engineering)
Includes bibliographies and indexes.
Contents v. 1. General introduction to molecular sciences.
1. Molecules. I. Maruani, Jean, 1937-
II. Series.
QC173.M645 1988 539'.6 88-6811

ISBN-13978-94-010-7781-1 e-ISBN:978-94-009-2849-7
DOI: 10.1007/978-94-009-2849-7

Published by Kluwer Academic Publishers,
P O Box 17, 3300 AA Dordrecht, The Netherlands

Sold and Distributed in the U S A and Canada
by Kluwer Academic Publishers,
101 Philip Drive, Norwell, MA 02061, U S A

In all other countries, sold and distributed
by Kluwer Academic Publishers Group,
P O Box 322, 3300 AH Dordrecht, The Netherlands

Table of Contents

EVOLUTION AND ORGANIZATION OF MOLECULAR SYSTEMS

MODELLING AND ESTHETICS OF MOLECULAR STRUCTURES

Introduction to the Series

The Series 'Topics in Molecular Organization and Engineering' was initiated by the Symposium 'Molecules in Physics, Chemistry, and Biology', which was held in Paris in 1986. Appropriately dedicated to Professor Raymond Daudel, the symposium was both broad in its scope and penetrating in its detail. The sections of the symposium were: 1. The Concept of a Molecule; 2. Statics and Dynamics of Isolated Molecules; 3. Molecular Interactions, Aggregates and Materials; 4. Molecules in the Biological Sciences, and 5. Molecules in Neurobiology and Sociobiology. There were invited lectures, poster sessions and, at the end, a wide-ranging general discussion, appropriate to Professor Daudel's long and distinguished career in science and his interests in philosophy and the arts.

These proceedings have been arranged into eighteen chapters which make up the first four volumes of this series: Volume I, 'General Introduction to Molecular Sciences'; Volume II, 'Physical Aspects of Molecular Systems'; Volume III, 'Electronic Structure and Chemical Reactivity'; and Volume IV, 'Molecular Phenomena in Biological Sciences'. The molecular concept includes the logical basis for geometrical and electronic structures, thermodynamic and kinetic properties, states of aggregation, physical and chemical transformations, specificity of biologically important interactions, and experimental and theoretical methods for studies of these properties. The scientific subjects range therefore through the fundamentals of physics, solid-state properties, all branches of chemistry, biochemistry, and molecular biology. In some of the essays, the authors consider relationships to more philosophic or artistic matters.

In Science, every concept, question, conclusion, experimental result, method, theory or relationship is always open to reexamination. Molecules do exist! Nevertheless, there are serious questions about precise definition. Some of these questions lie at the foundations of modern physics, and some involve states of aggregation or extreme conditions such as intense radiation fields or the region of the continuum. There are some molecular properties that are definable only within limits, for example, the geometrical structure of non-rigid molecules, properties consistent with the uncertainty principle, or those limited by the neglect of quantum-field, relativistic or other effects. And there are properties which depend specifically on a state of aggregation, such as superconductivity, ferroelectric (and anti), ferromagnetic (and anti), superfluidity, excitons, polarons, etc. Thus, any molecular definition may need to be extended in a more complex situation.

Chemistry, more than any other science, creates most of its new materials. At least so far, synthesis of new molecules is not represented in this series, although the principles of chemical reactivity and the statistical mechanical aspects are

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, ix—x.

included. Similarly, it is the more physico-chemical aspects of biochemistry, molecular biology and biology itself that are addressed by the examination of questions related to molecular recognition, immunological specificity, molecular pathology, photochemical effects, and molecular communication within the living organism.

Many of these questions, and others, are to be considered in the Series 'Topics in Molecular Organization and Engineering'. In the first four volumes a central core is presented, partly with some emphasis on Theoretical and Physical Chemistry. In later volumes, sets of related papers as well as single monographs are to be expected; these may arise from proceedings of symposia, invitations for papers on specific topics, initiatives from authors, or translations. Given the very rapid development of the scope of molecular sciences, both within disciplines and across disciplinary lines, it will be interesting to see how the topics of later volumes of this series expand our knowledge and ideas.

WILLIAM N. LIPSCOMB

- a -

- b -

- c -

- d -

Preface to *Molecules in Physics, Chemistry, and Biology*

When we decided to organize an International Symposium dedicated to Professor Daudel, a question arose: on which themes should such a Symposium bear? After having reviewed all the themes on which Professor Daudel has worked during his long career, Imre Csizmadia and myself were somewhat at a loss; these themes ranged from Atomic Physics to Molecular Biology, with a stress on Theoretical Chemistry. Then I recalled a conversation I had in 1968, when I was in Vancouver, with Harden McConnell, on leave from Stanford. I asked him why he had switched to Biology; he answered: "I'm often asked this question. But I don't feel I've ever switched to Biology. I have always been interested in molecules, just molecules: in Physics, Chemistry, and Biology". I felt this flash of wit would make a perfect title for a Symposium dedicated to Professor Daudel, who has also been interested in molecules in Physics, Chemistry, and Biology, but from a theoretical viewpoint.

However, when it came to preparing a content appropriate to this title, we ended up with a several-page program, which defined what could have been some kind of an advanced-study institute, involving most of Physical Chemistry and parts of Molecular Biology. We announced the Symposium on that pluridisciplinary basis and then started receiving answers from invited speakers and proposals for communications. While classifying the letters, it appeared to us that a few key themes had emerged, which seemed likely to constitute 'hot topics' of the Molecular Sciences in the late 1980's and early 1990's. Indeed there are fashions in Science too, whether these are induced by the natural development of the sciences or by economic or cultural constraints. Afterwards we did our best to fill

---

LEGENDS TO THE PHOTOGRAPHS OF PLATE A
(Photographs by Miss Cristina Rusu)

— a — Minister of Research Alain Devaquet (on the left) awarding the Golden Medal of the City of Paris to Professor Raymond Daudel (on the right) in Paris City Hall. In the background, from left to right: Jean-Marie Lehn, William Lipscomb (between Devaquet and Daudel), Bernard Pullman, Jacques-Emile Dubois, Georges Lochak (all three wearing spectacles), Ernest Schoffeniels.

— b — William Lipscomb and Jean Maruani chatting after the ceremony. Also on the picture: Bernard Pullman (left), Jacques-Emile Dubois (center), Paul Caro (right).

— c — Senator Louis Perrein opening the closing banquet in the Senate House. From left to right: Alberte Pullman, Raymond Daudel, Jean-Pierre Changeux, Nicole D'Aggagio, Stefan Christov, Christiane Bonnelle.

— d — Composer and pianist Marja Maruani-Rantanen and Jean-Yves Metayer's string trio *I Solisti Europa* performing for participants in the Concordia Hotel.

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, xiii—xvii

what seemed to be gaps in the consistency of the emerging program. The main lines of the resulting program are recalled by Professor Lipscomb in his Introduction to the Series.

The Symposium gathered about 200 people, with interests ranging from the History and Philosophy of the Molecular Concept to Molecular Neurobiology and Sociobiology. A few social events were arranged, in order to help bring together participants with different interests, who otherwise would have tended to miss sessions not belonging to their own specialty. Miss Cristina Rusu recorded these oecumenical moments in photographs, a few of which are shown in Plate A.

During the nine months following the Symposium, I managed to gather together about 70% of the invited papers and 30% selected posters, as well as a few contributions not presented during the Symposium but expected to complete the Proceedings. The authors were requested to submit 'advanced-review' papers, including original material, and most of the manuscripts were refereed. The resulting arrangement of the topics is outlined in Table 1. In spite of the variety of the topics, there is a definite unity in the arrangement. This results from the specificity of the Molecular Sciences, which arises from the particular role played by the molecular concept in Science. In the hierarchy of structures displayed by Nature, molecules, supermolecules and macromolecules are situated just between atoms (which define the chemical elements) and proteins (which define biological

TABLE 1

Vol. I — General Introduction to Molecular Sciences
Part 1 — papers 01—03: History and Philosophy of the Molecular Concept
Part 2 — papers 04—06: Evolution and Organization of Molecular Systems
Part 3 — papers 07—11: Modelling and Esthetics of Molecular Structures

Vol. II — Physical Aspects of Molecular Systems
Part 1 — papers 12—13: Mathematical Molecular Physics
Part 2 — papers 14—15: Relativistic Molecular Physics
Part 3 — papers 16—17: Molecules in Space
Part 4 — papers 18—21: Small Molecular Structures
Part 5 — papers 22—25: Nonrigid and Large Systems
Part 6 — papers 26—28: Molecular Interactions
Part 7 — papers 29—33: Theoretical Approaches to Crystals and Materials

Vol. III — Electronic Structure and Chemical Reactivity
Part 1 — papers 34—40: Density Functions and Electronic Structure
Part 2 — papers 41—45: Structure and Reactivity of Organic Compounds
Part 3 — papers 46—49: Theoretical Approaches to Chemical Reactions

Vol. IV — Molecular Phenomena in Biological Sciences
Part 1 — papers 50—51: Biomolecular Evolution
Part 2 — papers 52—53: Biomolecular Chirality
Part 3 — papers 54—55: Topics in Molecular Pathology
Part 4 — papers 56—58: Topics in Biomolecular Physics
Part 5 — papers 59—63: Molecular Neurobiology and Sociobiology

specificity). In Physical Chemistry, indeed, there are thermodynamic, spectroscopic and diffraction data specifically related to molecular structure and dynamics.

Among the questions which arise in the Molecular Sciences, one may stress the following.

— How can a molecule be strictly defined with respect to the constitutive atoms, on the one hand, and the molecular gas, liquid, or solid, on the other? — Use of Topology and Fuzzy-Set Theory, Quantum and Statistical Mechanics, Effective Hamiltonian Operators and Reduced Density Matrices, X-ray and Neutron Diffraction, UV and IR Spectroscopy, etc. ('Molecular Phenomenology and Ontology').

— While hydrogen and helium constitute together 99% of the total mass of the natural elements (with, thank God! traces of heavier elements, including carbon), is molecular complexity a unique feature of the Earth or is it deeply related to the very structure of our Universe? Were Life and Man built into Nature or are they merely accidents? ('Molecular Cosmology and Evolution').

— What are the origin, nature and transfer of the information content packed in a molecular system? How can molecular information be extracted by the modelling of molecular structures? How can levels of information ordering be defined, and what are the relations between the information on simple substructures and that on complex superstructures? Can the higher levels of organization and functioning be understood in purely physicochemical terms? How do molecular assemblies cooperate to form organized or living structures? ('Molecular Organization and Cybernetics').

— Chemical laboratories and industries have created more molecules than there have been found in Nature, particularly pharmaceutics and polymers. Even such physical properties as superconductivity or ferromagnetism are no longer limited to classical metallic materials, but may also be found in molecular materials ('Molecular Synthesis and Engineering).

— Biological specificity and immunity are understood today basically as molecular phenomena related to the DNA and protein structures. Tiny structural modifications in these macromolecules may lead to metabolic deficiencies or other functional disorders ('Molecular Pathology').

— Communication within and between cells and organs in a living organism, as well as between individuals (particularly in sexual activity) in a species, and between species in an ecosystem, occurs very often through molecular interactions ('Molecular Communication').

Most of these and other related questions were dealt with in the Symposium, the Proceedings of which are published in this Series. Future volumes in the Series are expected to develop specific topics related to these questions.

The Symposium was sponsored by various bodies and companies, which are listed in Table 2. They are all gratefully acknowledged for their (material or moral) help, which made possible this gathering. The international honorary committee,

TABLE 2

SPONSORS

Ministère de l'Education Nationale
Ministère des Relations Extérieures
Ville de Paris

Centre National de la Recherche Scientifique
Commissariat à l'Energie Atomique
Institut National de la Santé et de la Recherche Médicale
Institut National de Recherche Pédagogique

Université Paris VI
Université Paris VII
Ecole Supérieure de Physique et Chimie Industrielles

World Association of Theoretical Organic Chemists
Fondation Louis de Broglie

Rhône-Poulenc
Moët-Hennessy
Amstrad France
Alain-Vaneck Promotion

COMMITTEES

Centre de Mécanique Ondulatoire Appliquée

and

International Honorary Committee

Sir D. Barton (*U.K.*)
J-P. Changeux (*France*)
M. Eigen (*F.R.G.*)
J. I. Fernàndez-Alonso (*Spain*)
K. Fukui (*Japan*)
G. Herzberg (*Canada*)
F. Jacob (*France*)
W. N. Lipscomb (*U.S.A.*)
P. O. Löwdin (*Sweden*)
H. M. McConnell (*U.S.A.*)
C. A. McDowell (*Canada*)
Sir G. Porter (*U.K.*)
I. Prigogine (*Belgium*)
B. Pullman (*France*)
M. Simonetta† (*Italy*)

† Deceased in 1986.

Local Organizing Committee

R. Acher (*Biological Chemistry*)
D. Blangy (*Molecular Biology*)
C. Bonnelle (*Physical Chemistry*)
P. Caro (*Inorganic Chemistry*)
P. Claverie† (*Theoretical Chemistry*)
I. G. Csizmadia (*Organic Chemistry*)
J-E. Dubois (*Molecular Systemics*)
A. Laforgue (*Theoretical Chemistry*)
R. Lefebvre (*Molecular Photophysics*)
J-M. Lehn (*Supramolecular Chemistry*)
G. Lochak (*Quantum Mechanics*)
P. MacLeod (*Molecular Neurobiology*)
J. Maruani (*Molecular Physics*)
P. Rigny (*Physical Chemistry*)
J. Serre (*Theoretical Chemistry*)

† Deceased in 1988.

also given in Table 2, involved fifteen distinguished scientists from ten different countries, including eight Nobel Laureates. May I express my gratitude to all of them, especially to those who managed to participate actively in the Symposium. The local organizing committee involved mostly French scientists belonging to different fields (Table 2), reflecting the interdisciplinarity of the meeting. They are all most gratefully thanked for their help and encouragement. Special thanks go to Prof. I. G. Csizmadia, who helped enormously in the early stages of the

organization of the meeting, and to Dr P. Claverie, recently deceased, who helped in the late stages of the organization and also in the selection of the papers for these volumes. Finally my thanks go to Bernard and Isabelle Decuypère, who prepared the indexes, and to the Staff of Kluwer Academic Publishers, for their pleasant and efficient cooperation.

I hope these books will prove to be of as much interest to the reader as the meeting was to the participants.

JEAN MARUANI

Preface to Volume 1: Molecules in the Cosmic Scale of Complexity

The title chosen for the Symposium, the first volume of which is introduced here, specifies the scope of the endeavour of the organizing committee. Its interdisciplinarity manifests their intention to relate the molecular concepts emerging from the highly different and specialized fields of *physical, chemical,* and *biological* sciences. In my mind, such a synthesis is best arrived at if we also bring into the picture the *astronomical* and *cosmological* point of view. This is what I will try to do here.

It is fair to say that the most important theme emerging from contemporary Cosmology is that of *matter organization.* The best and widely accepted theory of the Universe, the Big Bang, gives the temporal and spatial frame in which the various processes take place, which lead to the gradual build-up of structures of increasing complexity. The various sciences, mostly Physics, Chemistry, and Biology, describe these various processes with the concepts of their specific methodology.

Presently we believe that quarks and leptons are the elementary particles from which everything is made. This affirmation could be challenged when the next generation of accelerators, reaching energies of multi-TeV, becomes operational, in the next decades.

The initial stage of matter organization is identified with the so-called quark-hadron transition, at temperatures of 200 MeV or so, when the cosmic clock indicates a few micro-seconds. At this moment, all the quarks associate themselves, two by two, to form the pions, or, three by three, to form the nucleons. These processes are governed by the nuclear force, that is, by the action of an exchange particle called the gluon. The reactions take place uniformly throughout the entire space of the Universe (cosmological scale).

The second building step is the fusion of the nucleons into atomic nuclei. It is also governed by the nuclear force but in a much weaker version, quite analogously to the molecular binding in comparison with the atomic binding of electrons around a nucleus. After an early, brief episode of nuclear fusion on a cosmological scale — leading mostly to helium — the reactions forming heavier nuclei — all the way up to uranium — take place in the hot centers of the stars, throughout the entire life of the galaxies, such as our own Milky Way.

The following steps of matter organization involve the formation of atoms and molecules, by the association, first, of the nuclei with the electrons, and, second, of the atoms, to generate molecular structures. Since these processes involve the electromagnetic force, they cannot take place in the stellar cores where the nuclei

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, xix—xxi.

are formed. The electromagnetic binding energies are far too small (a few eV) to withstand the dissociating effects of the thermal stellar energies (several keV).

The electromagnetic building activity goes on in the matter ejected from the stars, mostly shortly after their death. In the case of small stars, the outer layers are progressively pushed away, leading to the formation of a glowing planetary nebula. For massive stars, the disruption is more catastrophic and leads to the explosive dispersal of most of the stellar matter, in a supernova.

From there on the set-up is the same; the ejected atoms cool rapidly to the extremely low temperature of the interstellar space. The ejecta become vast interstellar chemical laboratories where atoms and molecules result from an intense electromagnetic activity.

Because densities are low, the encounters of atoms are rare. Most of the molecules formed in these conditions are small, involving, at best, two, three, or four atoms. Nevertheless, larger molecules of more than ten atoms, and perhaps as large as 40 atoms, have been identified, through their specific millimetric radio emission.

One notable point is that *all* molecules involving more than three atoms incorporate some carbon atoms. This observation has far reaching implications. It shows that, throughout our galaxy, as well as in the neighboring galaxies where the same phenomenon is observed, carbon remains Nature's favorite building block. It is probably not incorrect to infer from there that, if life exists on other planets, it is carbon-built and not silicon-built as has often been suggested.

The search for new interstellar molecules goes on. Bigger specimens are very likely to be caught. However, it is quite unlikely that macromolecules of biological size will ever be found. The destructive effect of the various ionizing radiations, such as UV and cosmic rays, severely limits the duration of such molecules if they are ever formed.

In parallel with the formation of small molecules, the electromagnetic activity in stellar ejecta leads to the elaboration of another kind of atomic structure: interstellar dust particles. Here the building atoms, mostly oxygen, silicon, magnesium and iron, arrange themselves in a crystalline network. Astronomical photography of certain stars (such as the Pleiades) shows that they are surrounded by vast clouds of dust particles, on which the blue component of their light is reflected.

We believe today that these dust particles are the building blocks of solid planets such as our Earth. The agglomeration takes place in the equatorial disk surrounding a newborn star, where vast amounts of gas and dust are trapped by the gravitational field of the still collapsing stellar embryo. The final product (at least in our Solar system . . .) is a collection of solid bodies orbiting around the central star, some of them surrounded by a gaseous atmosphere and a liquid ocean....

Matter density in water is some 20 orders of magnitude larger than in the interstellar clouds. The rates of collision and atomic encounters are proportionally increased. In this natural chemistry laboratory of a new kind, we expect to

find much larger molecular aggregates than in space (especially if we also take into account the shielding effect of water layers to ionizing radiations).

We generally accept the idea that living processes have appeared as a result of millions of years of continuous chemical reactions, energized by the UV radiation of the Sun (in highly reduced proportions as compared to space conditions). However we must confess that we have still very little information on the exact paths followed from the first amino-acids and puric bases to DNA, the proteins, and the biomolecular machinery. . . .

From there on, we simply follow the road traced by biological evolution, leading to complex organisms and finally to ourselves, assembled here and talking about molecules. . . .

The ascent of complexity in Nature can be compared to the building-up of a pyramid made up of superimposed alphabets, as illustrated in the figure.

Atoms are words made of the three letters, *protons, neutrons, and electrons. Protons and neutrons* are words made of two letters, the *u* and *d* quarks. In the same sense, *molecules* are words made of the some *ninety atomic species. Cells* are associations of *molecules. Organisms* are associations of *cells*. . . .

The pyramid of complexity. The situation of the molecular structures in the organization of matter.

Organisms, ecosystems

Cells

Aggregates, polymers

Molecules

Atoms

Nucleons

Quarks, electrons, neutrinos

Our subject is the intermediate step between atoms and life. Physics looks *down* into the pyramid, Chemistry explores it *horizontally*, while Biology concentrates *on the upper levels.* By focusing our attention to the relationship between the various aspects of the molecular concept, we reach a new depth in the study of Nature's crucial endeavour: the rise in complexity.

HUBERT REEVES

History and Philosophy of the Molecular Concept

Phenomenology and Ontology of the Molecular Concept

E. SCHOFFENIELS
Laboratoire de Biochimie Générale et Comparée, Université de Liège, 17, place Delcour, B-4020 Liège, Belgium.

Introduction

As a biochemist, I have been interested in the history of the ideas giving rise to what could be termed today the system of biochemistry. The system of biochemistry rests upon a few basic concepts, some borrowed from the field of chemistry, others specific to the interpretation of biology and biological phenomena in terms of chemistry, physics and thermodynamics. Amongst the first category the concept of molecule is obviously of prime importance and this paper will deal essentially with the history of the ideas leading to the contemporary views relating to the organization of matter. I shall only mention as belonging more narrowly to biochemistry the specificity of the catalytic processes leading to the organization of metabolic pathways in catenary sequences of reactions, their integration and control, the existence of pairs of relata, auto-organization and self reproduction, etc., all processes subtended by concepts such as those of macromolecule, molecular architecture or molecular anatomy, molecular anatomy of cells and the like. As far as my topic is concerned I shall therefore refer here more extensively only to the concept of macromolecule. In the first part of this paper I shall recall the long and arduous path leading ultimately to the idea of ordered entities at the microscopic level of dimension — the term microscopic being taken here in the sense of the physicists. Since order is also the rule in biology, from the ecosystem down to the molecule, and since this order is mainly the result of the existence of pairs of relata, one member of the couple being often, if not always, a macromolecule, I shall also describe how the concept of macromolecule was identified, starting with the work of Berzelius and Graham on the colloids and ending with the more recent concept of molecular biology, thus closing the circle since it brings us back to a field more akin to organic chemistry than biology.

If one wishes to understand how chemistry as a science has evolved, it is necessary to provide some historical background starting with the period reaching from the Greek classicists to the alchemists during what could be called the prescientific era. However, if one excludes the daily work of the craftsmen involved in such technological practical aspects of chemistry as distillation, melting of metals, preparation of dyes, reagents and remedies, the ideas that were really seminal to the development of chemistry as a science are rather few. Most of the explanations of natural events had a mythological and magical character. Therefore, despite

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, 3—24.

the vivid interest in reviewing that part of history, not forgetting the oriental approach, it seems to me more adequate, given the time alloted, to very briefly sum up the prescientific era and to devote most of the time to an examination of the ontology of the molecular concept in a transition period and in the scientific era.

A. The Organization of Matter as Viewed by the Greek Philosophers

The view that the ultimate structure of matter is discrete rather than continuous is ascribed to Democritus (420 BC) according to whom the only existing things are the atoms and empty space, "all else is mere opinion". From this point of view, qualities such as smell, taste, colour, etc. are secondary. They cannot be associated with the individual atoms described solely in terms of motion and geometry i.e. position, shape, size. As a scientific explanation as we see it today the position of Democritus, though speculative, is rather profound and avoids the pitfall of a straight and naive reductionism. Indeed, to say that a substance is red because its atoms are red would not offer an explanation of any consequence. This primitive atomic theory was shared by Epicurus whose ideas were transmitted by Lucretius in his De Rerum Natura and was also echoed much later by Giordano Bruno (1548—1600), Francis Bacon (1561—1626), René Descartes (1596—1650) and Isaac Newton (1642—1727).

However this completely materialistic theory of Democritus was strongly opposed by more mystical philosophers from the school of Pythagoras whose thinking became influential under the leadership of Socrates, Plato and Aristotle and who set the standard of scientific thoughts until the XVIIth century.

B. The Emergence of the Concept of Molecules

1. THE ORGANIC MOLECULES OF BUFFON AND THE INTEGRAL MOLECULES (MOLÉCULES INTÉGRANTES) OF HAÜY

When Buffon writes the word *molecule* he clearly means an extremely small material particle. This is evident when we look at the way he envisages the formation of a crystal since he uses the expression "stony particles" detached by water from glossy or calcareous materials, that are thereafter aggregated [3].

More interesting is the qualification of "organic" that he gives to other "molecules" the association of which, according to some scheme, gives rise to the various animal species.

These "organic molecules" are endowed with special properties that make them different from inert material, and more specifically they are endowed with motion.

Haüy in 1784 developed a theory of crystal structure based on a three-dimensional repeat of elementary geometry. He built models of crystals and the units that were repeated in space were what he called the integral molecule [14].

It is here important to emphasize the fact that the mobility of the "molecules" of Buffon was considered to be an essential feature of living systems: they were living particles and this was a strong argument in favor of the vitalist theory since no external physical cause appeared to be responsible for the observed displacements.

2. THE BROWNIAN MOTION

At this point, it is adequate to recall the observations made by the British botanist R. Brown (1773—1853) who, in 1827, while examining fertilization in plants under the microscope, discovered particles in rapid movement thus adducing a rather convincing argument in favor of the ideas of Buffon regarding the organic molecules. These were indeed assumed to be highly concentrated in the semen of plants and animals. Brown also observed the phenomenon with crushed glass, mineral powders and the like. Brongniart in 1826 had also made the same type of observations on grains of pollen and one had to wait until 1877 when Delsaux established an analogy with the kinetic theory of gases and explained the Brownian motion as being due to a collision of the solvent molecules with the suspended solid particles.

Therefore the vitalist theory received the first serious blow and the difference between the organic and inorganic world as expressed mainly by the motion was no longer tenable.

3. DALTON'S ATOMIC THEORY

The concept of "element" had been clearly stated in 1661 by Robert Boyle. In his book entitled *The Skeptical Chymist* he defines chemical elements as those substances that cannot be further resolved into other substances by any means. As a matter of fact, by the end of the XVIIIth century some 30 substances conforming to the definition had been described. The law of conservation of mass in chemical reactions, carefully confirmed by Lavoisier, also gave strong support to the idea that all chemical changes are just the reorganization of unaltered basic units. Also the law of definite proportions — stating that every pure chemical compound contains fixed and constant proportions by weight of its constituent elements — was formulated by Proust (1799). These various propositions could be justified by Dalton in 1803—1804 and were given a unified expression in his book *A New System of Chemical Philosophy* (Part 1, 1808; Part 2, 1810; Part 3, 1877). The main features of Dalton's atomic theory of 1803—04 were already exposed systematically by Thomas Thomson's *System of Chemistry* (1807). They are:

1. Matter is made of indivisible atoms.
2. All the atoms of a given element are identical in weight and in every other property.
3. Different elements have different kinds of atoms.

4. Atoms are indestructible.
5. Chemical reactions are merely rearrangement of atoms.
6. A complex substance is made of its elements through the formation of compound atoms containing a definite but small number of atoms of each element.

What is remarkable in Dalton's theory is the clarity and precision of the statements rather than their fundamental novelty. It was therefore relatively easy to put its predictions to quantitative tests. To strength his position Dalton added two other principles, the rule of greatest simplicity (later shown to be incorrect) and the law of multiple proportions. Thus, for example, in Dalton's day the only known compound of oxygen and hydrogen was water, formed by the reaction of about 7 parts by weight of oxygen and 1 part by weight of hydrogen. The ratio 7/1 was the result of some inaccuracy of analysis and the rule of simplicity would specify a formula of the type OH thus leading to the idea that 1 atom of hydrogen should weigh 7 time less than 1 atom of oxygen.

Dalton also considered but rejected the hypothesis that equal volumes of gases contain equal numbers of atoms and the idea that elementary substances might exist as polyatomic molecules did not occur to him.

Despite this sort of inaccuracy the ingenuity of Dalton must be recognized.

As to the law of multiple proportions, Dalton stated that whenever two elements combine in more than one proportion by weight, the different proportions bear a simple ratio to one another. Thus it was known that carbon and oxygen formed two distinct compounds: the first compound contained 28% by weight of carbon and 72% by weight of oxygen while the second compound contained 44% carbon and 56% of oxygen (Table I). When looking at the ratio of oxygen to carbon it is seen that it is almost exactly twice as great in the first compound as it is in the second. This was also one of the great achievements of Dalton, to show that it is so.

The results of Table I can be interpreted in two ways. If it is assumed that the first compound is CO, the second compound must be C_2O since the analysis shows that it contains half as much oxygen relative to the same amount of carbon. On the other hand, if the second compound is CO, the first must be CO_2. This ambiguity was resolved by the study of gases which also provided the first definite estimates of atomic size and weight.

Table I. Proportion by weight of carbon and oxygen in two compounds.

	% carbon	% oxygen	$\frac{\%\text{ oxygen}}{\%\text{ carbon}}$
First compound	28	72	2.571
Second compound	44	56	1.272

By applying Dalton's rules, if the first compound is CO the second must be C_2O. But if the second compound is CO, then the first is CO_2.

4. GAY-LUSSAC LAW OF COMBINING VOLUMES AND AVOGADRO—AMPÈRE LAW

The Gay-Lussac law of combining volumes (1808) — i.e. volumes of combining gases are in the ratio of small integers — could be recognized as a counterpart of the law of definite proportions, concerned only with the weights of the reacting substances. The simplicity of these volumetric relationships led Avogadro to propose in 1811 that equal volumes of different gases at the same pressure and temperature contain equal numbers of particles.

If this hypothesis is correct, the fact that 2 volumes of hydrogen react with 1 volume of oxygen means that 2 particles of hydrogen react with one particle of oxygen. If the particles in each cases are a single atom then:

$$2\,H + O = H_2O$$

But, according to the same hypothesis, this would imply that to each volume of oxygen reacting there would only be one volume of water produced, in contradiction with experiment, which yields two volumes.

If now, it is assumed that the smallest particle of hydrogen is a single atom while in the case of oxygen it is made of two atoms, one could write:

$$2\,H + O_2 \rightarrow 2\,HO$$

But other reactions were known to Avogadro where three volumes of hydrogen combine with one volume of nitrogen to form two volumes of ammonia. This cannot be explained except by assuming that hydrogen as well as nitrogen particles are made each of two atoms. This led to the proposal that the reaction between hydrogen and oxygen is of the type:

$$2\,H_2 + O_2 = 2\,H_2O$$

which is easily explained by assuming a polyatomic structure for the elements.

In 1814, Ampère sent to *Annales de Chimie* a letter "Sur la détermination des proportions dans lesquelles les corps se combinent d'après le nombre et la disposition respective des molécules dont leurs particules intégrantes sont composées" [1].

In this paper, Ampère rediscovered the same concept independently of Avogadro [20]. Today we refer to it as the Avogadro—Ampère law. This concept was rather well accepted, but later in the century, up until about 1858, the chemists still believed that the formula of water is OH. Marcelin Berthelot (1827—1907) was still using this formula in 1891! Its very success somehow puts into oblivion another important idea also presented in the same paper: molecular geometry is defined in terms of simple polyhedra in which atoms are placed at the vertices. Of course, to make sense today, what Ampère called "particle" should be read (as pointed out by Laszlo [20]) "molecule" and what he called "molecule" should be understood as meaning "atom".

5. THE SHAPE OF MOLECULES ACCORDING TO AMPÈRE

In his letter to *Annales de Chimie* Ampère formulates with clarity and precision the idea of *representative shape of a molecule* (forme représentative de la particule). It is based on some considerations on chemical bonding as well as on the foundations laid by crystallographers and more specifically on the conception of R. J. Haüy (1743—1822) regarding the crystal structure viewed as the three-dimensional repeat of an elementary geometry [14]. According to Ampère [1]:

Des conséquences déduites de la théorie de l'attraction universelle, considérée comme la cause de la cohésion, et la facilité avec laquelle la lumière traverse les corps transparents, ont conduit les physiciens à penser que les dernières molécules des corps étaient tenues par les forces attractives et répulsives qui leur sont propres, à des distances comme infiniment grandes relativement aux dimensions de ces molécules.

Dès lors leurs formes, qu'aucune observation directe ne peut d'ailleurs nous faire connaître, n'ont plus aucune influence sur les phénomènes que présentent les corps qui en sont composés, et il faut chercher l'explication de ces phénomènes dans la manière dont ces molécules se placent les unes à l'égard des autres pour former ce que je nomme une particule. D'après cette notion, on doit considérer une particule comme l'assemblage d'un nombre déterminé de molécules dans une situation déterminée, renfermant entre elles un espace incomparablement plus grand que le volume des molécules; et pour que cet espace ait trois dimensions comparables entre elles, il faut qu'une particule réunisse au moins quatre molécules. Pour exprimer la situation respective des molécules dans une particule, il faut concevoir par les centres de gravité de ces molécules, auxquels on peut les supposer réduites, des plans situés de manière à laisser d'un même côté toutes les molécules qui se trouvent hors de chaque plan. En supposant qu'aucune molécule ne soit renfermée dans l'espace compris entre ces plans, cet espace sera un polyèdre dont chaque molécule occupera un sommet, et il suffira de nommer ce polyèdre pour exprimer la situation respective des molécules dont se compose une particule. Je donnerai à ce polyèdre le nom de forme représentative de la particule.

Ampère then proceeds to define elementary geometries that should account for the structure of matter:

Si nous considérons maintenant les formes primitives des cristaux reconnues par les minéralogistes et que nous les regardions comme les formes représentatives des particules les plus simples, en admettant dans ces particules autant de molécules que les formes correspondantes ont de sommets, nous trouverons qu'elles sont au nombre de cinq: le tétraèdre, l'octaèdre, le parallélépipède, le prisme hexaèdre et le dodécaèdre rhomboïdal.

Les particules correspondantes à ces formes représentatives sont composées de 4, 6, 8, 12 et 14 molécules; les trois premiers de ces nombres sont ceux dont nous avons besoin pour expliquer la formation des particules des gaz cités tout à l'heure; j'ai montré dans mon Mémoire que le nombre 12 est celui qu'il faut employer pour exprimer la composition des particules de plusieurs combinaisons très remarquables, et que le nombre 14 rend raison de celle des particules de l'acide nitrique, comme il serait si on pouvait l'obtenir sans eau, de celle des particules du muriate d'ammoniaque, etc.

It should be noted here, with Laszlo [20], that the five polyhedra considered by Ampère differ from the five Platonic solids, which are the tetrahedron (4), the octahedron (6), the cube (8), the icosahedron (12) and the pentagonal dodecahedron (20). Also in the models proposed by Ampère, it is evident that the underlying notion is that of valence. One has however to wait much longer to gain a more precise definition of the chemical bonding to replace the vague idea of atomicity.

One may wonder why such a critical contribution as that of Ampère in his paper of 1814 was overlooked until the middle of the century. As remarked by Pauling [24] in conjunction with Avogadro's contribution:

> The value of Avogadro's law remained unrecognized by chemists from 1811 until 1858. In this year S. Cannizzaro (1826—1910), an Italian chemist working in Geneva, showed how to apply the law systematically and immediately the uncertainty regarding the correct atomic weights of the elements and the correct formulas of compounds disappeared. Before 1858, many chemists used the formula OH for water and accepted 8 as the atomic weight of oxygen; since that year, H_2O has been accepted as a formula for water by everyone. The failure of chemists to accept Avogadro's law during the period from 1811 to 1858 seems to have been due to a feeling that molecules were too "theoretical" to deserve serious consideration.

This comment of Pauling applies also to the Ampère paper of 1814 and one reason explaining a lack of immediate recognition could well be related to the use by Avogadro and Ampère of the word "hypothesis", laden with romantic flavors.

6. TOWARDS A MOLECULAR ARCHITECTURE: THE STRUCTURE OF AROMATIC COMPOUNDS ACCORDING TO KEKULÉ

In 1858, Kekulé, simultaneously with A. S. Couper, showed that two postulates were sufficient to the description of organic compounds: the tetravalence of carbon and the ready ability of this element to form long chains, the binding between two carbons atoms requiring either one or two "units of affinity". Until then, the modern notion of valence was not in use; one spoke of affinity or atomicity. Though chemical analysis had greatly progressed, a consensus regarding the exact meaning of Dalton's theory of atoms was far from being reached. If it was clear that the combination of various "particles" were achieved according to well-defined proportions, the exact meaning of the observation was not clear. Many interpretations were available. For instance the *theory of types* according to which, from simple *types* one could obtain various compounds by substituting the hydrogen atom with more or less complex radicals. This idea was expanded by the chemist C. F. Gerhardt (*Traité de Chimie organique*, 1853—1856) who proposed four inorganic types (water, ammonia, hydrochloric acid, hydrogen) from which in principle all the organic derivatives could be obtained. Gerhardt defined the "atomicity" of an acid as being a number equal to the number of water molecules from which this acid is derived within the framework of the theory of types. Of course this led to many dificulties and as a result the whole idea was soon rejected. At that time it was not even sure that the chemical symbols designated real entities. Were they really representing the compound under study?

More specifically the problem was concerned by the way atoms are bound together and by their exact mass. Thus some attributed a mass of 6 for carbon while other proposed 12. Also for oxygen it was either 16 or 8. Let me also recall that for most of the chemists the structure of water was still OH.

It is in this context that the ideas of Kekulé, together with those of Couper,

Butlerov, Frankland, Williamson, Olding and Wurz, who in the early 1850s set to work on a careful analysis of carbon compounds, blossomed in the notion of valence (1868) leading eventually to the ideas of carbon's tetravalence and the long chain-forming capacity of carbon atoms. This structural theory solved some basic problems since it became possible to rationalize the coexistence of several carbon atoms in a single molecule, an observation that had puzzled many chemists.

Moreover Kekulé was building models that he was using extensively, even in his teaching. Therefore we can trace back to Ampère and Kekulé a practice that became such a powerful tool in contemporary science and more specifically in our approaches to the structure of macromolecules.

Benzene was discovered in 1825 by Faraday and thoroughly studied by such chemists as Wöhler, Dumas, Berzelius, Liebig and Laurent. Thus "aromatic" chemistry was founded. Also the terpenes, a name coined by Kekulé in 1864 for another family of compounds rich in carbons were studied. But due to the lack of a general consensus regarding the validity of existing concepts, the meaning of chemical symbol and even the very existence of atoms and molecules were in a state of great chaos. Therefore the ideas of Kekulé regarding the notion of valence as applied to carbon was certainly of great significance. Kekulé worked at the structure of benzene during the years 1862—65. At the meeting of the Société chimique in Paris, on January 27, 1865 Kekulé presented his paper "Sur la constitution des substances aromatiques" [16] in which he suggested that all the aromatic compounds were made of a ring of six carbon atoms. Interestingly enough, as published, the formula was cyclic and planar. It had localized double bonds as in the representation of 1866 [17]. In 1872, Kekulé proposed that the three double bonds in benzene were not fixed but rather oscillate between two adjoining positions [18]. Therefore two formulas must be used to characterize the structure of benzene (Fig. 1).

Thus the stage is now set to elaborate more deeply on the structure of organic molecules and to account for the observations of chemists in terms of molecules with definite geometries. The properties of compounds made of carbon, hydrogen, nitrogen and oxygen could be understood in terms of characteristic valencies.

The key notion of asymmetric carbon in relation to the optical activity of a compound was discovered by Le Bel and van't Hoff in the 1870s and soon the notion of planar aromatic molecules would include heterocyclic systems. Now a few words about the dream of Kekulé. Much has been written about this dream in which Kekulé seems to have seen a snake biting its own tail. Three interpretations seem to be favored today. According to Partington [23], a French chemist, A. Laurent had already in 1854 used a hexagonal structure to represent benzoyl chloride, but he also used a hexagon to designate ammonia, stating that in both compounds the atoms are organized according to hexagonal figures. The work of Laurent was certainly known to Kekulé, since he referred to it in 1858. Therefore, it could be possible that unconsciously the hexagon of Laurent could have become, in the "dream" of Kekulé, the snake biting his tail.

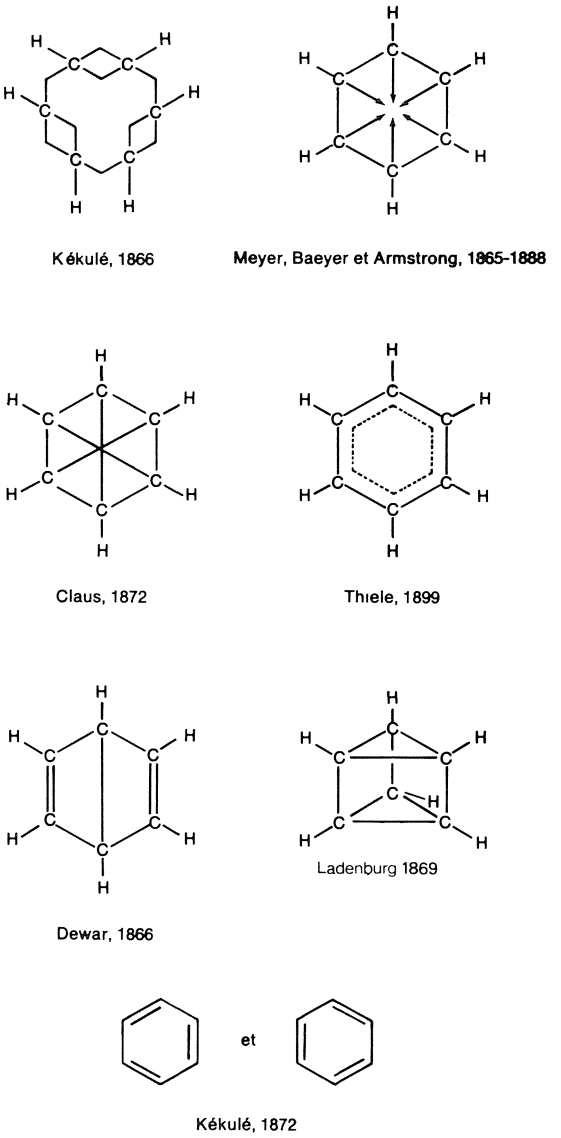


Fig. 1. Some formulae proposed for benzene at Kekulé's time. (After Laszlo, 1979).

According to Laszlo [19], when Kekulé told the story of his dream in the 1890s he could have been influenced by the representation of Ouroboros, one of the central alchemical symbols that he could have seen in the contemporary writings of Berthelot on alchemy.

For Thuillier [28] it seems fair to accept the stories told by Kekulé about his dreams of atoms, snakes and the like even if they have been a bit embellished.

7. THE EMERGENCE OF SPECTROSCOPY

From the work of Newton on the decomposition of sunlight into seven colors (1664) to the work of Ångström (1868) measuring the Fraunhofer lines and expressing them in units of 10^{-10} meter, a great number of observations were made, first on the spectral distribution of heat from the sun (Herschel, 1800) as well as the effect of spectral light upon silver salts (Ritter, 1801) leading to the discovery of the infrared and ultraviolet spectra. Then the relation between color and wavelength was established by Thomas Young (1802) who calculated the approximate wavelengths of the 7 colors of Newton.

By modifying the Newton experiment with regard to the solar spectrum, Joseph von Fraunhofer was able to show that the solar spectrum is interrupted by many hundreds of dark lines; those that were measured later by Ångström.

In 1859, Kirchhoff formulated the law connecting absorption and emission of light by showing that each species of atom has a uniquely characteristic spectrum. Two years later, Kirchhoff and Bunsen laid the foundation for spectrochemical analysis and for astrophysics by comparing the solar spectrum with the flame or spark spectra of the purest elements available, thus presenting the first chemical analysis of the sun's atmosphere.

These data have served in making chemical identification and since 1885 have contributed to the development of quantum theory and to that of fruitful hypotheses concerning atomic and nuclear structure, thus bringing new evidences as to the structure of matter and molecules.

8. STRUCTURE-FUNCTION RELATIONSHIPS

The notions of Democritus, according to whom the bodies in nature are made of atoms with a well-defined shape explaining some of their properties, were maintained up to the XVIIIth century. The complementarity between bases and acids was thus explained as expressing the fact that the atoms of an acid are needle-shaped, thus giving rise to the sour taste, while bases were made of spongy, porous particles in which the needles of the acid could be inserted. Therefore the idea of a relation between some microscopic structural characteristics of a compound and some of its properties is rather precocious. However one had to await the long and arduous maturation of the scientific concepts of atoms and molecules to introduce the necessary rationalization. As a matter of fact one has to wait until the XIXth century.

a. *The Maturation of the Concept of Molecule*

From Dalton's contribution up to the work of Kekulé and others, the notion of molecule is more and more accepted thanks to careful studies performed on the structure of crystals, the evaporation of a molecule crystal, its vapor pressure together with the studies on gases and liquids (vapor pressure, boiling point,

freezing point and the like). Also the discovery of electricity (W. Gilbert, 1540—1603), of the electron (G. J. Storrey in 1874; J. J. Thomson (1856—1940), and that of X-rays (Röntgen, 1895) and natural radioactivity (Becquerel, 1896) were certainly seminal to the establishement of the reality of the concepts of atoms and molecule and they now attract a wide acceptance among scientists. Therefore previous attempts to relate specific properties of a compound to a specific architecture of the constituting atoms had to be taken seriously by the scientific community. To illustrate the idea that there is a relation between structure and function, it seems to be relevant to me as a biochemist to consider as a starting point the work of Würtz, Mayer and Duclaux related to the explanation of enzyme action, as well as that on the fermentation of sugars by Pasteur, and on sugar's stereochemistry by Fischer. This aspect of the problem has been thoroughly discussed by Debru [5].

b. *Rationalization of Catalytic Phenomena*

Physical chemistry and chemical kinetics emerge from the work of Guldberg and Waage, Van't Hoff, and Arrhenius. I am of course speaking of the law of mass action, of the influence of temperature on reaction rate and of the theory of electrolytic dissociation.

In this context, Würtz in 1881, explains the action of soluble enzymes by proposing the hypothesis of a temporary and renewed fixation of the enzyme on the substrate. Some time later, in 1882, Mayer concludes that while the reaction proceeds, the decrease in enzyme activity is explanable in terms of product accumulation, though the catalyst remains intact. Later, O'Sullivan and Tompson (1890), by following the inversion of cane sugar by polarimetry, showed that the phenomenon obeys a logarithmic law such as that proposed by Wilhelmy,[1] thus associating the action of the enzyme with that of the acid catalysis and showing that it follows the law of mass action.

Duclaux, in studies performed between 1883 and 1898 on the same material, proposes different conclusions. In 1883 he shows that the same amount of enzyme produces identical effects whatever the quantity of sugar present. Therefore the enzyme acts as a constant force, which in a given time, can only produce a given amount of work. The action of the enzyme is thus related to time and not to the amount of sugar present. It is a linear, non-logarithmic relation, with zero-order kinetics. The linearity is observed at the beginning of the process when the accumulation of reaction products is not yet perceptible, otherwise the logarithmic behavior of O'Sullivan and Tompson is observed [7]. It should be noted that for Duclaux, his analysis does not imply the existence of an enzyme—substrate complex and does not rest on a theory of catalysis. This will be done later by Victor Henri.

[1] Incidently, when L. Wilhelmy showed in 1850 that the rate of hydrolysis of cane sugar could be calculated by a mathematical equation, it was one of the first attempts to use a mathematical formula to express a chemical process.

It is interesting to note that the existence of an enzyme—substrate complex is already apparent in the work of Fischer when he deals with the stereochemical aspects of various sugars as well as in the kinetic studies of A. Brown. Pasteur had already pointed out the important relation between the selectivity of yeast and the asymmetry of sugars. But it is clear that the work of Fischer on the structure of sugars, their synthesis, stereochemistry and isomeries was rather decisive in this respect. From 1894 in 1898 the stereochemistry of sugars based on the theory of asymmetrical carbon of Le Bel and Van't Hoff leads Fischer (1898—1899) to the hypothesis that the active components of the yeast must have a configuration that complements that of the sugars on which they act [8]. Thus the three asymmetrical carbons of *d*-fructose, *d*-mannose and *d*-glucose (Fig. 2) are equivalent, therefore explaining the ability of yeast to ferment them into alcohol.

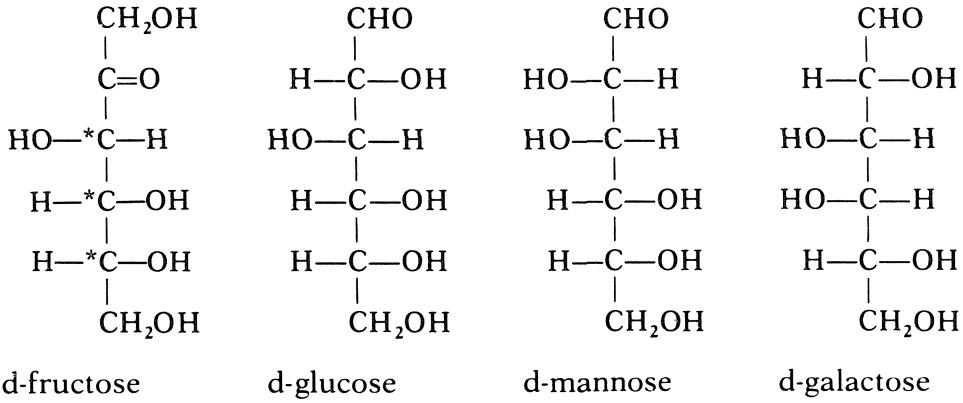


Fig. 2. Stereochemistry of some sugars.

d-Galactose is a poor substrate, as well as the α- and β-*d*-methylglucosides synthesized by Fischer. The specificity of the enzyme for its substrate must be as great as that of the intact cell of yeast leading henceforth to the concept of the key and lock (Schloss und Schlüssel) to explain the stereochemical adaptation. The important corollary to this hypothesis is that the organization in space of the atoms in a given molecule is an important factor of the specific properties of the compound.

Fischer was of course quick to notice that even if the exact nature of the biological catalyst is not yet known, their analogy with the "proteic matter" is so great that they should be considered as "molecular objects optically active and therefore asymmetrical" [8, 9].

c. *The Theory of Asymmetry and the Complementarity Principle of Pauling*

For Fischer, an enzyme, though a complex entity, can be understood using the basic concepts of organic chemistry. As a matter of fact, enzymatic specificity

could be used as a means to recognize stereochemical differences and this specificity was as meaningful for cell functioning as it was useful for experimental chemistry.

Enzyme specificity is thus formulated with reference to chiral substrates. The fact to be explained is why from two molecules, one being the mirror image of the other, only one of the two is hydrolyzed. Fischer thus introduces the way of reasoning of an organic chemist into the preoccupations of a biologist. Moreover the idea that asymmetry is typical of life processes is somehow outdated by the progress of organic synthesis of compounds of biological interest. And this turn of events had a profound philosophical impact since it helped to counteract the vitalism of protoplasmic doctrines by imposing the principles of a structural chemistry.

As formulated by Fischer the concept of asymmetry is a geometrical concept. This was further emphasized and expanded by Pauling [24] in his principle of complementarity in structure. I shall come back to this problem at the end of this paper, when dealing with the notion of pairs of relata, but it seems however adequate now to quote Pauling [24] to show how pregnant were the ideas of Fischer for the fields of both biology and chemistry. Pauling wrote in 1946:

> Despite the lack of detailed knowledge of the structure of proteins there is now very strong evidence that the specificity of the physiological activity of substances is determined by the size and shape of molecules, rather than primarily by their chemical properties, and that the size and shape find expression by determining the extent to which certain surface regions of two molecules (at least one of which is usually a protein) can be brought into juxtaposition — that is, the extent to which these regions of the two molecules are complementary in structure. This explanation of specificity in terms of "lock-and-key" complementariness is due to Paul Ehrlich, who expressed it often in words such as "only such substances can be anchored at a particular part of the organism which fit into the molecule of the recipient combination as a piece of mosaic fits into a certain pattern".

and a little further:

> The one general chemical phenomenon with high specificity is closely analogous in both its nature and its structural basis to biological specificity: this phenomenon is crystallization. There can be grown, from a solution containing molecules of hundreds of different species, crystals of one substance which are essentially pure. The reason for the great specificity of the phenomenon of crystallization is that a crystal from which one molecule has been removed is very closely complementary in structure to that molecule, and molecules of other kinds cannot in general fit into the cavity in the crystal or are attracted to the cavity less strongly than a molecule of the substance itself. Only if the foreign molecule is closely similar in size and shape and the location and nature of active (hydrogen bond-forming) groups to the molecule it is replacing, will it fit into the crystal; and it is indeed found that the tendency to solid-solution formation depends upon the same structural features (such as replacement of a chlorine atom by a methyl group) as the tendency to serological cross reaction.

The specificity and efficiency of enzyme catalysis indicated that the three-dimensional molecular structure was the primary determinant of the mechanism. This structure should recognize the preferred substrate, and energetic considera-

tions showed that the enzyme's active region should be complementary to the transition state of the substrate [13, 24], see also [6].

C. The Concept of Macromolecule, or Colloidal versus Macromolecular Chemistry

The preoccupations of molecular biology are derived from a current of ideas the origin of which is to be found in the notion of high polymer or macromolecule. The idea that proteins, cellulose, starch, etc. are polymers, i.e. associations of small units through covalence, is certainly due to Hlaziewetz and Habermann who, in 1871, published a paper in which they very clearly defined the concept. Unfortunately, Graham was at that time including these compounds amongst his "colloids", thus introducing a sterile confusion between macromolecules and true colloidal particles in which the molecules are associated by means of residual or secondary valencies. It is Staudinger [27] who should be credited for putting back biochemistry on the right track, by showing that the so-called colloidal properties of macromolecule solutions do exist whatever the type of solvent, contrary to what is observed with the micelles obtained by the association of small molecules through secondary valencies.

By insisting on the level of organization corresponding to the formation of macromolecules and the emergence of the resulting specificity, the pioneers of macromolecular chemistry were doing — as Molière's Monsieur Jourdain spoke prose — molecular biology, without knowing it.

One defines as a macromolecule a compound having a molecular weight above 10 000, in which the network of covalencies expands in a tridimensional structure. However this does not exclude other type of bindings at various points of the structure. This boundary, defined by a molecular weight, corresponds approximately to the dimensions of macromolecular aggregates that give rise in solution to the so-called colloidal properties. Thus the macromolecules are made of monomeric structures associated in a given order, through covalencies. These linear structures may also be associated through lateral bindings, also covalent in nature, giving a tridimensional structure to the whole.

1. THE CONCEPT OF MACROMOLECULE

In the 1920s the dominant paradigm was that molecules with a molecular weight above 10 000 did not exist. Polymers were therefore aggregates made of low molecular weight associated through secondary valencies. These aggregates gave to the solution the so-called colloidal properties. Clearly, in 1920, Staudinger [27] introduced the concept of macromolecule where one should only rely on normal valencies (i.e. covalence) to describe the properties and behavior of various polymerization products.

Wieland counteracted by stating that the crystallization of a polymer would lead to the demonstration that they are made of the association of low molecular weight subunits, through secondary valencies. Unfortunately the crystallization of some natural polymers obtained thereafter indicated the heterogeneity of synthetic polymers. According to Staudinger the colloidal character of a solution does not necessarily demonstrate the association of small molecules into aggregates but rather that the molecules are truly gigantic: these are the eucolloids. The demonstration came from the hydrogenation of rubber, which should give rise, according to Pummerer, to a volatile hydrocarbon. As a matter of fact, the product thus obtained behaved, as did the rubber itself, as a colloid.

2. THE COLLOIDAL-CHEMICAL STAGE

Berzelius, in 1833, defined a polymer as made of an aggregation of particles of the same type [2]. On this background, Graham [11] made a distinction between colloids and crystalloids, i.e. substances able to crystallize and to diffuse quickly through a membrane in opposition to the colloids. However, for Graham [12], colloids are only aggregates of crystalloids. Between 1862 and 1929, the period qualified by Pritykin [26] as corresponding to the colloidal-chemical stage, a large number of substances having high molecular weight were described: hemoglobin (16 700; 66 700) egg albumin (14 000—17 000; 73 000; 34 000); serum albumin (67 000) etc.[1]

Fischer, however, rejected the idea that proteins could have molecular weight greater than 5000, an argument based on the idea that there is no guarantee that the natural proteins are homogeneous [9].

Moreover, during the same period, the notion of micelle, as developed by Nägeli around 1858, was revived and quickly the cohesion of micelles was explained in terms of "partial valencies". Colloid chemistry received a further impetus thanks to the adhesion of Ostwald and the creation of not only a society devoted to the work of the adherents of the "new" chemistry but also of two journals: *Zeitschrift für chemie und Industrie des Kolloide* and *Kolloidchemische Beihefte*. It is on this background and against such respected personalities as Fischer and Ostwald that Staudinger introduced the concept of macromolecule as an ordinary object of organic chemistry.

3. TOWARDS MOLECULAR BIOLOGY

It seems that L. Pauling was the first one, in the 1930s, to use the expression "molecular biology". And this was on the occasion of filing a grant request with the NSF to do some work on hemoglobin. Maybe he did so in order to get some distance from biochemical methods, since he would mainly be concerned with the

[1] See [22] for a more complete account.

architecture of proteins as polypeptidic chains folded and held by hydrogen bonds, Van der Waals forces and other weak interactions.

Also Pauling was a model-builder and his approach was based on some comparison made with the structure of inorganic crystals. Therefore, he largely used the physical methods already in use in crystallography. This conceptual breakthrough, very apparent in the paper Pauling publishing in 1946 and already cited, opened a completely new field of investigation based not only on model building but also on the use of physical methods such as X-ray diffraction, and some techniques directly related to the field of microbiology. Thus, joining all these techniques within the framework of the concept of molecular architecture lead to the important discoveries of two fundamental patterns in protein structure, the α-helix and the β-sheet [25], the double helix of DNA [29], the structure of viruses [4], the contractile proteins (for a review see [15]) and so on.

D. Some Pairs of Relata in Contemporary Biochemistry

Since two parts of this work are devoted to the molecules in biological sciences I shall only refer here briefly to a few cases examplifying the key-and-lock (Schloss und Schlüssel) analogy of Emil Fischer as well as the principle of asymmetry and complementarity of Pauling.

1. THE CONCEPT OF BIOMOLECULAR RELATA

In his attempt to interpret the biosphere in terms of a biosemiotics Florkin [10] has introduced a few useful concepts borrowed to the general linguistics of De Saussure. As stated by Florkin: "The recognition of systems of *relata* at the integrative molecular level indicates that biomolecular order is governed by systems of signification which we may considered in a structuralist (intensive) perspective besides the thermodynamic viewpoint and the quantifying (extensive) viewpoint of the information theory." Thus the minimal configuration aspects, carriers (significant) of molecular signification (signified) are biosemes. At the level of the bioseme, the significant is an aspect of molecular configuration or architecture while the biological activity, the signified, comprises the sign i.e. the theme of signification.

Of course the signification system implies a sign (significant—signified) and a receptor and this is what is referred to by the expression couple or pairs of relata. On the other hand a biosyntagm is defined by Florkin as a unit of signification higher than a bioseme, and "composed of significant units in a relation of reciprocal solidarity, i.e. an associative configuration of biosemes". Clearly an enzyme or a metabolic sequence are biosyntagms. But in the case of many polypeptides, the distinctive configuration results from spatial folding and stabilization bringing together in close vicinity a number of specific biosemes, in this case amino acids. It

is therefore justified to define it as a tertiary biosyntagm. A quaternary biosyntagm will then be defined by the association of identical subunits, homopolymer, or different subunits, heteropolymer, of polypeptidic chains.

2. EXAMPLES OF PAIRS OF RELATA

The protein component of myoglobin is a good case in point to illustrate the concept of tertiary biosyntagm. It is defined as eight pieces of α-helices organized in a box-like fashion in which the heme is buried. This heme pocket is hydrophobic, while the surface of the molecule bears polar groups.

The *relata* of the biological activity of the myoglobin are the heme-protein molecule and the oxygen molecule. The biological signification i.e. the reversible binding of oxygen is related to the precise architecture of the whole structure: the heme iron octahedrally coordinated (four N atoms in the porphyrin ring of the heme, one N from His-F_8 of the globin moiety) and the sixth ligand is O_2. The existence and the maintenance of the oxygen combining site in the heme pocket is due to the eight pieces of α-helix, all right-handed and formed by 7 to 26 amino acid residues. Other forms of tertiary biosyntagms that also illustrate the concept of pairs of relata may also be found by considering the case of hydrolytic enzymes, hormone-receptor interactions, etc. . . .

I shall take as another brief example the case of the hydrolysis of acetylcholine by the enzyme acetylcholinesterase since the model proposed in the early 1950s by Nachmansohn was of great historical importance not only in the field of enzymology but also in neurobiology where it introduced a completely new way of reasoning.

Let us consider Fig. 3. The enzyme—substrate complex is presented. It is stabilized by Coulomb and Van der Waals forces at the anionic site and by covalent bond formation between the carbonyl carbon and the basic group of the esteratic site. The latter is symbolized by G (for basic group), H representing a dissociable hydrogen atom not involved in binding. It would take too long to describe the experimental data that lead Wilson and Nachmansohn to propose the structure of Fig. 3 to describe the Michaelis—Menten complex acetylcholine-acetylcholinesterase, together with the identification of the forces bringing about its stabilization. Further details may be found in the original papers quoted in two early reviews [21, 30]. It is only important to notice that the thermodynamic data referring to the enthalpies and entropies of activation of the various steps of the catalytic process indicate that the configuration of the enzyme is altered during the binding of the substrate: the tetrahedral structure of the quaternary nitrogen with the 3 methyl groups should be wrapped by the enzyme protein at the anionic site, thus explaining the contribution of all the *three methyl groups* to the stabilization of the complex.

The notion that the ionic conducting sites in a membrane are mainly made of proteins and that the conductance of the site is related to the configuration of the

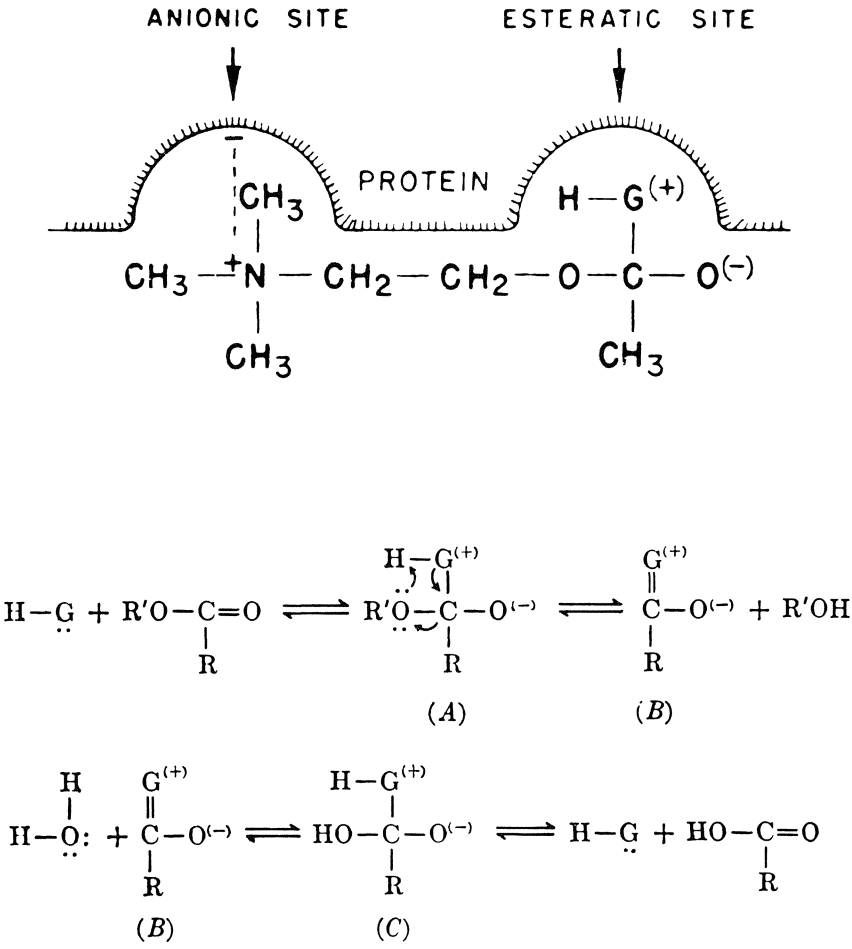


Fig. 3. Schematic presentation of the interaction of the acetylcholinesterase active group with its substrate acetylcholine. Forces stabilizing the Michaelis—Menten complex and, below, the hydrolytic process. (After Nachmansohn and Wilson, 1951).

protein are directly derived from the studies of Wilson and Nachmansohn of the acetylcholinesterase.

These ideas have been applied with great success to explaining the properties of the acetylcholine receptor at the synapses, thus showing how penetrating were the views of Nachmansohn more than 30 years ago, when the field of membrane permeability was still in the hands of the physiologists and dominated by a purely macroscopic phenomenology. We can also take the case of the acetylcholine receptor to illustrate the concept of a quaternary biosyntagm since, as will be discussed in the last volume of this work, it is known that it is a pentameric structure formed by the association of four different polypeptidic chains. Here

again the relata of the biological activity are the receptor and the acetylcholine molecule.

E. Conclusions

It seems to me that the consideration of the ontogeny of the concept of molecule illustrates a rather general and monotonous principle of the history of mankind, i.e. the complexity of the historical relationships between researchers, discoveries, extant paradigms, etc. . . . As a consequence, when reviewing the subject, a choice had to be made amongst the events, the importance of which is appreciated through the subjectivity of the writer. In the present situation the fact that I am a biochemist certainly introduces another bias. However, it seems to me evident from the consideration of the historical data that three periods may be taken into consideration (Table II). The pre-scientific era is dominated by the ideas of the Greek philosophers and by the mythical and magical practices of alchemists. However the development of alchemy from Greek philosophy, oriental technology and oriental mysticism in the Hellenistic city of Alexandria in Egypt, benefiting also from the practical experience inherited from Copper, Bronze and Iron Ages, led to the development of technological skills and therefore to a practical chemistry of good quality. The Transition Period that covers, in my analysis, the XVIth, XVIIth and XVIIIth centuries, sees the introduction and progressively more important *usage of the balance* in the quantitative study of reactions. As a consequence rational thinking becomes dominant in replacing magical and mythical interpretations and culminates in the oxygen theory of Lavoisier: thus bringing us to the Scientific Era that opens early in the XIXth century with Dalton's and Ampère's contributions. This is also evident in the technical innovation of *measuring the volume of gases* instead of weighing reagents and products. The development of what we call today the scientific method leads to a more precise quantification of the reactions and even to their mathematical description. But despite the so-called objectivity of science, one sees a kind of incommunicability installed between protagonists, as we see it today in the daily life of our researchers. Also, important discoveries that do not fit into the extant paradigm are neglected and need to be brought to light and rediscovered later: this is well illustrated by the work of Cannizaro and his revival of the Avogadro—Ampère hypothesis 50 years later. Here again the situation is familiar in contemporary science. The introduction of thermodynamic formalism brings about a better understanding of chemical reactions and, together with other observations, definitely establishes a clear-cut difference between atoms and molecules. Also the discovery of isomerism (Berzelius), the specificity of fermentation (Pasteur), the stereoisomerism of Van't Hoff and Le Bel and the key-and-lock theory of Fischer are landmarks that bring us to the complementarity principle of Pauling (1946), that is still so pregnant in contemporary chemistry and biochemistry. Thus, with the concept of macromolecule and the rejection by Staudinger of Graham's ideas

Table II. A very subjective synthesis of the ontogeny of the concept of molecule.

Prescientific era: Greek classicists; alchemy

Transition period: XVIth Century: mining, metallurgy, distillation (H. BRUNSCHWIK, BIRINGUCCIO, AGRICOLA, ERCKER)
Tria prima theory of Paracelsus (sulfur-mercury theory + salts = metals)
Iatrochemistry: Paracelsus, J. B. Van Helmont, Fransiscus Sylvius

XVIIth Century: 1 to 5 basic elements
Use of balance and quantitative study of reactions
Art of producing magisteries (reagents and their use)
Revival of Democritus (GASSENDI, DESCARTES)
The Sceptikal Chymist (1661): BOYLE (definition of element, corpuscular theory, rationalization)
BECHER and tria prima as vitreous, fatty & fluid earth; solid as only constituted of matter
CO_2 (VAN HELMONT) = gas. Only physical properties
Phlogiston (STAHL): fire principle
Metal = calx + phlogiston (immaterial principle)

XVIIIth Century: combustion — Lavoisier (1770—1790). Definition of element
Modern system of chemical nomenclature (with GUYTON DE MORVEAU)
respiration = combustion. Oxygen theory

Scientific era: XIXth Century: affinity, atomicity
Dalton's atomic theory
Gas volume measurement
Gay-Lussac: combining ratio of various gases
Avogadro (1811): equal volumes of gases = equal numbers of molecules
Ampère (1814): idem + shape of molecules (use of models) | disregarded for
Avogadro—Ampère: atoms ≠ molecules | 50 years
Berzelius, Dulong-Petit rule (atomic weight × specific heat = constant)
Table of atomic weight
1833: Berzelius: catalysis, isomerism, electrochemical theory of atomic combination. Separation of gases and metals with Voltaic pile
Organic chemistry: radicals as unit in chemical reactions
1850: Wilhelmy: cane sugar hydrolysis (equation)
± 1850: heats of reaction (HESS, BERTHELOT, THOMSEN) thermodynamics
Organic compounds: KEKULÉ, COUPER (tetravalence of C) ring structure (KEKULÉ, 1865) (use of models)
1860: CANNIZARO: revival of Avogadro hypothesis: atoms ≠ molecules
1863: law of mass action (GULDBERG and WAAGE)
1868: valence
1869—1871: MENDELEYEV and MEYER (Gallium 1875; Scandium, 1879; Germanium, 1886)
1869: HORSTMANN (entropy in chemistry)
1876—78: GIBBS: treatment of heterogeneous equilibria
1877: DELSAUX: explanation of Brownian motion (1st blow to vitalism)
1878: KUHNE: Enzyme

Spectroscopy: Newton (1664), spectrum of sun light (7 colours)
Herschel (1800), spectral distribution of sun heat → IR
Ritter (1801), effect of spectral light on silver salts → UV
Thomas Young (1802), substituted his wave theory of light for Newton's corpuscular theory, explained colours of thin films, calculated the approximate wavelengths of the 7 colours of Newton

Table II (continued)

Joseph von Fraunhofer (1814), solar spectrum interrupted by many hundreds of dark lines
Kirchhoff (1859), general law connecting absorption and emission of light, each species of atom has a uniquely characteristic spectrum
Kirchhoff and Bunsen (1861), 1st chemical analysis of sun's atmosphere laid foundation for spectrochemical analysis and astrophysics (spectrum of the purest elements available, discovery of Cesium and Rubidium)
Ångström (1868) measured ± 1000 Fraunhofer lines, expressed them in units of 10^{-10} meter (Å)
1800: Volta's pile
1853—1858: HITTORF: migration of ions (electrolysis)
1870: KOHLRAUSCH: electrolytic conductivity
1884: Arrhenius: electrolytic dissociation
1885—1888: textbook of physical chemistry (OSTWALD: but atoms may not exist at all!)
1895: X-rays (RÖNTGEN)
1896: radioactivity (BECQUEREL)
1897: electron (THOMSON)

1874: VAN'T HOFF — LE BEL: stereochemistry, tetravalence C, asymmetric C
1885: Pasteur fermentation and specificity
1890—1910: E. FISCHER, sugars and enzyme; asymmetry typical of life reproduced by synthesis (2nd blow to vitalism)
1897: E. BUCHNER and cell homogenate: destroys the protoplasm doctrines (3rd blow to vitalism)

1833: BERZELIUS: polymer: aggregation of identical particles
1861: GRAHAM: colloids — crystalloids
1858: micelles of NAGELI
1862—1929: colloidal — chemical stage (PRITYKIN)
1907: FISCHER: PM of proteins < 5000
1920: macromolecules of STAUDINGER
1946: PAULING and the complementarity principle
1948: PAULING and COREY: α-helix, β-sheet (use of models: cfr Ampère and Kekulé)
1953: Double helix

regarding the colloid state, the way is largely open to our understanding of biology in terms of molecular interactions (pairs of relata) and to the formalisation of the observations into the system of biochemistry.

This is also true of chemistry, where the power of the concept of molecule is well illustrated by its predictive value and its explanatory effectiveness.

References

1. A. M. Ampère: *Ann. Chim.* **XC**, 43—86 (1814).
2. J. Berzelius: *Jahresbericht ü. die Fortsch. Chemie* **12**, 1—70 (1833).
3. G. Buffon: *Oeuvres complètes*, Vol. 3, p. 223, Duménil, Paris (1835).
4. D. L. D. Caspar and A. Klug: *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1—24 (1962).
5. C. Debru: *L'esprit des protéines.* Hermann, Paris, p. 365 (1983).
6. M. L. J. Drummond: *Prog. Biophys. Molec. Biol.* **47**, 1—29 (1986).
7. E. Duclaux: *Ann. Inst. Pasteur* **XII**, 96 *et seq.* (1898).
8. E. Fischer: *Hoppe-Seyler's Zeitschr. Physiol. Chem.* **26**, 60—87 (1898).

9. E. Fischer: *Ber. dt. Chem. Ges.* **40**, 1754—1767 (1907).
10. M. Florkin: *Concepts of molecular biosemiotics and of molecular evolution* (Comprehensive Biochemistry v. 29A, Eds M. Florkin and E. Stotz), pp. 1—124. Elsevier (1974).
11. T. Graham: *Phil. Trans. Roy. Soc. London* **151**, 183—224 (1861).
12. T. Graham: *Liebigs Ann. Chem.* **121**, 1—71 (1862).
13. J. B. S. Haldane: *Enzymes*, p. 182. Longman, Green and Co. (1930).
14. V. Haüy: Essai d'une théorie sur la structure des cristaux appliquée à plusieurs genres de substances cristallisées (1784).
15. H. E. Huxley: *Muscle cells* (The Cell v. 4, Eds. Brachet and Mirsky) pp. 365—481. Academic Press (1960).
16. A. Kekulé: *Bull. Soc. Chim. France* **3**, 98 (1865).
17. A. Kekulé: *Liebigs Ann. Chem.* **137**, 129 (1866).
18. A. Kekulé: *Liebigs Ann. Chem.* **162**, 77 (1872).
19. P. Laszlo: *La Recherche* **104**, 1019—1020 (1979).
20. P. Laszlo: *Nouv. Journ. Chimie* **4**, 699—701 (1980).
21. D. Nachmansohn and I. B. Wilson: *Adv. Enzymol.* **12**, 259—339 (1951).
22. R. Olby: *J. Chem. Educ.* **47**, 168—174 (1970).
23. J. R. Partington: *A History of Chemistry*, v. IV, p. 553, Macmillan (1970).
24. L. Pauling: *Chem. Engin. News* **24**, 1375—1377 (1946).
25. L. Pauling and R. B. Corey: *Nature* **168**, 550—551 (1950).
26. L. M. Pritykin: *Isis* **72**, 446—456 (1981).
27. H. Staudinger: *Ber. dt. Chem. Ges.* **53**, 1073—1085 (1920).
28. P. Thuillier: *La Recherche* **175**, 386—390 (1986).
29. J. D. Watson and F. H. C. Crick: *Nature* **171**, 737 (1953).
30. I. B. Wilson and D. Nachmansohn: *The Generation of Bioelectric Potentials* (Ion Transport Across Membranes, Ed. H. T. Clarke) pp. 35—64. Academic Press (1954).

Emergence and Evolution of the Molecular Concept

MARIKA BLONDEL-MÉGRELIS
Centre National de la Recherche Scientifique, Centre d'Analyse des Formes et des Systèmes, Université Jean-Moulin Lyon-III, 74, rue Pasteur, 69007 Lyon, France.

1. Introduction. Problematic Qualities

To recount the history of the concept of molecule is, when you think about it, neither linear nor simple. For, however firmly it may be combined with the atomic theory, this concept has neither the same history nor the same determinations. It is actually:

- not a question of concerted progress: the molecule is not resolved by the atom and even less does the atom build the molecule;
- neither is it a question of polemical progress: the final, indivisible atomic concept was to be recognized only very late as essentially different from that superior unitary construction.

Moreover, this can be misleading: if the term 'atom' has many meanings, 'molecule' may have many more, and among them 'atom'. Under these conditions, some fervent anti-atomic scientists used this term, and one will have to try and make clear what they meant.

Finally, the concept of molecule suffered from the great number of imaginary constructions which were so pleasant to some authors, from Gaudin to Delafosse, not forgetting to mention Ampère: his *Mémoire* of 1814 [1] has to be saluted, but its geometrical constructions, which have been juged unnecessary and mar such a famous *Mémoire*, are regrettable.

We will see that the concept of molecule, hampered by the atomic hypothesis, had to find its construction *apart from* the atom, and then had to be defined *in contrast* with it. Finally, one had to have recourse to atoms, and thus to hypotheses. One would search in vain through the *Dictionnaire de Chimie* of Wurtz [2], to find the entry 'molecule'; Wurtz is precisely the author of the famous *Théorie atomique* [3], a work in which he tries to demonstrate the soundness and the richness of this 'idea', against his still numerous detractors. No more doubt: this concept is eminently questionable; the term is found through the literature of the 19th century in physics, chemistry and crystallography, and it is only in 1913 that it will be possible to write: "It is thus becoming difficult to deny the objective reality of molecules" [4].

With the numerous possible ways to tackle the concept, we have to give up any claim to be exhaustive. Our purpose is only to try and understand how it was

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, 25—44.

capable of becoming a tool so particularly prolific in the explanations that it allows, as well as in the discoveries that it caused.

2. Small Masses

It is possible to say very roughly that the origin of the term 'small masses' goes back to Newtonian philosophy. When Descartes wants to characterize matter by the extent it occupies: "the nature of matter only consists in being just a substance extended in length, width and depth" [5], this is because he holds that there is a measurable quantity and because its degrees of quantity can be expressed in numbers. Newton characterizes matter by the 'quantity of matter', and the number expressing it is directly obtained by weighing. But, in order to give an account of the 'specific gravity' of the different bodies, he assumes the following postulate: "all spaces are not equally full; for if all spaces were equally full (...), all bodies would be of the same density" [6]. So all bodies are constituted of "solid, massy, hard, impenetrable, moveable particles" [7], small masses between which "pores, a void, space, or vacuum" [6] exist.

This is thus approaching a representation which is similar to the opinion of atomic scientists, Lucretius or Gassendi. But, as Hélène Metzger points out, the reasons are not at all identical in both cases. The small indivisible bodies of atomic scientists are a product of their imagination: the point is that man must build a stable and finite world which would not be strange to it. On the contrary, the point is for Newton to ensure the initial proposition: "all substances have a mass which can be measured with a balance and which, disregarding the pores, is proportional to its volume." [8]

The hypothesis of atomic scientists is thus the result of their imagination; as Leibnitz put it, "imagination was cheerful. We limit our research there: we fix deep thoughts as if we were nailing them down, we think that we have reached the first elements" [9]. Newton's proposition is the *condition of possibility* of the constancy of the quantity of matter, or of mass, through and under the different modifications modifying it.

This intention of giving an account of the variety of matter by means of these small corpuscular masses — sorts of unit of matter — is brilliantly illustrated in crystallography.

Romé de l'Isle characterizes the mineralogical species in the form of his 'molécule intégrante' (which, besides, is resulting from the form of 'molécules constituantes' — completely unknown and probably never to be known! [10]).

R. J. Haüy, whose project is not to distribute the mineralogical species but to establish a comparison between the different crystalline forms, characterizes the species by:

- the figure of its 'molécules constituantes', which is a piece of experiment, in this case cleavage;
- their arrangement.

In his *Traité de Cristallographie*, he limits to *three* the number of the forms of the 'molécules intégrantes' which have been adopted by crystallization; they "are supposed to be the smallest solids that can be extracted from the mineral without altering its nature" [11]. If they are not exactly the ultimate result of mechanical division, at least they are some 'equivalents' of it, *simple* and *prolific* forms from which, as they are applied, the laws of crystallization will produce the vast variety of nature's forms: "what is beyond the point where I stop seeing, resembles what I had seen until then". Beyond this, it is outside the field of crystallography, the nature of the mineral is changed; the 'molécules-principes' or élémentaires' concern only the chemist. As for the 'molécule soustractive', it is "the unit on which the theory acts".

Now, if Haüy's 'molécule intégrante' becomes the 'operational' concept of crystallography — in consideration of the amendments that Delafosse [12] made — the notion of *smallest masses* (the ultimate term of separation), carrier of individuality, is removed from chemistry in the great catharsis that Lavoisier implemented:

All that can be said about the number and the nature of the elements only amounts to purely obscure and metaphysical discussions (. . .). I will content myself with saying that if we intend to refer by the name of elements to the simple and indivisible molecules forming bodies, it is likely that we do not know them; if, on the contrary, we attribute to the name of elements, or of principles, the meaning of the last term *that the analysis reaches*, all the substances that we have not yet decomposed are elements *for us* [13].

On the other hand, it is known what use Lavoisier was to make of Newton's postulate concerning the constancy of the quantity of matter, 'the preservation of mass'. Mass is the quantity directly measurable which allows us to get fixed to an island of invariance, in the flood of turbulences and variations in matter; it allows us to check that what was put in at the source can actually be found in the end, to make a count, to classify in tables, to establish connexions and equivalences.

2.1. EQUIVALENCES

Mass is sufficient to develop Lavoisier's chemistry. It expresses especially well the typical reaction of this chemistry: neutralization.

In a stemmed glass, as Wurtz describes it [14], there is a

corrosive liquid with a strong savour, sulphuric acid. If I pour a few drops of that blue liquid called tincture of litmus, the colour immediately turns to vivid red. If now I pour, drop by drop, a liquid with a strong, lixivial and alkaline savour called potash, after some time the red colour of the acid solution turns blue. I stop exactly at that point: the liquid does not act anymore on litmus, does not have the corrosive and acid savour of the potash, but it has a salted savour.

It can be seen that if there is a given *quantity* of sulphuric acid, the same and invariable quantity of potash is always obtained to proceed to neutralization. If

$$40\ \text{g} : H_2SO_4 \rightarrow 47\ \text{g} : KOH$$

The quantities of other bases able to neutralize these 40 g H_2SO_4 are also deter-

mined: it is said that the equivalents of potash, soda, lime are 47, 31, 28. In the same way, the nitric acid is able to neutralize the equivalent of KOH; the equivalents of sulphuric acid and nitric acid are 40, 54 . . . Finally, as the equivalents of potash, soda and lime all contain 8 g of oxygen, it will be said that the 39 g of potassium, 23 g of sodium and 20 g of calcium are the equivalents of these metals.

Besides this work, and from the experiment of Gay-Lussac and Humboldt in 1805 about the volumetric composition of water, a whole series of research is performed, aiming to fix the relations between the volumes of the gases which are combined, and between the volumes of the gas (or the vapour) of the compound which has been obtained:

2 volumes of hydrogen combined with 1 volume of oxygen
→ 1 volume of water vapour
2 volumes of nitrogen combined with 1 volume of oxygen
→ 2 volumes of nitrogen protoxide
1 volume of nitrogen combined with 1 volume of oxygen
→ 2 volumes of nitrogen dioxide
1 volume of nitrogen combined with 2 volumes of oxygen
→ 2 volumes of nitrogen peroxide

In the volumes according to which gases combine, *permanent relations* are thus encountered that are noticed between the weights of the elements being combined, and *simple relations* existing between the weights of the same elements in the multiple combinations of these elements [15].

Experiment, and experiment alone, is thus allowing the establishment of tables of relations expressed in volume relations, equivalents, proportional numbers, parts or 'atoms'.

And in 1843 Gerhardt can still write "atoms, equivalents and volumes are synonymous" [16].

2.2. DIFFICULTIES

All the same, many problems were beginning to arise, among which emerge, very schematically:

— confusion and lack of arbitration in the choice of the proportional numbers. According to the proportional number chosen for H (for 100 parts of oxygen combine with 12.5 or 6.25 parts of hydrogen), formulas OH or OH^2 were obtained for the same water. So, and after the corrections of Dumas and Stas about the atomic weight of carbon [17] in 1841, a body as simple as the formic acid can be written

$C^4\ H^4\ O^2$	Dumas
$C^2\ H^4\ O^2$	Berzelius
$C\ H^2\ O$	Gerhardt

— the absolute lack of *efficiency* of the formulas whereas the development of organic chemistry requires that some relations existing between bodies are brought out with the formula: such is the demand of Charles Gerhardt. Some relations exist between water, sulphuric hydrogen and succinic acid: in all the cases, it is possible to replace with metal either 1 atom and 2 atoms of hydrogen. This is not possible for benzoic acid. And dualistic formulas hide these relations.

— polybasicity has been acknowledged, and, since almost everybody agrees to represent acids with a radical plus water, since the degree of 'basicity' of a polybasic acid (as they said then) is linked to the number of 'molecules' of water that it contains, it is enough to pass from a system of equivalence to another, or still to double its formula, in order to turn a monobasic acid into a bibasic acid. A famous example is *citric acid*, for many people work on it. According to Berzelius, citric acid is:

$$\dot{H}\,\overline{Ci} \quad (\overline{Ci} = C^4\ H^4\ O^4).$$

It is monobasic [18]. According to Liebig, it is tribasic [19] and can be written:

$$Ci,\quad 3\,H_2O \quad (Ci = C_{12}\ H_{10}\ O_{11}).$$

The controversy has been developing for some time [20]. Hence the question: does polybasicity depend on substance or on conventional use? Is it a property of matter or is it a formal definition? [21]

— Finally, there is the problem of those organic substances such as methylene ($C^4\ H^4$), olefiant gas ($C^8\ H^8$), quadricarbonated hydrogen ($C^{16}\,H^{16}$); their centesimal compositions are similar, but "it is obvious that their molecule contains different quantities of matter", as Dumas writes [22].

2.3. ATTEMPTS

A certain number of requirements thus come to light, and French chemists such as Charles Gerhardt and Auguste Laurent undertake to get chemistry out of the impasse where it had been.

2.3.1. *Work of Formal Unification*

It is the choice of a *unity shared by* mineral chemistry and organic chemistry. This is the *Reform of equivalents* that Charles Gerhardt made in two stages:

i. It always has to be based on *experiment.* The equivalents generally used for the three products of decomposition,

$$C^2O^2 = 275$$
$$H^2O = 112.5$$
$$Az^2H^6 = 193.75$$

are not always in accordance with facts. The proportion of carbonic acid elimi-

nated or fixed in the reactions is always represented by $C^4O^4 = 550$, and water by $H^4O^2 = 225$. He then proposes relating all the equivalents to a common measure corresponding to *4 volumes* of vapour. This virtually amounts:

* to keeping all the formulas of Berzelius with 4 volumes H^2Cl^2 . . . as well as most of the formulas in organic chemistry;

* to doubling all formulas with *2 volumes*: water, carbonic acid, ether, oxalic acid . . . .

ii. *Experiment*, again, shows numerous abnormalities. To write water: 'HO' had actually led to 'RO' for metallic oxides. 'A quantity of double metal' has thus always been taken 'as a basis', so that if formulas of bibasic acids are generally well represented, formulas of monobasic acids should be divided in two. He relates to two volumes mineral substances

H^2O
H^2S
SO^2
AzH^3

and organic matters

$C^2H^4O^2$	acetic acid
C^2 (HCl^3) O^2	chloracetic acid
C^2 (H^3Ag) O^2	silver acetate

As oxygen is *100*; the equivalent is the unit corresponding to 2 volumes of vapour. It is only later that Gerhardt will adopt O = 16 and H = 1.

In his two 'Mémoires', Gerhardt does not use the term *molecule* but always the term *equivalent*. On the other hand, he uses the term *atom*, but without giving it a corpuscular meaning. There is always a comparison between masses, volumes. From this standpoint, he remains in the tradition of the empiricist chemists who wanted to build science on solid bases and leave aside any hypothesis. This principle will still drive Berthelot in 1869.

2.3.2. *Search for Molecule as an Elemental Unit*

A. Laurent, in his *Recherches sur les combinaisons azotées* (1846) [24], poses the problem: "on what will we rely to discover the simple or compound particle on which chemical actions are exerted?" [25]. "The *definition* of equivalent proves that $N^2H^2O^6$ is equivalent to SH^2O^4, but it does not demonstrate in any way that the *molecule* of sulphuric acid is not $S^2H^4O^8$ or $S^4H^8O^{16}$", he had specified above. It is thus no longer a question of the standardization of chemistry, a formal question, that preoccupied him. But it seems decisive to him to determine the particle on which chemical forces are exerted, to penetrate the reactional mechanisms. Now the determination of equivalents (case of polyatomic acids), or of volumes (abnormalities) in itself is inadequate; it will be necessary, in any case, to

leave it up to the 'judgment of *chemists*' and to determine the proportional number which "*generally* goes with volume, form, specific heat, specific volume, etc. . . . and which moreover allows to represent the reactions and formulas of all bodies in the most simple way."

In his 'Mémoire', A. Laurent arrives at a major distinction for the chemist, the distinction between atom and molecule, based on the consideration of chemical reactions. He makes there a break with the older science: Lavoisier's chemistry was the science of elements. From now on, and in perfect harmony with the definition that Ch. Gerhardt will give, chemistry becomes the science of metamorphosis: "the very special aim of chemistry is to seize bodies in that moment of transition when they cease to be themselves, when they turn into other bodies" [26].

> Gerhardt's atom represents the smallest quantity of a simple body that may *exist in a combination* (. . .). My molecule would represent the smallest quantity of a simple body that has to be used in order to *carry out a combination* [27].

Two major matters are to be considered following Laurent's work:

i. It can be said that, from this date on, *molecule* becomes *the chemical object* par excellence, even if the scientific community does not seem to have received the message. It is no longer just a relation of equivalence in mass or in volume, but a *unit* on which chemical forces have an effect, even if this unit remains, in this 'Mémoire', a unit of mass. "Our aim is to discover the weight of particles." Laurent then goes further than Gerhardt.

ii. The first step is taken in order to write chemical reactions; the writing will only take place very late in chemical literature: the sign remains = (identity) and does not become → (transformation).

3. Arrangement

The consideration of the *arrangement* of 'atoms' in order to describe, explain and characterize a substance already presupposes a particular composition.

3.1. BINARY STRUCTURE

The first famous attempt was made by Berzelius, and is based on Gay-Lussac's law: "the volumes of the combining gases are in very simple ratios with each other; the volume of the combination formed is linked in the same simple way with the sum of the constituent gases". This is interpreted with the help of the hypothesis of Avogadro and Ampère, and, in the light of the decomposition of the battery performed by Davy, with the help of Richter's 'Mémoires'.

> We admit it easily: the opinion of former physicists — that bodies are composed of indivisible atoms — has often been accompanied by absurd inventions, but a sounder argument has rejected them for a long time (. . .). As bodies are composed of elements that cannot be broken down, they have to be composed of particles whose quantity cannot be divided later, and can be called *particles, atoms, molecules, chemical equivalents*, etc. . . . I will preferably choose the appellation of 'atoms' [28].

Berzelius carries out the change from equivalents to the indivisible atoms of Dalton.

From the combination of the elemental atom A with the elemental atom B, which is an electric phenomenon to him, there results a compound atom AB which in turn can combine and produce a "second-rate compound atom".

— It can be seen that, in the action of combination, there is no composition but a juxtaposition of two parts under the effects of electrochemical forces. There is thus no real creation of an *upper* unit.

— Berzelius expresses his dualistic notion of the 'intimate nature' of particles in his Formulas: it will be the *rational formula*, 'expression of his idea' about the structure of bodies. And this is new. For example:

SO_3, H_2O or $\dddot{S} + Aq$
SO_3, K_2O or $\dddot{S} + \dot{K}$

The structure is also binary for organic bodies. Every body is composed of 2 parts, the compound organic radical being simply of a more complex kind. The result will be the famous theory of organic radicals of Justus Liebig, who will admit oxygen inside the radical. Berzelius will grant his request for a time, and even enthusiastically. But experience will teach him caution and, after having written, in 1832 and like Liebig [29]:

benzoic acid : Bż (Bz = $C^{14}H^{10}O^2$: benzol radical)
hydrated acid: Bż, Aq [30]

he will regain a wiser standpoint in the 5th edition of his *Traité de Chimie* [31]: "we have no idea of its rational composition:

crystallized acid : $\dot{H}\,\overline{Bz}$ ($\overline{Bz} = C^{14}H^{10}O^3$)
potassic benzoate: $\dot{K}\,\overline{Bz}$

In reaction against this dualistic theory and the numerous theories stemming from it, two trends arise:

i. The first and classical one is to go back to *facts* only. It is futile to try and conduct a survey on the intimate nature of the chemical phenomena [32]. Atoms and their arrangement are only harmful speculations, and chemists have to deal only with what is measurable [33] and visible [34]. Any rational or dualistic formula has to be abandoned (both terms became synonyms), the arrangement of atoms must not be expressed anymore but the formula has to express the only empirical ratios between bodies. Such is notably Gerhardt's attempt: "I use the word 'radical' in the sense of ratio and not in the sense of isolable or isolated body" [35]. In fact, Laurent will demonstrate to his friend that, without his knowing, and if they intend to express ratios of analogy, his formulas tell *more* than the facts do [36].

ii. The other trend consists of an attempt at giving an account — through the arrangement of atoms — of some properties of molecules that no other data can explain.

Phenomena of substitution are thus very much being noticed; experiments of Gay-Lussac (action of chlorine on wax), of Dumas (action of chlorine on alcohol [37]), and more generally action of chlorine on hydrogenated organic substances, lead to the first speculations about atomic mechanisms (and no longer just 'radicalaire') which are likely to happen inside the chemical unit. It seems that chlorine, as Laurent will say, *takes the place* and *plays the role* of hydrogen when there is an 'equivalent substitution' [38]. Dumas's position on this question will appear to be more fluctuating: he will soon go back to strict considerations of ponderal equivalence [39].

3.2. UNITARY STANDPOINT

Two main points emerge from Laurent's doctoral thesis, right to the end of his work:

A proposition: the arrangement must be the same between a compound and its substituted and equivalent derivative. As a result of this similarity in the arrangement, compounds must be analogous (properties, transformations, crystallographic forms) [40].

A plan: if these compounds are analogous, the formula has to *mention* it, and formulas must be analogous too. The writing will thus be:

naphthalene	$C^{40}H^{16}$
'chloronaphtalase'	$C^{40}H^{14}Cl^{2}$
'chloronaphtalèse'	$C^{40}H^{12}Cl^{4}$

but

hydrochloride of 'chloronaphtalase' $C^{40}H^{14}Cl^{2} + H^{2}Cl^{2}$
hydrochloride of 'chloronaphtalèse' $C^{40}H^{12}Cl^{4} + H^{4}Cl^{4}$ [41].

This, of course, is inadmissible for dualistic chemistry: chlorine, which is situated at the opposite extremity in the scale of electronegativities, could not but deeply modify the properties of the substance, if it was replacing hydrogen. "I am mistaken in thinking that the nature of combinations depends on the nature of the elements which compose them," writes Berzelius [42], then forced, in order to give an account of these phenomena, to imagine sector-based theories being more and more complicated and 'speculative', for instance the 'copula' theory [43].

It is in the *thesis* of Auguste Laurent about the 'theory of organic combinations' that the representation of the chemical *molecule* is expressed as being an architecture of physical atoms ("molecule = combination of atoms" as he writes in a footnote) [44], a geometrical structure possessing at the same time:

— its spatial and chemical individuality; this is common to the representations of Avogadro and Ampère who both think that particle is devised like a polyhedron of which each molecule constitutes an apex.

— and the property to partially or totally disorganize during the reaction, to go through "these architectural revolutions that chemical reactions are" [45], and this

according to some laws of modification whose pattern is manifestly crystallographic.

For the first time:

— one tries to understand what happens inside the molecule from a spatial point of view;

— one tries to seize the structure at the very moment when it modifies;

— one chooses a kind of middle course between the hypothesis of a complete upheaval, "absolute dissociation of the elemental atoms of the components", as Delafosse will take it up [46], and the supposition of a single bringing together, while keeping the relative places, in accordance with the electrochemical theory [47]. Chemical action is *real* but *limited*;

— one settles a theory liable to give an account of it and which, if it does not 'naturally' ensue from 'facts', must be subjected to the test of facts.

For one believes that it is possible to seize an invariant in the arrangement, thanks to chlorinated substitutions which occupy the years 1830—40 and which, as they keep their chemical and crystallographic properties, allow the restitution of "something common to the arrangement of the atoms with one another." [48] Laurent's hypothesis is at the same time moderate and bold:

i. *Bold* because it is once again a matter of 'speculations' on atoms, and we know what that caused: the deadlocks that chemistry reached. So that they could come out of it, chemists, and especially French ones, wish they could be positivists: "chemistry has a method as old as the hills: it is a blind submission to the power of facts. . . . Chemists want to see through the eyes of the body before they see through the eyes of the mind." [49]

The atomic hypothesis was badly thought of in the opinion of chemists. They had still not made a clear distinction between *atom* and *molecule*, and were obtaining unacceptable results when it came to the application of Dalton's hypothesis to gaseous volumes [50]; in particular they were considering every simple gas as being composed of *atoms*, keeping the term *molecule* for compound bodies [51]. Moreover, the synthesis of water, considered under the ratio of volumes, compelled them to cut the atom of oxygen in two, which is a contradiction in terms. This leads Dumas to make a distinction between physical atoms and chemical atoms, the first being subjected to physical forces, the latter to chemical forces!

> That is to say masses which are undividable for the first forces and other masses which are undividable for the latter. . . . In the cases of chlorine and hydrogen, chemistry was cutting the atoms that physics could not cut [52].

Later, Lothar Meyer will tackle this distinction,

> which is anything but clear and should be abandoned. A molecule is neither 'chemical' nor 'physical' only; it has both characteristics at the same time, i.e. it really exists; or it has none of these characteristics, i.e. it does not really exist [53].

Avogadro had raised the problem in 1814. Using Gay-Lussac's 'mémoire' [54] and linking the volumes of gases and the number of *simple or compound* molecules making them up, he put forward his famous hypothesis [55]. He had supposed that the unity of volume of the simple gases does not represent the *final* particles, and that *simple* gases or vapours, as *compound* gases or vapours, are composed of elemental molecules: but in one case, these elemental molecules are of the *same* kind, in the other they are of *different* kinds. As for Ampère, he had also, in 1814, made the distinction between *particles* (responsible for the state of aggregation of the body), *molecules* (whose number is equal in a same volume) and *atoms* (whose number and position determine the molecule) [56]. But these distinctions go unheeded and the honour goes to Laurent to have done credit to Avogadro's hypothesis — however he only quotes Ampère — undoubtly after the publication of his 'mémoire' about 'atomic volumes' and of the 'Annales' [58], greatly criticized by Regnault.

ii. *Moderate*: Laurent attempted to solve the problem of molecular arrangement *in several parts*, trying to delimit areas of permanence by means of experiment and with a *minimum* hypothesis. He thus determines limits to the possible knowledge [59]. He does not treat of atoms but only of arrangement: "I come to an important conclusion about the arrangement (do not be afraid of the word) of atoms . . ." he writes to Gerhardt [60], and he always makes it clear: "I think that it is impossible for us to know this arrangement" [61]. But, in the vast field of the unknown, he tries to delimit areas of lesser non-knowledge; we can yet know if, in a certain body, it is the same as in another.

I admit thus that the atomic arrangement is the same in benzoic acid and in nitrous benzoic acid, in acetic acid and in chloracetic acid, in isatin and in chlorisatin, in benzamide and in benzoic ether, even if I cannot say, in any case, what this arrangement is.

He gives himself the means to begin a resolution which perfectly suits his subject and is really ahead of the definition of organic chemistry that Lespieau will give in 1926: "It is the study of the reactions which leave the major part of the molecular structure intact." [62]

Finally, Laurent gives himself a means to try and solve the problem of the arrangement of atoms: it is the *Formula*, but defined in a new way. It is not the expression of the 'truth', a representation of the real (?) arrangement of atoms [63] which would be an illusion, but a way of representing his opinion about the arrangement of atoms in compound bodies. This is not finally very far from the main directives of Berzelius, if they were understood in their strict meaning: "Formulas must only be used when the ideas they are to express are based on certified truths" [64].

However, according to Laurent, formula, certainly inspired by experiment, even if it is only a picture of the mind [65], is heavily ballasted by reality; but on the other hand, if one can say: formula must rejoin reality, and more than rejoin it, it

must *foresee* it, or rather *foretell* it: formula leads to the set up of experiments which will sanction theory with intransigence [66].

Right from 1837, Laurent thus anticipates, thanks to theoretical speculations being materialized by formulae, the existence, and sets up the preparation of oenanthic ether [67]. Formulation is the active *method*, leading chemist from theory to experiment.

Laurent's attempt failed, but it "paved the way to a new direction." The method described will be the one used by later chemists who will claim to belong to the group of the 'theoretician chemists', an expression which was a contradiction in terms, some decades before [68].

As can be seen, Laurent had always deeply associated *role* with *place*, which had allowed him, when choosing certain types of privileged metamorphosis, to link 'speculations' about atomic arrangements to what could be given prominence through experiment: analogies of property, provided that some had influence upon others.

Later chemists will be bolder: they will resort to atoms themselves. But, since the 'topographic position of atoms' is absolutely impossible to know, they will dissociate role and place, and will focus their speculations only on the chemical role, which is newly defined and can be related to experiment. It amounts to the saving of a hypothesis about physical atoms which does not yet seem to have "become essential for our science" [69]. It is always the application of the old adage of Dumas: it is not sufficient for a theory to be rigorously possible, it must be necessary. [70]

The atomic hypothesis is thus niggardly reintroduced in stages and accurately controlled, after having forbidden any work on the molecule. To mention chemical atoms is to reject every corpuscular idea, and it may not be by accident that, after Avogadro and after Laurent, Couper is an author who will be forgotten.

4. The Role

In 1858, as he was in Paris, and was working in this international team created by Wurtz, and in his laboratory at the Faculty of Medicine in Paris, A. Couper presented a report that Wurtz rather left lying about on his desk [71]. He insisted in it on the necessity of resorting to atoms themselves. For science must not fix its analysis "at the very moment when an explanation will be needed" [72]:

> For the unity of science and for the improvement of research, it is absolutely necessary to consider these bodies that have been called *radicals* as being derived and containing no ultimate hidden force, to admit that their properties are a direct consequence of the individual properties of the elements composing them

This obviously recalls the resolution of compound atoms by Berzelius, but it is here no longer a matter of the permanent and determined *nature* of the ultimate elements: their electrochemical nature. For the first time, element is considered

according to its specific 'chemical nature', its capacity for establishing some connection or other in order to make up a molecule.

'Forces and properties' of the most simple elements must allow a knowledge of the combinations they produce. With Couper, molecule gains a unity, a stability and a capacity for changing, of which the combinations between the elements give an account, because of this property shared by every element: *chemical affinity.* He thus makes a distinction between:

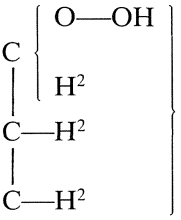
— the *elective affinity* between carbon and oxygen;

— the *affinity of degree* responsible for C^2O^2 and C^2O^4.

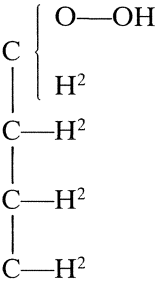
And he demonstrates with a simple argument based on substitution that carbon possesses precisely the property to unite with oxygen, that "it is not hydrogen which serves as a link". This property, which is fundamental to organic chemistry, "is, I think, indicated here for the first time."

Affinity, this mysterious attraction, this rather magic force [73], at the same time an object of fascination [74] and an imponderable being eminently suspect in the eyes of the scientist [75], that chemistry tried to expel and to reintroduce in turn [76] — for it was impossible to do without it [77] — is subjected here to a kind of transmutation: it becomes a constructive concept.

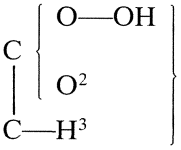
A representation of molecules is obtained:



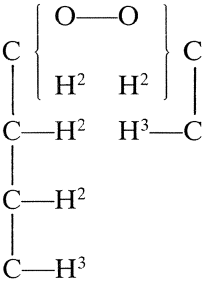
propylglycol



butylic alcohol



acetic acid



mixed butyl-alcoholic ether

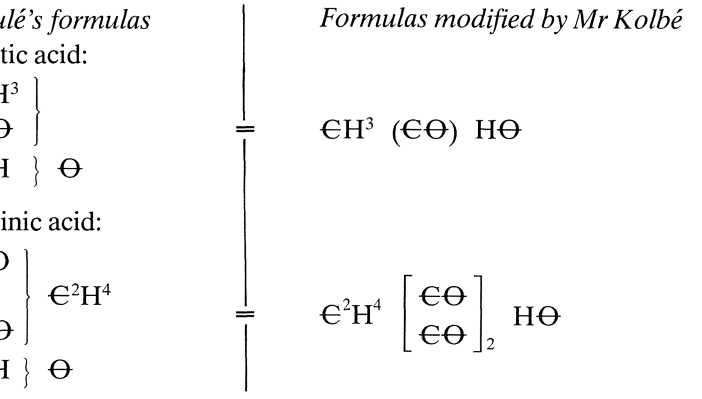
a representation which expresses unity and the reason for this unity: the way "atoms are chemically very close to each other."

The 'report' of Kekulé, in the same year 1858, is more famous [78]. The German chemist, also in Paris and at Wurtz's laboratory, develops there ideas which are similar, but more grounded, on the considerations about metamorphosis. Atoms are still considered from a quantitative point of view, according to Gerhardt's and Laurent's definition in 1846. Every atom has a fixed number of units of combinations (4 for carbon, 2 for oxygen, etc. . . .), which are 'used to form' atoms, or compound radicals. No doubt that this language and this cast of mind are more pleasant to Wurtz.

From that, many chemists will be interested in the *chemical relation* of elemental atoms in the molecule: Kolbé, Kekulé, Butlerov. Says Butlerov:

> Messrs Kolbé and Kekulé both want to determine the way elemental atoms are chemically linked in a compound molecule. Mr Kolbé refers to this way by the word '*composition*'; as for me, I would rather name it '*chemical structure*', an appellation to which different significations were not yet attached, and which, because of that, is less liable to give rise to misunderstandings [79].

Therefore, writes Butlerov, Mr Kolbé's formulas express, as well as Mr Kekulé's, the chemical structure of the bodies:



We will notice that the authors always use the 'double atoms' of Berzelius, represented by the chemical symbols crossed by a horizontal bar.

5. The Place

5.1. THE PLANE

There was no question, except maybe in the case of Kolbé, of expressing, through the rational formula, the real grouping of atoms in space [80], the 'topographical position of atoms': "when it comes to *chemical connection* of atoms in a molecule,

it is of no use *either* to know their position, *or* even to acknowledge the existence of physical atoms", says Butlerov [81].

The principle of *chemical structure*, an undeniable progress in the explanation of the molecule, indicates also the *ultimate attempt of chemists* to avoid the atomic hypothesis. "As long as it will not have become essential to our science", it is advisable to avoid to resort to these physical atoms.

Wurtz, at the same time, begins his speculations about the positions of atoms, which is *inseparable* from the relation that atoms can have with each other. If formulas

$$\left\{\begin{array}{l}\text{C}^4\text{H}^7\\ \text{C H}^3\end{array}\right. \quad \text{and} \quad \left\{\begin{array}{l}\text{C}^3\text{H}^5\\ \text{C}^2\text{H}^5\end{array}\right.$$

express the mode of formation of these carbonaceous hydrogens (this is the general meaning of typical formulas), they are powerless to indicate "some molecular differences being characteristic of a state of isomerism" [82]. Hence formulas of *composition* which will express more: without intending to express the exact positions of the atoms, it is possible to indicate, via the formula

$$\begin{array}{l}\text{CH}^2\\ \text{CH}^2\\ \text{CH}^2\\ \text{CH}^2\\ \text{CH}^2,\end{array}$$

"that three atoms of carbon are in close relation with each other and with the other two, whereas it is not the same for the last two atoms (extremes), between which we can suppose there is a great gap."

Hence the theoretical explanation of isomerism, through the positions of these gaps: the first spatial representation of the molecule is a representation of its *gaps*.

The four different molecular arrangements, the four isomeric states of the molecule are indicated through the following formulas of composition:

$$\begin{array}{llll}(\text{CH})'' & (\text{CH}^2)' & (\text{CH}^2)' & (\text{CH}^2)'\\ \hline \text{CH}^2 & (\text{CH})' & \text{CH}^2 & \text{CH}^2\\ \text{CH}^2 & \hline & (\text{CH})' & \text{CH}^2\\ \text{CH}^2 & \text{CH}^2 & \hline & (\text{CH})'\\ \text{CH}^3 & \text{CH}^3 & \text{CH}^3 & \text{CH}^3,\end{array}$$

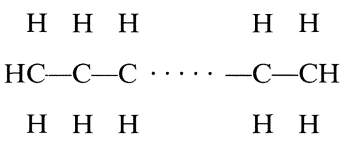
that the following typical formulas express 'only to a certain extent':

$$\left.\begin{array}{l}\text{C H}\\ \text{C}^4\text{H}^9\end{array}\right\} \quad \left.\begin{array}{l}\text{C}^2\text{H}^3\\ \text{C}^3\text{H}^7\end{array}\right\} \quad \left.\begin{array}{l}\text{C}^3\text{H}^5\\ \text{C}^2\text{H}^5\end{array}\right\} \quad \left.\begin{array}{l}\text{C}^4\text{H}^7\\ \text{C H}^3\end{array}\right\}.$$

To imagine the positions of atoms in space (the plane), even if these positions are only relative, is thus a step further in the degree of hypothesis. Butlerov still denies it in 1866 [83]: "I speak, of course, of the chemical relations of atoms, not of their positions in space"; and he is able to anticipate two possible isomerisms for the formula C^4H^{10}, and many more for its upper homologues. The same principle of chemical structure also anticipates the isomerism of ketones by Popoff [84].

As for Kekulé, his molecular conception seems rather vague for a time. If, in the years 1858—60, no spatial consideration seems to go into his reasoning, considerations on some cases of isomerism in 1863 [85] bring him to speak of arrangement (a term dear to Laurent), of a *gap* (like Wurtz) that hydrogen would fill, of a *place* that bromine could take. In 1866, he does some memorable spatial speculations: he draws the benzene hexagon and gives an account of the benzene derivatives of substitution, through probability considerations on the place occupied by the bromine which has been progressively substituted, using for that the notion of 'sphere of action' [86].

As far as we know, it is Friedel and Ladenbourg who for the first time use the developed formulas where the well-known dashes appear:



"Saturated hydrocarbons are composed of a kind of continuous chain of carbons, at the links of which hydrogen fastens" [87].

5.2. THE SPACE

In 1874, Le Bel [88] and Van't Hoff [89], independently of each other, give to the molecule the last dimension that was lacking: they imagine atoms in space, and carbon at the centre of a tetrahedron, in order to give an account of certain cases of isomery which had already occupied Pasteur [90], and of which plane formulas could not give any account.

Molecule is then considered as a geometrical structure which is certainly not rigid but which makes up a chemical and spatial unit, and of which the elements and the physical atoms, through their properties and their positions, allow to *give an account* of the stability of the structure, and to *anticipate* the possible transformations of this structure.

It will be necessary to wait a long time before understanding that one needed to go further than Couper thought in the resolution of the parts within the atom: "We think of a molecule, in the broad sense of the term, as a stable system composed of nuclei and electrons", Professor Daudel will write [81].

6. Conclusions

But what is going on? For some time, our account has been dealing with formula more than with molecule. Did we go from the problem which was proposed to us, concept, to the problem of representation? Or wouldn't it rather be the authors who, without our knowing, would have changed the initial problem, dealing no longer with the existence and the composition of a real or imaginary object, but with the representation of a theoretical object? Without our knowing?

6.1. IMAGINARY OBJECT. THEORETICAL OBJECT

As long as molecule will not be a 'visible' object giving rise to *measurement*, it will remain a 'product of the human mind'. But so that it becomes a 'necessary hypothesis', a 'fecund concept', it has to go from the status of object of our imagination to the status of theoretical object. Many people put forward conjectures about the 'intimate nature of phenomena', views which are more or less probable of the way molecular structures are built, many people could 'sense' the relative arrangement of atoms: Laurent said of Baudrimont "He is a dreamer who does not lack ideas". But there is a long way from this form of aesthetic satisfaction to the structuring of a thought around a hypothesis (or a hypothetical system) which has been put forward about a certain number of observations and will mark the starting point of scientific inquiry. "Research is an adventure just beginning".

6.2. THE CHANGE OF METHOD

This transformation — which is essential though sometimes not very discernible — was possible at the expense of a change, or almost of an *inversion*, of the chemical method. Lavoisier's method only proceeds by "going from what is known to what is unknown", "begins only with facts and then gradually goes up to the degree of generality" that science requires. If it has ever been rigorously applied, it had to be reformed. One must not begin with facts anymore, "experiment in the hand", to reach a theory which, in the best case, would become a "truth of fact and observation"; but one must come close to the concrete from an intellectual elaboration — even if it is inspired by facts. Theory would thus correspond to a question that man addresses to things, and which is formulated so that things can answer: this is the setting up of an 'experimental research'. The construction of formulas by Kekulé, Butlerov, Wurtz, was leading to the prediction of cases of isomerism that experiment had to undertake to confirm. The concrete thus firmly ballasts the abstract and prevents "the free flight of imagination". For "if all our knowledge begins *with* experiment, it does not prove that it all derives *from* experiment", Kant had already said.

6 3 THE CHANGE OF OBJECTIVE

In the chemical method, the change from the phase of theoretical elaboration to the phase of the experimental checking of the judiciousness of this elaboration, needs a mediator This interlinking function is provided by the formula The formulation of the theoretical object — the molecule — allows one to verify the appropriateness of this object, with respect to molecular reality, and this is because the formula of the molecule must not only give an account of properties, but also anticipate phenomena This capability of anticipation, through formula, measures the real fecundity of the molecular concept Laurent, Kekulé, Butlerov, Wurtz have been very sensitive to that

At the same time, the objective of science had to be

— bolder to dare to speak, to 'speculate' about atoms, their arrangement, to use the atomic hypothesis as a hypothesis, to conceive affinities,

— and more simple not to have anymore the ambition — at least for some time — to 'penetrate the inner nature of things", and not to have the plan to reproduce the reality of the molecule, but only to express an 'idea', a 'thought' relating to 'chemical composition', refining it thanks to experiment

Thus, the chemical project had to be no longer the *discovery* of truth, but the *representation* of a theoretical object more and more subjected to an inflexible experiment

References

1 J M Ampere *Ann Chim et Phys* **90**, 43—86 (1814)
2 Ad Ch Wurtz *Dictionnaire de Chimie Pure et Appliquee*, Hachette, Paris, 1869—76 1er supplement (1886)
3 Ad Ch Wurtz *La Theorie Atomique*, Germer Baillière, Paris (1879)
4 J Perrin *Les Atomes*, p 151 Alcan, Paris (1913)
5 R Descartes *Les Principes de la Philosophie*, Amsterdam (1644) 2e partie, Des Principes des choses materielles, n°4
6 I Newton *Mathematical Principles of Natural Philosophy* and *System of the World*, Book **III**, p 414 Greenwood, reprint (1969)
7 I Newton *Opticks*, p 375 3rd Ed (1721)
8 H Metzger *Newton, Stahl, Boerhaave et la Doctrine Chimique*, p 28 A Blanchard Paris (1930)
9 G W Leibnitz *Apostille* a la lettre a la Princesse de Galles du 12 mai 1716
10 Rome de L Isle *Essai de Cristallographie* (1722) Cited by H Metzger *La Genese de la Science des Cristaux*, p 66 Blanchard, Paris (1969)
11 R J Hauy *Traite de Cristallographie* **I**, pp 50—53 Bachelier et Huzard, Paris (1822)
12 G Delafosse *Memoires des Savants Etrangers* **8**, 641—690 (1843)
13 A Lavoisier *Traite elementaire de Chimie*, Cuchet, Paris (1789) Discours preliminaire XVij
14 Ad Ch Wurtz *Leçons elementaires de Chimie Moderne*, pp 16—17 Masson Paris (1884)
15 Ad Ch Wurtz *La theorie atomique, op cit*, p 25
16 Ch Gerhardt *Ann Chim et Phys* **8**, 245 (1843)
17 J Dumas and A Stas *Ann Chim et Phys* **1**, 5—59 (1841)
18 J J Berzelius *Traite de Chimie Minerale Vegetale et Animale* Vol V pp 113—116 2nd Fr Ed , Firmin Didot Paris

19 J Liebig *Traité de Chimie Organique*, Vol II, p 38 Trad Fr Ch Gerhardt, De Fortin, Paris (1840—44)
20 *Cf* for example *Compte-rendus de l'Académie des Sciences* **V**, 863, **VI**, 629, 747, 823 (1838), **VII**, 352 (1839) *etc*
21 A Laurent *Méthode de Chimie*, pp 74—75 Mallet-Bachelier, Paris (1854)
22 J B Dumas *Leçons sur la philosophie chimique*, p 309 Bechet Jeune, Paris (1837)
23 Ch Gerhardt *Ann Chim Phys* **7**, 129—143 (1843), **8**, 238—245 (1843)
24 A Laurent *Ann Chim Phys* **18**, 266—98 (1846)
25 *Ibid*, 288
26 Ch Gerhardt *Introduction a l'étude de la chimie par le systeme unitaire*, p 7 Chamerot, Paris (1848) *Cf* also p 9
27 A Laurent *Ibid*, 296 Underlined by the author
28 J J Berzelius *Essai sur la théorie des proportions chimiques et sur l'influence chimique de l'électricite*, pp 22—23 Johnson reprint (1972) 1st Ed 1819
29 J Liebig *Traité de Chimie organique op cit* **1**, 239
30 J J Berzelius *Ann Chim et Phys* **51**, 308—314 (1832)
31 J J Berzelius *Traité de Chimie, op cit* **5**, 285
32 A Comte *Cours de philosophie positive*, Hermann, Paris (1975) 38e leçon, Sept 1835
33 J B Dumas *Leçons sur la philosophie chimique, op cit*, 419
34 *Cf lettre* de A Laurent a Ch Gerhardt, 4 juillet 1844, in H Tiffeneau, *Correspondance de Ch Gerhardt*, Masson, Paris (1918)
35 Ch Gerhardt *Traite de Chimie organique*, Vol 4, p 568 Firmin Didot, Paris (1853)
36 *Cf lettre* de A Laurent, in *Corr, op cit*, 10
37 *Cf* for example J Dumas *Ann Chim et Phys* **56**, 125—150 (1834)
38 For example A Laurent *Compte-rendus* **10**, 409 (1840)
39 J Dumas *Compte-rendus* **6**, 695 (1838)
40 For example A Laurent *Revue Scientifique et Industrielle du Dr Quesneville* **11**, 361—374 (1842)
41 A Laurent *Recherches diverses de Chimie Organique*, in *Theses de Chimie et de Physique* presentees a la Fac des Sc de Paris, le 20 dec 1837, p 85
42 J J Berzelius *Rapport annuel sur les Progres de la Chimie*, 155 (1841)
43 J J Berzelius *Rapp ann, op cit*, 226—234 (1847)
44 J Jacques *Bulletin de la Societe Chimique de France*, D31—39 (1954)
45 J P Daudey, J P Malrieu, D Maynau *Le Courrier du CNRS* **64**, 51 (1986)
46 *Cf* Delafosse *Mem Savants etrangers*, 546—547 (1848)
47 A Laurent in J Jacques, *op cit*, D34—35
48 A Laurent *Meth de Chim, op cit*, 400
49 J B Dumas *Leçons , op cit*, p 4 Iere leçon, Apr 1836
50 *Cf* Wurtz *La Theorie atomique, op cit*, 26—27
51 *Cf* Wurtz *Dict chim, op cit*, 'Atomicite' **I**, 451 (1869)
52 J Dumas *Leçons , op cit*, 7eme leçon, 265
53 L Meyer *Les theories modernes de la chimie et leur application a la mecanique chimique*, pp 209—210 Trans Bloch, G Carre, Paris (1887)
54 Gay-Lussac *Memoires de la societe d'Arcueil* **2** (1813)
55 A Avogadro *Journal de Physique* **33**, 58 (1813)
56 Ampere *Cf* [1]
57 *Cf* Clara de Milt *Chymia* **4**, 107 (1953)
58 A Avogadro *Ann Chim et Phys* **14**, 350 (1845)
59 A Laurent *Meth de Chim, op cit*, 321—332
60 A Laurent *lettre* a Ch Gerhardt, *Corr, op cit*, 81
61 A Laurent *Meth de Chim, op cit*, 392
62 R Lespieau *La molecule chimique*, p 4 Alcan, Paris (1926)
63 *Cf* Chevreul *Recherches chimiques sur les corps gras d'origine animale*, pp 2—4 F G Levrault, Paris (1823)
64 J J Berzelius *Ann Phys Chim* **51**, 314 (1832)

65 A Laurent *These, op cit*
66 A Laurent *Ann Chim Phys* **63**, 207—219 (1836)
67 A Laurent *These, op cit*, 34—35
68 *Cf* A Butlerov *lettre* a Ad Wurtz, Feb 1882, in G V Bykov and J Jacques, *Revue d'Histoire des Sciences*, 130—134 (1960)
69 A Butlerov *Bulletin Société Chimique*, note 2, 102 (1864)
70 J B Dumas *Leçons , op cit*, 345
71 *Cf* J Jacques *Bull Soc Chim*, 528—530 (1953)
72 A S Couper *Ann Chim Phys* **53**, 469—489 (1848)
73 *Cf* W Goethe *Les affinités électives*
74 *Cf* for example Millon *Compte-rendus* **19**, 799—806 (1844)
75 *Cf* A Lavoisier *Traité élémentaire de chimie*, Discours preliminaire
76 *Cf* H Metzger *Les doctrines chimiques en France du début du 17e a la fin du 18e siecle*, Blanchard, Paris (1922)
77 *Cf* E Mitscherlich *Ann Chim Phys* **4**, 67 (1842), **7**, 5 (1843)
78 A Kekule *Ann der Chem u Pharm*, 129 (1858) *Bull Soc Chim*, 20 (1859)
79 A Butlerov *Bull Soc Chim*, 100—128, (1864)
80 Ad Wurtz *Bull Soc Chim*, 365 (1860)
81 A Butlerov *op cit*, 102
82 Ad Wurtz *Bull Soc Chim*, 51—70 (1863)
83 A Butlerov *Bull Soc Chim*, 25—36 (1866)
84 A Popoff *Bull Soc Chim*, 36—42 (1866)
85 A Kekule *Bull Soc Chim*, 37—40 (1863)
86 A Kekule *Bull Soc Chim*, 40—46 (1866)
87 C Friedel and A Ladenbourg *Bull Soc Chim*, 67 (1867)
88 J Le Bel *Bull Soc Chim*, 337 (1874)
89 J M Van't Hoff *Bull Soc Chim*, 295 (1875) *Stéréochimie*, Paris (1892)
90 L Pasteur *Leçons sur la dissymétrie moléculaire*, Paris (1861)
91 R Daudel *Les fondements de la Chimie théorique*, p 31 Gauthier-Villars, Paris (1956)

Quantum Theory and the Molecular Hypothesis

R. G. WOOLLEY
Department of Physical Sciences, Trent Polytechnic, Clifton Lane, Nottingham NG11 8NS, U.K.

Today we realize that the whole of chemistry is one huge manifestation of quantum phenomena.
C. J. Ballhausen: *J. Chem. Ed.* **56**, 357 (1979).

1. Introduction

Over the last few years there has been a growing awareness that the traditional formulation of quantum chemistry does not exhaust the possibilities for the application of quantum theory to chemical problems. This awareness has come about through a re-examination of the foundations of theoretical chemistry to which many have contributed in the last decade. Ten years ago, I encountered considerable hostility to my suggestion that the programme of conventional quantum chemistry is not just a simple consequence of setting out a molecular quantum theory if one starts from the Schrödinger equation for a system of interacting electrons and nuclei; today that is a much less controversial statement, and it is now widely recognized that classical molecular structure is problematic for a quantum theory of molecules. It is of course true that most of us use the procedures of conventional quantum chemistry in most of our work; nevertheless under the impact of new experimental techniques, especially those based on the use of laser radiation sources and supersonic molecular beams, new theoretical methods have come into quantum chemistry. An overriding concern with the determination of molecular structures has receded somewhat, and some new theoretical techniques in chemical physics make little or even no reference at all to molecular structure. Perhaps this increase in abstraction is no more than an illustration of the process foreseen for the development of theoretical science by the late P. A. M. Dirac in the opening pages of his famous book, *The Principles of Quantum Mechanics.*

The purpose of this article is to review some of these new concepts in theoretical chemistry; the article has three main sections. Section 2 is devoted to a short résumé of the origins and development of the modern science of chemistry that culminates in the classical stereochemistry of Van't Hoff based on the atomic/molecular conception of matter. It is important to recognize the metaphysical content that chemistry thereafter incorporated since the pioneers of quantum chemistry founded their quantum mechanical treatment of chemical phenomena squarely on the classical structural conception they had inherited from Van't Hoff (Section 2.3).

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, 45—89.

Section 3 is concerned with a re-examination of the possibilities for a theoretical structure that could realize the grand conception of chemistry contained in the quotation above from Ballhausen *without* essential recourse to an underlying classical framework. We begin with the formalism and interpretation of quantum theory, focusing our attention on the holistic nature of quantum theory that is implied by the general occurrence of the Einstein—Podolsky—Rosen (EPR) correlations. Next, the Second Law of Thermodynamics and the irreversible evolution of chemical reaction mixtures towards an equilibrium state are considered, for they are central to chemistry and must be accomodated in a general theoretical framework; a major effort in this direction has been made by Ilya Prigogine and his co-workers. Their main conclusion is that irreversibility in quantum dynamics must be based on a *quantum field theory*; in my view this invites a reconsideration of the atomic/molecular conception of matter because the notions of 'atom' and 'molecule' must be found a meaning in this new theoretical chemistry based on a quantum field theory of chemical substances.

I approach this question in two ways. In the last (and longest) part of Section 3 I adopt the definition of a molecule as a collection of a specified number of electrons and nuclei and examine the consequences of a quantum theory based on this model; this 'molecular quantum mechanics' turns out to be *different* from conventional quantum chemistry in some important respects. The application of the Generator Coordinate Method (GCM) to theoretical molecular spectroscopy and a reworking of Van Vleck's theory of polar gases illustrate the possibilities for using *space-fixed axes* in molecular problems. We shall see that potential energy surfaces are constructs that are built into particular models (classical chemistry, quantum chemistry); on the other hand, as far as molecular quantum mechanics is concerned a potential energy (hyper)surface has *no* fundamental significance. The transition from this model to the usual quantum chemistry is achieved by means of singular transformations such as the Born—Oppenheimer approximation and the Eckart theory, which is briefly reviewed. Finally, and in preparation for the consideration of the development of a quantum field theory for chemical substances, the failure of molecular quantum mechanics to account for natural optical activity is explained.

Section 4 is written in two parts; 4.1 gives an elementary introduction to a field description and its quantization; several familiar examples of fields and their quantization are discussed. Quantum field theory is the quantum theory of macroscopic systems; superconducting metals, superfluids and lasers are easily recognized as 'large' quantum systems. Although many practical results can be derived from a classical treatment of the electromagnetic field, the only consistent treatment of the interaction of matter and radiation is based on the application of quantum laws to both 'parts' (quantum electrodynamics) because the *principle of gauge invariance* implies that there can be no meaningful, complete separation of matter and radiation. Recently we have come to appreciate that we encounter 'quantum behaviour' in more of physical science than these relatively esoteric examples. In Section 3, the Second Law of Thermodynamics and natural optical activity are

discussed using quantum field theory; these are physical aspects of macroscopic matter that cannot be obtained as the sum of microscopic 'parts'. More generally, the same remark applies to states of matter that exhibit any broken symmetry. It follows from this discussion that it is a mistake to suppose that classical theory always applies to macroscopic objects — those things apprehended by our senses — whereas quantum theory is restricted to the microscopic domain. This traditional view can no longer be sustained.

Section 4.2 gives a heuristic account of the application of quantum field theory to the description of chemical substances in which 'atoms' and 'molecules' are identified as the composite 'quasi-particles' of the interacting electron and nuclear fields. The *molecular structure hypothesis* reappears as a device (and a powerful metaphor) for facilitating the construction of the quantum states of the quasi-particles without having the 'full' solution of the quantum field theory. From this point of view molecular structure is seen to arise as a reflection of macroscopic factors (the interacting fields). This interdependence of microscopic and macroscopic properties is a consequence of quantization and it suggests that a purely molecular account with atoms and molecules regarded as 'building-blocks' — the programme of classical reductionism — is insufficient for a theory that encompasses the general aspects of chemistry. Finally, in Section 5, some concluding remarks are presented.

2. The Historical Perspective

C'est de la chimie organique — la chimie du carbone — que naquit cette belle théorie de l'atomicité, qui permet de représenter la molécule comme un groupement d'atomes liés entre eux d'après certaines lois, et formant un système complet et stable.

J. H. Van't Hoff: *La Chimie dans l'Espace*, Bazendijk, Rotterdam (1875).

2.1. ORIGINS

Chemistry is concerned with the composition and properties of matter, and with the transformations that can occur spontaneously or under the action of heat, light or other sources of energy, when different kinds of matter are brought together. The modern science of chemistry can be conveniently taken to date from the publication (1789) [1], by Antoine Laurent de Lavoisier of *Traite élémentaire de chimie* in which chemical knowledge is described in an entirely new way that we recognize as a genuine precursor of the chemistry of our times. The distinguishing feature of Lavoisier's work lies not so much in the experiments he conducted, but rather in the *theoretical* approach he developed; one of the most revealing signs of this is his revision of chemical nomenclature which became a completely new instrument of thought, adapted to his conception of chemistry as an analytical science. In this enterprise Lavoisier was strongly influenced [2] by the linguistic philosophy of the Abbé de Condillac.

The modern 'scientific' approach that emerged in the seventeenth century aimed

to describe the natural world through analytical procedures of classification and systematization; the underlying mechanical philosophy was firmly grounded in a picture of the world as a world of physical objects endowed with well-defined fixed properties that can be described in mathematical terms. It can be seen as a return to the mathematical ideals of the Pythagoreans, and of Plato, and a renewal of the ideas of the early Greek atomists, for example Demokritos; there was quite explicitly a movement against the still prevailing Aristotelian system. The scientific revolution represented a fundamental shift from the organism to the machine as the model in terms of which the natural world would be understood [3].

Nevertheless, as far as chemistry is concerned, this change in outlook did not take place until more than a century later. Prior to Lavoisier's work, chemical phenomena had been understood in terms of a quite different tradition that had originated in Antiquity. This tradition is derived from a perception of substances as possessing inherent qualities that govern their chemical and physical behaviour; alchemy belongs to this human mythical tradition which is directly contrary to the analytical, scientific approach. The mythical perception of the world is synthetic in outlook, seeing the world without distinction between the animate and the inanimate as a dynamic living whole. Things have dynamic qualities that can be accounted for by analogy with human emotional states, or through personification; the mythical account of the physical world may be seen as an extension of human life in the world, and is based on perceptions that are totally alien to the modern world of science and technology [4]. Although alchemy was regarded with increasing scepticism throughout the eighteenth century due to the accumulation of empirical evidence that spoke against transmutation of metals, the decisive break in favour of the modern approach to chemistry can be associated with the publication of Lavoisier's book.

In the years that followed, Lavoisier's classification was refined — for example, heat and light were dismissed from his list of elements — and the chemical elements were given a microscopic interpretation in terms of Dalton's atomic hypothesis (1808) [5]. Henceforth the elements were to be regarded as being composed of microscopic building-blocks — *atoms* — which were indestructible and had invariable properties (notably weight) characteristic of the individual elements. Progress was slow however, and nearly fifty years elapsed before Cannizzaro gave a clear explanation of the significance for analytical chemistry of the earlier ideas of Avogadro (1811) and Ampère (1814) [6, 7].

2.2. CLASSICAL CHEMISTRY

After the resolution of the difficulties of analytical chemistry in the mid-nineteenth century, chemists developed a kind of chemical language — a system of signs and conventions for their use — which gave them a representation of their fundamental postulate that 'atoms' are the 'building-blocks of matter'. When we speak of a molecule in classical chemistry we mean a semi-rigid collection of atoms held

together by chemical bonds; thus molecules are built up using atoms like the letters of an alphabet. The 'laws' that govern the relative dispositions of the atoms in ordinary 3-dimensional space are the classical valency rules which therefore provide the syntax of chemical structural formulae.

The development of an interpretation of chemical experiments in terms of the molecular model was a highly original step for chemists to take. In the nineteenth century the only known forces of attraction which might hold atoms together in molecules were the electrical and gravitational forces, but these seemed useless for chemical purposes and so were given up in favour of a basic *structural* principle [8]. This obviously marked a break with the then known physics, a rupture that would not be healed until the discovery of the quantum theory. Nevertheless in one important respect chemists made a change that brought their thinking much more into line with the customary approach in physics; from the 1860's onwards, inductive argument was replaced by a deductive model based on the formulation and testing of hypotheses [9].

The structural idea was given a definite philosophical content by J. H. Van't Hoff, the founder of classical stereochemistry; Van't Hoff proposed that atoms and molecules were microscopic material objects in the ordinary 3-dimensional space of our sensory experience; molecules could not be directly perceived simply because of the limitations inherent in our senses. For Van't Hoff, stereochemistry was part of an argument to give a 'proof' of the physical reality of molecules [10]. In the second half of the nineteenth century the attribution of a physical reality to atoms was highly controversial because of its metaphysical nature; while the realist position was advocated strongly by chemists such as Van't Hoff, and physicists such as Maxwell and Boltzmann, it was strongly attacked by other noted scientists such as Duhem and Ostwald whose scientific philosophy was related to the positivism of Ernst Mach; for them atoms were fictions of the mind, and they preferred to restrict their discussions to the macroscopic domain. Yet again, others preferred to maintain a sharp distinction between what they regarded as objective knowledge and what was only probably known or speculative. For example, Kekulé did not share the strong convictions of his student Van't Hoff about the structural model, but Kekulé was nevertheless an effective user of the model [11].

By the early twentieth century however, the successes of the structural model and developments in physical chemistry such as Perrin's study of Brownian motion [12], and the Bragg's investigation [13, 14] of the X-ray diffraction properties of alkali halide crystals, had confirmed classical realism as the dominant metaphysics in chemistry. From that time onwards throughout the twentieth century, classical chemistry has been characterized by its own central dogma:

CENTRAL DOGMA OF CLASSICAL CHEMISTRY

All chemical experiments will be rationalized in terms of the relative spatial dispositions of the atoms in molecules, conceived as microscopic material bodies in ordinary space.

Thus, if chemistry had begun with Lavoisier as the science of the transformation of chemical substances, by the time of the formulation of the quantum theory (1925—1926) it had enjoyed a long tradition as the science of *molecular* transformations.

2.3. QUANTUM CHEMISTRY

As chemistry made rapid progress in the late nineteenth century, physicists began to unravel sub-atomic structure. Perhaps the most important idea that followed from the early investigations was the recognition that understanding the structure of the atom seemed to involve the solution of a problem in dynamics. Out of various well-known dilemmas grew the old quantum theory of Bohr and Sommerfeld, and subsequently the quantum mechanics of Heisenberg, Schrödinger and Dirac [15—17]. Both theoretical schemes were quickly applied to chemical problems, essentially in the same fashion; the classical structural model of a molecule can be described in mechanical terms and, if regarded as a closed system, can be associated with a well-defined molecular Hamiltonian which can then be investigated as a problem in dynamics. The historical record shows that an important set of concepts which persist in molecular spectroscopy and quantum chemistry today, for example, the idea of body-fixed coordinates (Eckart), potential energy surfaces, and the Born—Oppenheimer separation of electronic and nuclear motion, were originally developed within the framework of the old quantum theory applied to the classical molecular model [18].

Among many notable achievements of the period 1916—1930 we may mention the following: the separation of molecular energies into electronic, vibrational and rotational contributions which is crucial for the interpretation of molecular spectra in terms of the bond-lengths and bond-angles of the molecular model [19, 20]; the use of potential energy surfaces for describing chemical reactions, and the Franck-Condon principle [21—23]; the angular momentum classification of molecular electronic orbits of diatomic molecules [24, 25] which was subsequently reinterpreted in wave mechanical terms as the molecular orbital model of the electronic structure of diatomic molecules [26, 27]. Moreover although the attempts to apply the Bohr theory to the problem of the chemical bond were not very successful [28, 29] the subsequent quantum mechanical 'Valence Bond' calculation on the hydrogen molecule due to Heitler and London [30], which is usually taken as the starting point of quantum chemistry, was based on the same general idea that the nuclei provide an effectively fixed framework about which the electrons move, so that the dynamical problem is to determine the electron distribution appropriate to a given nuclear framework (= molecular structure). The celebrated analysis of electron-nuclear separation due to Born and Oppenheimer [31] was intended as a justification for an already well-established procedure; it amounted to a reworking

in terms of the new quantum theory (wave mechanical perturbation theory) of earlier ideas developed by Born and Heisenberg using the action-angle perturbation methods of the old quantum theory [32].

The close relationship between the methods developed using the old quantum theory and the wave mechanical formalism of Schrödinger influenced decisively the subsequent development of quantum chemistry; Professor Daudel has played a prominent role in this story which is well known and need not be detailed here. The techniques of quantum chemistry are still the usual procedures that chemists, myself included, resort to for solving chemical problems; numerous examples of these techniques will be found in these volumes. In recent years however, unexpected and fundamental questions about the relationship between chemical theory and quantum mechanics have been raised, and it is to this topic that I shall devote the remainder of this review.

3. Chemistry and Quantum Mechanics

It turns out to be very difficult to predict precisely what will happen in a given chemical reaction; nevertheless the deepest part of theoretical chemistry must end up in quantum mechanics.

R. P. Feynman: *Lectures on Physics*, Vol. 1, Ch. 3 (1963).

3.1. QUANTUM THEORY

The technical machinery of quantum theory is well enough known that only a brief summary need be given here to establish some notation (cf. Löwdin's contribution to this volume). The mathematical structure of quantum theory is founded on the use of *linear operators* acting on a *Hilbert space* $\mathfrak{H}$. Some of these operators have classical analogues e.g. the Hamiltonian operator $\hat{H}$ and the coordinate operator $\hat{q}$ correspond to the classical variables for energy and position respectively; however quantum theory also uses operators that have no classical counterparts e.g. the operators describing spin angular momentum [33]. The so-called 'Correspondence Principle' linking classical variables and quantum operators is now of mainly historical interest; in the modern view, quantum theory is founded on group theory and the invariance principles associated with the laws of nature, and is characterized by its algebraic structure [34, 35].

The *states* of a system can be described in two ways, one more general but probably less familiar to chemists than the other; *wavefunctions*, $\{\psi\}$, are vectors in the Hilbert space ($\psi \in \mathfrak{H}$) and are also referred to as 'pure states'. On the other hand the density matrix, $\hat{\rho}$, is an operator on the Hilbert space ($\hat{\rho} : \mathfrak{H} \rightarrow \mathfrak{H}$) and is the quantum analogue of the Gibbs distribution function in classical statistical mechanics; the density matrix is often called a 'mixed state' [36]. The states and appropriate operators are used to form expectation values which provide the connection between the theory and experimental measurements; for a pure state ψ

we have to evaluate a scalar product, written in Dirac notation as

$$\langle \mathrm{A} \rangle = \langle \psi | \hat{A} | \psi \rangle \tag{1}$$

while for a density matrix the expectation value of $\hat{A}$ is written as a Trace over $\hat{\rho}$ and $\hat{A}$:

$$\langle A \rangle = \mathrm{Tr}[\hat{\rho}\hat{A}]. \tag{2}$$

The wavefunctions of a quantum system $\{\psi_n\}$ form an orthonormal basis for $\mathfrak{H}$, and we can use them to write a density matrix in its *spectral representation*:

$$\hat{\rho} = \sum_n w_n |\psi_n\rangle\langle\psi_n|; \tag{3}$$

here the $\{w_n\}$ are statistical weights or *probabilities*, so

$$\sum_n w_n = 1; \qquad 0 \leqslant w_n \leqslant 1. \tag{4}$$

As an example consider a dilute gas at equilibrium in a box at temperature T for which the $\{w_n\}$ are given by Maxwell—Boltzmann statistics,

$$w_n = \frac{\exp(-E_n/k_B T)}{\Sigma_n \exp(-E_n/k_B T)} \tag{5}$$

where the $\{E_n\}$ are the energies of the system corresponding to the states $\{\psi_n\}$ and k_B is Boltzmann's constant.

If we use the spectral representation of $\hat{\rho}$, an expectation value $\langle A \rangle$ can be written as a sum of scalar products involving the wavefunctions $\{\psi_n\}$ weighted by the probabilities $\{w_n\}$,

$$\begin{aligned} \langle A \rangle &\equiv \mathrm{Tr}[\hat{\rho}\hat{A}] \\ &= \sum_n w_n \langle \psi_n | \hat{A} | \psi_n \rangle. \end{aligned} \tag{6}$$

In the special case: $w_k = 1$; $w_n = 0$, $n \neq k$, $\hat{\rho}$ reduces to a *projection operator*:

$$\hat{\rho} \Rightarrow \hat{\mathrm{O}}_k = |\psi_k\rangle\langle\psi_k|; \qquad \hat{O}_k^2 = \hat{\mathrm{O}}_k \tag{7}$$

and we get

$$\langle A \rangle \Rightarrow \mathrm{Tr}[\hat{O}_k \hat{A}] = \langle \psi_k | \hat{A} | \psi_k \rangle. \tag{8}$$

In this special case $\hat{\rho}$ describes a pure state and has no advantage over the wavefunction ψ_k.

The time dependence of wavefunctions is generated by the Hamiltonian according to the time-dependent Schrödinger equation,

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{H}\psi; \tag{9}$$

the corresponding equation of motion for the density matrix is the Liouville equation

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = [\hat{H}, \hat{\rho}] \equiv \hat{L}\hat{\rho} \tag{10}$$

which defines the linear *Liouville* operator, $\hat{L}$.

A characteristic feature of the quantum mechanics of systems with a *finite* number of degrees of freedom is that, up to unitary equivalence, all possible states of the system lie in a single coherent Hilbert space $\mathfrak{H}$; a general wavefunction for such systems can always be represented as a superposition of any set of wavefunctions that form an orthonormal basis for $\mathfrak{H}$ [33],

$$\Psi = \sum_n c_n \psi_n. \qquad \text{Superposition principle} \tag{11}$$

The traditional expositions of quantum theory for chemists have been based almost exclusively on the use of wavefunctions; it is important to know that the density matrix formulation is more general (see equations (7) and (8) which show the wavefunction to be a special case of the density matrix formalism) and in certain cases may be indispensable; that is, for some quantum systems it may not be possible in principle to give a description in terms of pure states. We shall come to an example shortly.

Something must now be said about the interpretation of quantum theory as compared with the usual mechanistic view of chemistry. One extremely important idea that follows from the structure of quantum mechanics, and distinguishes it from classical science, is that, in the quantum mechanical view, Nature is *non-separable i.e.* the whole cannot be constituted out of parts [34, 37, 38]. Quantum mechanics is a logically consistent, holistic theory of the physical world — logically consistent because it has a well-defined mathematical structure that is free of contradictions, and holistic because of the lack of separability. As far as we know, it is consistent with experimental evidence so we have no real grounds for altering its structure; having said that, it is obvious that the claim that the fundamental physical theory of matter is holistic is likely to be seen as deeply offensive to the traditional molecular conception of chemistry which it contradicts. In recent years there has been a growing awareness among theoretical chemists of this paradoxical situation, and I shall review some of their attempts to come to terms with it.

The lack of separability comes about in the following way: if we analyse the quantum states of some system in terms of the quantum states of subsystems (its 'parts') we almost always find correlations in the *absence* of interactions — the so-called Einstein—Podolsky—Rosen (EPR) correlations. It should be recognized that EPR correlation is not confined to a few special experimental arrangements (2-photon cascades, electron—positron annihilation, etc.). On the contrary, the mathematical structure of quantum theory shows that EPR correlations are general.

The restriction of an EPR correlated state to a single subsystem is not usually a pure state, and so it cannot describe a 'real object' having its own *context-free* individuality, for an object should have a pure state description at every instant.

Another central feature of quantum theory is the *superposition principle* for wavefunctions, equation (11); its applicability is limited by the occurrence of superselection rules associated with conservation principles (conservation of mass, conservation of electric charge and so on) but apart from such restrictions, superposition would seem to rule out an interpretation of quantum theory in terms of well-defined objects. On the other hand, a literal interpretation of quantum theory as a holistic theory in terms of the Hamiltonian and states of 'the universe' (whatever that means) is manifestly not of interest to science. The way out of this dilemma has been described by Primas: objects are created by abstraction from universally existing EPR correlations. In physically interesting theories we make the traditional distinction between a 'system' and its 'environment' and accept that some EPR correlations are necessarily ignored; if *all* EPR correlations are ignored we shall have a purely classical theory. There is therefore no single correct description of Nature; different descriptions are not just possible but are inevitable and will usually contain mutually incompatible features [37, 38].

The above résumé of quantum description shows that a decisive change in the philosophy of science has occurred. The classical view is based on separability — the notion of an 'object' is stable and well-defined; objects are generally correlated of course, but the correlations are attributed to dynamical interactions and an *isolated* object is a valid and useful *ideal* limiting case. The state of a whole collection of objects, or subsystems, is determined once the interactions and the states of the individual subsystems are known.

An immediate conclusion that follows from this clear distinction between classical and quantum theories is that an atomic or molecular account of chemistry is by no means a *necessary* consequence of adopting a quantum theory basis for theoretical chemistry. For molecular quantum mechanics we must deliberately chose to ignore the EPR correlations between a molecule and its environment [38]; some implications of such an approach will be reviewed shortly and compared with conventional quantum chemistry with which it does not coincide. In the meantime it is pertinent to note that quantum theory is also consistent with a theory of chemistry based on *chemical substances*; there is a non-trivial logical distinction between the categories of 'molecule' and 'substance' that must be recognized, and a quantum field theory of chemical substances possesses features that are not found in, and are incompatible with, a purely molecular theory [39]. Thus the situation is that we have several different (and incompatible) *quantum* theories concerned with chemical phenomena, and these coexist with the purely *classical* stereochemistry that still enables chemists and life scientists to ignore quantum theory completely for many purposes [40, 41], for example in the practice and rationalization of much synthetic chemistry and biochemistry.

3.2. "THE DEEPEST PART"

Richard Feynman [42] doesn't tell us what "the deepest part of theoretical chemistry" actually is but there are surely two absolutely basic questions for chemists that need to be integrated with the theoretical framework provided by quantum mechanics, namely:

(i) the equilibrium problem — what is meant by the molecular structures associated with stable chemical substances,
(ii) the kinetic problem — how we describe the general facts of chemical kinetics when substances change, one to another, in a chemical reaction and finally evolve to an equilibrium state.

The very essence of chemistry is that when different substances are brought together, chemical reaction takes place and the combined system transforms into a *new equilibrium* mixture *i.e.* eventually there is no further change in time provided the external conditions remain fixed. For the theoretician this *irreversible* behaviour is a very striking fact since both the classical and quantum mechanics of closed systems are mathematical structures that do not distinguish between the future and the past *i.e.* they are formally invariant under time-reversal. Hence we come to the problem of irreversibility.

An example of an equilibrium system from which we can learn some important lessons [43] is a gas in a container with walls held at a constant temperature T. As first discussed by Einstein [44] the equilibrium state at this temperature is the gas bathed in black-body radiation with the gas molecules having a Maxwell—Boltzmann velocity distribution. An essential part of the evolution towards, and maintenance of, the equilibrium state is the continuous absorption and emission of electromagnetic radiation by the gas, along with collisions between gas molecules and collisions involving the walls of the container.

For the Hamiltonian of the combined system of matter and radiation we refer to quantum electrodynamics (QED) and write

$$\hat{H}_{\mathrm{QED}} = \hat{H}^{\mathrm{RAD}} + \hat{H}^{\mathrm{MATTER}} + \hat{H}^{\mathrm{INT}} \tag{12}$$

where $\hat{H}^{\mathrm{INT}}$ describes the matter-radiation interaction, and $\hat{H}^{\mathrm{RAD}} + \hat{H}^{\mathrm{MATTER}}$ refers to the idealized situation of isolated radiation and isolated matter with no coupling between the two parts. Very sophisticated arguments to do with gauge invariance tell us that only $\hat{H}_{\mathrm{QED}}$ has physical signifiance, and that matter and radiation can never be completely separated [39]. We expect that the Hamiltonian $\hat{H}_{\mathrm{QED}}$ can describe irreversible evolution towards equilibrium for suitable initial states, although a proof without simplifying assumptions remains a formidable task in mathematics. Still, we see that we do not have well-defined stable entities because the energy in the system is divided between $(\hat{H}^{\mathrm{RAD}} + \hat{H}^{\mathrm{MATTER}})$ and $\hat{H}^{\mathrm{INT}}$; photons — the quanta of the radiation field — come and go, being continually

absorbed and emitted by the material system; such dynamic processes are an essential ingredient of the evolution to thermal equilibrium.

The Maxwell—Boltzmann velocity distribution describes chaotic, uncorrelated motion of the gas molecules, and this may be taken as the signature of thermal equilibrium. On the other hand, correlations imply departure from equilibrium. Turning this round and interpreting uncorrelated motion as implying the maximum amount of independence consistent with a macroscopic equilibrium state, we come to a definition of physical particles: *physical particles are those entities which exhibit chaotic, uncorrelated motion in an equilibrium state.* The quantum states of such physical particles have *finite* lifetimes because of the persistent interactions — in the case described above, between gas molecules and radiation — and because of collisions which maintain the equilibrium state.

How can this picture of equilibrium be related to the quantum theory? It is perhaps easiest to start by saying how we will *not* get to this picture. Suppose we are given a Hamiltonian operator, $\hat{H}$, containing an interaction term, $\hat{V}$, for example $\hat{H}_{\mathrm{QED}}$,

$$\hat{H} = \hat{H}_0 + \hat{V}. \tag{13}$$

The conventional goal of quantum theory is to solve the Schrödinger equation for the stationary states of $\hat{H}$,

$$\hat{H}\psi_n = E_n\psi_n. \tag{14}$$

This is equivalent to finding an operator $\hat{\Lambda}$ that belongs to a family of unitary transformations and brings $\hat{H}$ to a new representation in which it is diagonal,

$$\begin{aligned} &\hat{\Lambda}\hat{H}\hat{\Lambda}^{+} \rightarrow \text{diagonal form} \\ &\hat{\Lambda}^{+} = \hat{\Lambda}^{-1}. \end{aligned} \tag{15}$$

In the process of transforming away all the interactions (the result of the diagonalization) we lose contact with the notion of physical time and the distinction between the past and the future. Evidently we need to go outside the usual framework of quantum theory, recognizing that irreversibility and equilibrium are intimately associated with persistent interactions which must not be transformed away.

Ilya Prigogine and his collaborators [46—50] have made an elaborate theoretical study of irreversible evolution to equilibrium and the Second Law of Thermodynamics. Their work highlights the fact that irreversible processes in quantum mechanics are only possible in the limit of a continuous spectrum; an immediate consequence of this restriction is that no *finite* quantum system, for example a molecule or a finite collection of N ($< \infty$) molecules with intermolecular interactions, can show irreversible behaviour, and the Second Law of Thermodynamics cannot be applied to such systems. This is a clear example of the need for theoretical chemistry to go beyond a purely molecular theory. Prigogine's theory is based on a special quasi-particle representation of a quantum field theory in which

physical 'particles' have finite lifetimes. He introduces a notion of 'microscopic entropy' and constructs a transformation operator $\hat{\hat{\Lambda}}$ such that

$$\hat{\hat{\Lambda}}\hat{H}\hat{\hat{\Lambda}}^{+} \rightarrow \left\{ \begin{array}{l} \text{A representation describing 'physical} \\ \text{entities', with lifetimes } \tau, \text{ that interact} \\ \text{incoherently among themselves.} \end{array} \right. \tag{16}$$

The microscopic entropy is related to the lifetime, τ, of the particles which is a new observable only found in theories in which the Hamiltonian has a continuous spectrum. For finite systems and certain kinds of field theories (e.g. non-interacting fields) this representation does not exist because the construction reduces to the unitary transformation that diagonalizes the Schrödinger equation [50].

According to Prigogine and George [50] the new representation, if it exists, shows irreversible behaviour in the following sense: there are initial states (density matrices) $\hat{\rho}_0$ which evolve into the future ($t \rightarrow +\infty$) towards an equilibrium state, even though the initial state did not develop from an equilibrium state in the remote past ($t \rightarrow -\infty$). Thus time-reversal symmetry is broken because $t = \pm\infty$ play different, asymmetrical roles. A further consequence of the non-unitary time evolution is a severe limitation on the validity of the concept of the *wavefunction*; even if we start with a pure state, $\hat{\rho}_0$, it will be eventually transformed by an irreversible process into a mixed state [50]. Hence it is essential to describe the states of the system using density matrices and the Liouville operator, $\hat{L}$, (the operation of commutation with the Hamiltonian $\hat{H}$) in place of the rather more familiar theory based on wavefunctions and the Schrödinger equation.

The general analysis of irreversibility made by the Brussels school is of a rather formal nature and not all their conclusions can be given rigorous proofs for physically interesting quantum field theories e.g. ones with electrostatic or electrodynamical interactions. It is as well to recognize that there are considerable mathematical difficulties associated with quantum field theory such that one cannot yet achieve the degree of mathematical rigour that is possible for the quantum mechanics of systems with a finite number of degrees of freedom. That aside, the essential feature is that the occurrence of chemical equilibrium when substances are brought together in a chemical reaction vessel must be associated with a Hamiltonian operator possessing a continuous spectrum. If, with Primas, we abstract from the universal EPR correlations [37, 38], it seems that the smallest 'system' we can usefully separate from the 'environment' for a description of evolution towards equilibrium is the contents of the reaction vessel *i.e.* the reacting substances; in other words, the Hamiltonian operator referred to in the Prigogine theory should be based on a quantum field theory describing chemical substances.

A similar interpretation is possible for a far-from-equilibrium system; consider something like the Zhabotinski—Belousov reaction which may lead to striking spatio-temporal colour oscillations between red and blue due to redox processes involving cerium sulphate. If we imagined this as proceeding through elementary reactions involving chaotic collisions of molecular ions going on at random we

could only expect a uniform colour with possible occasional flashes of red and blue due to fluctuations. However under far-from-equilibrium conditions which lead to periodic oscillations, in space and time, of red to blue, the system is *acting as a whole* — its behaviour is coherent over macroscopic distances and times, and is not explicable purely in terms of short-range intermolecular forces involving molecular ions and their near neighbours [48]. Actually the chemical instabilities in reactions of the Zhabotinski—Belousov type can be characterized as examples of the phenomenon of 'broken symmetry'; not surprisingly such typical quantum field theory notions as 'Goldstone mode' and 'order parameter' have been utilized [51] in their description. On the other hand it seems possible to identify the quasi-particles of the quantum field theory with our usual notions of atoms and molecules; this idea [39] will be described in more detail in Section 4.2. In the next section I shall discuss the question of molecular structure starting from the purely molecular theory.

3.3. MOLECULAR QUANTUM MECHANICS

The notion of an *isolated molecule* has the same sort of status in molecular quantum mechanics as that of a body undisturbed by external forces introduced in the foundations of mechanics by Galileo [4]. Such entities are possible, rather than actual, and in a sense not even possible, because no such completely isolated object exists in the natural world. Nevertheless, without these quite unreal conceptions, Galileo could scarcely have developed his new science of mechanics — so it is here, except that the notion of 'isolation' also requires the deliberate neglect of the non-dynamical EPR correlations between the molecule and its environment as well as the neglect of interactions.

On this basis we have the following characterization of molecular quantum mechanics [18, 52]:

(i) Atoms and molecules are bound aggregates of electrons and nuclei;
(ii) The interactions of electrons and nuclei are governed by the Schrödinger equation for such aggregates;
(iii) The Hamiltonian, $\hat{H}_{\mathrm{MOL}}$, for any given molecule can be obtained directly by reference to the molecular formula.

To realize (iii) in a specific application we assume purely electrostatic interactions between electrons and nuclei together with Galilean kinematics, and we then have enough information to fix $\hat{H}_{\mathrm{MOL}}$ unambiguously. In a space-fixed reference frame $\hat{H}_{\mathrm{MOL}}$ may be written as a sum of kinetic and Coulombic potential energy terms for all the particles,

$$\begin{aligned} \hat{H}_{\mathrm{MOL}} &= \sum \hat{T}_a + \sum \hat{V}_{ab} \\ \hat{T}_a &= \hat{p}_a^2/2M_a \quad \text{per particle} \\ \hat{V}_{ab} &= \frac{Z_a Z_b}{4\pi\varepsilon_0 |\hat{\mathbf{q}}_a - \hat{\mathbf{q}}_b|} \quad \text{per pair } (a, b) \text{ of particles.} \end{aligned} \tag{17}$$

This is all very straightforward and familiar — we are after all dealing with the notional starting point of quantum chemistry — but it leads at once to a paradox. The full, spin-free, non-relativistic molecular Hamiltonian, $\hat{H}_{\mathrm{MOL}}$, describes the interactions of an assembly of electrons and nuclei; *it does not describe a particular molecular species* [53]. As an example, the chemical formula C_8H_8 is associated with 8 carbon nuclei, 8 protons and 56 electrons; the Hamiltonian $\hat{H}_{\mathrm{MOL}}$ for this collection of charges is shared by cubane, cyclooctatetraene, vinylbenzene and any other system where the given number of electrons and nuclei are interacting. Nevertheless once $\hat{H}_{\mathrm{MOL}}$ is given, some development of an *ab initio* theory of its eigenvalues and eigenfunctions is possible, and it is of interest to review these general results.

Firstly, there are the symmetries of the molecular Hamiltonian which are essentially the same as those of an atomic Hamiltonian; thus $\hat{H}_{\mathrm{MOL}}$ is invariant under continuous translation and rotations (so giving stationary state quantum numbers $\mathbf{k}$ and J, M respectively), space and time inversions (parity, time reversal), and permutations of identical particles (Bose, Fermi statistics) [18]. It is conventional to make the translational and rotational symmetry explicit by choosing three coordinates for the molecular centre-of-mass $\mathbf{X}_{\mathrm{C.M.}}$ and three coordinates as Euler angles, Ω. This is a transformation to a new representation with overall translations and rotations referred to the molecular centre-of-mass,

$$\begin{aligned} \hat{H}'_{\mathrm{MOL}} &= \hat{\Lambda}\hat{H}_{\mathrm{MOL}}\hat{\Lambda}^{+} \\ &= \hat{P}^2_{\mathrm{C.M.}}/2M + \hat{T}' + \hat{V} \end{aligned} \tag{18}$$

where $\hat{T}'$ depends on how the Euler angles are introduced, $\hat{\mathbf{P}}_{\mathrm{C.M.}}$ is the momentum operator conjugate to the centre-of-mass coordinate $\hat{\mathbf{X}}_{\mathrm{C.M.}}$, and M is the total molecular mass. The remaining coordinates are referred to as *internal* coordinates; since the Coulomb interaction only depends on the distance between the particles it may be written at once in terms of the internal coordinates; the main concern therefore is in the treatment of the kinetic energy, $\hat{T}'$.

We can always find transformation operators $\hat{\Lambda}$ that preserve the inversion and permutation symmetries and are unitary ($\hat{\Lambda}^{+} = \hat{\Lambda}^{-1}$); then the discrete energy eigenvalues of $\hat{H}_{\mathrm{MOL}}$ and $\hat{H}'_{\mathrm{MOL}}$ calculated for zero total linear momentum are the same. The eigenfunctions $\{\psi_a\}$ of $\hat{T}' + \hat{V}$ are the wavefunctions for the molecule with its overall translational motion separated out — these are the *spectroscopic states* of the molecule. Arguments along these lines are well-known in atomic physics and in the physics of light nuclei.

The spectral properties of $\hat{H}_{\mathrm{MOL}}$ can be characterized in general as follows: consider again $\hat{H}_{\mathrm{MOL}}$ for the cubane family. In the rest-frame of the system there is a continuous spectrum lying above the set of bound stationary states of the cluster (C_8H_8). In the complex energy plane the continuous spectrum is a branch cut on the real E axis starting at some E_0 and extending to $+\infty$; it is rich in structure having numerous sets of other bound states embedded in it. These bound states correspond to the states of *all* the possible clusters one can make using some, but not necessarily all, of eight carbon nuclei, eight protons and fifty-six electrons, not

just those of the form C_nH_n, $n < 8$ (the union of any choice of cluster decomposition must of course add up to 'C_8H_8' because the system is closed). The fragment bound states give a discrete spectrum in the rest-frame of their own fragment. As the energy increases the clusters diminish in complexity (mass) until eventually we reach a completely dissociated plasma of electrons and nuclei. A new cluster fragment does not appear until the energy has exceeded a characteristic threshold value. These spectral features can be made explicit using the complex coordinate rotation techniques of Combes *et al.* [54—56] applied to the N-body Hamiltonian $\hat{H}_{\mathrm{MOL}}$. The above description of the spectrum is evidently what one would expect intuitively from the facts of mass spectrometry.

The stationary state wavefunctions of $\hat{H}_{\mathrm{MOL}}$ form an orthonormal basis for a Hilbert space $\mathfrak{H}$ which, for our example, we may label as $L^2(C_8H_8)$. This Hilbert space may be constructed as a tensor product of any independent set of fragment Hilbert spaces provided only that all the particles of 'C_8H_8' are accounted for. Irrespective of the choice of fragments (electrons and nuclei, ions and electrons, neutral atoms and/or molecular fragments of C_8H_8 etc.) the resulting constructions are all *isomorphic* representations of $L^2(C_8H_8)$ and differ at most by unitary transformations. $L^2(C_8H_8)$ is a coherent Hilbert space in which every wavefunction corresponds to a physically possible excitation of the system [33].

The stationary states $\{\psi_\alpha\}$ of the molecular Hamiltonian $\hat{H}_{\mathrm{MOL}}$ have formal properties closely analogous to the stationary states of atoms. For example, *time* has no role for these states — there is no possibility of change and no true distinction between past and future. The molecular stationary states describe eternal entities, lacking any extension in time (*cf.* Section 3.2). Even worse, these molecules have no size, no extension in ordinary space and consequently there can be no contact with the ordinary chemical notions of molecular structure [18]. This remark can only be understood if it is appreciated that there are two quite distinct notions of structure which must be distinguished. In the words of Pauli,

> no description of the motion of the electron in an atom in space and time is given . . . this is clear from the fact that outside the domain of validity of geometrical optics i.e. the semiclassical limit of quantum theory it is impossible to construct 'rays' . . . that can be considered as orbits of particles [57].

Pauli's insight is every bit as applicable to nuclear motion associated with the molecular eigenstates as to electrons in atomic stationary states. The electrons and nuclei are of course strongly correlated because of their Coulomb interactions and this correlation can be extracted from the spectroscopic states $\{\psi_\alpha\}$ by calculation of appropriate reduced density matrices; this 'quantum structure' must not be confused with the traditional picture of a molecule as a semi-rigid framework of atoms connected by bonds that rotates and translates in ordinary space as time elapses, what Claverie and Diner [58] refer to as 'classical structure'. This *classical* molecular structure cannot be associated with a spectroscopic state of $\hat{H}_{\mathrm{MOL}}$, and seems to have nothing to do with quantum theory *unless* we make the transition to the classical limit [43, 52].

It is now appropriate to summarize the discussion so far. The most important result is that from the point of view of molecular quantum mechanics there does not appear to be any room to assign a fundamental significance to molecular structure [18, 43, 53]; this conclusion is consistent with a recent remark by Löwdin [35], who reminds us that we have no *ab initio* quantum theory of molecular structure. In my view the real motivation for the conventional methods of quantum chemistry arises from a powerful 'felt need' by chemists to make contact with the classical idea of *molecular structure* [43]; this has led to the use of singular 'approximations', for example the Born—Oppenheimer approximation, and modifications of $\hat{H}_{\mathrm{MOL}}$, which amount to qualitative changes in the theoretical formalism. We come to the conclusion then that molecular quantum mechanics, in the sense defined here, is not congruent with quantum chemistry, and reiterate that this fact does not imply that one approach is 'right' and the other 'wrong'. As emphasized by Primas, the logical structure of quantum theory invites the coexistence of incompatible theories referring to the same phenomena [37].

Qualitative changes in the formalism can greatly affect the conceptual structure of the theory without causing large quantitative changes. We might imagine starting with a covariant molecular theory that displays the symmetries of the Lorentz group (no matter that we scarcely know where to begin with such a theory); the limit $c \to \infty$ for the Lorentz group is a *singular* limit that leads to a Galilean invariant theory which, in a molecular theory, must be based on $\hat{H}_{\mathrm{MOL}}$. Finally we might consider the Born—Oppenheimer limit $\kappa \to 0$ corresponding to infinite nuclear mass. Now it is often the case that adiabatic Born—Oppenheimer calculations give energy eigenvalues that are close to the results of a fully non-adiabatic theory; for example, this has been demonstrated for the neutral hydrogen molecule and its cation, including the isotopic variants. We also think that 'relativistic corrections' to these energy levels are likely to be small, so one would expect a covariant molecular theory to yield energy eigenvalues close to those obtained in the non-adiabatic theory based on $\hat{H}_{\mathrm{MOL}}$ (and often its adiabatic approximation). Thus as far as energy levels are concerned we conclude that the non-adiabatic treatment and its relativistic extensions give only small improvements to the traditional quantum chemical approach. But from the conceptual point of view the changes are dramatic; in a model in which the nuclei are treated classically one can describe a classical molecular structure. In the non-adiabatic theory based on $\hat{H}_{\mathrm{MOL}}$ there is no classical molecular structure although, as noted above, the reduced density matrices display the correlations we call 'quantum structure'. Even this is eliminated in relativistic quantum mechanics because *we cannot define* a position representative wavefunction $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ and *a fortiori* there is no reference to *any* kind of structure.

If we forego the usual calculational procedures of quantum chemistry and try to work directly with the isolated molecule model based on $\hat{H}_{\mathrm{MOL}}$, we can still make progress, albeit more limited than with the traditional theory. We are, after all, at an early stage and as yet there are few experimental results that might prompt the quantum chemistry community to look to new methods. Nevertheless, the model

based on $\hat{H}_{\mathrm{MOL}}$ is a fully specified quantum theory and it is of interest to explore its consequences; my attitude is that its failures can be just as instructive as any of its successes.

One area where working directly with $\hat{H}_{\mathrm{MOL}}$ has already produced useful results is theoretical molecular spectroscopy. In conventional terms, working directly with $\hat{H}_{\mathrm{MOL}}$ means that we are interested in a *non-adiabatic* approach that does not involve the Born—Oppenheimer and/or Born adiabatic approximations. One such method that has been brought into molecular quantum mechanics from theoretical nuclear spectroscopy is the Generator Coordinate Method (GCM). The GCM has been described at length by Lathouwers and Van Leuven [59, 60], and only a brief account will be given here.

The GCM is a novel approach for chemical physics not least because it is based on *space-fixed* axes — there are no troublesome transformations to 'body-fixed coordinates' in the manner of Eckart (see below) — and it makes use of molecular structure only in the intermediate steps of the construction of an approximate molecular wavefunction for use in a variational calculation of the energy eigenvalues of the full molecular Hamiltonian $\hat{H}_{\mathrm{MOL}}$. This is in contrast to more traditional methods in molecular spectroscopy that are aimed at finding exact stationary states of *modified* molecular Hamiltonians. In the GCM the usual spectroscopic constants for term formulae emerge as matrix elements of well-defined operators.

Let r and R stand for all electronic and nuclear coordinates respectively. In the GCM for molecules, a trial wavefunction is written in the form

$$\Psi(r, R)_{\mathrm{GCM}} = \int \mathrm{d}\alpha \, f(\alpha)\phi(r, \alpha)\chi(R, \alpha) \tag{19}$$

where α stands for a set of n integration variables called *generator coordinates.* In the approach described by Lathouwers *et al.* [59, 60] for molecules containing N nuclei, n is equal to the number of nuclear degrees of freedom, $3N$, and the α are real scalars. Provided the generator coordinates are real, any specific set of α-parameters can be interpreted as describing a structure in the $3N$-dimensional generator coordinate space. These are conventional assumptions that help to define a molecular model; for example, provided that the value of the integral is not altered the α may just as well be chosen as *complex* numbers, so a structural interpretation is not a necessary consequence of working with the GCM [61].†

For the integrand of (19) $\phi(r, \alpha)$ has been taken as an electronic wavefunction

† One should also note that provided the integral (19) exists the value of n is arbitrary; for example in principle one can associate generator coordinates with some or all of the *electronic* degrees of freedom and further enhance the flexibility of the wavefunction (19). In practice these choices are guided by convenience and physical intuition with the aim of developing a useful practical variational method. Thus the molecular GCM does not lead to a picture of molecules; rather it utilizes an *a priori* conception of the dynamics of molecules.

for the fixed-nucleus Hamiltonian for the (real) configuration α, while the nuclear functions $\chi(R, \alpha)$ are assumed to be sharply peaked functions about $R = \alpha$; the functions $f(\alpha)$ are the analogues of the expansion coefficients in LCAO type theories, and must be determined through a Rayleigh—Ritz variational calculation based on $\hat{H}_{\mathrm{MOL}}$ and the trial wavefunction Ψ. The overall translational and rotational symmetry of the isolated molecule can be incorporated in the trial wavefunction by appropriate choices of the $f(\alpha)$ which can be factorized in the form

$$f = f_{\mathrm{trans}} \cdot f_{\mathrm{rot}} \cdot f_{\mathrm{internal}}. \tag{20}$$

As explained by Lathouwers *et al.* [60] by making certain further assumptions about the trial wavefunction Ψ the variational calculation leads directly to the Dunham series for the molecular energy levels of diatomic species correct to terms in κ^8 where $\kappa = (m_e/M_N)^{1/4}$ is the usual Born—Oppenheimer expansion parameter; for a semi-rigid polyatomic molecule, a similar calculation leads to the usual spectroscopic term formula,

$$E_{\mathrm{MOL}} \approx E_{\mathrm{TRANS}} + E_{\mathrm{EL}} + E_{\mathrm{VIB}} + E_{\mathrm{ROT}} + \text{'small terms'}. \tag{21}$$

Given the resolution and accuracy of modern laser spectroscopy, it can be claimed that the GCM derivation of molecular term formulae is a significant improvement over the traditional adiabatic approximation; for example the adiabatic approximation yields a Dunham series correct [60] only to terms in κ^6. From the conceptual point of view, we see that the separation of molecular energies (21) that characterizes much of molecular spectroscopy neither implies the separability of the molecular wavefunction (Ψ_{GCM} is not a simple product of electronic and nuclear terms) nor involves the use of a potential energy surface. *A potential energy surface cannot be given the status of a fundamental molecular property within molecular quantum mechanics*; it is an artefact of the quantum versions of the traditional molecular models. Thus the GCM provides a formal quantum mechanical derivation of molecular term formulae based on space-fixed axes and without essential reference to classical molecular structure [52, 60, 61].

Another topic where molecular quantum mechanics gives a satisfactory interpretation of experimental results that is different from the traditional view of quantum chemistry is the dielectric properties of simple gases. A clear-cut distinction between *polar* and *non-polar* substances can be established *without* recourse to 'permanent dipole moments'; this is done by direct calculation of the temperature dependence of the bulk electric susceptibility [18].

The classical calculation of Langevin and Debye [62] shows that the average scalar electric susceptibility, χ, of a gas at temperature T can be written in the form

$$\begin{aligned} \chi(T) &= \mathrm{Tr}(\boldsymbol{\chi}) \\ &= a + b/T \end{aligned} \tag{22}$$

where a and b are constants characteristic of the gas, and $\boldsymbol{\chi}$ is the susceptibility tensor. For *non-polar* molecules b is zero and is essentially temperature independent; for *polar* molecules, $b \neq 0$ and χ is inversely proportional to the temperature.

The quantum mechanical result for the electric susceptibility of a gas is due to Van Vleck [63]. Let $\hat{d}$ be the electric dipole moment operator for a molecule; then the *electric polarizability* tensor for the spectroscopic state ψ_α is

$$\boldsymbol{\alpha}(\alpha) = \sum_\beta \frac{\langle \psi_\alpha | \hat{\mathbf{d}} | \psi_\beta \rangle \langle \psi_\beta | \hat{\mathbf{d}} | \psi_\alpha \rangle}{E_\alpha - E_\beta} \tag{23}$$

and its equilibrium thermodynamic average is the molar electric susceptibility tensor, $\boldsymbol{\chi}(T)$,

$$\varepsilon_0 \boldsymbol{\chi}(T) = N_\mathrm{A} \frac{\sum_\alpha \boldsymbol{\alpha}(\alpha) \exp(-E_\alpha / k_B T)}{\sum_\alpha \exp(-E_\alpha / k_B T)} \tag{24}$$

where E_α and E_β are the energies of the spectroscopic states ψ_α, ψ_β, and N_A is Avogadro's constant.

According to Van Vleck, the Langevin—Debye formula can be recovered from these equations provided that we can distinguish two classes of spectroscopic states $\{\psi_\alpha\}$ which are identified by two inequalities involving the thermal energy $k_B T$. Van Vleck's analysis leads to the following conclusions:

(i) *Polar* molecules are those for which there are states $\{\psi_\beta\}$ with

$$\langle \psi_\alpha | \hat{\mathbf{d}} | \psi_\beta \rangle \neq 0 \quad \text{when} \quad |E_\alpha - E_\beta| \ll k_B T; \tag{25a}$$

(ii) *Non-polar* molecules are those for which

$$\langle \psi_\alpha | \hat{\mathbf{d}} | \psi_\beta \rangle \neq 0 \quad \text{when} \quad |E_\alpha - E_\beta| \gg k_B T. \tag{25b}$$

Notice that with respect to ordinary ambient temperatures the conditions in (i) imply that the molecule has fully allowed electric dipole transitions in the microwave region of the electromagnetic spectrum, whereas molecules that satisfy (ii) have fully allowed transitions at higher frequencies in the infrared, or visible/U.V. region, in agreement with the experimental results.

The constant 'b' in the Langevin—Debye formula is related to an expectation value

$$b \sim \langle \psi_\alpha | \hat{\mathbf{d}} \cdot \hat{\mathbf{d}} | \psi_\alpha \rangle \tag{26}$$

averaged over the rotational quantum numbers. There is no 'dipole moment' as such because $\hat{\mathbf{d}}$ changes sign under space-inversion and $|\psi_\alpha|^2$ does not, so [18]

$$\langle \psi_\alpha | \hat{\mathbf{d}} | \psi_\alpha \rangle = 0. \tag{27}$$

The selection rules governing matrix elements of the dipole operator for *diatomic*

molecules are largely fixed by the permutation symmetry of the nuclei and Van Vleck's theory can be fully worked out. The general result is that we expect:

(i) *homonuclear* diatomic molecules to be *non-polar*, while
(ii) *heteronuclear* diatomic molecules should be *polar*,

as is the case of course. This classification also covers isotopically distinct species, so that one expects e.g. isotopically mixed hydrogen, HD, to be weakly 'polar' because the two nuclei are different; this quantum mechanical prediction is in agreement with electric deflexion experiments on HD using a thermal molecular beam. Analogous experimental results have been found for other isotopically substituted species such as HC_2D; the effect is small essentially because the terms that break the permutation symmetry are of order κ^4 (κ is the usual Born—Oppenheimer parameter).

Moss and Sadler [64] have recently shown that the effects of isotopic substitution may be made explicit in an effective Hamiltonian. They consider the ion HD^+ and construct a unitary transformation of the Hamiltonian,

$$\begin{aligned} \hat{H}' &= \exp(i\hat{S})\hat{H}\exp(-i\hat{S}) \\ &= \hat{H} + i[\hat{S}, \hat{H}] + \ldots \end{aligned} \tag{28}$$

where $\hat{S}$ is chosen so that its commutator with $\hat{H}$ cancels precisely the non-adiabatic coupling term $-\frac{1}{2}\mu_a^{-1}\nabla_{\mathbf{r}} \cdot \nabla_{\mathbf{R}}$ in the original Hamiltonian. The leading terms of $\hat{H}'$ then take the form of a Born adiabatic Hamiltonian in which the two nuclei have *different effective nuclear charges*, $Z \pm \frac{1}{4}\mu_a^{-1}$, where $\mu_a = (1/M_H) - (1/M_D)$; the new operator $\hat{H}'$ provides a simple route to the estimation of 'non-adiabatic' corrections for high-lying vibration-rotation levels of HD^+. It seems reasonable to suppose that an analogous transformation can be made for neutral HD, and perhaps also for the general diatomic case.

This quantum mechanical calculation removes a difficulty found in the classical interpretation of polar molecules; although the nuclei in HD are different, they preserve *charge symmetry* and so would not be expected to lead to any 'polarity' in the HD bond. As we have just seen, explicit consideration of the electron-nuclear coupling shows that the molecule behaves *as though* the nuclei had different charges. Van Vleck's theory has not been used for polyatomic species; indeed apart from his own classic text of 1932 [63], the beautiful calculation summarized in equations (23)—(27) has practically vanished from quantum chemistry texts (the analogous calculation for the *magnetic* susceptibility is much better known). The conventional account of polarity depends on the use of *body-fixed* coordinates so as to make contact with the notion of classical molecular structure and the classical interpretation of Langevin and Debye. In such coordinates a molecule at rest may have a permanent non-zero 'dipole moment' which of course has zero rotational average in accordance with (27). Thus HD is conventionally said to have a dipole moment $\mu \approx 10^{-2}$ Debye; in Van Vleck's quantal approach this is the value of $\sqrt{b}$,

equation (26). As we shall see shortly, body-fixed coordinates in the polyatomic case are problematic in quantum theory.

These two examples of quantum theories based on $\hat{H}_{\mathrm{MOL}}$ in space-fixed axes are to be contrasted with traditional quantum chemistry and theoretical molecular spectroscopy based on body-fixed coordinates (Eckart theory) and the Born—Oppenheimer approximation. While the Born—Oppenheimer approximation has for many years provided (and of course still provides) a powerful basis for molecular calculations, we should recognize that it is more than a mere calculational device. The fixed nucleus approximation which leads to the characteristic product (separable) wavefunction form

$$\Psi_{\mathrm{BO}} = \psi(r)_{\mathrm{electronic}} \cdot \phi(R)_{\mathrm{nuclear}} \tag{29}$$

is achieved by taking a *singular* (asymptotic) mathematical limit ($\kappa \to 0$) [65] and this gives a qualitatively *new* theory [34].

A wavefunction written as a product of electronic and nuclear functions *i.e.* a pure state description for the electrons and for the nuclei considered as separate subsystems, has no EPR correlations between the two groups of particles; such a model can support some classical features including a well-defined molecular structure related purely to the distribution of nuclei. The states of molecular quantum mechanics, for example Ψ_{GCM}, equation (19) cannot in general be decomposed into pure states for electrons and nuclei. Now, I am not suggesting that one or other of these approaches is 'wrong'; what is emphasized here is that they are *different* and that quantum theory based on $\hat{H}_{\mathrm{MOL}}$ does not pass smoothly into quantum chemistry. We are free to choose the traditional Born—Oppenheimer approach, and if that is productive well and good; however when we do so we renounce any claim to have an *ab initio* quantum mechanical theory of molecular structure starting from $\hat{H}_{\mathrm{MOL}}$.

The quantum mechanics of polyatomic molecules based on $\hat{H}_{\mathrm{MOL}}$ is much less developed than the corresponding theory for diatomic species where few surprises can be expected. Whether there is anything useful to be gained by pursueing a direct approach based on $\hat{H}_{\mathrm{MOL}}$ for polyatomic species is an open question; implicit in this remark is a recognition that the usual theory of a vibrating-rotating molecule based on the Eckart Hamiltonian is a *model* distinct from the isolated molecule model based on $\hat{H}_{\mathrm{MOL}}$.

The Eckart theory [66, 67] takes as its starting point the classical picture of a molecule as a quasi-rigid body translating and rotating about its centre-of-mass but with internal degrees of freedom as well (vibrations). The first thing that is done is a transformation of $\hat{H}_{\mathrm{MOL}}$ to a set of rotating coordinates referred to the centre-of-mass; this is always possible with the aid of an orthogonal matrix $\mathbf{C}$ that depends on a set of three Euler angles, Ω:

$$\begin{aligned} \hat{H}' &= \hat{\Lambda}_E(\mathbf{C})\hat{H}_{\mathrm{MOL}}\hat{\Lambda}_E(\mathbf{C})^+ \\ &= \hat{P}^2_{\mathrm{C.M.}}/2M + \hat{h}_{\mathrm{Eckart}}. \end{aligned} \tag{30}$$

The problem then is to decide how to choose the matrix **C**, and hence $\hat{\Lambda}_E$, so as to achieve the stated aim of Eckart's programme. Eckart's solution is based on the introduction of a *reference nuclear framework*; the definition of the matrix **C** depends only on the nuclear coordinates, so we need not make explicit reference to the electronic coordinates. Here's how to do it.

Any nuclear position vector $\mathbf{R}_i$ $(i = 1, \ldots, N)$ can be written

$$\mathbf{R}_i = \overline{\mathbf{X}} + (\mathbf{C}\mathbf{r}_i) \tag{31}$$

where $\overline{\mathbf{X}}$ is the centre of nuclear mass vector. The vectors $\{\mathbf{r}_i\}$ are related to those of the reference nuclear configuration $\{\mathbf{a}_i\}$ by

$$\mathbf{r}_i = \mathbf{a}_i + \sum_k l_{ik}\mathbf{q}_k, \qquad k = 1, \ldots, (N-2) \tag{32}$$

where the $(N-2)$ vectors $\{\mathbf{q}_k\}$ are internal coordinates for vibrations (i.e. $3N-6$ degrees of freedom); the vectors $\{\mathbf{r}_i\}$ and the coefficients $\{l_{ik}\}$ have to satisfy a set of so-called *Eckart conditions* to ensure consistency.

Eckart assumed that the $\{\mathbf{a}_i\}$ should refer to the *classical equilibrium structure* of the molecule; however these equations apply to (almost) any nuclear configuration, and suffice to determine the matrix **C** as follows [67]. Define a new matrix **F** by

$$\mathbf{F} = \sum_i m_i(\mathbf{R}_i - \overline{\mathbf{X}})\mathbf{a}_i \tag{33}$$

where the $\{m_i\}$ are the nuclear masses. Then the equation for the transformation matrix **C** is

$$\mathbf{C}^T\mathbf{F} - \mathbf{F}\mathbf{C}^T = 0 \tag{34}$$

i.e. $\mathbf{C}^T$ must commute with **F**. Therefore by polar decomposition we have

$$\mathbf{C} = \mathbf{F}(\mathbf{F}^T\mathbf{F})^{1/2} \tag{35}$$

so that **C** *is well-defined provided* **F** *is positive definite* [67].

This is all standard Eckart theory; in historical terms the importance of these results for molecular spectroscopy can hardly be overemphasized; however for a characterization of the relationship between quantum chemistry and the molecular Hamiltonian $\hat{H}_{\mathrm{MOL}}$ it is essential to recognize that the Eckart approach has several difficulties. Classically, there are singular configurations of the nuclei for which the condition $\mathbf{F} > 0$ is not valid, so that **C** is not defined everywhere in the nuclear configuration space. One must therefore introduce *ad hoc* constraints to keep the molecular motions away from the singular configurations. In quantum mechanics this difficulty takes a different form.

The transformation to the Eckart representation (i.e. the introduction of a fixed reference framework) is known from early work of Hirschfelder and Wigner [68] to introduce singularities in the transformed kinetic energy operator, $\hat{T}'$, that behave as q^{-2}. In numerical calculations based on the body-fixed Hamiltonian,

a judicious choice of a finite basis of expansion functions can make these singularities harmless; however from the theoretical point of view their presence means that $\hat{h}_{\text{Eckart}}$ is no longer strictly self-adjoint in the sense of Kato [69, 70]. Finally, Eckart's method does not give a set of Euler angles Ω that are invariant under permutations of like particles; permutations among identical nuclei can lead to new and generally different Eckart frames [67]. These facts suggest that $\hat{H}'$ is not a unitary transform of $\hat{H}_{\text{MOL}}$ i.e. that the transformation to body-fixed coordinates induced by $\hat{\Lambda}_E(\mathbf{C})$, which is intimately related to the concept of molecular structure, is a non-unitary transformation. It is interesting to note that Prigogine's theory of irreversibility also requires the introduction of a non-unitary transformation and it may be conjectured that the quasi-particles with finite lifetimes implied by his theory should be identified with molecules with structures.

In the example just considered we found that the molecular Hamiltonian $\hat{H}_{\text{MOL}}$ was modified in the course of the development of the conventional Eckart theory, but there was no suggestion that a direct calculation based on $\hat{H}_{\text{MOL}}$ would lead to any major differences for the low-energy part of the spectrum which has been the main interest of spectroscopists up to now. There would of course be differences in interpretation due to the use of, or absence of, potential energy surfaces and/or body-fixed coordinates. To conclude this section however I would like to describe a case where there is a striking failure of the molecular theory based on $\hat{H}_{\text{MOL}}$ — the example of *natural optical activity* [39]; there are some similarities with the discussion of irreversibility in Section 3.2.

To investigate natural optical activity, we use *plane polarized* light as a probe, and examine the polarization state of light scattered in the forward direction (see Figure 1). We are looking for a rotation, θ, in the plane of polarization. The empirical fact of natural optical activity requires the state of an optically active medium to be lacking in space-inversion symmetry; in the language of condensed matter physics the rotation angle, θ, is the order parameter associated with the broken space-inversion symmetry. Recall that the rotation angle is a pseudoscalar (odd-parity) observable, i.e. it changes sign under inversion of coordinates just as the sign of the pitch of a helix depends on the handedness of the helix.

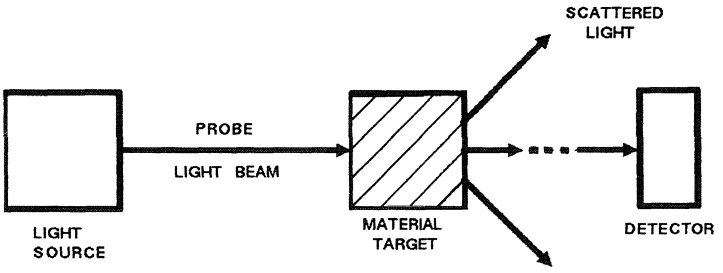


Fig. 1. Schematic light scattering experiment. To investigate natural optical activity, we use *plane polarized light* as probe, and examine the polarization state of light scattered in the forward direction. We are looking for a rotation, θ, in the plane of polarization.

Achiral states of matter are described by density matrices, $\hat{\rho}$, with the property

$$\hat{\rho}' = \hat{P}\hat{\rho}P^{+} \equiv \hat{\rho} \tag{36}$$

where $\hat{P}$ is the space-inversion (parity) operator, i.e.

$$[\hat{\rho}, \hat{P}] = 0. \qquad \textit{Achiral system} \tag{37}$$

Since any measurement is related to the mean value of a suitable operator,

$$\langle A \rangle = \mathrm{Tr}[\hat{\rho}\hat{A}] \tag{38}$$

we see that *achiral* matter cannot have odd-parity observables; in particular $\theta = 0$. In the particular case of natural optical activity, the operator $\hat{A}$ is chosen as the *electric polarization field* $\hat{\mathbf{P}}(\mathbf{x}, t)$ for the material target and $\hat{\rho}$ refers to the density matrix for the combined system of polarized light beam and the material target (see Figure 1). Then the radiation arriving at the detector can be related to the expectation value

$$\langle \mathbf{P}(\mathbf{x}, t) \rangle = \mathrm{Tr}[\hat{\rho}\hat{\mathbf{P}}(\mathbf{x}, t)] \tag{39}$$

where the trace is taken over the degrees of freedom of the target only. A perturbation theory evaluation of this mean value yields the rotation angle, θ, expressed in terms of matrix elements involving the target states [39].

To fix the ideas, I shall only consider the optical activity of an ordinary fluid, an isotropic and homogeneous medium; an optically active crystal such as quartz is a different case because it is solid rather than fluid. The conclusion of Van't Hoff [10], like Pasteur [71], was simply that a macroscopic optically active fluid medium is composed of a very large number of elementary objects (molecules) which must themselves carry the space-inversion dissymmetry. Suppose however that we try to elucidate the origin of natural optical activity and, at the same time, take seriously the formalism of quantum theory. Then we must recognize that the classical account of enantiomerism contains some deep-lying assumptions about the molecular nature of ordinary matter that need to be reconsidered in the quantum mechanical account. Van't Hoff's conviction [10] that optically active substances were made up of molecules with dissymmetric structures has no obvious place in the quantum theory. There are two assumptions here:

(i) molecules have structures in physical space that can be dissymmetric; we have not found classical molecular structure in molecular quantum mechanics so the loss of space-inversion symmetry may well have to be accounted for in a much more abstract fashion.

Prior to (i) however there is another assumption, namely

(ii) the universal principle that substances are composed of atoms and molecules. The notion of a molecule is a *possible* rather than a necessary feature of a quantum mechanical account of chemical phenomena; for example, Prigogine's

quantum theory of irreversible evolution to equilibrium [46—50] seems to call into question the conventional molecular picture of substances, and the traditional ideas about molecules may need to be reconsidered.

Suppose we accept initially that the usual notions of 'atom' and 'molecule' (collections of electrons and nuclei) make an appropriate starting point for a chemical theory (the claim of quantum chemistry as well as classical chemistry), and base our theory of optical activity [39] on the molecular Hamiltonian, $\hat{H}_{\mathrm{MOL}}$. Then since the interactions between the charges are Coulombic we have $[\hat{H}_{\mathrm{MOL}}, \hat{P}] = 0$, where as before, $\hat{P}$ is the space-inversion operator.

Let $\hat{\rho}_0$ be any initial density matrix for the molecule at an arbitrary time $t = 0$, subject only to the constraint $[\hat{\rho}_0, \hat{P}] = 0$. Then at any later time t, the density matrix is given by

$$\hat{\rho}(t) = \exp(-i\hat{H}_{\mathrm{MOL}}t)\hat{\rho}_0 \exp(+i\hat{H}_{\mathrm{MOL}}t). \tag{40}$$

This is a *unitary* transformation and accordingly $\hat{\rho}(t)$ commutes with the space-inversion operator $\hat{P}$ if $\hat{\rho}_0$ does. Setting $[\hat{\rho}_0, \hat{P}] = 0$ does not beg the question; the aim is to elucidate the predictions of a fully specified molecular model. The correct conclusion is that chiral states are outside the scope of this model which is therefore of no use for the problem at hand. It is vital not to confuse the desired answer (agreement with empirical data) with what can be *demonstrated* within a given, fully specified model.

We can also consider the intermolecular interactions $\hat{V}$ between n molecules, originally isolated. Then the system Hamiltonian is

$$\hat{H}_n = \sum_i^n \hat{H}_{\mathrm{MOL}}^{(i)} + \hat{V} \tag{41}$$

and $[\hat{H}_n, \hat{P}] = 0$ provided we take, as usual, $[\hat{V}, \hat{P}] = 0$. Again, any initial state $\hat{\rho}_0$ must satisfy $[\hat{\rho}_0, \hat{P}] = 0$ because there is *nothing in the model* to lead us to assume otherwise. The time evolution of the density matrix under $\hat{H}_n$ is unitary provided $n < \infty$, so we again come to the conclusion that $\theta = 0$ if $\hat{\rho}_0$ commutes with $\hat{P}$.

This discussion of natural optical activity does not sensibly break down when the perturbation due to weak interactions between electrons and nuclei are considered. It is true that if we put

$$\hat{H} = \hat{H}_{\mathrm{MOL}} + \hat{V}_{\mathrm{WNC}} \tag{42}$$

where $\hat{V}_{\mathrm{WNC}}$ is the operator for neutral current interactions, then $[\hat{H}, \hat{P}] \neq 0$; however we know that $\hat{V}_{\mathrm{WNC}}$ produces optical rotation in all matter that is orders of magnitude too small to account for the phenomena of interest here. For example, according to Harris [72], the optical rotation in molar thallium vapour due to the weak interaction effects is of the order of 10^{-8} degrees cm^{-1}. Hence $\hat{V}_{\mathrm{WNC}}$ can be safely neglected in this discussion in comparison with electrodynamical interactions for which $[\hat{H}_{\mathrm{MOL}}, \hat{P}] = 0$ *is* valid. This is not to deny that

$\hat{V}_{\mathrm{WNC}}$ can have a discriminating effect [73] on optically active enantiomers; the point made here is that the essence of the phenomenon of natural optical activity must be sought elsewhere. Where should we look?

There seem to be two possible answers. One can always put in the broken symmetry by hand by choosing an initial density matrix $\hat{\rho}_0$ such that $[\hat{\rho}_0, \hat{P}] \neq 0$. This is what is done in quantum chemistry which asserts that an optical isomer is a nonstationary state of the molecule which is a superposition of two states of opposite parity [74—76]. In effect, the conventional wisdom simply declares that the structure hypothesis is required by the facts, and modifies the theory based on $\hat{H}_{\mathrm{MOL}}$ accordingly; however it leaves a fundamental question unanswered, namely: what is the dynamical origin of such a nonstationary state?

The other approach has some similarity with the theory of irreversibility and may supply the answer lacking from the quantum chemistry account; part of the difficulty is that unitary time evolution (more generally unitary transformation) cannot provide the requisite symmetry breaking. However, consider the Hamiltonian $\hat{H}_n$, equation (41) in the limit $n \to \infty$, system volume $V \to \infty$, but with (n/V) finite; we can still have *operators* commuting with space-inversion, as in the molecular theory,

$$[\hat{H}_\infty, \hat{P}] = 0 \tag{43}$$

$$[\hat{\rho}_0, \hat{P}] = 0 \tag{44}$$

but now the time evolution of an initial density matrix $\hat{\rho}_0$ satisfying (44) is *not necessarily unitary* because $\hat{H}_\infty$ has a continuous spectrum, and this makes it possible for the symmetry breaking to be generated dynamically. This means that if certain conditions on $\hat{H}_\infty$ are satisfied, we can obtain 'left-' and 'right-handed' states $\hat{\rho}_L(t)$, $\hat{\rho}_R(t)$ interrelated by the space-inversion operator, and these can support a non-zero rotation angle θ,

$$\hat{\rho}_R(t) = \hat{P}\hat{\rho}_L(t)\hat{P}^+. \tag{45}$$

The limit $n \to \infty$ means that we pass from a molecular theory to a quantum field theory of a macroscopic system; such a theory gives a *formal* description of the experimental facts of optical activity without invoking 'enantiomeric molecules' at the outset [39]. If further assumptions are made the formalism can be reduced to the usual Rosenfeld theory [77]; nothing new is thereby offered to replace the usual practical route to numerical calculations of optical rotation, but we may understand better what it is we have calculated. From the standpoint of a molecular theory the essential ingredient here is that the molecules are coupled to an 'environment' that is described by a quantum field theory. This 'environment' may be given a traditional representation as a 'reaction field' [78], or all other molecules in the substance apart from a reference molecule, but it may also be the quantized electromagnetic field [79], or other boson fields (*e.g.* phonons) [80].

Earlier, we noted that the molecular Hamiltonian $\hat{H}_{\mathrm{MOL}}$ does not describe a

particular molecular species so that, for example, the Hamiltonian related to the chemical formula C_8H_8 does not distinguish vinylbenzene from cyclooctatetraene or cubane, or any other system describable as 'C_8H_8'. A quantum field theory of chemical substances involves an enormous generalization of this conclusion, namely: *all* substances that contain the *same* chemical elements, in whatever proportions by weight, share the *same* quantum field theory Hamiltonian constructed from the Heisenberg field operators for electrons and the nuclei of the specified elements. That this is so may be seen from the following heuristic approach 'from below' which establishes a connection between the molecular theory based on $\hat{H}_{\mathrm{MOL}}$ and the quantum field theory.

We start again with $\hat{H}_{\mathrm{MOL}}$ for the cubane family, 'C_8H_8', but now recognize explicitly that we could just as well have chosen any C_nH_n for the description of the spectral properties and cluster decomposition of the N-body molecular Hamiltonian. In particular we can regard the eigenstates of 'C_8H_8' as being embedded in the continuous spectrum of the Hamiltonian for every system $(CH)_n$ with $n > 8$ (this obviously includes all clusters of the type C_nH_m with $n \neq m > 8$ as well). Once we adopt this point of view it is obvious that the 'molecular picture' should pass over to the 'substance picture' when n is allowed to increase without limit or, more precisely, $n \to \infty$ with (n/V) finite. We then have $(CH)_\infty$ as a symbol for hydrocarbon matter and, associated with it, a quantum field Hamiltonian $\hat{\mathscr{H}}_{\mathrm{CH}}$. Thus, instead of there being a fundamental Hamiltonian for, say, benzene — regarded as a *substance* — $\hat{\mathscr{H}}_{\mathrm{CH}}$ carries the burden of describing the dynamics and interrelationships of *all* hydrocarbon substances.

The states of all these different hydrocarbon materials must be located in the state space of this one field Hamiltonian; assuming Galilean invariance they may be classified in incoherent (non-combining) subspaces by distinct values of mass parameters (The 'Law of Conservation of Mass' follows automatically from the assumption of Galilean kinematics in this quantum theory, and need not be introduced as a separate postulate). Furthermore, when we consider the state space of this macroscopic limit, formally described by the Hilbert space $L^2[(CH)_\infty]$, we have a very different mathematical situation from that described earlier for $\hat{H}_{\mathrm{MOL}}$. According to Von Neumann's classical theorem a field theory has an infinite number of inequivalent representations of its Hilbert space; whereas for the molecular Hamiltonian every possible cluster decomposition of $(CH)_n$ led to isomorphic representations of the coherent Hilbert space $L^2(C_nH_n)$, in the $n \to \infty$ limit this is no longer true. Different choices of cluster decompositions are expected to lead to distinct representations that may not be related by unitary transformation. In such cases there may be *no* matrix elements connecting different sets of fragment states that share a common mass value. This suggests a conjecture — there is a hint that the field theory account of chemical substances *may* restore some fundamental quality to structure and isomerism which we have not found in molecular quantum mechanics based on $\hat{H}_{\mathrm{MOL}}$. Accordingly we now turn to the outline of a quantum field theory for chemical substances [39].

4. Quantum Field Theory for Chemical Substances

In this theory we must on many occasions renounce mathematical rigour and resort to heuristic considerations, which are frequently justified solely by the agreement of their results with experiment.
Iwo & Zofia Bialynicki-Birula: *Quantum Electrodynamics*, Pergamon (1975).

4.1. INTRODUCTION TO QUANTUM FIELD THEORY

Quantum field theory is best known as the theoretical basis for high-energy physics so it is as well to make clear at the outset what is meant by reference to 'quantum field theory' in a context where typical chemical energies are relevant. Such models of the low-energy properties of ordinary matter are related to the fundamental theories of elementary particle physics — the properties of matter at high energies — only through possessing a mathematical structure based on a Hilbert space $\mathfrak{H}$ parameterized by *functions* in place of the position and momentum coordinates of wave mechanics, and using functional integration [81]. Models based on functional methods provide the only known basis for accounting for *broken symmetry* phenomena which are of such interest in elementary particle physics, as well as condensed matter physics where the concept originated [82]. Irreversibility (Section 3.2) and natural optical activity (Section 3.3) were discussed earlier as examples of broken time reversal and broken space inversion symmetry respectively.

On the other hand, there is no general agreement that developments in high-energy physics have any direct relevance to condensed matter physics or chemistry. A discussion based on quantum field theory concerned with matter at 'ordinary' (*i.e.* chemical energies) is in no way confined to the tiny corrections to atomic energy levels derived from quantum electrodynamics such as the Lamb shift [83] and relativistic spin-spin interactions [84]. The field of application of functional methods is much broader (electronic structure of condensed matter for example) although little of this has yet made much of an impression on theoretical chemistry. Much of the theory of condensed matter is concerned with a quantum theory of electron and nuclear fields and could be expected to have some relevance to chemical theory. The precise chemical nature of materials has usually not been of much interest in general theories of some physical property which is exhibited by many different materials; equally, for a quantum description [85] of the physical properties of bulk matter (electrical conductivity, optical properties, magnetism, ferroelectricity, diffraction experiments *etc.*) it is not necessary to take the further step to the *classical* picture in terms of atoms and molecules. As a result the question of isomerism, which is of crucial significance for chemistry, is not confronted in such discussions. It seems quite possible however that quantum field theory will provide the basis for a future theoretical chemistry which does address such questions.

So, to begin: what *is* quantum field theory? Well, let's first look at the idea of a *field* in classical terms: afterwards we can bring in quantization. Fields are used

to describe physical quantities that are defined at all points in space and time; classically there are two equivalent ways of doing this with mathematics:

(i) use *continuous functions* of space (**x**) and time (t),
(ii) use an *infinite set of discrete variables.*

Examples of classical fields are not hard to find:

1. A fluid can be described by its density $n(\mathbf{x}, t)$, the pressure $p(\mathbf{x}, t)$ and the temperature $T(\mathbf{x}, t)$ at time t and at each point in the fluid located by the vector **x**.
2. An elastic body can be described by a displacement field $\mathbf{u}(\mathbf{x}, t)$ specifying the vector displacement of the part of the solid that in equilibrium occupies the position **x**.
3. The electromagnetic field is described by the electric and magnetic fields, $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{B}(\mathbf{x}, t)$ respectively; as with the displacement field **u** in 2, $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{B}(\mathbf{x}, t)$ have magnitudes *and* directions at each point **x**, t — they are 'vector-valued' functions. In the absence of charges, the energy of the field is

$$\mathscr{E} = \tfrac{1}{2}\varepsilon_0 \int \mathrm{d}^3\mathbf{x}\{\mathbf{E}(\mathbf{x}, t)\cdot\mathbf{E}(\mathbf{x}, t) + c^2\mathbf{B}(\mathbf{x}, t)\cdot\mathbf{B}(\mathbf{x}, t)\} \tag{46}$$

where c is the speed of light and ε_0 is the vacuum permittivity.

The use of an infinite set of discrete variables can be conveniently illustrated by reference again to the free electromagnetic field. A monochromatic light-wave has a circular frequency ω and a wave-vector **k** with magnitude $k = |\mathbf{k}| = (\omega/c)$. A given value of **k** can be used to label what we call a *field mode.* Since $0 \leqslant \omega \leqslant \infty$ there are an infinite number of modes. We can associate a classical *coordinate* $\mathbf{Q}(\mathbf{k})$ and momentum $\mathbf{P}(\mathbf{k})$ with each mode, and then the energy, $\mathscr{E}$, of the free field may also be written as [45]

$$\mathscr{E} = \frac{1}{2}\sum_{\mathbf{k}}\{P(\mathbf{k})^2 + \omega^2 Q(\mathbf{k})^2\} \tag{47}$$

i.e. the mode description of the electromagnetic field is an infinite collection of independent harmonic oscillators with frequencies $\omega = kc$. *After* quantization of the field oscillators we arrive directly at the 'light quantum' or 'photon' picture of electromagnetic radiation — photons of course are purely quantum mechanical entities and make no appearance in Maxwell's electrodynamics.

The distinction between the two approaches to fields illustrated in the examples is strongly reminiscent of Schrödinger's use of continuous wavefunctions, as compared with the infinite dimensional discrete matrices of Heisenberg, to describe the Hilbert spaces of quantum mechanics in equivalent ways. In a sense, the use of continuous functions emphasizes *classical wave aspects* of a physical field, while the use of discrete variables is related more naturally to a *particle description.*

Nevertheless (46) and (47) are completely equivalent equations for the energy of the radiation field, and in any case after quantization there is no real distinction to be made between waves and particles which are strictly classical concepts.

Field quantization is most easily approached by considering again the oscillator representation of the free radiation field. The prescription of canonical quantization is to reinterpret classical canonical variables as non-commuting operators and we write

$$\begin{aligned} &\mathbf{P}(\mathbf{k}) \Rightarrow \hat{\mathbf{P}}(\mathbf{k}); \qquad \mathbf{Q}(\mathbf{k}) \Rightarrow \hat{\mathbf{Q}}(\mathbf{k}) \\ &[\hat{Q}(\mathbf{k})_r, \hat{P}(\mathbf{k}')_s] = i\hbar\delta_{rs}\delta(\mathbf{k}-\mathbf{k}') \end{aligned} \tag{48}$$

where δ_{rs} is the Kronecker delta ($\delta_{rs} = 1$ if $r = s$, and 0 otherwise) and $\delta(\mathbf{k} - \mathbf{k}')$ is the Dirac delta distribution with the property

$$\int \mathrm{d}^3\mathbf{k} f(\mathbf{k})\delta(\mathbf{k}-\mathbf{k}') = f(\mathbf{k}').$$

The classical energy formula can then be reinterpreted as the Hamiltonian operator, $\hat{H}^{\mathrm{RAD}}$, for the free radiation field:

$$\hat{H}^{\mathrm{RAD}} = \frac{1}{2}\sum_{\mathbf{k}} \{\hat{P}(\mathbf{k})^2 + \omega^2\hat{Q}(\mathbf{k})^2\} \tag{49}$$

where the sum over $\mathbf{k}$ is understood as $V/(2\pi)^3 \int \mathrm{d}^3\mathbf{k}$ in the limit of an infinite system ($V \to \infty$). Solving the Schrödinger equation is easy because the mode oscillators are independent of each other, and the quantum theory of a harmonic oscillator is standard bookwork in quantum mechanics texts. Each mode has an infinite set of discrete energy levels with associated oscillator wavefunctions that can be labelled by integer valued quantum numbers $N(\mathbf{k})$,

$$E_{N(\mathbf{k})} = (N(\mathbf{k}) + \tfrac{1}{2})\hbar\omega; \qquad \psi_{N(\mathbf{k})}(\mathbf{Q}(\mathbf{k})). \tag{50}$$

The quanta of these mode oscillators are just the 'photons' of the radiation field. In this representation the electromagnetic field is specified by giving the values of the $N(\mathbf{k})$ — the mode 'occupation numbers' — i.e. the numbers of photons with given energy and momentum $\mathbf{k}$:

$$E_N^{\mathrm{RAD}} = \sum_{\mathbf{k}} E_{N(\mathbf{k})}; \qquad \Psi_N^{\mathrm{RAD}} = \prod_{\mathbf{k}} \psi_{N(\mathbf{k})}(\mathbf{Q}(\mathbf{k})). \tag{51}$$

The ground-state of the field — the 'vacuum' — has all $N(\mathbf{k}) = 0$; there is an infinite zero-point energy, which cancels out of all excitation energies, while the state is described by the *wavefunctional*

$$\Psi_0^{\mathrm{RAD}}[\mathbf{Q}] = \exp\left\{-(c/2\hbar)\sum_{\mathbf{k}} k Q(\mathbf{k})^2\right\} \tag{52}$$

as follows from (51) and the usual formula for the ground-state wavefunction of an harmonic oscillator with unit mass and frequency kc.

The infinite number of degrees of freedom in the field in general requires the use of functional mathematics; in effect the use of the oscillator representation for the free radiation field disguises this fact and enables us to use familiar elementary methods to obtain (51) and (52). Equation (52) can be obtained directly if we choose a representation labelled by functions of the mode coordinates $\{\mathbf{Q}(\mathbf{k})\}$ and interpret the operator $\hat{\mathbf{P}}(\mathbf{k})$ as a functional derivative operator $\hat{\mathbf{P}}(\mathbf{k}) \Rightarrow i\hbar\delta/\delta\mathbf{Q}(\mathbf{k})$; equation (52) then follows as the ground-state solution of the functional Schrödinger equation based on $\hat{H}^{\mathrm{RAD}}$ with this realization of the operators $\hat{\mathbf{Q}}(\mathbf{k})$ and $\hat{\mathbf{P}}(\mathbf{k})$.

There is always an equivalent 'wave' formulation; it is much more convenient than the field quanta representation for calculations involving *fluctuations* of the field variables. Quantization of the field description in terms of continuous functions requires us to replace such functions by *operators* that will act on functionals defined at all points in space and time. For the free electromagnetic field we have electric and magnetic field operators that are effectively defined by their equal-time commutation relation,

$$\begin{aligned} &\mathbf{B}(\mathbf{x}, t) \Rightarrow \hat{\mathbf{B}}(\mathbf{x}, t); \qquad \mathbf{E}(\mathbf{x}, t) \Rightarrow \hat{\mathbf{E}}(\mathbf{x}, t) \\ &[\hat{E}(\mathbf{x}, t)_r, \hat{B}(\mathbf{x}', t)_s] = i\hbar\varepsilon_0^{-1} \in_{rst} \nabla_t \delta(\mathbf{x} - \mathbf{x}') \end{aligned} \tag{53}$$

where $\nabla_t = \partial/\partial x_t$, and $\in_{rst}$ is the antisymmetric Levi-Civita symbol which vanishes if two of its indices are the same and has values of ± 1 that alternate under permutation of pairs of indices. In the '**B**'-representation (*i.e.* with respect to wavefunctionals that depend on the magnetic field $\mathbf{B}(\mathbf{x})$) the electric field operator is realized as the functional derivative

$$\hat{E}(\mathbf{x})_r \Rightarrow i\hbar\varepsilon_0^{-1} \in_{rst} \nabla_t \delta/\delta B(\mathbf{x})_s. \tag{54}$$

In terms of these field operators we can write

$$\hat{H}^{\mathrm{RAD}} = \tfrac{1}{2}\varepsilon_0 \int \mathrm{d}^3\mathbf{x}\{\hat{\mathbf{E}}(\mathbf{x}, t) \cdot \hat{\mathbf{E}}(\mathbf{x}, t) + c^2\hat{\mathbf{B}}(\mathbf{x}, t) \cdot \hat{\mathbf{B}}(\mathbf{x}, t)\} \tag{55}$$

which, with (54), again gives rise to a functional Schrödinger equation; this form of $\hat{H}^{\mathrm{RAD}}$ is completely equivalent to the one based on mode oscillators [45]. It has been used by Power [86] to discuss the long-range retarded intermolecular forces between neutral molecules which are interpreted as being due to perturbations of the vacuum state of the electromagnetic field due to the presence of the polarizable matter.

A quantum mechanical theory of the vibrations in a solid body can be based on the use of the displacement field $\mathbf{u}(\mathbf{x}, t)$ as a dynamical variable. For small displacements the corresponding Hamiltonian has the form of equation (47); the quanta of this field are described in the same way as the electromagnetic field and are known as *phonons*. Attempts have been made to describe ordinary matter in

terms of its density $n(\mathbf{x}, t)$ and an appropriate conjugate variable (the matter 'current') but these so-called *current algebras* have led to relatively meagre results. For this reason the quantum field theory of macroscopic matter rests on quantum electrodynamics with radiation coupled to quantized electron and nuclear fields.

Free fields do not cause mathematical difficulties essentially because they can be represented as an infinite collection of independent modes as was done for the free electromagnetic field above; however physically interesting interesting interacting fields pose many open questions because field quantization in the general case is an incompletely understood mathematical topic. This means that heuristic arguments and physical insights are likely to be very important in the development of field theories which cannot be based wholly on rigorous proofs. Such an argument leading from a quantum field theory of chemical substances — pictured as coupled electron and nuclear fields — to the familiar description in terms of 'atoms' and molecules' will be described in the next section. The whole approach is based on the premise that the quantum mechanical considerations that are used routinely in other areas of physics can equally well be used for chemical substances — all that really distinguishes the different regimes of the physics is the *energy* at which one is working. When the true quantum nature of substances has been recognized, the field theoretical context may be put in the background to be used if and when convenient or, as in the cases of irreversibility and natural optical activity, necessary. In the 'particle' representation the basic building blocks of a quantum field theory are Heisenberg operators $\hat{\psi}_n^+(x)$, $\hat{\psi}_n(x)$ that respectively create and destroy particles of type n at the space-time point $x = \mathbf{x}, t_{\mathbf{x}}$. For the purposes of chemistry we can take the index n as e for electrons and α for nuclei only; their fields $(\hat{\psi}_n^+, \hat{\psi}_n)$ carry electric charge which is absolutely conserved. Accordingly, one cannot physically 'create' and 'annihilate' electrons and nuclei; rather charge is *carried* by a particle from one space-time point to another. Charge conservation appears in the theory through the requirement that physically significant quantities must be *gauge-invariant*; the bare Heisenberg operators $\hat{\psi}_n^+$, $\hat{\psi}_n$ are not gauge-invariant and should not be given a physical interpretation [87].

4.2. A QUASI-PARTICLE THEORY OF ATOMS AND MOLECULES

The formulation of the quantum field theory requires the specification of both the system Hamiltonian in terms of the Heisenberg operators $\hat{\psi}_n^+$, $\hat{\psi}_n$ and their algebraic properties. We shall not make any explicit reference to the spins carried by these fields beyond noting that odd-integral spins require Fermi statistics, so that for Fermi fields we have equal-time *canonical anticommutation relations* (CARS)

$$\begin{aligned} \{\hat{\psi}_n^+(\mathbf{x}), \hat{\psi}_{n'}(\mathbf{x}')\}_+ &= \delta_{nn'}\delta(\mathbf{x}-\mathbf{x}') \\ \{\hat{\psi}_n^+(\mathbf{x}), \hat{\psi}_{n'}^+(\mathbf{x}')\}_+ &= \{\hat{\psi}_n(\mathbf{x}), \hat{\psi}_{n'}(\mathbf{x}')\}_+ = 0 \end{aligned} \tag{56}$$

whereas integer spin fields require Bose statistics and hence the field operators satisfy *canonical commutation relations* (CCRS)

$$[\hat{\psi}_n^+(\mathbf{x}), \hat{\psi}_{n'}(\mathbf{x}')] = \delta_{nn'}\delta(\mathbf{x}-\mathbf{x}')$$
$$[\hat{\psi}_n^+(\mathbf{x}), \hat{\psi}_{n'}^+(\mathbf{x}')] = [\hat{\psi}_n(\mathbf{x}), \hat{\psi}_{n'}(\mathbf{x}')] = 0. \tag{57}$$

The Hamiltonian for the coupled electron and nuclear fields can then be written in the usual way as a sum of kinetic ($\hat{\mathscr{T}}$) and potential ($\hat{\mathscr{V}}$) energy terms,

$$\hat{\mathscr{H}} = \hat{\mathscr{T}}_e + \sum_\alpha \hat{\mathscr{T}}_\alpha + \hat{\mathscr{V}}_{ee} + \sum_\alpha \hat{\mathscr{V}}_{e\alpha} + \sum_{\alpha\neq\beta} \hat{\mathscr{V}}_{\alpha\beta} \tag{58}$$

where, for example,

$$\hat{\mathscr{T}}_n = \frac{1}{2m_n}\int \mathrm{d}^3\mathbf{x}\,\nabla\hat{\psi}_n^+(\mathbf{x})\cdot\nabla\hat{\psi}_n(\mathbf{x}) \tag{59}$$

is the kinetic energy for field n, and

$$\hat{\mathscr{V}}_{nn'} = \frac{1}{2}\int \mathrm{d}^3\mathbf{x}\int \mathrm{d}^3\mathbf{x}'\,\hat{\psi}_n^+(\mathbf{x})\hat{\psi}_{n'}^+(\mathbf{x}')V_{nn'}(\mathbf{x}-\mathbf{x}')\hat{\psi}_n(\mathbf{x}')\hat{\psi}_{n'}(\mathbf{x}) \tag{60}$$

describes the potential energy of interaction between fields n and n'. We take the potential $V_{nn'}(\mathbf{x}-\mathbf{x}')$ to be a Coulomb potential,

$$V_{nn'}(\mathbf{x}-\mathbf{x}') = \frac{Z_nZ_{n'}}{|\mathbf{x}-\mathbf{x}'|}. \tag{61}$$

This Hamiltonian is Galilean invariant, as is appropriate for chemical phenomena, rather than Lorentz invariant. We shall not deal explicitly with the coupling to the electromagnetic field (*cf.* Section 3.2) beyond noting that the Hamiltonian $\hat{\mathscr{H}}$, equation (58), is the field representation of $\hat{H}^{\mathrm{MATTER}}$ in equation (12).

Quantum field theory justifies the assumption that to a first approximation the *structure* of the energy level spectrum of $\hat{\mathscr{H}}$ obeys the same principle as that of the energy levels of an ideal gas [88]. In other words, any energy level can be obtained as the sum of energies of a certain number of "quasi-particles" or "elementary excitations", with momenta $\mathbf{p}$ and energy $\varepsilon = \hbar\omega(\mathbf{p})$, moving in the volume occupied by the system. Each different kind of elementary excitation is associated with a quantum field, described by annihilation and creation operators $\hat{\chi}_n(\mathbf{p})$, $\hat{\chi}_n^+(\mathbf{p})$ which may satisfy either boson or fermion statistics, and are such that their equations of motion are free (non-interacting) field equations,

$$[\hat{\mathscr{H}}, \hat{\chi}_n^+(\mathbf{p})] = i\hbar\omega(\mathbf{p})\hat{\chi}_n^+(\mathbf{p}) \tag{62}$$

$$[\hat{\mathscr{H}}, \hat{\chi}_n(\mathbf{p})] = -i\hbar\omega(\mathbf{p})\hat{\chi}_n(\mathbf{p}) \tag{63}$$

and

$$\hat{\chi}_n(\mathbf{p})|0\rangle = 0 \tag{64}$$

if $|0\rangle$ is the ground-state wavefunctional. Before describing the technical machinery used to transform the field theory to the quasi-particle picture it is instructive to consider an example.

An important model in the electronic structure theory of condensed matter is based on the quantum field theory of an electron field in the presence of a compensating positively charged background to ensure overall electroneutrality (an approximation to equation (58)). This is a model of ordinary matter that arises in the electron-gas model which uses a smeared out, uniform background charge, and in the adiabatic model based on the Born—Oppenheimer approximation for the nuclei. The quanta associated with the electron field $\hat{\psi}_e(x)$ are called 'bare electrons'; they are always highly correlated because of their charge and because of the operation of the Pauli Exclusion Principle. The result is that there is a region of space around each bare electron from which other (bare) electrons are excluded — the 'exchange-correlation hole'. A key insight is that this many-particle system of [bare-electron + exchange-correlation hole] can be treated as a one-particle-like entity obeying Fermi statistics, which we call a 'quasi-electron' or simply 'an electron'. The motion of this complicated entity can be described by a Schrödinger-like equation using the coordinate $\mathbf{r}$ of the bare electron at its centre. The solutions $\{\phi_i(\mathbf{r})\}$ of the familiar one-electron SCF equations that we use in quantum chemistry such as the Hartree—Fock and local density equations are approximations to the wavefunctions for the quasi-electron. Residual interactions cause quasi-electrons to decay into electron—hole pairs which are more complicated quasi-particles (this is the solid state physics interpretation of *electron correlation*).

There are also *collective excitations* due to coherent oscillations of all the quanta of the electron field. They are described by harmonic oscillator Hamiltonians with frequencies $\omega(\mathbf{p})$, so these excitations have energies $\varepsilon(\mathbf{p}) = \hbar\omega(\mathbf{p})$; they obey Bose statistics and are called *plasmons*. Plasmons are used to describe the excitations seen in electron energy-loss experiments in which a low-current electron beam is passed through a thin sample of material and the energy loss is monitored — a kind of absorption spectroscopy using electrons instead of light. Plasmons also decay, either into quasi-electrons or into electron-hole pairs depending on the residual interactions.

From this example we learn that the elementary excitations are the result of the collective interactions of the bare field(s) in the system, and therefore pertain to *the system as a whole*. Moreover the elementary excitations can be identified with the 'physical particles' we 'see' in spectroscopic experiment; generally excitations are observed to have some spread in energy corresponding to a finite lifetime, and equations (62)—(64) are only realized approximately. The elementary excitations may be described as wave-packet states associated with the field Hamiltonian $\hat{\mathscr{H}}$ *i.e.* as superpositions of large numbers of exact stationary states of $\hat{\mathscr{H}}$ with a narrow spread in energy. Thus the transformation to the quasi-particle picture does not go as far as diagonalization of the field Hamiltonian $\hat{\mathscr{H}}$; in most cases we do not know how to diagonalize $\hat{\mathscr{H}}$, and because the mathematics of functional methods is not completely understood, it is rarely possible to show that physically inter-

esting field theories possess 'exact' solutions. In any case even if we could find exact stationary states of a field Hamiltonian it is not clear what use they would be, for everything would be *timeless* and we would have no way of describing physical change. Thus the heuristic transformation to a quasi-particle theory achieves what we require for a physically useful theory; the similarity of these considerations to the discussion in Section 3.2 should be evident.

The transformation to a quasi-particle theory can be summarized as follows; first of all we have to recognize that a fundamental difference between quantum field theory and the ordinary quantum mechanics of point-particle systems with finitely many degrees of freedom, is that in the latter, the Hamiltonian $\hat{H}$ can be defined unambiguously (in the sense of unitary equivalence) on a Hilbert space $\mathfrak{H}$. In quantum field theory the existence of inequivalent representations of the CCRS (CARS) means that *we must choose* the representation of Hilbert space we are going to work in. A natural, physical procedure is to formulate the theory in the Fock space of the elementary excitations only, since it is these that we can relate to our experiments [89]. We only introduce the Heisenberg operators $\hat{\psi}_n^+$, $\hat{\psi}_n$ so as to be able to state the fundamental Hamiltonian $\hat{\mathscr{H}}$, and hence the equations of motion of the bare fields. The bare operators $\{\hat{\psi}_n^+, \hat{\psi}_n\}$, and the Hamiltonian, $\hat{\mathscr{H}}$, are then to be interpreted as operators in the Fock space of the physical elementary excitations: we can set up an equivalence between the matrix elements of the bare operators in this Hilbert space, and the matrix elements of the quasi-particle field operators. Such relations may conveniently be written as "weak" equalities between bare and quasi-particle operators with the symbol $\approx$, and with the bra and ket omitted. If we write the bare fields $\{\hat{\psi}_n^+, \hat{\psi}_n\}$ and the quasi-particle fields $\{\hat{\chi}_n^+, \hat{\chi}_n\}$ collectively as $\hat{\Psi}$ and $\hat{\mathrm{X}}$ respectively (*e.g.* as column vectors), then one seeks only those solutions of the Heisenberg equations of motion for $\hat{\Psi}$ that can be expressed in terms of a complete set (an irreducible operator ring [90]) of free quasi-particle fields, $\hat{\Psi} \approx \hat{\Psi}(\hat{\mathrm{X}})$.

It is vital to note that we do not know *a priori* how many different elementary excitations constitute such a complete set. One way of dealing with this problem is to set up a self-consistent quantum-field theory [91] in which one is guided by physical considerations in the initial choice of elementary excitations in order to begin a self-consistent treatment. Of course there are many other approaches to the transformation of quantum field theory to a quasi-particle theory [92, 93]. The self-consistent procedure is particularly convenient [89] in situations where 'composite' quasi-particles are of interest and so is appropriate for a scheme leading to atoms and molecules. In the first step of the calculation we choose the "incoming particles" (in scattering theory language) associated with the Heisenberg fields as the quasi-particles, that is, we use the weak limit,

$$\lim_{t \to -\infty} \hat{\Psi}(x) \approx a\hat{\mathrm{X}}(x) \tag{65}$$

to define $\hat{\mathrm{X}}(x)$. We then expand $\hat{\Psi}(x)$ in terms of *normal products* of the quasi-

particle operators $\hat{\mathrm{X}}$,

$$\hat{\Psi}(x) \approx a\hat{\mathrm{X}}(x) + \int \mathrm{d}^4 y_1 \int \mathrm{d}^4 y_2 g(x - y_1, x - y_2) : \hat{\mathrm{X}}(y_1)\hat{\mathrm{X}}(y_2) + \ldots \tag{66}$$

where the notation $:\hat{\mathrm{X}}(y_1)\hat{\mathrm{X}}(y_2)$ means the normal product in which all the annihilation operators stand to the *right* of all the creation operators. The $+\ldots$ stands for higher-order normal products with more factors. The time integrations in this equation extend from $-\infty$ to t_{x} in accordance with the choice of (retarded) incoming fields as the quasi-particles. The normal product expansion (66) is substituted in the equation of motion for $\hat{\Psi}$ obtained from $\hat{\mathscr{H}}$, and one determines the coefficients g such that the quasi-particle operators $\hat{\mathrm{X}}$ satisfy free-field equations (62—64): in the course of this process the energy spectra of the quasi-particles are also determined. When the equations for these coefficients have no solutions, we modify the initial set of free fields and repeat the computations; such a modification frequently involves the introduction of additional quasi-particle fields, for example, for *composite particles.* This process is continued until a self-consistent set is determined. One can then write the Hamiltonian $\hat{\mathscr{H}}$, equation (58), in terms of the quasi-particle fields (in the sense of a weak equality) as a sum of free (non-interacting) Hamiltonians: of course this can only be done to a certain accuracy, and as mentioned above, the residual terms in the Hamiltonian are responsible for decay processes of the quasi-particles [91].

In terms of Green functions, what we have just described can be expressed by saying that, for example, by studying the equation of motion of the single-particle Green function,

$$G^{(1)}(\mathbf{x}, \tau) = -i\langle 0 | T\{\hat{\psi}(\mathbf{x}, \tau)\hat{\psi}^{+}(0)\} | 0\rangle \tag{67}$$

(where T is the Dyson chronological operator), under the many-body Hamiltonian $\hat{\mathscr{H}}$, one can show that $G^{(1)}$ has an approximate representation, after Fourier transformation, in the form [33]:

$$G^{(1)}(\mathbf{p}, \varepsilon) = \frac{Z_p}{\varepsilon - \varepsilon(\mathbf{p}) + i\Gamma(\mathbf{p})} + \text{small terms.} \tag{68}$$

One then sees that for times $t \lesssim \Gamma^{-1}(\mathbf{p})$ it is useful to speak of a single-particle-like object with energy $\varepsilon(\mathbf{p})$ in the many-body system: this is the one-particle elementary excitation associated with the field $\hat{\psi}(\mathbf{x})$. If we express the field operator $\hat{\psi}(\mathbf{x})$ in terms of a complete set of single-particle states,

$$\hat{\psi}(\mathbf{x}) = \sum_n \hat{a}_n \phi_n(\mathbf{x}) \tag{69}$$

where $\hat{a}_n$ is an annihilation operator for particles in state n, then the $\{\phi_n\}$ associated with $G^{(1)}(\mathbf{x}, \tau)$ are, in the simplest approximation, the solutions of the Hartree or Hartree—Fock equations.

Composite quasi-particles arise in the same way from higher-order Green functions, $G^{(n)}$, involving $n > 1$ pairs of annihilation and creation operators for the various fields involved in the composite (the $G^{(n)}$ are also known as n-point correlation functions). At finite temperatures the expectation value $\langle 0| \ldots |0\rangle$ must be replaced by a trace over the statistical operator $\hat{\rho}_T$ at temperature T. The new feature of n-particle Green functions ($n > 1$) as compared with $G^{(1)}(\mathbf{x}, \tau)$, is that their energy spectra contain "internal" excitations as well as the momentum $\mathbf{P}$ ("centre-of-mass" motion). As with the single-particle Green functions one can define a set of states for the composite quasi-particle as the solutions of some Schrödinger-like equation [94]. As a simple example, consider the two-particle Green function which, in a certain approximation can be shown to satisfy the Bethe—Salpeter integral equation [94—96]. The inhomogeneous form of this integral equation describes scattering solutions and is related to the response functions of the many-body system; on the other hand the homogeneous integral equation describes bound states of the two-particle system (in the approximation considered). If one takes the $\hat{\psi}_n$ operators to refer to electron and proton fields then the two-particle Green function, obtained from the Bethe-Salpeter equation can be used to describe hydrogen atoms; by induction one can extend the argument to any number of particles and the n-particle Green function $G^{(n)}$ satisfies a well-defined equation [96] that is determined purely by the field Hamiltonian $\hat{\mathscr{H}}$. The important point here is that quantum field theory leads one to expect the existence of long-lived composite quasi-particles possessing internal excitations as well as centre-of-mass motion that can be described in terms of the quantum states of a system with a *finite* number of degrees of freedom; these states can be used to build up a Fock space as a tensor product.

A physical argument that is consistent with this picture has been given by Essén [97]; application of the virial and Hellmann—Feynman theorems to a many-body system of electrons and nuclei suggests that the ground-state and low-lying excitations will be dominated by configurations of the particles in which each positively charged nucleus Z_α is surrounded by about Z_α electrons on average since this minimizes the Coulomb energy. This idea can be described mathematically by transforming the many-particle Hamiltonian into a representation that describes the motion of certain collective and individual coordinates that are approximately decoupled (the separation is valid locally rather than globally in the configuration space) [61]. Each strongly interacting complicated collection of electrons about a nucleus can be treated as a quasi-particle, an 'atom'; as with quasi-electrons, its motion can be described using the coordinate $\mathbf{R}$ of the nucleus at its centre (more precisely the centre-of-mass $\mathbf{X}_{\text{C.M.}}$ of the nucleus and electrons). Unlike the electron case however, the atomic motion in a molecule or solid is always strongly correlated, and their collective motions are the vibrational excitations of the molecule (phonons in a solid).

There is thus a natural association between the composite n-particle elementary excitations of the macroscopic system, and a finite dimensional Hamiltonian

operator on the Hilbert space $L^2(\mathbb{R}^{3n})$. For a composite system of electrons and nuclei, this is the usual molecular Hamiltonian, $\hat{H}_{\mathrm{MOL}}$, we discussed earlier. In special circumstances, the quantum states of these elementary excitations can be identified with sufficient accuracy with the stationary states supported by $\hat{H}_{\mathrm{MOL}}$; this is true for example in high-resolution spectroscopy of dilute gases of low molecular weight, for which the Generator Coordinate Method is useful. Moreover at finite temperatures in such systems one can choose the Gibbs state $\hat{\rho} = \exp(-\hat{H}_{\mathrm{MOL}}/k_B T)$ as the statistical operator for calculating expectation values, as in the discussion of polar molecules (Section 3.3). In general however, the atomic and molecular quasi-particles cannot be completely decoupled; the quantum states of composite quasi-particles are determined by the dynamics of the macroscopic system, so that an *ab initio* theory of molecular states requires the full solution (*assumed* to exist!) of the quantum field theory for a macroscopic system of given density. From the theoretical point of view the existence of such solution is speculative, and we certainly do not have any practical techniques which come anywhere near this fearsome demand.

A decade ago I proposed [18] that this dilemma is 'resolved' by the *molecular structure hypothesis* which suggests that we should relate the quantum states of the composite quasi-particle to the stationary states of a Hamiltonian $\hat{h}_{\mathrm{model}}$ for an individual molecule with a definite equilibrium structure. This statement discloses a meaning for the molecular structure hypothesis in quantum mechanical terms. We make a classical mechanical model of a molecule as a quasi-rigid material object with some equilibrium structure i.e. we make use of the legacy of Van't Hoff; this leads to a classical Hamiltonian which can be quantized, and we then calculate the eigenstates of *this* model Hamiltonian:

$$(\hat{h}_{\mathrm{model}} - E_n)\phi_n = 0. \tag{70}$$

The eigenstates $\{\phi_n\}$ of $\hat{h}_{\mathrm{model}}$ are not stationary with respect to $\hat{H}_{\mathrm{MOL}}$, and their symmetries *may* be different from the molecular stationary states of $\hat{H}_{\mathrm{MOL}}$ depending on the model chosen: with this we return to the basis for the programme of quantum chemistry. There are claims in the literature that the atomic/molecular model of matter and 'molecular structure' can be deduced *ab initio* from quantum theory; in my opinion these arguments are not based on the *same* premises as I have used, so different and contradictory conclusions are not too surprising. Even so, I maintain that my original suggestion that molecular structure must be put into quantum theory in order to give the usual quantum chemistry has not been refuted.

I will close this section with one of many possible examples that could be given to show that quantum field theory is the general and fundamental approach to the description of the properties of matter, whether one is interested in solids or molecules, nuclei or elementary particles. If we disregard electronic phenomena and specialize to the case of a single kind of composite quasi-particle the dynamical map $\hat{\Psi} \approx \hat{\Psi}(\hat{\mathrm{X}})$ allows us to write the Hamiltonian in the form of (58) for a single quasi-particle field $\hat{\chi}(x)$ with mass parameter M, and an 'inter-

molecular' interaction term $U(\mathbf{x} - \mathbf{x}')$ which is of much shorter range than the original Coulomb interaction (61) if the composite quasi-particles are overall neutral. If the interaction term is initially neglected there remains the Hamiltonian for a free field of mass M:

$$\begin{aligned} \hat{\mathscr{H}}^0 &= (1/2M) \int \mathrm{d}^3\mathbf{x} \nabla \hat{\chi}(\mathbf{x})^+ \cdot \nabla \hat{\chi}(\mathbf{x}) \\ &= \sum_{\mathbf{p}} \varepsilon(\mathbf{p}) \hat{a}(\mathbf{p})^+ \hat{a}(\mathbf{p}) \end{aligned} \tag{71}$$

where $\varepsilon(\mathbf{p}) = p^2/2M$; $\hat{a}(\mathbf{p})^+$ and $\hat{a}(\mathbf{p})$ are creation and annihilation operators respectively. It is a straightforward matter in statistical mechanics to show that this field Hamiltonian accounts for the properties of an Ideal Gas, e.g. one can derive the Equation of State and the usual thermodynamic relations. With U retained we have the basis for the quantum statistical mechanics that describes 'Intermolecular Forces'. A Hamiltonian of this type has even be used [98] to give a quantum field theory description of perfect crystals and their defect structures and also their mechanical properties by further transformation to a set of field operators for new elementary excitations ('quasi-atoms or quasi-molecules', 'phonons') that once again approximately diagonalize $\hat{\mathscr{H}}(\hat{\chi}^+, \hat{\chi})$. Many other applications of such a formalism are known in statistical physics, but to describe them would take us too far afield in this review (see *e.g.* [99].)

5. Concluding Remarks

Quantum theory thus provides us with a striking illustration of the fact that we can fully understand a connection though we can only speak of it in images and parables.

Attributed to Niels Bohr (1952).

This review has been concerned with the status of the molecular hypothesis within the framework of quantum theory; this is not something that is usefully discussed in blunt 'right' or 'wrong' terms; rather, the concern of this debate, to which many have contributed in the past decade (for bibliography see [100]), is to make sense of the conventional structural representation of molecules in terms of quantum mechanics, preferably the quantum theory of the 1980s. Since this representation is achieved through symbols (structural formulae) there is an obvious connection with metaphor (Section 4.2) and in my view molecular structure is best understood as a powerful metaphor, not as a description of 'the way molecules are'. What has to be decided then is how the symbols of chemistry function as *signs*; what rules of convention or habitual association are there between say the chemical symbol C_6H_6 and its object, the substance we call benzene? Classical chemistry, quantum chemistry and quantum theory all give answers to this sort of question, and in my view it is essential to recognize that these answers are not the same.

In Van't Hoff's stereochemistry, well-defined, self-contained and timeless objects

— *atoms* — are seen as *figure* in empty or otherwise qualitatively different space which serves as *ground*, to borrow a terminology familiar in aesthetics. Such a model, animated through classical realism, is easily visualized and exerts a powerful emotional attraction [43]. Moreover, no one doubts that molecular structure is a key concept in chemical science — perhaps the key concept. Gilbert Newton Lewis once wrote:

> No generalization of science, even if we include those capable of exact mathematical statement, has ever achieved a greater success in assembling in a simple way a multitude of heterogeneous observations than this group of ideas which we call structural theory [101].

Few chemists will dissent from this judgement; if however we separate the metaphysics (classical realism) from the science (structure as a concept for solving chemical problems) we see that, irrespective of any metaphysical stance, what really matters about the molecular structure hypothesis is that it offers a *representation of chemical phenomena* that has enabled chemists to grasp fresh and significant relationships in their experimental findings. This aspect of the structural principle was recognized at the time of its first use; for example Gerhardt wrote (1856):

> Les formules chimiques comme nous l'avons dit, ne sont pas destinées à representer l'arrangement des atomes, mais elles ont pour but de rendre évidèntes, de la manière la plus simple et la plus exacte, les relations qui rattachent entre eux sous le rapport des transformations [102].

Until rather recently such a view would have seemed to be of purely historical interest, the material reality of molecules being apparently 'incontestable' [103, 104]; we now know that it *is* possible to say useful things about chemical substances without slavish adherence to the '*central dogma*' (Section 2), in some cases without reference to molecular structure at all. This has come about because of new experiments and a renewed interest in the foundations of chemical theory. Today we recognize that those foundations must be built on quantum theory (Sections 3 and 4).

In Section 2.3 I argued that the historical record suggests that the pioneers of quantum chemistry simply borrowed the conceptual framework of classical chemistry and proceeded to make the best of it that they could. That enterprise, stimulated initially by Mulliken and Pauling, Slater and Van Vleck and developed in the post-war years by *inter alia* Coulson, Daudel and Löwdin, has been brilliantly successful; nevertheless, as we have seen (Sections 3 and 4) conventional quantum chemistry does not exhaust the possibilities of application of quantum theory to chemical problems.

It is an elementary point of logic that the success of the structural model in accounting for experimental observations, with or without quantum refinements, *does not imply* that the molecular structure hypothesis is a true account (or even probably true account) of the nature of chemical substances as they 'really' are: to go down that road is to become entangled in the mirages of metaphysics. Truth of

a hypothesis is a matter of *fit* — fit with a body of theory, and fit of hypothesis and theory to the data at hand; goodness of fit takes a two-way adjustment — of theory to facts *and* of facts to theory [105]. So while we may all agree readily that new experimental data may require the rebuilding of a theoretical scheme, the claim may also be made that a radical change in theory may lead to a completely new look to the 'facts', even to the extent of determining what may be called a 'fact' [106]. This is a matter that requires vigilance because our beliefs help to define the unthinkable — those thoughts that are ordinarily put out of mind — and hence limit what we regard as possible. Progress in science may well require the acceptance of what hitherto has seemed 'unthinkable'; thus for example, and central to the discussion of the molecular hypothesis from the viewpoint of quantum theory, there is no *necessity* for us to interpret *every* chemical experiment in terms of classical molecular structure [107]. The quickening transformation of much of 'physical chemistry' into 'chemical physics' is witness to this fact.

None of the usual arguments for the 'reality' of atoms as material entities are logically compelling; all so-called 'experimental proofs' can be described in terms of physical theory that does not require this materialist content, which is metaphysical in character. Irrespective of what one believes about atoms 'existing as the ultimate material units of the elements' [5], it is still the case of course that, for example, crystals of sodium chloride give Bragg X-Ray line spectra and diffraction spots; it is up to *us* to construct theoretical frameworks for such observations. Thereby we do not 'discover or prove the atomic constitution of the world', a conception which strictly speaking belongs to metaphysics rather than science [108].

In the classical approach, diffraction of X-rays by a regular crystal is interpreted as the scattering of the radiation by a lattice of atoms. Quantum chemistry takes over this picture and, by treating an atom as a system of a nucleus plus a specified number of electrons, can make a quantitative estimate of the scattering factors. In general quantum theory however the description of the crystal is based on a coupled electron—nuclear wavefunction that evolves in time. If one adopts the GCM technique (Section 3.3) to describe the non-adiabatic wavefunction, it is not difficult to construct a semi-empirical theory in which the lattice is associated with the space of generator coordinates $\{\alpha\}$, rather than the electrons and nuclei. The two descriptions seem incompatible, but from the perspective advocated here there is no necessity to declare one 'right' and the other 'wrong'.

The concept of a molecule in physics, chemistry and biology no longer has an all embracing meaning; instead there is a plurality of meanings that are more or less appropriate depending on the context of the problem under consideration. In biology the classical notions of structure, mechanism and function remain extremely important and quantum theory is conspicuous by its absence. My impression is that classical stereochemistry is now of diminishing importance in chemistry which in recent years has seen a marked increase in the utilization of qualitative arguments based on *electronic structure theory*. Finally, in chemical

physics, reference to molecular structure is increasingly recognized as being optional; as we have seen (Sections 3 and 4) some progress can be made without reference to it at all, and the very notions of 'atom' and 'molecule' are now under review in the theoretical community.

Acknowledgements

The financial support of The Royal Society, London and C.N.R.S., Paris which made possible my participation in the International Symposium 'Molecules in Physics, Chemistry and Biology' is gratefully acknowledged.

References

1. A. L. de Lavoisier: *Traité élémentaire de chimie* in *Oeuvres de Lavoisier*, Imprimerie Impériale, Paris (1864).
2. O. Hannaway: *The Chemists and the Word: The Didactic Origins of Chemistry*, Johns Hopkins Univ. Press, Baltimore, p. 156 (1975).
3. S. Sambursky (Ed.): *Physical Thought from the Presocratics to the Quantum Physicists*, Hutchinson, London (1974).
4. E. Cassirer: *An Essay on Man*, Yale Univ. Press (1944).
5. J. Dalton: *A New System of Chemical Philosophy*, Manchester (1808).
6. A. Avogadro: *Jour. de Physique de Delamétherie* **73**, 58 (1811).
7. A. M. Ampère: *Ann. de Chimie* **90**, 43 (1814).
8. A. J. Rocke: *Ambix* **30**, 1 (1983).
9. A. J. Rocke: in *Robert Schofield Festschrift* (1986).
10. J. H. Van't Hoff: *La Chimie dans l'espace*, Bazendijk, Amsterdam (1875).
11. R. Auschutz: *August Kekulé*, Verlag Chemie, Berlin (1929).
12. J. Perrin: *Ann. de Chimie et Physique* **18**, 1 (1908).
13. W. L. Bragg: *Proc. Cambridge Phil. Soc.* **17**, 43 (1913).
14. W. L. Bragg: *Scientific American* **219**, 58 (1968).
15. A. Hermann: *The Genesis of Quantum Theory*, MIT Press (1971).
16. F. Hund: *The History of Quantum Theory*, G. G. Harrap, London (1974).
17. W. Heisenberg: *The Physical Principles of the Quantum Theory*, University of Chicago Press (1930).
18. R. G. Woolley: *Adv. Phys.* **25**, 27 (1976).
19. K. Schwarzschild: *Sitzungsber. Preuss. Akad. Wiss.* **1**, 548 (1916).
20. T. Heurlinger: *Z. Phys.* **20**, 188 (1919).
21. A. Marcelin: *Annalen Phys.* **3**, 158 (1915).
22. J. Franck: *Trans. Faraday Soc.* **21**, 536 (1925).
23. E. U. Condon: *Phys. Rev.* **27**, 640 (1926); *ibid.* **28**, 1182 (1926).
24. R. S. Mulliken: *Phys. Rev.* **26**, 561 (1925).
25. R. S. Mulliken: *Proc. Nat. Acad. Sci. U.S.A.* **12**, 144, 151, 158, 338 (1926).
26. R. S. Mulliken: *Phys. Rev.* **32**, 186, 761 (1928); *ibid.* **41**, 49 (1932).
27. F. Hund: *Z. Phys.* **51**, 759 (1928); *ibid.* **73**, 1 (1931).
28. W. Pauli: *Annalen Phys.* **68**, 177 (1922).
29. L. Nordheim: *Z. Phys.* **19**, 69 (1923).
30. W. Heitler and F. London: *Z. Phys.* **44**, 455 (1927).
31. M. Born and J. R. Oppenheimer: *Annalen Phys.* **84**, 457 (1927).
32. M. Born and W. Heisenberg: *Annalen Phys.* **74**, 1 (1924).
33. P. Roman: *Advanced Quantum Mechanics*, Addison-Wesley (1965).

34. H. Primas: *Chemistry, Quantum Mechanics and Reductionism*, Lecture Notes in Chemistry, Vol. 24, Springer, Berlin (1982).
35. P. O. Löwdin: *Int. J. Quant. Chem.* **29**, 1651 (1986).
36. E. B. Davies: *Quantum Theory of Open Systems*, Academic Press (1976).
37. H. Primas: *Chimia* **36**, 293 (1982).
38. H. Primas: *Chemie in unserer Zeit* **19**, 109, 160 (1985).
39. R. G. Woolley: *Structure and Bonding* **52**, 1 (1982).
40. F. J. Dyson: *Disturbing the Universe*, Ch. 23, p. 248, Harper & Row (1979).
41. G. Feinberg: *Solid Clues*, pp. 49—51, Heinemann (1985).
42. R. P. Feynman: *Lectures on Physics*, Vol. 1, Ch. 3, Addison-Wesley (1963).
43. R. G. Woolley: *J. Amer. Chem. Soc.* **100**, 1073 (1978).
44. A. Einstein: *Z. Phys.* **18**, 121 (1917); English translation in B. L. Van der Waerden (Ed.): *Sources of Quantum Mechanics*, Dover Publications, pp. 63—71 (1968).
45. R. G. Woolley: *Adv. Chem. Phys.* **33**, 151 (1975).
46. I. Prigogine and C. George: *Int. J. Quant. Chem.* **12** (**S1**), 177 (1977).
47. I. Prigogine, F. Mayné, C. George and M. de Haan: *Proc. Nat. Acad. Sci. U.S.A.* **74**, 415? (1977).
48. I. Prigogine: *Tanner Lecture*: 'Only an Illusion', delivered at Jawaharlal Nehru University, New Delhi, India, December 1982.
49. B. Misra and I. Prigogine: in *Long-Time Prediction in Dynamics*, C. W. Horton, L. E. Reich and A. G. Szebehely (Ed.), Wiley (1983).
50. I. Prigogine and C. George: *Proc. Nat. Acad. Sci. U.S.A.* **80**, 4590 (1983).
51. P. H. Richter, I. Procaccia and J. Ross: *Adv. Chem. Phys.* **43**, 217 (1980).
52. R. G. Woolley and B. T. Sutcliffe: *Chem. Phys. Letters* **45**, 393 (1977); see also R. G. Woolley: *Isr. J. Chem.* **19**, 30 (1980).
53. S. Aronowitz: *Int. J. Quant. Chem.* **14**, 253 (1978).
54. J. Aguilar and J. M. Combes: *Comm. Math. Phys.* **22**, 269 (1971).
55. E. Balslev and J. M. Combes: *Comm. Math. Phys.* **22**, 280 (1971).
56. W. P. Reinhardt: *Ann. Rev. Phys. Chem.* **33**, 223. (1982).
57. W. Pauli, quoted by B. L. Van der Waerden in: *The Physicist's Conception of Nature*, J. Mehra (Ed.), p. 282, Reidel (1973).
58. P. Claverie and S. Diner: *Isr. J. Chem.* **19**, 54 (1980).
59. L. Lathouwers, P. Van Leuven and M. Bouten: *Chem. Phys. Letters* **52**, 439 (1977).
60. L. Lathouwers and P. Van Leuven: *Adv. Chem. Phys.* **49**, 115 (1982).
61. R. G. Woolley: *Chem. Phys. Letters* **55**, 443 (1978).
62. P. Debye: *Polar Molecules*, Dover (1929).
63. J. H. Van Vleck: *The Theory of Electric and Magnetic Susceptibilities*, Oxford: Clarendon Press (1932).
64. R. E. Moss and I. A. Sadler: *Chem. Phys. Letters* **132**, 190 (1986).
65. J. M. Combes and R. Seiler in: *Quantum Dynamics of Molecules: The New Experimental Challenge to Theorists*, R. G. Woolley (Ed.), NATO ASI **B57**, Plenum (1980).
66. C. Eckart: *Phys. Rev.* **47**, 552 (1935).
67. B. T. Sutcliffe in: *Quantum Dynamics of Molecules: The New Experimental Challenge to Theorists*, R. G. Woolley (Ed.), NATO ASI **B57**, Plenum (1980).
68. J. O. Hirschfelder and E. P. Wigner: *Proc. Nat. Acad. Sci. U.S.A.* **21**, 113 (1935).
69. T. Kato: *Perturbation Theory of Linear Operators*, Springer, Berlin (1966).
70. T. Kato: *Isr. J. Math.* **13**, 135 (1972).
71. L. Pasteur: *Comptes Rend. hebd. séances Acad. Sci. Paris* **26**, 535 (1848).
72. R. A. Harris in: *Quantum Dynamics of Molecules: The New Experimental Challenge to Theorists*, R. G. Woolley (Ed.), NATO ASI **B57**, Plenum (1980).
73. S. F. Mason: *Nature* **311**, 19 (1984).
74. L. D. Barron: *Chem. Phys. Letters* **79**, 392 (1981).
75. R. G. Woolley: *Chem. Phys. Letters* **79**, 395 (1981).
76. E. B. Wilson: *Int. J. Quant. Chem.* **S13**, 5 (1979).
77. L. Rosenfeld: *Z. Phys.* **52**, 161 (1927).

78. P. Claverie and G. Jona-Lasinio: *Phys. Rev.* **33A**, 2245 (1986).
79. P. Pfeifer: *Chiral Molecules — A Superselection Rule Induced by the Radiation Field*, Doctoral Thesis, Swiss Federal Institute of Technology, Diss. ETH No. 6551 (1980).
80. E. B. Davies: *Comm. Math. Phys.* **64**, 191 (1979); *ibid.* **75**, 263 (1980).
81. S. F. Edwards: *Physica* **96A**, 212 (1979).
82. P. W. Anderson: *Science* **177**, 393 (1972).
83. L. Landau and E. M. Lifshitz: *Quantum Electrodynamics: The course of theoretical physics*, Vol. 4, Part 1, First Edn, Pergamon (1974).
84. R. G. Woolley: *Mol. Phys.* **30**, 649 (1975).
85. P. W. Anderson: *Basic Notions in Condensed Matter Physics*, Benjamin/Cummings: Addison-Wesley (1984).
86. E. A. Power: *Phys. Rev.* **A10**, 756 (1974).
87. R. G. Woolley: *J. Phys.* **A13**, 2795 (1980).
88. A. A. Abrikosov, L. P. Gorkov and I. E. Dzyaloshinskii: *Methods of Quantum Field Theory in Statistical Physics*, Dover Publications (1975).
89. L. Leplae, F. Mancini and H. Umezawa: *Phys. Reps.* **10C**, 151 (1974).
90. R. Haag and B. Schroer: *J. Math. Phys.* **3**, 248 (1962).
91. H. Umezawa: *Nuovo Cimento* **40**, 450 (1965).
92. D. Pines: *The Many-Body Problem*, Benjamin (1962).
93. P. W. Anderson: *Concepts in Solids*, Benjamin (1963).
94. L. D. Landau and E. M. Lifshitz: *Relativistic Quantum Mechanics, Course of Theoretical Physics*, Vol. 4, Part 2, pp. 103, 122 (1974).
95. M. Gell-Mann and F. Low: *Phys. Rev.* **84**, 350 (1951).
96. S. Mandelstam: *Proc. Roy. Soc.* **A233**, 248 (1956).
97. H. Essén: *Int. J. Quant. Chem.* **12**, 721 (1977).
98. H. Wadati: *Phys. Reports* **50**, 88 (1979).
99. E. M. Lifshitz and L. P. Pitaevskii: *Statistical Physics*, Part 2, Pergamon (1980).
100. R. G. Woolley: *Chem. Phys. Letters* **125**, 200 (1986).
101. G. N. Lewis: *Valence and the Structure of Atoms and Molecules*, Chemical Catalog Company, New York (1923).
102. W. Gerhardt: *Chimie Organique*, Vol. IV, p. 566, Paris (1856).
103. M. J. S. Dewar: *Introduction to Modern Chemistry*, University of London Athlone Press, pp. 1—9 (1965).
104. P. W. Atkins: *Physical Chemistry*, 2nd Edition, OUP, Ch. 1 (1982).
105. N. Goodman: *Languages of Art*, 2nd Edition, Harvester Press, p. 264 (1976).
106. W. Heisenberg: *Physics and Beyond*, George Allen & Unwin, p. 63 (1971).
107. R. G. Woolley: *J. Chem. Educ.* **62**, 1082 (1985).
108. M. Planck: *Phil. Mag.* **28** (1914).

Evolution and Organization of Molecular Systems

Encapsulating Time?

I. PRIGOGINE
Université Libre de Bruxelles, Service de Chimie Physique II, CP 231, Campus Plaine, B-1050 Brussels, Belgium

and

University of Texas at Austin, Center for Statistical Mechanics and Thermodynamics, Austin, Texas, 78712 U.S.A.

1. Introduction

I am very pleased to contribute to this volume honouring Raymond Daudel; we have been friends for so many years. Indeed, I have to go back forty years, to the years 1946—1947. I was staying in Paris with a grant from the French Government; I remember vividly a lecture I heard in 1946 on the relations between thermodynamics and biology. I confess that I forgot the name of the lecturer; but I remember very clearly the criticisms expressed by Edmond Bauer, who chaired the meeting.

Professor Bauer suggested, in brief, that the lecture could be divided into two parts: a classical one, containing nothing new, and an original one, which to him was incomprehensible. He concluded that thermodynamics was a science of the XIXth century, to which not much new had to be added; more interesting topics were becoming ripe for investigation.

I felt concerned by this opinion, as I was starting at that time the exploration of a new branch of thermodynamics, the branch which is today called non-equilibrium thermodynamics. During the discussion, I tried to explain this new approach, and was very pleased to notice that Professor Bauer and his coworker Michel Magat seemed to be interested. In fact, this appears to me now as having been the start of my international career in science.

It is during those weeks that I was to meet Raymond and Pascaline Daudel. I was very impressed by their work on quantum chemistry, and I remember how interested I was in their work on the relation between quantum mechanics and carcinogenesis. Forty years have elapsed since and Raymond is still busy, improving the methods of quantum chemistry, while I continue to be fascinated by non-equilibrium processes.

In this brief paper, I would like to present a few remarks on the role of irreversibility in connection with the production of information-rich molecules. But let me start first with some historical comments.

2. Entropy, Irreversibility and Creation of Structures

Over the last few decades we have witnessed impressive developments in non-

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, 93—104.

equilibrium physics and chemistry. We can now make more precise the requirements to be satisfied for an evolutionary view of nature [1]. Such a view should reveal itself to be a better frame for our understanding of biological or social phenomena than the classical, atemporal view, based on deterministic and time-reversible laws, which was proposed as an ideal for all sciences after the astounding achievements of classical mechanics in the XVIIth—XVIIIth centuries.

One of the great heritages of the XIXth century is the Darwinian theory of biological evolution. Already at that time, many scientists understood that an evolutionary view of the world requires a revision of physics in order to incorporate the idea of evolution there too. It is enough to mention people like Ludwig Boltzmann or Charles Sanders Peirce.

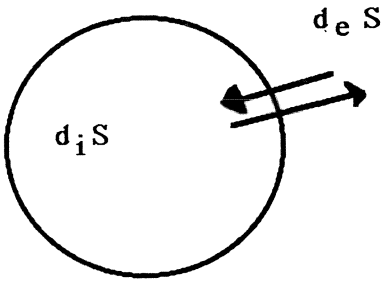
It is indeed this "Darwinian revolution in physics" through which we are going now, and this is the outcome of three convergent trends which, as far as I can see, play in this respect an essential role: the discovery of the instability of elementary particles; the new historical perspectives open in cosmology; and the surprising developments of non-equilibrium physics, and specifically the discovery of non-equilibrium spatio-temporal structures. In this short note, I shall limit myself to the macroscopic aspects of non-equilibrium physics, in spite of the fact that I see it only at a first step towards a revision of the basic laws of microscopic physics, in order to include on a more fundamental level the concept of an evolutionary universe.

It may probably be stated that an evolutionary view of nature should include at least three aspects: irreversibility, probability, and coherence.

The necessity to speak about *irreversibility* is obvious; in order to speak about evolution we cannot but distinguish between Past and Future. Next, the evolutionary view of the world has to include *probability.* It is not meaningful to speak about evolution if we are only to use deterministic equations to compute trajectories. It is only when we are to meet some choice — described by probabilities — that we can speak about evolution. In addition, we should be able to describe the apparition of new structures, involving *coherence* among large populations of molecules. At all levels of modern science, we see now that the world is made of many units assembled in coherent structures; the question for the physicist is to conceive mechanisms susceptible to display the long-range correlations that produce these structures.

Let us start with irreversibility as expressed by the second law of thermodynamics. While this law was formulated during the XIXth century, it is only now that we begin to understand its fundamental role in the evolution of structures. From the classical point of view of equilibrium thermodynamics, order was associated with equilibrium and non-equilibrium with disorder. As we shall see, the situation is now inverted; non-equilibrium leads to long-range coherence between the units involved in the microscopic description, be they atoms or molecules entropy has therefore an important, constructive role in evolution.

We consider a system which exchanges matter and energy with the outside



$$d_i S \geq 0$$

$$dS = d_e S + d_i S$$

Fig. 1. An open system, exchanging entropy with its environment.

world (Fig. 1). The change of entropy dS is equal to the entropy flow d_eS with the outside world, plus the internal entropy production d_iS, entailed by all irreversible processes going on inside the system. The characteristic feature of the second law is that this entropy production, due to irreversible processes, is always positive. So, if the system would be isolated, entropy could only grow, and reach its maximum for a sufficiently long time.

The first question then is: what are irreversible processes? Irreversible processes are processes which present a broken time symmetry. If I write the "heat equation", which describes the temporal evolution of the temperature, it contains a first derivative with respect to time, and therefore has a broken time symmetry: if we change t to −t, we will obtain a different equation. In contrast, let us consider the wave equation in the vacuum: it contains a second derivative with respect to time, and is therefore a time-reversible equation.

It has often been said that the XIXth century was the century of evolution; indeed, entropy corresponds to the evolutionary approach of physics, introduced by engineers and physical chemists, to the amazement of mathematicians and theoretical physicists, who were unhappy with this new concept and tried to play it down — some try to do this even today.

But what is entropy production? The first idea was that entropy production was something uncontrollable, like noise or waste. (Figure 2.)

Consider a system of two boxes which communicate; suppose you heat one of them while you cool the other; put a mixture of two types of molecules — say hydrogen and nitrogen — into the system. Because of the temperature difference, you will have more hydrogen in the warmer box and more nitrogen in the cooler one. Every physical chemist knows about this "thermal diffusion" effect. Now, in

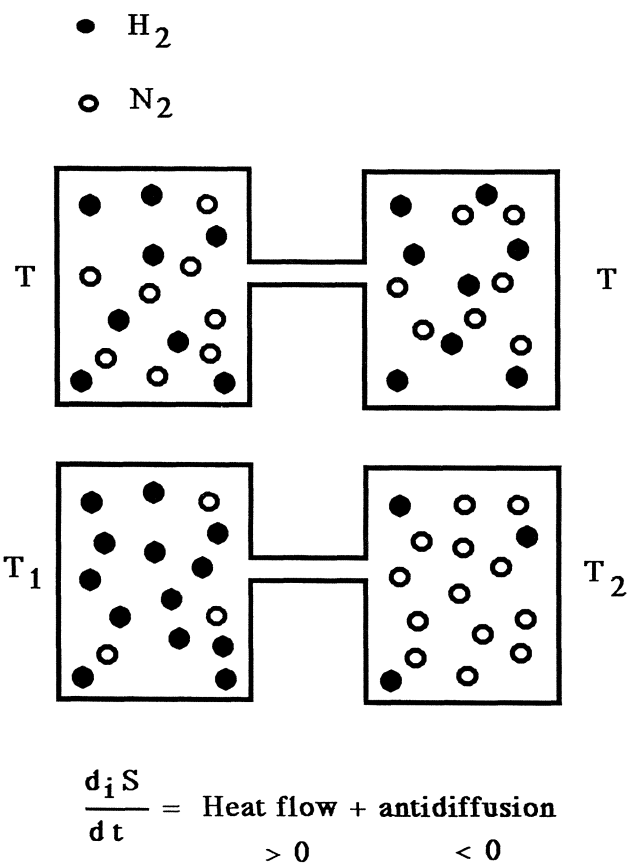


Fig. 2. Two coupled processes, producing respectively disorder and order: thermal diffusion and anti-diffusion.

this example, the entropy production is due for one part to heat conduction, (this is a positive contribution to entropy production); but it is also related to "anti-diffusion", because hydrogen goes to the region where it is already more concentrated. Therefore, we have indeed a positive total entropy production, yet one of the processes (conduction) corresponds to the creation of disorder, while the other corresponds to the creation of order. We see that order is created at the expenses of disorder; this is something we find in many experiments today. (Figure 3.)

A very popular example is the Bénard instability: if one heats a liquid from below, for some critical value of the temperature gradient imposed upon the system, beautiful convection patterns will appear. Again, one has to heat it and dissipate some energy, but as a result patterns are produced; the price one has to pay for the order manifested in these patterns is entropy. Very recently, beautiful experiments in molecular dynamics [2] have been performed to simulate the Bénard instability. These experiments show very clearly that the velocities of two

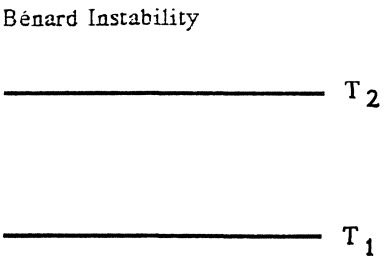


Fig. 3a. In order to instaure Bénard instability, the inferior (T_1) and superior (T_2) interfaces are submitted to thermal constraints.

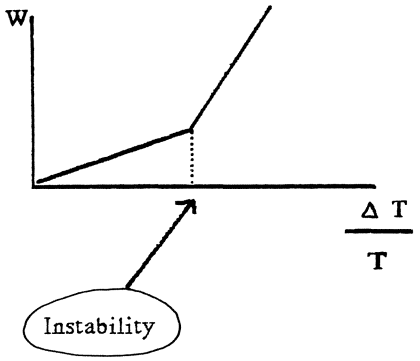


Fig. 3b. For a critical value of the control parameter, a new regime of convection is installed.

molecules studied at some given distance are independent variables at equilibrium, while they are correlated over macroscopic distances when the temperature gradient is introduced. The macroscopic structures we observe are the outcome of these molecular correlations.

One can imagine more elaborated examples: the production of biomolecules such as DNA is coupled with the destruction of other molecules, the total process leading to a positive entropy production.

Another reason why far-from-equilibrium systems are of great interest is because there we have to take into account the non-linearity of the equations which describe the proceses involved: mass transport, chemical reactions and so on. As everybody knows, non-linear equations have more than one solution; this multiplicity of non-equilibrium solutions in hydrodynamical or chemical systems is now studied in many laboratories. I am still amazed by the variety and the beauty of these structures. A very simple example corresponds to chemical clocks. Everybody has heard about the Belousov—Zhabotinski reaction, in which molecules would become "red", then "blue" and so on. The *Scientific American* and many other journals have had pictures illustrating this reaction on their front page. And indeed, I do believe this is one of the greatest experiments performed in our

century. A great experiment is one in which something unexpected is to be learned. If we think about a chemical reaction according to the classical paradigm, we imagine molecules travelling like dust particles in all directions, colliding at random. This would imply no coherence, no macroscopic oscillations. And indeed, at equilibrium, we would have no coherence at all. Some small regions could be red, another blue, but one would not expect the whole vessel to become blue or red. Again, it is far from equilibrium that coherence may appear. This leads to a revision of our traditional views about order and disorder. According to these views, order was associated to equilibrium, and disorder to non-equilibrium; indeed, look at a crystal: what could be more ordered than this equilibrium structure? What could be more disordered than turbulence?

The present view runs oppositely: a crystal may be described as a superposition of normal modes of waves; and these waves are incoherent at equilibrium. This is what we have to take into account when we measure the specific heat of a solid. On the contrary, we know today that turbulence is a highly ordered structure.

Now, non-linearity implies the possibility of many states; transitions between those states occur at bifurcation points, in the neighbouring of which one can expect a probabilistic behaviour [3, 4] (Figure 4).

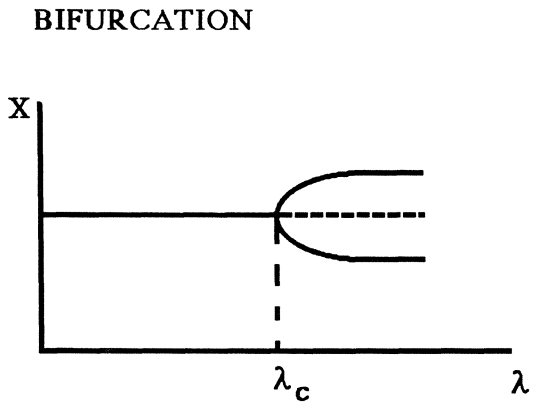


Fig. 4. Bifurcation: for small values of the control parameter λ, the stationary branch remains at the value of equilibrium; for larger values of λ, two stationary states are accessible.

This has indeed been observed; we may check the outcome of experiments of bifurcations, and test that they correspond to some simple probabilistic distributions. Macroscopic physics illustrates therefore today the emergence of probabilistic schemes, which are no more specific of microscopic, quantum physics.

The decisive feature of dissipative systems is that they are able to forget perturbations, which is not the case for conservative systems. If one imposes some temporary perturbation to the orbit of the earth, the ancient orbit will never be

restored. In contrast, dissipative systems show attractor states, whose distribution may be diverse. Simple attractors include the point-attractor (for example the equilibrium state); other may be line-attractors, such as displayed by chemical oscillations, which obey to a limit cycle. Other may be less coherent. The interest of researcher has shifted today towards "dynamical chaos" arising in deterministic systems far from equilibrium. I cannot go here into much detail. Let us only mention that chaos corresponds to situations where phase points describing the state of the system are attracted to regions which are formed by complex sets of points, which have an effective dimension which may not be an integer but a fractal number). We cannot be but impressed by the large number of applications dynamical chaos has found today for the understanding of our environment. For example, very interesting papers have been published about the evolution of terrestrial climate, showing that climatic equations should have this type of complexity [5]. How could a classical physicist see the succession of Ice Ages? He may think "This is so because the temperature on emerged continents is the outcome of many, say 120 variables, each of them fluctuating, and thus it has to display some Gaussian distribution". Recent analysis has shown that things are not so; in fact, the (fractal) dimension of the system involved is only of the order of 3.2, which means that the climatic system could be described by 4 independent variables. The world around us displays an intrinsic instability, leading to a higher complexity. In a paper which some of my co-workers have just published [6], they show that brain waves are also characterized by such a fractal dimension, except in epileptic fits, during which the signals are described in a low-dimensional space (in other words, here brain activity becomes much more regular). In biology, we encounter both regulation of enzymatic reactions described by limit cycles, but also instability, through which small effects may be amplified; this seems to happen with the brain, which requires an amplification mechanism.

3. The Role of Chemistry in Non-Equilibrium Phenomena

There exists an important distinction between hydrodynamics and chemistry in respect to irreversible processes. If we release the constraints which are imposed upon a hydrodynamical system, such as those leading to the Bénard instability, non-equilibrium structures will disappear after some characteristic relaxation time. This is not so in chemistry, where even after constraints have been taken away, the structure generated may record the non-equilibrium conditions in which they were formed. In this sense, we may "encapsulate" time into matter through chemical processes. Let us consider a simple example.

Let us consider two monomers $\{X, Y\}$, involved in some reaction, which are responsible for the formation of some polymer chain (Figure 5).

More precisely, we may assume that when X crosses a given level L_X with a positive slope, a new process is switched on, as a result of which X is rapidly

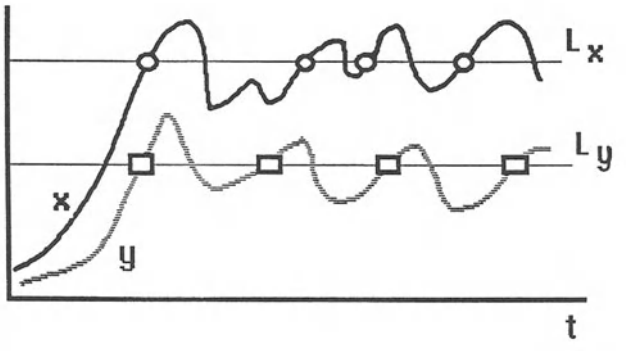


Fig. 5. Characteristic thresholds for a 3-variable system.

precipitating or diffusing outside the reaction space, being then "recorded on tape", and becoming the next segment of the polymer.

Near equilibrium, we may expect that this chain will be highly disordered, and no long-range correlation will exist between the X and Y. Indeed, their concentrations at equilibrium fluctuate according to some 'Poisson-like' distribution, and we should then obtain a very disordered chain:

$$\ldots \text{XYYXXXYXYYXXXYXXXX} \ldots \tag{3.1}$$

which would be essentially a "Bernoulli-type" of chain, provided that the successive digits (N_i = X or Y) are independent. Now, let us assume that $\{X, Y\}$ are the components of a chemical clock; both $\{X, Y\}$ would be then represented by a periodic function of time: the concentrations of $\{X, Y\}$ evolve toward a limit cycle. We may therefore expect, and verify through numerical simulation, a periodic polymer.

$$\ldots \text{XYXYXYXYXYXYXYXYXY} \ldots \tag{3.2}$$

which is reversible, as it reads the same in either direction. This is a non-equilibrium structure, and we may freeze it and transport it from one place to another. In this sense, we may state that we have encapsulated time, incorporating it in a chemical structure. However, obviously, no "information" has been generated in this way. This remains true even if we would work with a chemical clock involving three components. A new element appears there, but still no genuine information. We could with a three-components reaction, generate the simple chain:

$$\ldots \text{XYZXYZXYZXYZXYZ} \ldots \tag{3.3}$$

which describes an absolutely asymmetrical object; still, if by introducing abbreviated symbols we rewrite this type of chain through:

$$\ldots \{\text{XYZ}\} = \alpha \tag{3.4}$$

sequence (3.3) reads:

$$\ldots \alpha\ \alpha\ \alpha\ \alpha \ldots \tag{3.5}$$

and thus loses its space asymmetry; we can hardly speak of information at all here, since we are confronted by the two extremes of insignificance: zero-correlation (Bernoulli chain) or redundancy (limit cycles). Can we go further? Can we imagine that in this way biomolecules (or precursors of biomolecules) have been produced? Some qualitative reasons for this can be produced.

We know that much of what we call *information* is often carried by a one-dimensional spatial structure which displays an absolute asymmetry in the form of a preferred polarity. This asymmetrical structure is the DNA, and the information carried is the genetic code. The genetic code depends on two elements: a particular sequence of codons along the genetic material; and the possibility of "reading" this sequence unidirectionally. Although perfectly reproductible from one generation to the next, the codon sequence is basically unpredictable in the sense that its global structure cannot be inferred from the knowledge of a part of it, however large. It can thus be regarded as a stochastic process, and it is this possibility that allows us to speak of "information". Significantly, in all known biosynthetic reactions in which the information is revealed, there are start signals, and the reading proceeds down the messenger, three nucleotides by three in a fixed direction from the start point. Can we produce at least a "caricature" of this situation?

4. Chaotic Dynamics and Generation of Information

We follow here a model proposed by the Nicolis and the Subba Rao [7], which have studied this question by coupling polymerization with a chaotic attractor. Let us consider a 3-variable system described by the equations

$$\mathrm{d}X/\mathrm{d}t = -Y - Z \tag{4.1}$$

$$\mathrm{d}Y/\mathrm{d}t = X + aY \tag{4.2}$$

$$\mathrm{d}Z/\mathrm{d}t = bX - cZ + XZ \tag{4.3}$$

where we have choosen simple threshold values of polymerization:

$$L_X = L_Y = L_Z = 3.0 \tag{4.4}$$

For parameter values $a = 0.38$, $b = 0.3$ and $c = 4.5$, this model evolves through a chaotic attractor, known as the Rössler attractor [8, 9] (Figure 6).

Starting from initial conditions

$$X_0 = Y_0 = Z_0 = 1.0 \tag{4.5}$$

one can generate a sequence of symbols of the form:

$$\ldots \mathrm{ZYXZXYXZXYXZYXZXYXZYXZYXZXZYXZYXZXYXZYX} \ldots \tag{4.6}$$

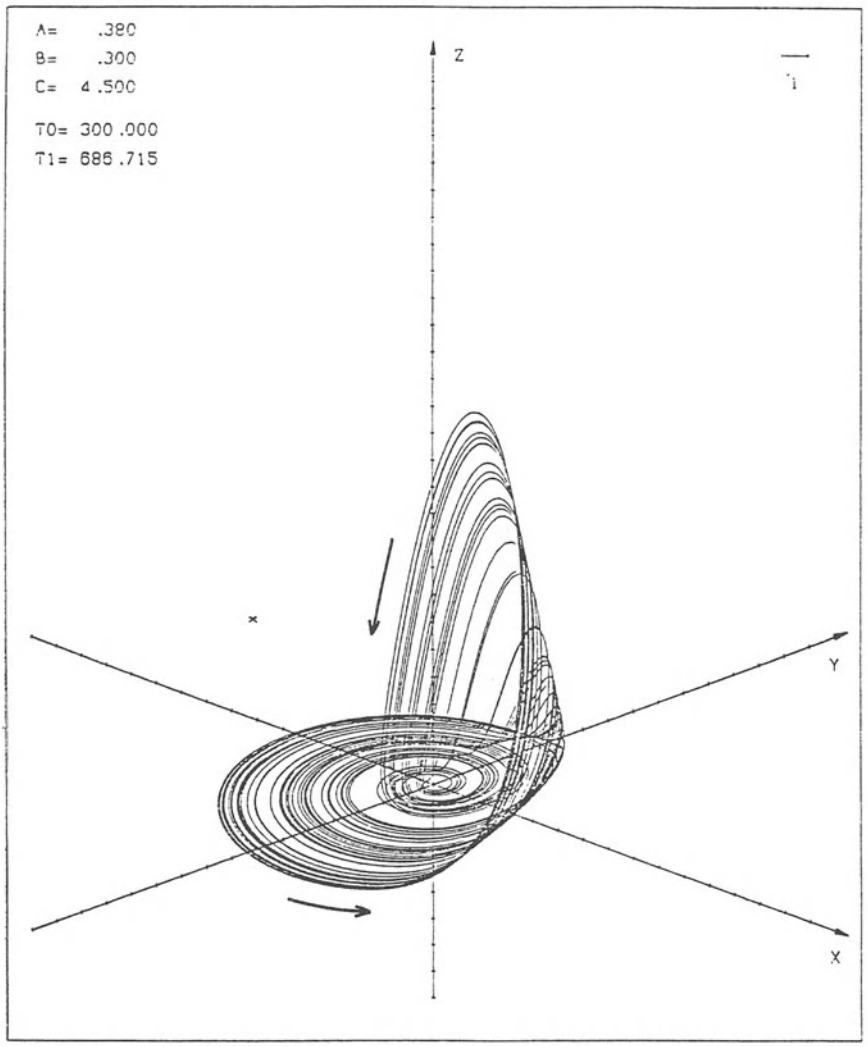


Fig. 6. Rössler attractor, displaying chaotic behavior of the chemical system (Equations 4.1, 4.2, 4.3).

where printing of the symbols has begun only after a sufficiently long time for transients to die out has elapsed. In summary, the results of the Subba Rao and the Nicolis show that a polymer has been generated in which we may observe long-range correlations, expressed by a high-order Markov chain. It can be checked that this sequence can be entirely rewritten by introducing the hypersymbols

$$\alpha = \mathrm{ZYX} \quad \beta = \mathrm{ZXYX} \quad \gamma = \mathrm{ZX} \tag{4.7}$$

the result for sequence (4.6) being then:

$$\ldots \alpha\, \beta\, \beta\, \alpha\, \beta\, \alpha\, \alpha\, \gamma\, \alpha\, \alpha\, \beta\, \alpha \ldots \tag{4.8}$$

This interesting property has to be attributed to the deterministic origin of the mechanism giving rise to the sequences. It suggests the existence of strong correlations in the succession of the symbols {X, Y, Z}, that is to say a high order Markov process. To check this, we may evaluate the statistical properties of the resulting sequence as follows. The system of equations ((4.1), (4.2), (4.3)) is integrated up to about 40 000 time units, generating in this way about 10 000 symbols and 3000 hypersymbols. Results are recorded from $t = 300$, in order to have all transients die out. Then we count the numbers of observed singlets (X, Y, or Z), doublets (XX, XY, etc.), triplets, etc. . . . The string presents some striking regularities. The number of formally possible "words" of length seven formed with the three hypersymbols α, β and γ is 2187. However, only 21 of these are observed in the numerical experiment. Moreover, for about one half of these the conditional probability of the last symbol, given the five preceding ones, turns out to be equal to 1: everything happens as if the system is endowed with a set of "grammatical" rules, followed automatically as a result of the dynamics. The dynamical system generating chaos acts as an efficient selector, rejecting the vast majority of random sequences and keeping only those compatible with the underlying rate laws. What is perhaps equally important, the irreversibility incorporated in these laws gives rise to a preferred direction of reading, and allows for the existence of attractors enjoying asymptotic stability and thus reproducibility. All this suggests a strong analogy with the production of biomolecules. Certainly no claim can be made that we have now understood the origins of biomolecules. Still, some progress has been achieved: "non-trivial" chains can be produced in far-from-equilibrium conditions, through the coupling with a chaotic reaction. Irreversibility has thus been encapsulated in a Markov chain. Here, non-equilibrium in a chaotic reaction has generated one of the essential aspects of biomolecules: information. Time has been transformed to information.

It is clear that all this cannot be the whole story, since meaningful information leads to the question: "information for what?". Biological information has to be defined in terms of the molecules involved in its management: nucleic acids *and* proteins; this meets the program of Eigen and Schuster [10].

We may conclude that much can be learned from theoretical and experimental study of polymerization under non-equilibrium constraints. We may only hope these results may encourage experimentalists to start studying such problems. We regret that due to lack of time, we could not envisage here other aspects of non-equilibrium chemistry, such as the problem of sensitivity to fluctuations near bifurcation points, which gives us the hope to produce selective mechanisms separating left and right structures, in spite of the very small energetic difference between these structures. As shown by recent work by D. Kondepudi [11], non-equilibrium may give rise to a new chemistry of weak forces.

Acknowledgements

This is a written version (revised by S. Pahaut) of a lecture given at the Interna-

tional Symposium on Molecules in Physics, Chemistry, and Biology, dedicated to Raymond Daudel.

References

1. I. Prigogine and I. Stengers: *Order out of Chaos*, New York, Bantam Books, London, Heinemann (1984).
2. A. Garcia: *Phys. Rev. A* **34**, 1454 (1986); M. Malek-Mansour and A. Garcia: *Ph. Rev. Lett.*, to appear in March, 1987; M. Mareschal and E. Kestemont, *Nature* **329**, 427 (1987).
3. P. Glansdorff and I. Prigogine: *Thermodynamic Theory of Structure, Stability and Fluctuations*, New York, Wiley (1971).
4. G. Nicolis and I. Prigogine: *Self-Organization in Non-Equilibrium Systems*, New York, Wiley (1977).
5. C. Nicolis and G. Nicolis: *Nature* **311**, 529 (1984).
6. A. Babloyantz and C. Nicolis: *Phys. Letters* **111A**, 152 (1985).
7. G. Nicolis, G. Subba Rao, J. Subba Rao and C. Nicolis: 'Generation of spatially asymmetric, information-rich structures in far-from-equilibrium system', *Working Paper* (1987), Service de Chimie Physique II, Université Libre de Bruxelles, CP 231, Campus Plaine, B-1050 Brussels, Belgium.
8. O. Rössler: *Ann. New York Acad. Sci.* **316**, 376 (1979).
9. P. Gaspard and G. Nicolis: *J. Stat Phys.* **31**, 449 (1983).
10. M. Eigen and P. Schuster: *The Hypercycle*, Berlin, Springer (1979).
11. D. Kondepudi and G. Nelson: *Phys. Rev. Lett.* **50**, 1023 (1983).

Entropy Variation and Molecular Organization

J. TONNELAT
Laboratoire de Biologie Physico-chimique, Université Paris-Sud, 91405 Orsay Cedex, France.

1. Introduction

When it is possible to calculate an entropy variation and when it is established that the entropy has increased, it is generally concluded that the disorder of the system has increased. This fact is frequently interpreted by saying that the number of degrees of freedom of the molecules has increased.

Inversely, when the number of ways in which a peptidic chain, say, can take shape is calculated, one expects to obtain a value of the entropy utilisable in thermodynamic relations by multiplying this value by Boltzmann's constant.

The origin of this idea can be traced back to the work of Clausius [1]. This author established that a new quantity which he introduced, the entropy, was constantly increasing in the *spontaneous* evolution of an isolated system. Clausius also stated, as a matter of course, that the disorder of a spontaneously evolving system increases. Therefore, this idea was very easily accepted by everyone at that time. Indeed, this seems only too evident when one looks at what has happened to men's undertakings left to their own fate! Clausius made this idea a scientific dogma by saying that an increase in disorder is linked to an increase in entropy. But everyone is aware that spontaneous appearance of order is also observed, e.g. in crystals. It was admitted that this appearance of order is compensated by an increase in disorder in other components of the surroundings. But how a general rule could be locally transgressed remained to be explained.

The idea of a tendency towards disorder was confirmed by Boltzmann [2] who, in the final part of his 'Lessons on Gas Theory', stated that a system has more chance of tending towards a disordered state because such states are more numerous than ordered ones. Now, this assertion is in contradiction with his fundamental hypothesis that not all states have the same probability. This inadvertancy reveals that the idea of a tendency towards increasing disorder is a preconceived idea, not the result of reflection.

If one admits, as do Boltzmann and all workers in statistical thermodynamics, that, in a spontaneous evolution, a system tends towards more and more probable states, one necessarily admits that it tends towards the most probable one or its neighbourhood. Once the system has reached this state, it will have a tendency to return to it if it is shifted away for any reason. This state is thus an equilibrium state. A state can be an equilibrium state only if there is compensation between the various forces which act upon its components. Now, the realization of a compensa-

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, 105—110.

tion between forces requires the components to be arranged in a particularly way: they cannot be arranged at random, in a chaotic manner. Therefore, Clausius's assertion of a tendency toward as great as possible a disorder is certainly not correct.

2. Order and Disorder

When can it be stated that a certain order exists in a system?

The existence of a spatial order is characterized by the fact that one can lay down a few rules governing the arrangement of the components in space.

When a system contains many components distributed in very different manners, many rules are necessary to describe their distribution. The system will be said to be very poorly ordered. If it is not possible to lay down any rules, the system is said to be completely disordered. There are many sorts of intermediate situations. It frequently happens that observers do not agree in saying that a system is more ordered than another. Nevertheless, when two systems contain the same components, observers generally agree in saying that the one in which the arrangement of the components can be characterized by the least number of rules is the most ordered one.

It may happen that, even for a system containing only a few components, a great number of rules are necessary to describe the system, because the connections between the components are difficult to define. In such a case, the system is said to be complicated. There is not always a clearcut difference between a complicated system and a disordered system. But, in most cases, the distinction is very clear. The wiring of a computer is at the same time very complicated and very ordered.

3. Perfect Gases

A perfect gas is a gas in which no interaction exists between the molecules except during their collisions, at which time they can exchange energy. The molecular distribution in space of a perfect gas at equilibrium is quasi-uniform. The dogma of the tendency towards disorder leads us to consider this distribution as the most disordered possible. A very ordered distribution would, for example, be a distribution in regularly spaced rows, as in a cubic crystal. The molecules of a gas cannot remain in any defined distribution since they are perpetually moving. The distribution they adopt is a constantly changing, quasi-uniform distribution. This distribution is the closest to that in a cubic crystal, as we shall see. This distribution in space cannot be shown on a plane, but the distribution projected onto a thin slide at a given time is displayed in Figure 1. A parameter of order can be chosen as being the sum of the distances of a point representing the projection of a molecule to the position of a point of the regular squaring, each point being considered only once. There are many ways of performing this operation. The one which gives the minimal value for the order parameter will be chosen. For

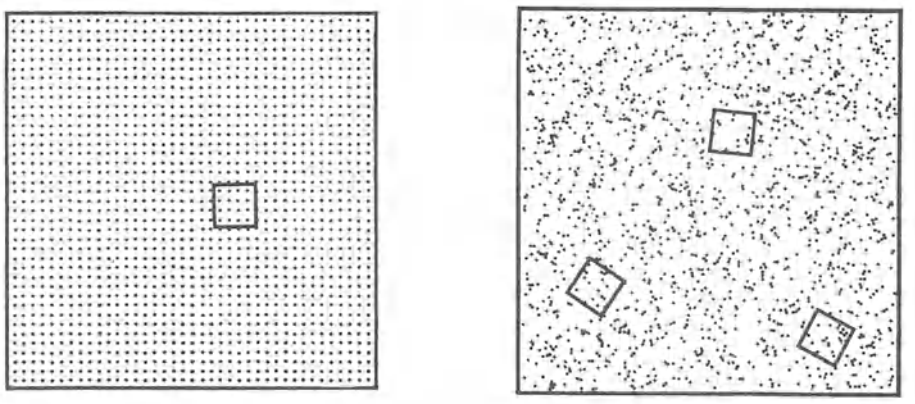


Fig. 1. *Left*: distribution of points in regular rows as in a crystal. *Right*: projection of the quasi-uniform distribution of perfect-gas molecules in a thin parallelepipedic slide. In a square whose side is 1/10th that of the great square, the same number of figurative points is observed for both distributions.

molecules which are perpetually moving, this sum will be minimal if there is, on the average, the same number of molecular figurative points as there are points in a tiny surface element in the network, i.e., if the distribution is quasi-uniform [3].

A quasi-uniform distribution is not a disordered distribution. On the contrary, this distribution is the most ordered possible for perpetually moving molecules.

4. Particles with Interactions in Fluid Surroundings

When molecules are in surroundings which permit them to move around and when long-distance interactions exist between them, those which attract each other tend to move closer together and those which repel move away from each other. There will thus be no tendency towards some quasi-uniform distribution. On the contrary, in a spontaneous evolution, there will be a tendency to the formation of what can be called structures.

Questions might be asked concerning the first appearance of interactions at the time of the Big Bang. But once interactions exist (following a fluctuation?), the origin of symmetry breaking need no longer be sought.

5. Increases in Complexity

When a system evolves, it sometimes happens that the nature of its components does not change and that their interactions remain the same. Such a system is called a conservative system. This system will tend towards its most probable state, that is, an equilibrium state, a state which therefore presents a certain order.

It frequently happens that, in a spontaneous evolution, components react chemically with one another, splitting or aggregating, and new interactions appear

between them. These transformations occur because they correspond to the passage to more probable states. The system does not remain conservative. Irreversible transformations take place in it [4].

If the new states are more probable, it is because the number of new components is greater, or because new energy levels become accessible, or because new interactions appear. The situations are then more varied: the system is diversifying. The description of these diversified situations will require more expressions than the former situation. The system becomes more complicated.

6. Entropy

We shall now return to the entropy concept.

In classical thermodynamics, this concept is originally derived from the study of the energetic yield of steam engines. The variation of entropy for a system is defined by the relation $dS = dQ/T$, where T is the absolute temperature, which corresponds to the mean energy of thermal agitation of the molecules, and dQ is the heat exchange with the surroundings at this temperature. No attention is paid to the variation of the distribution of the molecules in space. This relation is just as valid for a gas turbine as for a steam engine. It concerns only the energetic viewpoint [4].

In thermodynamics of irreversible processes, entropy is also defined by this relation without any reference to the distribution of the components in space.

For Boltzmann [2], who was the founder of statistical thermodynamics, entropy is a function of the distribution probability of the system components in space and between the energy levels, $S = k \ln W$, where W is the number of microscopic states corresponding to a macroscopic state. The microscopic states may differ in microscopic differences for the distribution of the components in space and between the energy levels. A macroscopic state is more probable than another one if it corresponds to a greater number of microscopic states. For establishing the relation $dS = dQ/T$, Boltzmann considers the case of a gas whose molecules are assumed to be quasi-uniformly distributed in space. Then, the variations in entropy concern only the possibilities of distribution between the energy levels.

In modern statistical thermodynamics, the entropy function takes fully into account the distribution in space. Entropy is defined by the relation $S = \Sigma\, p_i \ln p_i$, where p_i is the probability of each microscopic state corresponding to the macroscopic state of equilibrium. Everybody agrees that this definition is equivalent to Boltzmann's definition, $S = k \ln W$. The relation between an elementary variation in entropy and a heat exchange is deduced from the study of bringing two isolated systems, A and B, into internal equilibrium, under so-called reversible conditions (5, 6). To obtain equilibrium, entropy must present a maximum. *A priori*, an entropy variation must correspond to a variation in the number of ways of distributing the components both in space and between the energy levels. But, if the energy of the whole system is always the sum of the energies of the system components, on the other hand the entropy of the whole system can be the sum of

the partial entropies only if bringing in contact the subsystems does not induce modifications in the interactions between their components. In order that this condition be satisfied, the component distribution in space must remain unmodified. This distribution is characterized by a given number of parameters, called external parameters. A variation of these parameters may modify the entropy of each system. The relations between both variations may be very different, depending on the nature of the systems. For example, when a perfect gas confined in a given volume receives heat, the spatial distribution of its molecules is unmodified; when a protein is denatured by heating, the configuration of its peptidic chain is modified: charged groups are unveiled and the arrangement of water molecules is disturbed. A *general* condition for reaching the maximum of entropy can be obtained only if the variations in the distribution in space are not taken into account. Such a condition will concern only the variations in the distribution between the energy levels [3]. Usually, it is expressly notified that the external parameters are assumed to be constant, but the authors do not make clear that this condition implies that eventual variations of the distribution possibilities in space are not taken into account.

That part of the ln W value which results from the possibilities of distribution between the energy levels is a function of the energy of each system. In the imposed conditions, we have

$$\ln W_{ab}(U_{ab}) = \ln W_a(U_a) + \ln W_b(U_b).$$

The necessary condition for the entropy to have a maximum is that the sum of the derivatives of ln $W_{a/b}$ with respect to $U_{a/b}$ be zero, that is:

$$\frac{\mathrm{d} \ln W_a}{\mathrm{d}U_a} + \frac{\mathrm{d} \ln W_b}{\mathrm{d}U_b} = 0.$$

Remembering that the energy exchange is a heat exchange and that the heat acquired by one system is lost by the other, one obtains:

$$\frac{\mathrm{d} \ln W_a}{\mathrm{d}Q_a} = \frac{\mathrm{d} \ln W_b}{\mathrm{d}Q_b}.$$

One notices that if two systems are in equilibrium the value of this ratio is the same for both. This ratio is characteristic of the systems which can exchange heat at equilibrium. Everybody agrees that this ratio is equal to $1/kT$. Thus, one obtains

$$kT \, \mathrm{d} \ln W = \mathrm{d}Q,$$

which is the same as

$$\mathrm{d}S = \mathrm{d}Q/T.$$

Clausius's relation is obtained again, and this relation is established only for the variation of the distribution of the components between the energy levels. It concerns only the energetic viewpoint, as in classical thermodynamics.

In every case, a variation in entropy has a quantitative link only with variations

in the distribution between the energy levels (however, an exception must be made for the expansion of perfect gases; because of a coincidence due to their singular properties, the variation of the "entropy of configuration" has the same expression as the variation due to a work exchange). This conclusion is clearly expressed only by few authors [4, 7].

When an entropy variation is calculated for a transformation, almost by the intermediary of free energy, the relation $dS = dQ/T$ is, in fact, used: this relation concerns only the variations in the molecular distribution between the energy levels.

I think the notion of entropy should no longer be used [4]. Entropy has the dimensions of an energy without having the physical properties of energy. Energy can neither be created nor annihilated, while entropy can appear from nothing and disappear without leaving a trace. Entropy has the dimensions of an energy because Clausius, following Carnot [8], was led to consider several ratios dQ/T and compare them which each other. If he had come after Boltzmann, he might as well have considered the ratios dQ/kT. If entropy is divided by Boltzmann's constant k, a pure number is obtained, which I have called "complexity" [4]. Indeed, we have seen that what is increasing when a system is evolving is not its disorder, but its complexity. The fact that complexity is a pure number will help avoid dangerous misinterpretations.

It frequently happens that an increase in the number of ways of distributing the components among the energy levels corresponds to an increase in the number of components. In such a case, it is likely that the number of ways of distributing the components in space will be greater. This can be interpreted by saying that the number of degrees of freedom has increased. However, this number does not necessarily increase, and the amplitude of its variation has no numerical relationship to the entropy increase; on the other hand, an increase in this number may perfectly well correspond to the possibility of formation of ordered structures. It is thus clear that no link can be established between an entropy variation and a variation in the internal organization of a molecule or in the organization of a whole set of molecules.

References

1. R. Clausius: 'Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie'. *Ann. d. Physik u. Chemie Pogg.* **CXXV**, 400 (1865).
2. L. Boltzmann: *Vorlesungen über Gastheorie*. Barth, Leipzig. I. Teil (1896). II. Teil (1898).
3. J. Tonnelat: 'Sur l'incorrection de plusieurs raisonnements fondamentaux en thermodynamique'. *Fundamenta Scientiae* **3**, 339 (1982).
4. J. Tonnelat: *Thermodynamique et Biologie; I. Entropie, Désordre et Complexité*. Maloine (1978).
5. Berkeley Physical Course, vol. 5, *Statistical Physics*. Educ. Dev. Center, Newton, Mass. (1967).
6. R. Castaing: *Thermodynamique statistique*. Masson (1970).
7. D. Gabor: 'Light and information'. *Progress in Optics* **1**, 109 (1964).
8. S. Carnot: *Réflexions sur la puissance motrice du feu*. Bachelier, Paris (1824).

Molecular Evolution and Biological Self-Organization

H. ATLAN
Department of Medical Biophysics, Hadassah University Hospital, Kiryat Hadassah, 92120 Jerusalem, Israel.

The evolution of molecular structures in nature can be observed at two different levels. First, the prebiotic evolution, i.e., the set of reactions which are assumed to have taken place in the primitive soup on the Earth 5 billion years ago, leading to the synthesis of aminoacids and nucleotides and then to their polymerization in polypeptides and polynucleotides. These polymers, of course, are the ancestors of the macromolecules which carry the biological information in living organisms, namely, the proteins and nucleic acids. The second level is the organism itself: evolutionary changes can be observed on some macromolecules, essentially proteins, which are present in very old species as well as in recent ones. These changes consist in substitutions of aminoacids, so that a "distance" can be measured between two polypeptidic chains as the number of aminoacids by which they are different.

Now, the questions asked at these different levels, the prebiotic evolution and the biological evolution, are different. At the first level, the question is how the reactions in the primitive soup were able to synthesize self-reproducing information-carrying molecules with increasing specificity. These two features, reproduction and specificity (which go together with complexity or high information content), are necessary to characterize the first microscopic living organisms from which it is postulated that biological evolution could start. On the other hand, at the second level of evolution, that of the organism, the question is what is the relationship between the evolutionary changes observed in the molecules carrying the genetic information and macroscopic changes observed in the organism itself, i.e., roughly speaking the changes in phenotypic characters which accompany the appearance of new species.

It is well known that the accepted general framework to address these questions to is the neo-Darwinian theory of evolution: random mutations occur in the molecular structure of the genes and the natural selection operates on the mutants in such a way that the populations of organisms which are the better adapted to a given environment (i.e., have survival advantages such as the ability to reach food easily, resistance to aggressions and/or better conditions for reproduction) reproduce faster than the others.

This scheme was established in the twenties as the so-called general synthetic theory long before the discovery of the molecular structure of the gene and of the mechanisms of its expression. At the beginning of molecular biology it was expected that these discoveries would confirm the theory since they uncovered the

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, 111–126.

molecular substrata and the actual mechanisms of the mutations. In fact, the situation today is very different from what was expected and it is a bit ironic: neo-Darwinian models of evolution succeed in accounting for the prebiotic evolution, while this was not their concern to begin with since they were built long before the discovery of a possible molecular prebiotic evolution; on the other hand, they fail to answer the questions raised by the biological evolution itself, which they were meant to explain, so that other mechanisms such as genetic drift and molecular drive must be added to natural selection to account for the data on evolutionary changes. These additional mechanisms transform completely our picture of the dynamics of biological evolution.

As far as prebiotic evolution is concerned, the works of Schuster and Eigen [1], for example, show remarkably well how it is possible to successfully apply neo-Darwinian models and even to solve one of their classical difficulties: the definition of the adaptability or selective value is no longer circular as is often the case when survival of the fittest is the outcome of some property called adaptability (or fitness to an environment, or selective value), which itself does not have any definition other than the ability to survive. Here, this kind of circularity no longer exists, since the selective value of replicating molecules is defined in terms of chemical kinetics as a velocity of replication.

I cannot delve into the details of this theory now. Suffice it to say that either by analytical analysis or by computer life games, one can see how an optimum amount of mutations, i.e., of errors in replication processes, is necessary to produce molecules with increasing selective value. When applied to the combination of catalytic molecules, (the proteins), and of replicating molecules, (the nucleic acids), in what Eigen calls hypercycles (the proteins catalyze the replication of nucleic acids which govern the synthesis of the proteins), it is even possible to show how the system will eventually be stabilized with a selection of reactions which will necessarily lead to some kind of genetic coding. Thus, according to this model, although the present day genetic code is not the result of a pre-determined history and did not have, a priori, to evolve as it is today, nevertheless one sort of genetic code was necessary to appear as the result of interactions between catalytic polypeptides and replicating polynucleotides with selective pressures acting on the mutants produced by errors in replication.

My argument now is that whereas relatively simple neo-Darwinian models can thus account and be very useful for prebiotic evolution, they fail, at least in their original form, to answer the questions raised today by biological evolution. This is due to the fact that selective pressures can act on the phenotype of the organisms only (and not directly on their genes), whereas what is transmitted from one generation to the other is the genome and not the phenotype. Darwin stated his theory by saying that "individuals who have any advantage over others have a better chance to survive and to transmit their own type." What is definitely wrong with this view is that the individuals do not transmit their type. Even the very fact of sexual reproduction prevents them from doing so, since they transmit only half

of their genotype, and their offspring will be new original individuals. According to neo-Darwinian theories, following Mendel's laws and molecular genetics, this difficulty was overcome by taking into account the probability of parental gene associations in the offspring; but this was possible only on the assumption that selective advantages of phenotypes would be the expression of a selective advantage of the transmitted genes. A kind of dogma which was stated at the beginning of molecular biology supported this assumption: it was summarized into a formula: one gene — one enzyme — one character. In other words, a one-to-one correspondence was assumed between a gene and a phenotypic character. This assumption was based on some observations in bacteria, but it turned out to be more an exception than the rule, since in most cases, several genes, sometimes far apart from one another in the genome, are involved in the determination of a character, and the same gene may be involved in the determination of several characters. In fact, gene regulation in eukaryotic cells has appeared to be very different from what it is in bacteria. More generally, it is only recently that the hierarchical nature of biological organization with several levels of integration has been taken into consideration as an important factor for the models of evolution. This, together with recent unexpected findings concerning the rates of changes in the molecular structures at the genetic level, has created a real challenge to our understanding of the mechanisms of evolution. A very good review can be found in a paper on Hierarchy and Evolution by Eldredge and Salthe [2], published two years ago.

This challenge to the classical synthetic theory of evolution began with the discovery of neutral mutations. The direct measurement of mutation rates in the genes of evolving species showed that there is no reason to believe that the survival of the mutant genes is oriented by selective pressure. Just the opposite is correct: the data are easily accounted for if one assumes that the non-lethal mutations are neutral, i.e. all the mutations compatible with the life of the organism have an a priori equal selective value. This conclusion, originally drawn by Kimura [3], was based on the observation of mutation rates constant in time, whereas the changes in phenotypic characters in a given evolutionary line of species occur with an increasing rate, meaning that they are more and more favored by environmental selective pressures, therefore not "neutral". Thus, in addition to natural selection, a second mechanism called genetic drift was considered to account for a kind of random fixation of mutant genes in a population, with no higher selective value than others.

The discovery of repetitive DNA with multiple copies of genes and of non coding DNA sequences in eukaryotic cells, has made the phenomenon even more complicated, whether it is interpreted as so called "selfish DNA" [4] or as a genome redundancy to be used eventually as a source of genome complexification. Under the influence of the idea of "selfish genes", a third mechanism has been proposed, called "molecular drive" [5] to account for large modifications of genomes involving collective cooperative effects of large numbers of genes. More-

over, the discovery of "transposons" [6] has shown the existence of a large mobility in the genome with continuous turnover of genes within a given population. Finally, this diversified set of data has not been simplified by the discovery of master genes postulated for a long time but identified only recently, first in drosophila, then in other species. These genes regulate the expression of batteries of other genes through a common short DNA sequence (called "homeobox") which codes for a common polypeptide able to repress and derepress several genes so that they are active or inactive together. The situation is only more challenging because, as W. J. Gehring [7], who discovered these common DNA sequences, wrote:

> ... even if the homeobox containing genes regulate many other genes, they must be regulated themselves. Finding out how the regulators are regulated will be another significant accomplishment, and one that could lead to the identification of the factors in the egg cytoplasm that provide the positional information.

In other words we are back to the notion of a delocalized "program" [8] not limited to DNA sequences and extended in fact to the whole cellular machinery. This includes the cytoplasmic network of coupled biochemical reactions and transport mechanisms which create some spatial heterogeneity — e.g. in the form of concentration gradients — to serve as a support for positional information.

All these facts recently discovered at the molecular level of organisms make it more difficult to understand how evolution takes place at the phenotypic level of the species since, for example, a model of evolution today must be able to account for the different rates of changes observed at the different levels. More and more, evolutionary biologists have come to realize in the last few years that one has to seriously take into consideration the known hierarchical organization of organisms and populations and try to understand the kind of interactions which take place between different levels of organizations. As S. J. Gould wrote recently:

> A new and general evolutionary theory will embody this notion of hierarchy and stress a variety of themes either ignored or explicitly rejected by the modern synthesis: punctuational change at all levels, important nonadaptive changes at all levels, control of evolution not only by selection, but equally by constraints of history, development and architecture — thus restoring to evolutionary theory a concept of organism. [9]

What I would like to do now is to show the kind of logic with unexpected and apparently paradoxical results which is involved in the analysis of multilevel hierarchical organizations. For this purpose, I will present a very simplified model suggested by the study of random networks made of Boolean automata. These networks simulate highly nonlinear interactions between large numbers of genes and interactions between levels. Of course, the simulation is still too crude to be directly applicable to actual biological structures whether they be molecular or organismic; however, it does exhibit generic properties which help in understanding similar properties of the evolutionary processes. In other words, these

networks do not pretend to be realistic models of what exists in nature, but rather idealized simulations of conspicuous and unexpected properties which happen to be generic to large classes of models and therefore lose some of their mystery. The two properties I plan to discuss are (1) a so-called buffered stability, which helps to build models of evolution with different rates of change at different levels, and in particular the appearance of new species in a discontinuous way according to the theory of punctuated equilibria set forth by Eldredge and Gould [10]. (2) Functional self-organization, which enables one to understand the emergence of new meaningful functions.

Let me first briefly recall how these networks are built and how the simulations are run.

We are dealing with automata connected in a matrix (Figure 1) in such a way that each receives two inputs from two of its four neighbors. The matrix is closed on itself as on a torus so that the end of a row is connected to its beginning, and the same holds true for a column. The automata work in parallel, in a synchronous

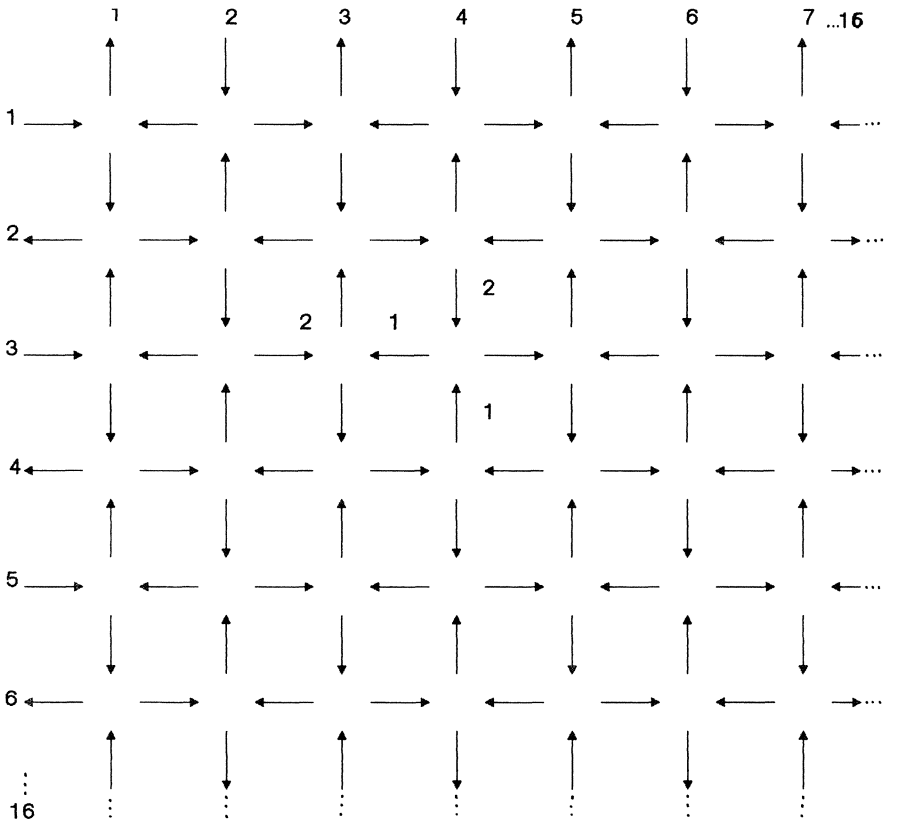


Fig. 1. Connection matrix. At each node an automaton computes the two inputs according to one of the functions of Figure 2. The inputs are numbered 1 and 2 for right and left, respectively, and for below and above.

way; i.e., at each discrete time interval, each automaton computes its state as a function of its two inputs and sends its two outputs equal to that state. The same procedure is reiterated one time interval after the other, and the state of the network, i.e., of all the automata, can be printed for each time, starting from given initial states.

Now the automata are boolean; i.e., they can be in only one of two possible states, 0 or 1, and they compute their state from their binary inputs by one boolean law. These boolean laws or binary functions are represented in Figure 2, where one can recognize the known logical functions (AND, OR, NAND, NOR,

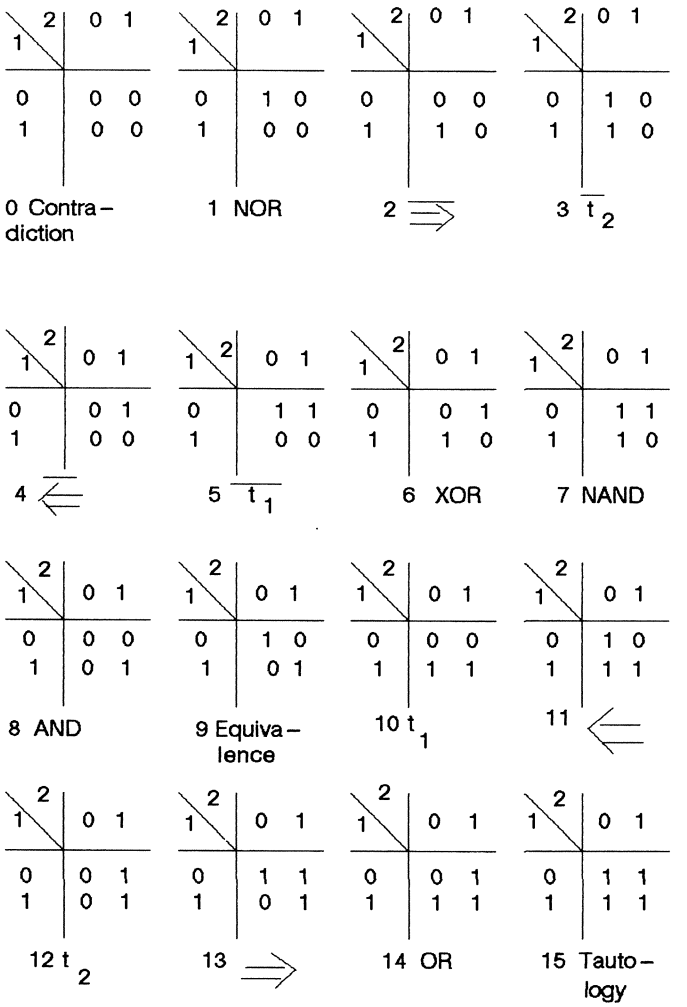


Fig. 2. Boolean functions of two variables are numbered 0 ... 15 according to the decimal representation of their table of values. Table (d b) is numbered: $1.2^3 + b.2^2 + c.2 + d$.
(c a)

XOR, etc.). The properties of such networks, their logical relationship with Turing machines, their dependence on the number of connections and on the classes of Boolean functions have been studied extensively by S. Kauffman [11], who initiated this research, and by our group, of whom I want to name G. Weisbuch, F. Fogelman, E. Goles, J. Salomon, E. Ben Ezra and S. Hoffer [12, 13, 14]. I now will merely summarize the results which are necessary to understand the following. Each automaton in a network is randomly assigned one of these Boolean laws except for the two constant functions which have been eliminated because of their too strong stabilizing properties. This given random distribution of laws, together with the connections, defines the microscopic structure of the network. Initial conditions are also defined randomly by the initial state of each automaton, so that the overall initial appearance of the network, i.e., its initial macroscopic structure, is a statistically homogeneous distribution of 0 and 1 over the matrix. Then, in a typical simulation experiment after a few tens of time intervals, the state of the matrix is checked periodically, and one observes that the network evolves rapidly towards a mixture of steady state and oscillating states which characterize a macroscopic spatiotemporal typical structure. As an example, Figure 3 shows the kind of macroscopic structure which emerged from one simulation: starting from a random homogeneous distribution of the initial states, the network defined by the distribution of laws shown in Figure 3a evolved towards a non-homogeneous macroscopic structure characterized by the appearance of subnets. Now it appears no longer homogeneous (Figure 3b) since it is divided into some areas where all the automata have become stable (denoted s in the figure), either in state 0 or in state 1; and these areas are surrounded by other subnets where all the elements undergo periodic changes, (denoted P in the figure), i.e., they go through a relatively short sequence of states which indefinitely cycles on itself. In other

1	3	8	13	3	14	13	12	3	8	10	11	9	14	12	2
11	3	12	13	7	2	5	2	14	6	9	2	7	1	6	7
4	7	8	6	13	4	12	14	9	13	1	3	9	5	14	5
4	13	3	5	2	13	12	8	4	1	9	11	1	14	7	5
12	12	10	8	12	13	12	1	10	8	6	14	3	3	14	4
12	5	6	12	10	12	12	10	3	13	2	9	13	8	6	10
4	10	2	9	2	13	7	3	3	11	9	3	5	13	8	8
10	4	7	13	13	14	1	14	7	5	6	13	5	11	8	9
8	13	8	10	1	3	3	11	8	2	8	14	11	11	9	3
4	8	10	12	8	2	14	4	13	7	9	14	3	3	13	14
6	10	13	5	14	7	4	10	13	7	9	12	13	4	10	10
2	9	13	7	13	6	5	6	14	5	9	11	6	3	9	9
10	14	9	3	3	13	8	2	14	13	1	6	9	10	2	6
5	8	12	7	9	5	9	7	8	3	8	10	14	9	11	12
2	12	13	6	11	7	13	5	11	6	11	12	5	4	1	2
10	6	11	4	8	5	4	5	13	3	3	9	10	3	7	8

Fig. 3a. Example of network structure. The numbers refer to Figure 2. The connection matrix of Figure 1 must be superimposed on it with the beginning of lines and columns in the upper left corner.

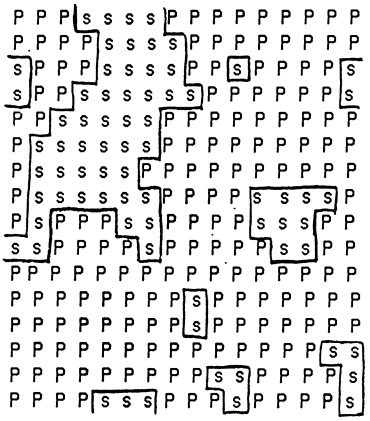


Fig. 3b. Display of the network of Figure 3a after it has reached its limit cycle, starting from a randomly set initial condition matrix. P: periodically oscillating; s: stable in 0 or 1. In this particular simulation the cycle length was equal to 96 time units.

words, the network has reached a limit cycle characterized by a spatio-temporal structuration of stable subnets separated by periodic oscillating ones.

One of the interesting properties of these networks is the kind of stability they exhibit towards changes in initial conditions and perturbations from their limit cycle. It is a stability which resembles the so-called buffered stability described by Austin and Crook [15] in ecological systems. The structure of the limit cycle in general is not assymptotically stable in the strict sense; i.e., the network after a perturbation or starting from slightly different initial conditions, does not go back to a limit cycle exactly identical to the previous one. However, if one measures a distance between the different limit cycles of a network, one finds that they are gathered in clusters, which means that the perturbed network goes back to a new limit cycle which, in general, is not very different, roughly speaking, from the previous one. Of course a given network may have several groups of neighboring limit cycles, as in dynamics where the attractors correspond to minima of a potential, one can have several local minima clustered within a large valley separated from other such valleys corresponding to other groups of local minima. In any case, the macroscopic pattern exhibited by a network in its limit cycle is mostly dependent upon its microscopic structure, i.e., the distribution of Boolean laws on the connection matrix.

This kind of behavior can be used for our model of evolution in the following way (Figure 4).

If the binary automata are considered to represent genes in an activated or non-activated state, this kind of dynamics may be viewed as a very crude simulation of a collective behavior of such interacting genes leading to macroscopic features appearing at the level of the whole network. In other words, if the microstructure of the network, i.e., the distribution of the boolean laws and

POPULATION	Reproduction Rate of Genomes regulated by an Adaptability Function of a Phenotypic Character (e.g. maximum length of limit cycle)
PHENOTYPE	Macroscopic Spatio — temporal Structure of the Limit Cycle
GENOME	Automata Network with a Specific Microscopic Structure (i.e. a Boolean law distribution and connections)

Fig. 4. Model of three levels involved in evolutionary theory by three corresponding levels in the observation of Boolean networks.

connections, would represent a genotype, its macroscopic final structure when it has reached its limit cycle would represent the phenotype produced by the interactions between the genes.

Thus, in a very crude sense, a Boolean network would be a model of phenotype viewed as a collective macroscopic structure and behavior determined by the dynamics of a large number of interacting genes. Of course, the model is very far from reality, as far as the complexity of feedback regulatory interactions is concerned between the actual phenotype and genotype of any organism. However, as oversimplified as it is, this model contains the basic elements of a hierarchical biological organization as described by Eldredge and Salthe [2] with at least three levels: in addition to the genome and the individual organism, the population is now considered with a law of multiplication of individuals, i.e. of network genomes, based on an adaptability or fitness parameter defined by some macroscopic property of the phenotype. In other words, a population is defined by a given number of similar networks which reproduce themselves according to a given dynamics, with a small given rate of mutations, i.e., of changes in their microscopic structure. At the same time, the rate of reproduction is itself a function of some phenotypic character. This character may be interpreted as a parameter of adaptability or fitness of the phenotype in a given environment, but this interpretation is not the only possible one: this parameter may also be a mere reproduction rate characteristic of a given genome. Since different mutants may result in different values of this parameter, their reproduction rates will be different and if resources are limited so that the total number of individuals is limited, competitions can occur between different genomes or groups of similar genomes.

A model of such population dynamics has been built by G. Weisbuch [16, 17] with very simple genomes containing only 6 genes with 4 possible alleles; i.e., networks with only 6 automata and 4 possible Boolean laws for each. The population dynamics obeyed to the set of equations represented in Figure 5 where, for a given genome population P_i, a_i is the reproduction rate parameter which is a collective phenotypic character and was chosen simply as the maximum cycle length produced by a given network. Then, starting from a few small populations of almost identical networks, i.e., of mutants different from one another by only one gene, the growth of these populations for large numbers of generations was computed with new mutants appearing and reproducing themselves at a rate modulated by their adaptability parameter, measured, as just mentioned, by their maximum cycle length. The results of one such simulation are represented in Figure 5.

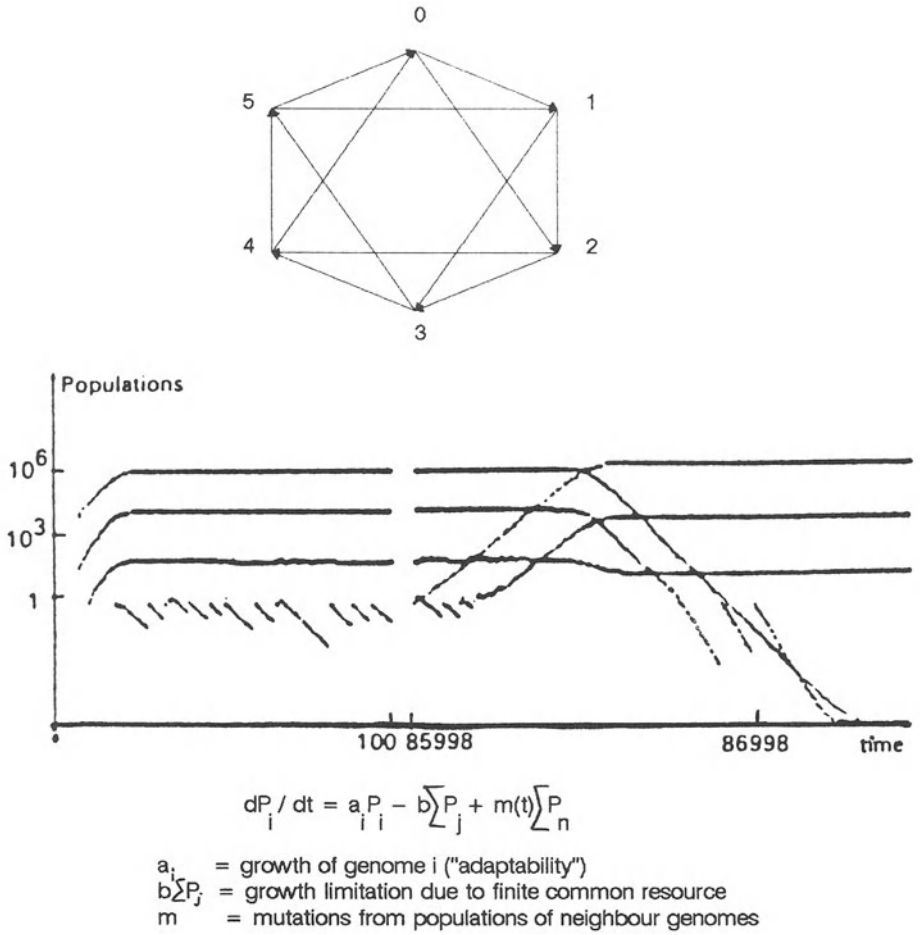


Fig. 5. (From G. Weisbuch [17]) (Top) Connexion graph of the network made of 6 Boolean automata. The Boolean functions are limited to the four laws AND, NAND, OR, NOR. (Bottom) Population dynamics.

In spite of the oversimplification of the model, one can find in these results at least two of the puzzling observations in the study of biological evolution. First, the existence of punctuated equilibria, as described by Gould and Eldredge [10], i.e., the existence of long periods of stability in the distribution of the different populations characteristic of a species, in spite of a constant mutation rate, followed by rapid changes in this distribution, characteristic of the disappearance of the species and its possible replacement by new ones, again with the same constant mutation rate. One can see in the model new mutants appearing but disappearing rapidly, because their phenotypic character is not better than that of the existing populations, until one of them has a better one and starts very rapidly to proliferate and replace the others.

The second observation, related to the first one, is the possibility of an accelerated rate of change of the phenotypic character, concomittant with a constant rate of mutations. As mentioned above, the finding was at the origin of the first puzzling problems encountered by the classical neo-Darwinian theory when Kimura measured mutation rates at the level of individual protein structures coded by individual genes in different species belonging to the same evolutionary line. An evolutionary line relates a set of species which appeared one after the other with a change in one obvious phenotypic character; e.g., the relative size of the brain in primates or the length of the feet in the line which led to the horse. In such a line the rate of change of the character increases in time as can be expected if one assumes that this change is oriented by selective pressures. However, the puzzle resulted from the fact that the mutation rate — which represents the actual changes taking place in the material which actually reproduces (i.e., the genome and not the phenotype), was found to be constant as if no selective pressure was acting on it. This finding was at the origin of the notion of neutral mutations and genetic drift, difficult to reconcile with the classical neo-Darwinian schemes based on the one gene — one character formula.

More generally, what we can learn from these models is that the hierarchical organization with non-linear interactions in the effects of one level on another is enough to explain the existence of completely different rates of evolution at the different levels, without giving up the causal determination of the behavior of one level by the other.

The last part of my discussion will be devoted to the presentation of another interesting property of these networks which is of help in approaching the question of *meaning*. It is a very difficult question, as well known in information sciences as in biology.

On the one hand, it is known that information theory does not deal explicitly with the meaning of information and that semantics in linguistics is what is most difficult to formalize. In biology, on the other hand, the translation of DNA molecule structures into protein structures is an instance of the transmission of information in a channel, in the precise technical sense of Shannon's information theory, and here also the question of meaning is left outside the scope of the

picture. However, this meaning exists: in fact, it is the output of the channel, the functional enzymatic activity of the proteins. More precisely, if we consider the cooperative activity of the genes, the meaning of the genetic information is the whole of physiological structures and functions which define the phenotype where this information is expressed. In other words, the phenotypic characters — which may not be limited to one individual organism in a population, as pointed out by Dawkins — are the meaning of the genotype.

Therefore, the behavior of cooperative assemblies of individual units, i.e., of molecules viewed as automata, raises the question of their biological meaning in the form of possible physiological functions *appearing as an emerging property* at a different level of organization. This is due to the fact that the meaning of information in natural organizations, not planned and manufactured by man, is very different from what it is in artifacts. When we see something which looks organized to us, like a machine, we ask about its finality, which task it performs, which problem it solves, what it is good for. In most artifical machines — except perhaps in some toys — the answer is quite simple because the finality of the machine has been set up at the beginning, in the planning and building of the machine. It is in relation to this known goal to be achieved that the meaning of the organization is appreciated, both as a whole and in its parts. Signals or connections, or different blueprints, make sense or are meaningless according to their contribution to the more or less efficient achievement of the goal. That is not the case in machine-like molecular organizations, built by nature, even with the help of natural selection, where the goal, if ever it exists, is not known and may change from one moment to the other.

That is why the meaning of organization in natural systems must be defined as an emerging property to the eyes of an observer, according to self-created criteria for what makes sense and does not make sense. And, of course, these self-created criteria may be very unexpected to the eyes of the observer, and even completely far-fetched a priori if one assumes this observer to be a rational human being who would never have imagined such meanings to begin with. However, a posteriori, they make sense. I would like to show you very briefly how our random Boolean networks help to illustrate this point by exhibiting emerging unexpected functional properties of pattern recognition. In addition to their structural self-organizing properties which I have described, these networks, after having reached their limit cycle, can be used to simulate processes of pattern recognition. The patterns they are able to recognize are classes of binary sequences like those represented in Figure 6. Some of these sequences have a hidden structure, not periodical and even not deterministically completely defined, but which nevertheless exists. A Boolean network is able to react in some way to the sequences which possess this structure and not to the other sequences. What is interesting is that the criterion for this distinction and classification is itself an emerging property of the network evolution until it has reached its limit cycle. Going back to an example of our Boolean network in its limit cycle (Figure 7), the process goes as follows.

```
a    0 0 1 0 1 0 0 0 1 0 0 0 0 0 0 1 1 0 1 0 0 0 0 1 1 0 0 0 1 0 0 1
b    1 0 0 0 1 0 0 1 0 0 0 0 1 0 0 1 0 0 1 0 1 0 0 1 0 0 1 0 1 0 0 0
c    0 1 0 0 1 1 0 0 1 0 0 0 1 0 0 0 1 0 1 1 0 1 0 0 0 1 0 0 0 1 0 1
d    1 0 1 0 0 0 0 1 1 0 1 0 1 0 0 1 0 0 1 0 1 0 0 1 1 0 1 0 1 0 0 0


     * 0 * 0 * 0 0 * * 0 * 0 * 0 0 * * 0 * 0 * 0 0 * * 0 * 0 * 0 0 *
```

Fig. 6. Out of the four strings represented, only a, b, d belong to the class of pseudo-periodic, partially random strings defined by the criterion *0 *0 **0 * and were recognized by the pathway described in Figure 7.

We choose an element of the network and, instead of letting it go through its normal sequence of states — either the same state if it is stable or a short cycling given sequence if it is periodic —, we impose on it an external perturbating sequence of states which have nothing to do with its normal, unperturbed states.

Then we look at what happens to the other elements. Most of them, in general, except for the immediate neighbors, are not perturbed. They remain stable if they were stable, periodically oscillating if they were oscillating. This is another instance of the kind of robustness and stability of these structures. However, some elements, sometimes far away from the perturbed one, become oscillating while they were stable in the absence of perturbation. Also, and maybe even more surprisingly, some elements which were oscillating, become stabilized by the perturbating sequence. It is this latter phenomenon which we chose to consider as a simulation of a recognition process: we have a channel of recognition with the element where the perturbating sequences are inserted being the input of the

```
   1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 (1)
 →0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  1
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0      *0*0*00*
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
   0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
 (1) 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0
```

Fig. 7a. Effects of permanent perturbating strings imposed on the element (2, 1) of the network represented in Figure 3 after it has reached its limit cycle. Elements labelled 0 did not change their status (stable or periodically oscillating) as a result of the perturbations. Elements labelled 1 were oscillating in the unperturbed limit cycle and stabilized by the effects of perturbating strings belonging to a class of pseudo-periodic, partially random strings as a, b, d in Figure 6.

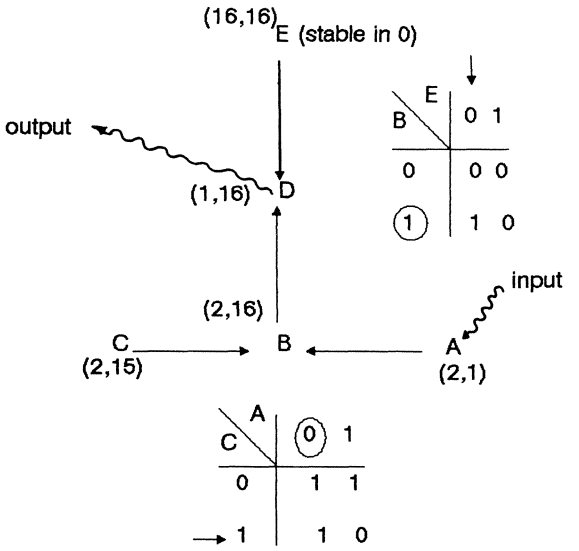


Fig. 7b. Magnified pathway of the recognition channel between the input element (2, 1) and the stabilized output element (1, 16). (Remember that the matrix is closed on itself as on a torus.) It shows how the criterion (*0*0*00*) for recognition by stabilization is generated: C is oscillating with a cycle length 8 and following sequence of states in the cycle: 0 1 0 1 0 1 1 0. D is stabilized in 1. Since E is in 0, for D to be in 1, B must be in 1. In order for B to be in 1, it is enough that when C is in 1, A must be in 0. When C is in 0, A is indifferent. As a result, in the input sequence, a 0 is necessary to match a 1 in C: * 0 * 0 * 00 *. The other bits (*) are indifferent.

channel and the element which is going to be stabilized or not depending upon the kind of sequence presented to the input, being the output. If it is stabilized, for example a light will be switched on, meaning that the sequence has been recognized; if it has not, it means that the sequence has not been recognized. For example, in one of our simulations, the criterion for recognition which appeared to us a posteriori and, of course, was not programmed explicitly at the beginning, is a pseudo-periodic or partially random structure represented in Figure 7, where only the 0 are determined and the stars can be indifferently 0 or 1 at random. By applying this criterion to the binary strings shown in Figure 6 one can see easily that it fits the strings a, b, and d but not c; which means that our network recognizes a, b, and d but not c.

The detailed mechanisms of these recognitions by stabilization of otherwise oscillating automata is a kind of resonance where a class of perturbating strings resonate with oscillating properties of a pathway of connected automata between the input and the output element, characteristic of the network in its limit cycle. It is quite a complicated phenomenon and more detailed analysis can be found in [13] and [14].

In general, a random boolean network will generate, together with its spatio-temporal macroscopic structure, a lot of pathways able to work as such channels

to recognize classes of binary strings by stabilization of an element at the end of the pathway when the string is inserted at its beginning. And, of course, the definition of the class sets up the criterion for the recognition: it simulates in a crude way what happens in a functional recognition system, like the immune system e.g., or the C.N.S., i.e. the definition by the system itself of which sequence will make sense and which will not. Since this definition is itself a result of the emerging macroscopic structure of the network in its limit cycle, one can see that the function of classification and recognition is also an emerging property of the collective dynamics of the network. It is in this sense that we can say that these simulations provide simplified models for creation of meaning in a kind of functional self-organization where what emerges from the collective behavior is not only a spatiotemporal morphology, but also a property which resembles some basic biological function. From then on, it is easy to speculate about a selective value of this function in a given environment which would give some selective advantage to the genome which produced it.

However, the main point of this discussion is that a genome is not a gene, and this is even more so in reality than in our models, since our model does not take into account the diversity introduced by sexuality. In our models, the genomes are assumed to be able to reproduce themselves as individual genes. In reality, only individual genes replicate, whereas the genome of each individual is not reproduced in its offsprings. Therefore, to state that a given genome as a whole produces a phenotypic character which has selective advantages amounts to saying that a given individual with a character which has a selective advantage may give to its genes some selective advantage in the population. This does not mean, however, that within other individuals the same genes will necessarily produce the same properties and have the same advantages: their context of expression being different, the resulting phenotype will be different. Again, this means that phenomena which look paradoxical from the point of view of classical neo-Darwinism must be expected simply as consequences of interactions between levels of organization, that is, the molecular one, the individual, and the population in evolution.

References

1. M. Eigen and P. Schuster: *The Hypercycle: a Principle of Natural Self-Organization*, Springer-Verlag (1979).
2. N. Eldredge and S. N. Salthe: 'Hierarchy and Evolution', *Oxford Surveys in Evolutionary Biology* **1**, 184—208 (1984).
3. M. Kimura and T. Ohta: *Theoretical Aspects of Population Genetics*, Princeton University Press (1971).
4. R. Dawkins: *The Selfish Gene*, Oxford University Press (1976).
5. G. Dover: 'Molecular Drive: a Cohesive Mode of Species Evolution', *Nature* **299**, 111—117 (1982).
6. G. M. Rubin: 'P-transposable Elements and their Use as Genetic Tools in Drosophila', *Trends in NeuroSciences* **8**, 231—233 (1985).

7. W. J. Gehring: 'The Molecular Basis of Development', *Scientific American* (Oct. 1985), 153—162.
8. H. Atlan: 'Two Instances of Self Organization in Probabilistic Automata Networks: Epigenesis of Cellular Networks and Self Generated Criteria for Pattern Discrimination', in *Dynamical Systems and Cellular Automata*, J. Demongeot, E. Goles, M. Tchuente (Eds.), pp. 171—186, Academic Press (1985).
9. S. J. Gould: 'Is a New and General Theory of Evolution Emerging?', *Paleobiology* **6**, 119—130 (1980).
10. S. J. Gould and N. Eldredge: 'Punctuated Equilibria: the Tempo and Mode of Evolution Reconsidered', *Paleobiology* **2**, 115—151 (1977).
11. S. A. Kauffman: 'Behaviour of Randomly Constructed Genetic Nets: Binary Elements Nets', in *Towards a Theoretical Biology*, vol. III, G. H. Waddington (Ed.), Edinburgh Univ. Press, pp. 18—37 (1970).
12. H. Atlan, F. Fogelman-Soulie, J. Salomon and G. Weisbuch: 'Random Boolean Networks', *Cybernetics and Systems* **12**, 103—121 (1981).
13. H. Atlan, E. Ben Ezra, F. Fogelman-Soulie, D. Pellegrin and G. Weisbuch: 'Emergence of Classification Procedures in Automata Networks as a Model for Functional Self-Organization', *J. Theoret. Biol.* **120**, 371—380 (1986).
14. H. Atlan: 'Self Creation of Meaning', *Physica Scripta* **36**, 563—576 (1987).
15. M. P. Austin and B. G. Crook: 'Ecosystem Stability: a Result From an Abstract Simulation', *J. Theoret. Biol.* **45**, 435—458 (1974).
16. G. Weisbuch: 'Modelling Natural Systems with Networks of Automata: the Search for Generic Behaviours', in *Dynamical Systems and Cellular Automata*, J. Demongeot, E. Goles, and M. Tchuente (Eds.), pp. 293—304, Academic Press (1985).
17. G. Weisbuch: 'Systèmes Désordonnés et Comportement Générique', in *Jeux de Réseaux*, F. Fogelman-Soulie (Ed.), C.N.R.S. Paris, pp. 81—89 (1986).

Modelling and Esthetics of Molecular Structures

Spatial Modelling and Simulation of Organized Macromolecular Structures by Means of Factor Analysis of Electron Microscopy Pictures

CLAUDE GAUDEAU
Laboratoire de Bio-Informatique, C.N.R.S. and U.E.R. de Médecine de Tours, 2 bis, Boulevard Tonnèllé, 37032 Tours Cedex, France.

MARIANNE BOSMA
Société de Bio-Informatique et de Bio-Technologie, 45, Rue Emile-Zola, 37000 Tours France.

JEAN-PAUL ROSSAZZA
Ecole Nationale Supérieure d'Electrotechnique, Electronique, Informatique et Hydraulique, 2, Rue Ch.-Camichel, 31071 Toulouse Cedex, France.

0. Introduction

In order to reconstruct a spatial modelization of the structure and behaviour of a self-organized molecular system, several steps must be followed.

1. ISOLATION OF THE MOLECULAR SYSTEM

We will consider two types of molecular systems:

i — the microtubules of the living cells [1—4];
ii — the hemocyanin molecule, a copper molecule containing oxygen carriers, which are common among arthropods and molluscs [5—8].

2. ACQUISITION AND PREPROCESSING OF THE MOLECULAR DATA

They are involved in the molecular assembling process: pictures from electron microscopy or molecular immuno-electron microscopy, X-ray crystallography, immuno-electrophoresis, etc....

The digital *preprocessing* takes into account the transfer modulation function of the acquisition system: image restoration, filtering and contour detection [9—14].

3. REDUCTION OF THE INITIAL MOLECULAR DATA

Necessary to obtain elaborated and structured data: for instance, compression procedures and spatial reconstruction of the whole molecule from a series of electron microscopy pictures will be concerned and will be performed by using factor analysis and correspondence analysis [15—17].

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, 129—157.

4. DEFINITION OF THE MOLECULAR MODEL

I.e. determination of the macro-molecular and molecular subunit structures with their active sites and the manner in which the structured and reduced data will be integrated to obtain the desired results: simulation of the molecular assembly and organized system [9, 15, 16].

5. PROCESSING OF THE STRUCTURED DATA

Analysis of the "distance" between the results of the model and the molecular data: these are known as the "object"; elaboration of the way in which the data will be restructured or modified to reduce the distance "*object—model*". Recent developments in artificial intelligence and expert systems would be useful instruments to improve these last two steps.

1. Isolation of the Molecular System

We must recall that in order to observe a molecular system, the system must be isolated by different biochemical methods.

The methods, used in systemics, only take into account the parts of the system which may be biochemically controlled at the input level and the parts which may be observed at the output level.

However, these methods are not complete, because in biomolecular systems potentialities or latent properties in the isolated system exist which are not observed (Figure 1.1).

We can isolate the subunits of the hemocyanin molecules or the tubulines α and β in microtubules, and observe them with X-ray crystallography for instance, but

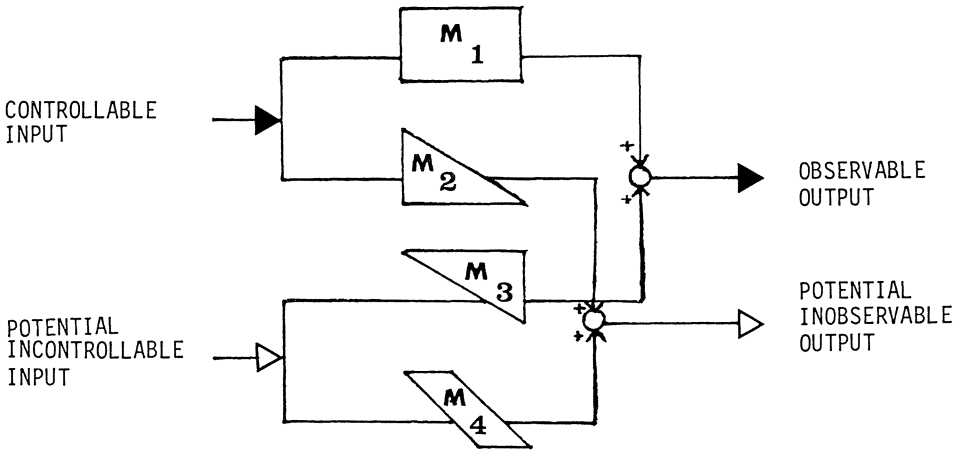


Fig. 1.1. Molecular system representation (Redon et al., INFAC, *Fuzzy information, knowledge representation and decision analysis*, Pergamon Press, p. 31, 1983).

we cannot observe latent properties at this level and by this means, which will emerge at higher levels of self organization. Although these properties cannot be observed at first, they may appear when several systems are associated together (whole hemocyanin molecule, microtubule network during mitosis, etc. . . .).

We can then say that there is an "emergence" of new potentialities.

According to the molecular-system classification of Polansky [9, 10], two classes exist:

— *Ordered systems*, in which the potential information can be more redundant than specific (from the biochemical point of view) and in which the order measure can be expressed by means of negative entropy;
— *organized systems*, in which the potential information is much more specific than redundant.

Two types of molecules can be studied: the microtubules, which are an example of ordered systems, and the hemocyanins, which are an example of organized systems [1—4].

The hemocyanins, which are more organized systems than the microtubules, are built up from polypeptide dimer subunits which are themselves already organized systems and are assembled in order to form complex macromolecules with more or less chemical specificity [5—8].

1.1. MICROTUBULES

Microtubules are cylindrical hollow tubes of proteins with an external diameter of 24—28 nm and an internal diameter of 15 nm.

The length varies approximately from one to ten nanometers (Figure 1.2).

A transversal section shows that, usually, 13 protofilaments of the tubuline polypeptides are aligned in a row and are arranged side by side around a central empty core (Figure 1.3).

More complex structures, such as ciliary axonomenes, consist of a sheet of nine doublet microtubules arranged in a ring around a pair of single microtubules.

At the base of the ciliar structure is a centriole: the cinetosome, which consists of a sheet of nine triplets of microtubules arranged in a ring. Molecular image processing, as we shall see in Section 2, has considerably increased our knowledge of the part played by cellular microtubules as cell transportation [21a, b, c].

1.2. HEMOCYANIN MOLECULES

The hemocyanins are high molecular weight copper proteins which serve as oxygen carriers [5—7]. They are found, freely dissolved, in the hemolymph of arthropods and molluscs.

The native molecules encountered can be considered to be built of successive dimerizations of building blocks, containing 2, 6, 12, 24 and 48 subunits with the same kidney shape, but with different biochemical properties (Figure 1.4). The

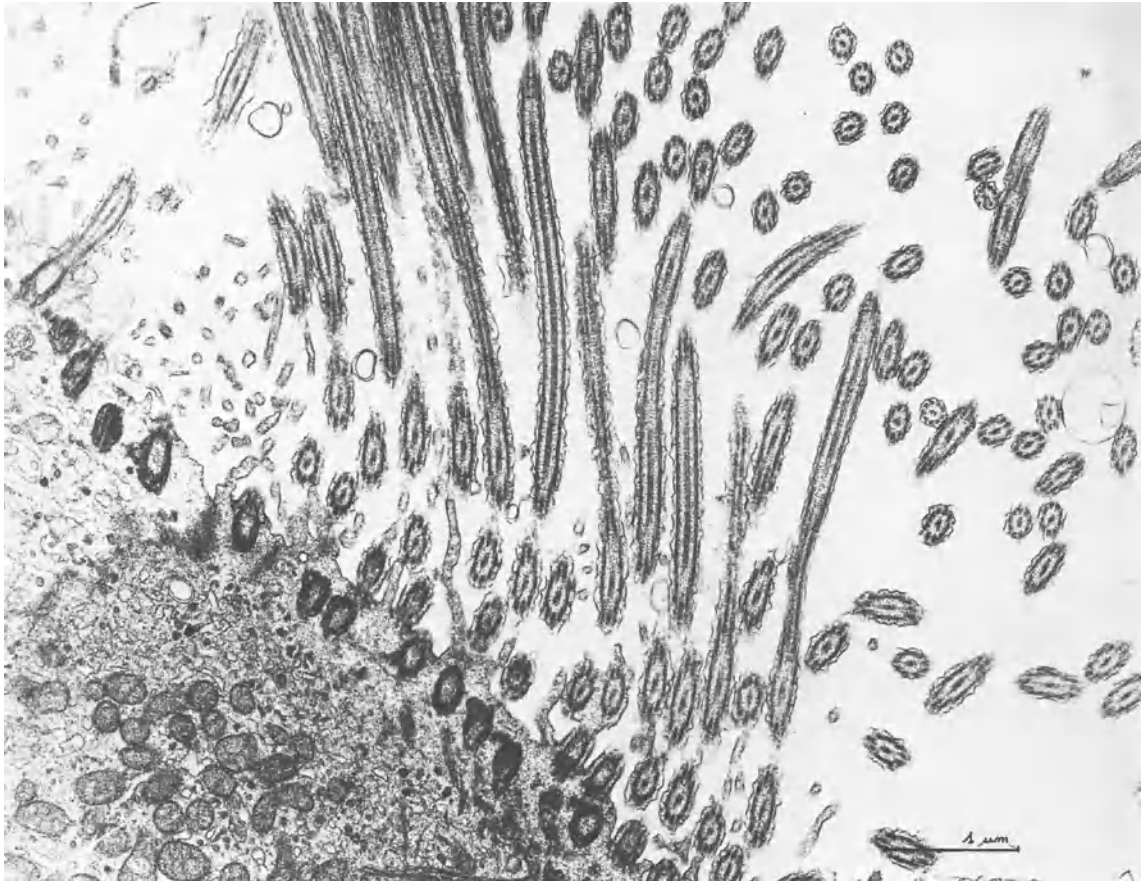


Fig. 1.2. Apical pole of a ciliar bronchic cell, in which can be observed: — basal corpuscules; — ciliar tiges; — axonems in longitudinal, tangential and transversal section: these differents views allow one to envision and to reconstruct the molecular structure in 3-D. (bronchic biopsy of a child; B. Arbeille, Unité de Microscopie Electronique, Tours).

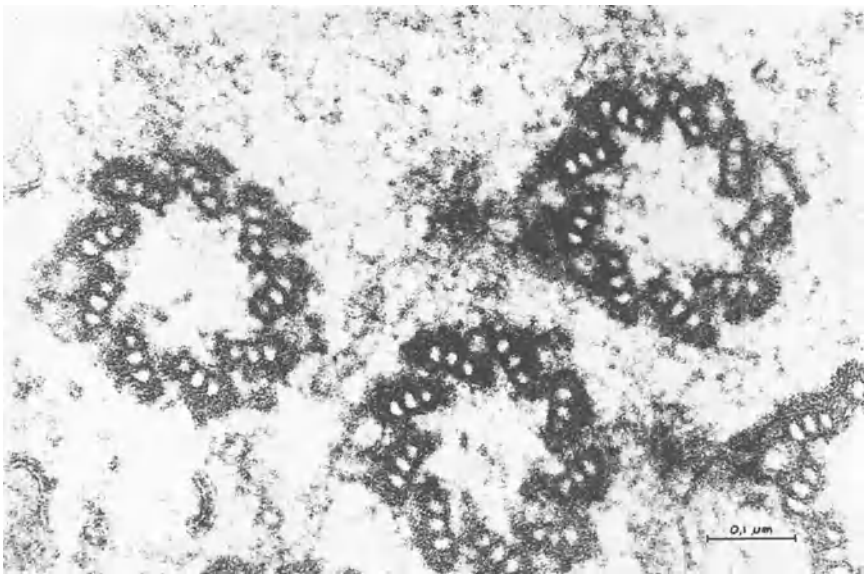


Fig. 1 3 a Transversal section of a basal corpuscule (or cinetosome) in which can be observed: — 9 × 3 triplets of microtubules A, B, C (the microtubule A is the nearest to the cylinder axis and has a circular section, the other two have a C structure), — nexin bridges between the microtubules triplets (B. Arbeille, Unite de Microscopie Electronique, Tours).

Limulus Polyphemus (*LP*) *hemocyanin* has 48 subunits containing eight immunological types of subunits (LP I, II, IIA, IIIA, IIIB, IV, V, V—VI) with different polypeptide chains [5, 6].

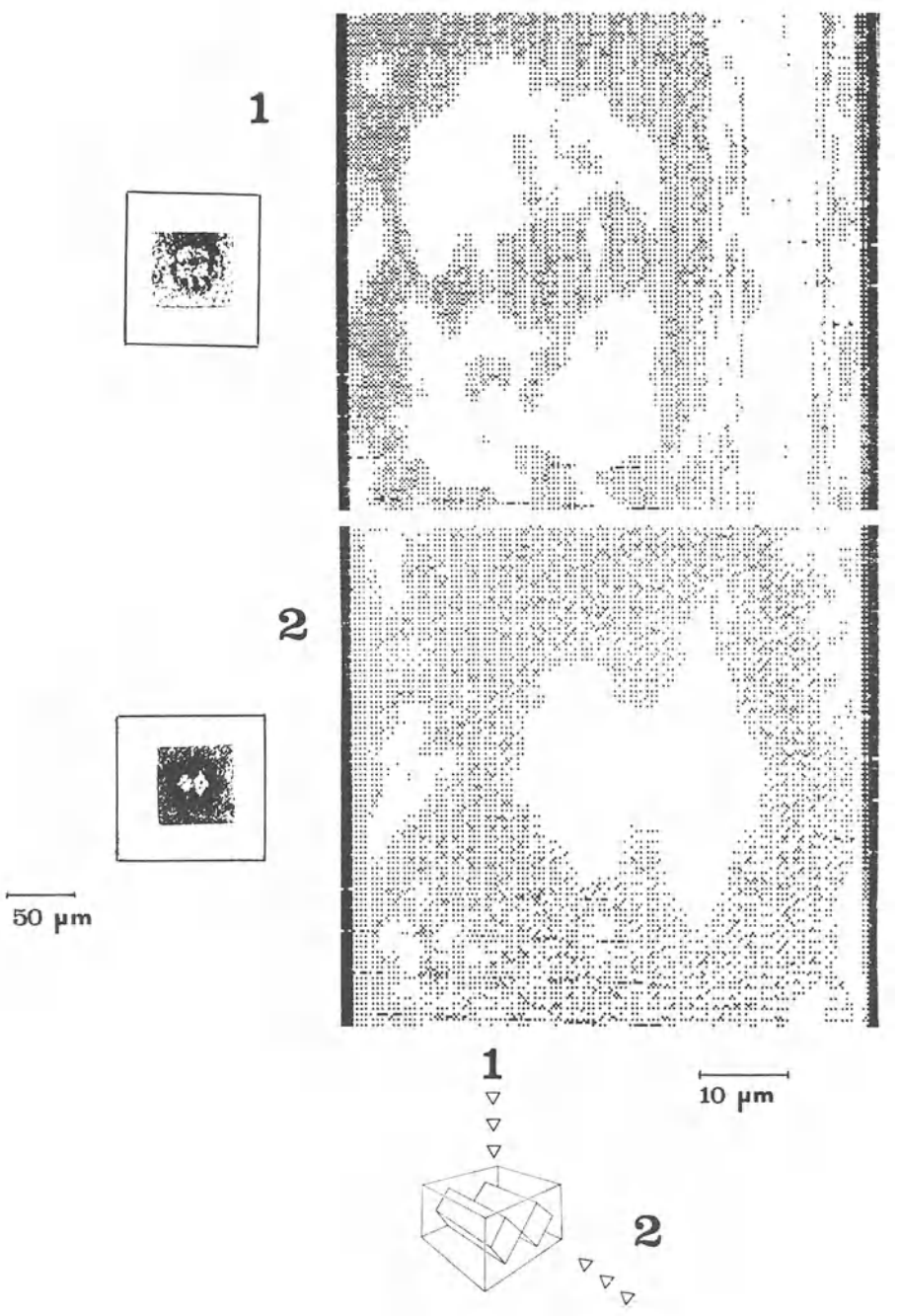
The *N*-terminal amino acid sequence of the various subunits or fractions have been determined [3, 4] and if two subunits have distinct polypeptide chains, an identical *N*-terminal sequence can result in additional evidence of homogeneity [8, 20].

2. Acquisition and Preprocessing of the Molecular Data

Different optical scanner systems have been developed on microcomputers (SBB 007 with video camera; SBB Lynx system (1024 × 1 photodiode matrix)).

Precision depends principally on the sampling rate and on the number of grey levels but it must be noted that great precision leads to a rapid saturation of the microcomputer's memory.

Preprocessing operators are transformations which filter the undesirable features and/or emphasize the informative features [17]. Preprocessing is a selective loss of the information of the picture. One of the first problems consists of optimizing this selective loss. We can consider three kinds of preprocessing operators:



'2X12, MERIC MODEL OF HEMOCYANIN AND ANDROCTONUS AUSTRALIS

Fig 1 4 Electron micrographs of soluble complex resulting from binding of Fab fragment specific for isolated subunits to *Androctonus australis hemocyanin* (bar is 50 nm) View of hemocyanin molecules labeled with anti-5A Fab and anti-3A Fab + anti-3B Fab These views represent the assemblage of two dodecamers (2 × 12 mers) (Lamy and coworkers, Laboratoire de Biochimie, Tours, *Biochemistry* **20**, p 1853, 1981)

— integration operators
— differentiation operators
— recursive operators

In order to reduce the noise, one must make use of integration operators, and in order to reinforce the necessary information, one must make use of differentiation operators. Optimization is an equilibrium between these two tendencies.

We must also discriminate between:

— context-free (or invariant) operators; and
— context-dependent operators.

Context-dependent operators themselves are either:

— global context-dependent; or
— local context-dependent (see 2.2).

If $M(i, j)$ is the digitalized picture and $Q(a, b; t(M(i, j)))$ an index function defined over a rectangle such that

$$-n < a < +n$$
$$(n, m) \in N$$
$$-m < b < +m$$

an output image $R(i, j)$ is formed by discrete convolution of the input function according to the relation, for a context-dependent operator $Q(h, l; M(i + h, j + l))$:

$$R(i,j) = \sum_{k=-n}^{k=+n} \sum_{l=-m}^{l=+m} Q(k, l; M(i+k, j+l)) \quad M(i+k, j+l) \tag{II.1}$$

The index function can be represented by a rectangular mask $(2n + 1 \times 2m + 1)$ whose values are functions of the local context of the picture.

An image may be subject to:

— electrical sensor noise,
— photographic grain noise,
— channel error,

which usually appears as isolated pixel variations that are not spatially correlated. For these types of noise context-free operators (integration operators) are generally applied.

2.1. CONTEXT-FREE OPERATORS

2.1.1. *Integration Operators*

This noise generally has a higher spatial frequency spectrum than that of normal

image components because of its spatial decorrelatedness. Hence, simple lowpass spatial filtering can be effective for noise smoothing [12].

The following masks have been tested:

$$\begin{array}{ccc} & & 0012100 \\ & 11111 & 0123210 \\ 111 & 12221 & 1246421 \\ 121 & 12321 & 2369632 \\ 111 & 12221 & 1246421 \\ & 11111 & 0123210 \\ & & 0012100 \\ \text{I} & \text{I}' & \text{II} \end{array} \tag{II.2}$$

with the relation:

$$R(i,j) = \frac{\sum_{K=-n}^{K=+n} \sum_{L=-m}^{L=+m} Q(k,l) M(i+k, j+l)}{\sum_{K=-n}^{K=+n} \sum_{L=-m}^{L=+m} Q(k,l)} \tag{II.3}$$

The use of masks of large dimension makes the digital picture more homogeneous and attenuates the fast transitions of grey levels (contours and often features).

Different authors [15—17] have compared the advantages and the disadvantages of different symmetric shapes for these masks.

Image sums can be classified as integration operators. This type of image transformation has provided evidence of molecular transport along the axon microtubules [21a—c].

2.1.2. *Differentiation Operators*

Differentiation operators have been widely applied to biomedical image processing for edge enhancement (generally without regard to edge direction).

Gradient operators and Laplacian operators have been widely applied [15].

As for integration operators, image subtractions can be classified as differentiation operators (Figures 2.1, 2.2).

2.1.3. *Recursive Operators*

Recursive filtering [15] is based on a recursive relationship between input and

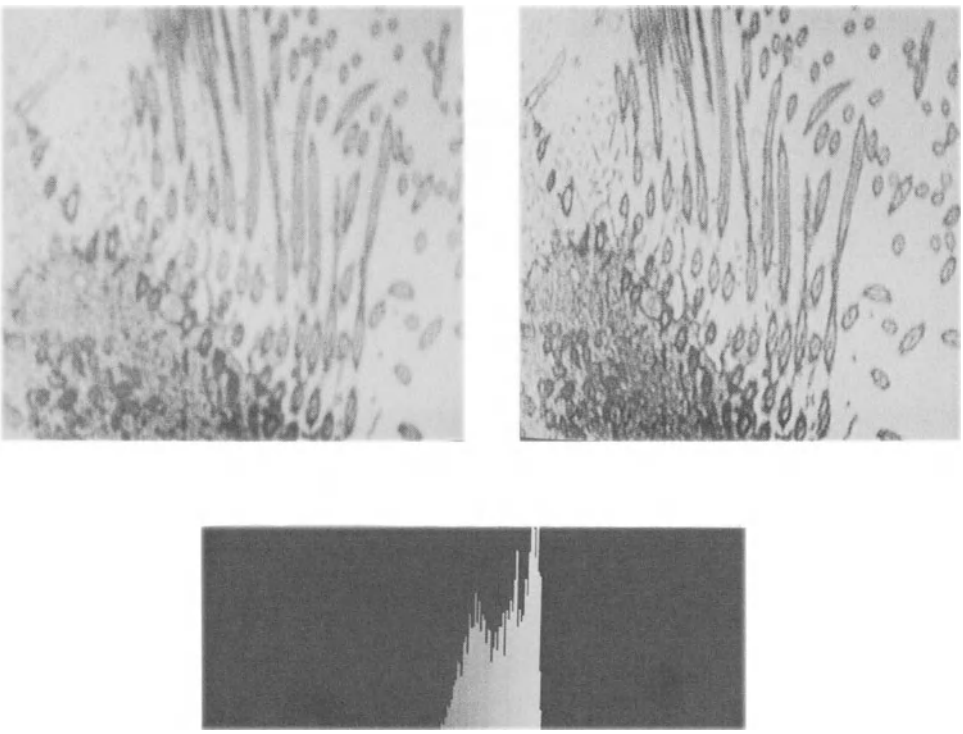


Fig 2 1 Processed electron microscopy picture obtained by using differentiation-operator, high-pass filter at left original digital picture, at right transformed picture (512 × 512 pixels), below grey level histogram (F Pathier, Societe Automatisme et Avenir Informatique, Mennery)

output images. The 2-D recursive filtering relation is defined by

$$R(l_1, l_2) = \sum_{J_1=1}^{J_1} \sum_{J_2=1}^{J} A(J_1, J_2) M(l_1 - J_2 + 1, l_2 - J_2 + 1) - \tag{II.4}$$

$$- \sum_{k_1=1}^{K_1} \sum_{k=2}^{K_2} B(k_1, k_2) R(l_1 - k_1 + 1, l_2 - k_2 + 1)$$

Recursive 3 × 3 and 5 × 5 filters have been applied advantageously to pictures (filters I, I′, II)

2 2 CONTEXT-DEPENDENT OPERATORS

Two classes of operators must be considered.

— global context-dependent
— local context-dependent

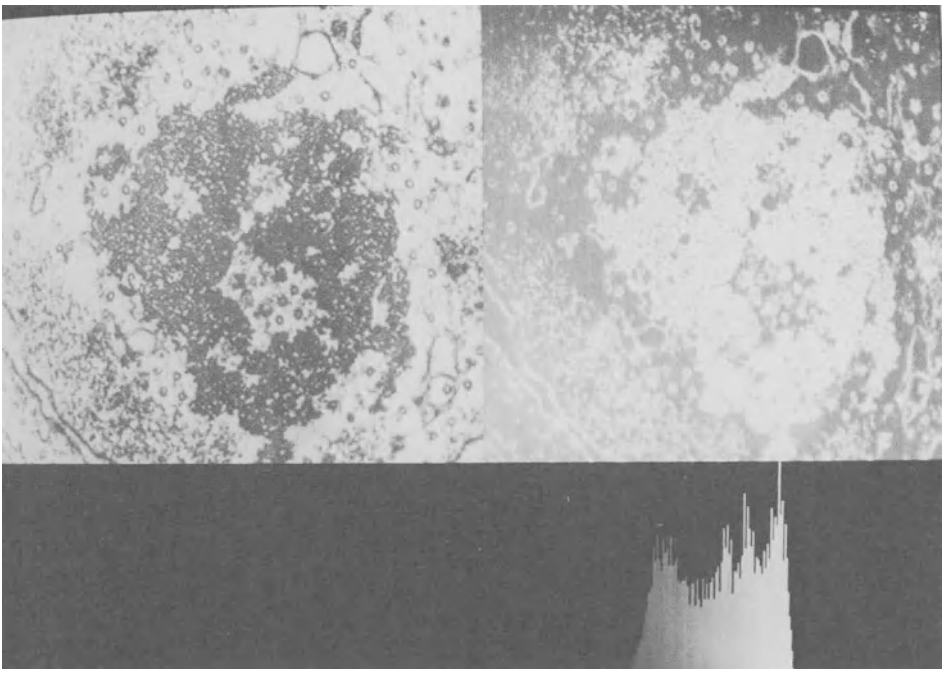


Fig 2 2 Image complementation at left original digital picture, at right transformed picture (512 × 512 pixels), below grey level histogram Black level around the microtubules shows empty area of cell structures, probably to permit the molecular transport along the microtubules during mitosis (F Pathier, Societe Automatisme et Avenir Informatique, Mennery)

2.2.1. *Global Context-Dependent Operators*

From equation II.1, we can see that the index function used in the transformation $Q(a, b; t(M(i, j)))$ is a function of the value of the pixel image. In a global context- dependent operator we will take into account the statistical properties in the picture (image 1st, 2nd, ... moments). These moments will permit us to evaluate the characteristics of the mask from the image auto-correlation function.

2.2.2. *Local Context-Dependent Operators*

The best example is the contour-follower automaton [14].

The automatic contour detection simulates the human visual perception, a domain which requires a great level of integration in the human brain, principally if the contour is a texture shade (texture difference detection automata). The automaton is functioning according to principles of structured recognition: in other words, by using a reference grammar the pertinent information of a picture appears as a closed and connected structure whose automaton output will result in an octal chain.

2.2.3. *Principle*

The pertinent information of a picture appears as a structure (a form) which relies on a non-specific element.

This assemblage must be described as an adapted formalism.

The theory of language brings together tools whose purpose is to describe these forms (sentences). The recognition of these forms can be accomplished by verifying the belonging of an unknown sentence to the language engendered by one of the reference grammars.

3. Reduction of the Initial Molecular Data

3.1. COMPRESSION OF MOLECULAR PICTURES

3.1.1. *Principle*

The compression algorithms [16] operate on 2-dimensional pixel blocks obtained with segmentation procedures of various sizes. Each image block, after transformation, is coded by a vector. Generally two modes of coding are available:

— *The fixed compression mode*

$N < k_1$;

— *The fixed error mode*

RMS $< k_2$.

3.1.2. *CHATUVERDI Algorithm Steps*

The CHATUVERDI algorithm consists of three discrete processing steps.

1st Step. After segmentation, the fixed image blocks $N \times N$ are operated upon by the unitary transform: the cosine transform which allows the image energy to be packed in a few low-ordered transform coefficients $(N - k) \times (N - k)$. To take advantage of the transformed energy packing properties of the fixed-image blocks, the sample must be sufficiently small enough in order to take into account statistical variations or unhomogeneity over the molecular picture. This is particularly true in the case of texture analysis.

2nd Step. The even, cosine transform is performed on each of the image blocks (Figure 3.1). The base function of the even, cosine transform is a class of discrete Chebyshef polynominals. Generally the energy of an image tends to concentrate towards the lower spatial frequencies.

Other transformations are also used, such as Hough's [15], which takes account of different orientations of molecules around various axes.

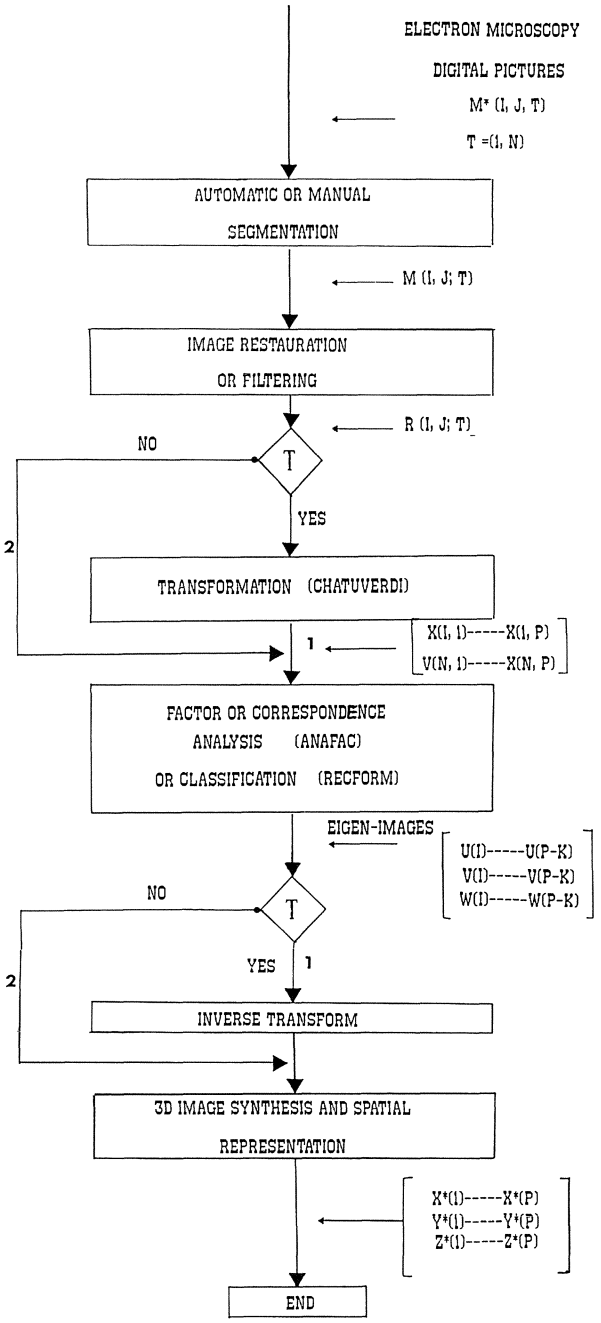


Fig. 3.1.

3rd Step. The two-dimensional transformed-coefficients matrix is arranged in a one-dimensional array by reading the matrix along the cross diagonal.

In this ordering, the transform coefficients show a trend of monotonical decrease in amplitude. Now, the molecular data is ready to be processed in the factor analysis programme.

3.2. FACTOR AND CORRESPONDENCE ANALYSIS

3.2.1. *Objective*

The high radiation sensitivity of molecules such as hemocyanins or microtubules forces one to use low electron exposures in the recording, and the signal to noise ratio is low [6, 25]. To increase the statistical significance of the molecular image, averaging a large number of individual images is an established technique, principally for molecules arranged in periodic structures such as crystals or helices. But biological macro-molecules have more orientational freedom as compared to the crystals or helices, and it was necessary to differentiate between the various molecular projections and rotations prior to operating. The principal method used was to determine the main orthogonal directions of the inter-image variance and to calculate the image coordinate in a system spanned by these newly determined axes.

The authors have generally used the correspondence analysis to obtain a large reduction in the total amount of data image. Each molecular image is coded with 64 density values and is characterized by a few factorial axis coordinates, which themselves are classified [25].

Thus the hemocyanin: images have been partitioned into four classes on the basis of clustering of the data in a two-dimensional system spanned by the first two axes.

3.2.2. *Method*

Consider the matrix $X(I, J)$ where each component can be either:

i. the transform coefficient J of an image I, or:
ii. the measured density value of each pixel J of an image I;

$J = (1, P)$, P = number of image pixels,
$I = (1, N)$, N = number of molecular images.

Row number I contains the densities of all P pixels of the image number I.

One can perform either factor (F.A.) or correspondence (C.A.) analysis of molecular pictures, in several ways:

(1) use of a transformation (such a CHATUVERDI transform) for data compression before application to the C.A. programme;
(2) direct correspondence analysis of each picture. Each molecular picture will thus by represented by one eigenvector (Figure 3.1).

3.2.3. *Interpretation*

The coordinate system of the R^N space, $X_1, X_2, \ldots, X_N$, is rotated, and possibly also shifted, so that the first unit vector (U) points towards the direction of the maximum inter-image variance (or dissimilarity between images).

The second unit vector (V) is perpendicular to point U in the direction of the maximum remaining interimage variance. The new coordinate system (U, V, W), calculated with eigenvalue and eigenvector analysis, is then adapted to the shape of the image cloud, i.e., to the properties of the image collection.

The first few unit vectors (U, V, W) describe the most significant part of the variance present in the molecular image set; the higher unit eigenvectors describe less and less of the total interimage variance and correspond to the noise [25].

Since the unit vectors in the rotated system correspond to a decreasing amount of interimage variance, the first few vectors characterize the most important differences in the image set. The unit vectors of the coordinate system (corresponding to a point) represent the eigenimages from which the original image set can be composed by linear transformations; the original images are represented by a few components (approximately 8, as opposed to, for example, 4096 density values).

In the Rossazza method, the total number of coordinates equal the number of transform coefficients selected.

The unit vectors of the new coordinate system represent the transformed eigenimages from which the original image set can be composed by linear transformation after inverse transformation (Figures 3.2—3.4).

For this texture analysis, it seems that the cellular texture around the microtubules is less compact than in other parts, maybe in order to permit molecular transports in the cell during mitosis (Figure 4.6b) and in the axon.

4. Definition of the Molecular Model

In the first part of this paper spatial representation of biological molecules and macro-molecules from electron microscopy has been considered [18, 19, 21, 22].

4.1. MOLECULAR ASSEMBLAGE REPRESENTATION

4.1.1. *Assemblage Rules*

The assemblage of dimers is realised in a plane. Every dimer becomes integrated into a block, which is formed before by other dimers. So, we exclude the fact that the blocks can connect with each other. These rules are at the basis of the structure of the software [3].

4.1.2. *Assemblage Method*

The method used involves the Freemann code (Figure 4.1).

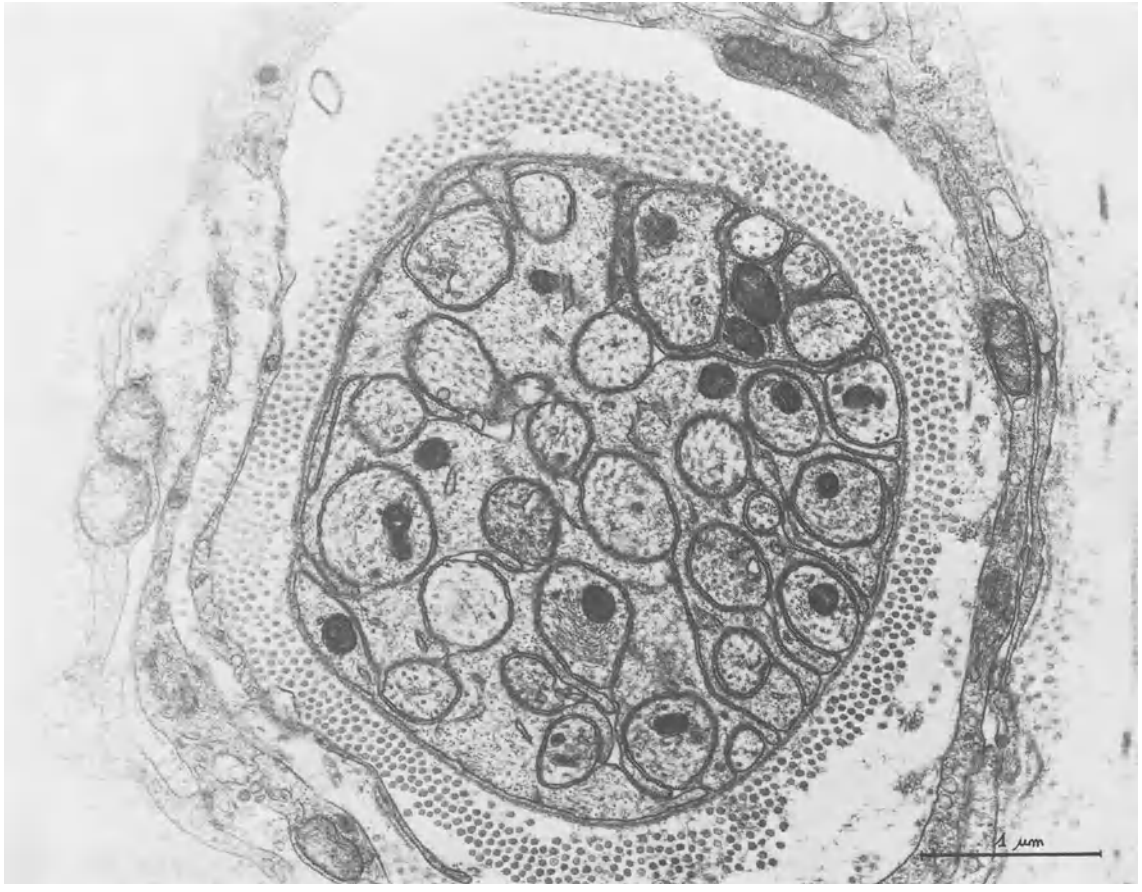


Fig 3 2 Transversal section of a peripheral bronchic nerve from part of which digitalisation and segmentation have been performed (by 2AI) to study the cellular environment of microtubules Microtubules serve to transport the different cellular constituents (B Arbeille, Unite de Microscopie Electronique, Tours)

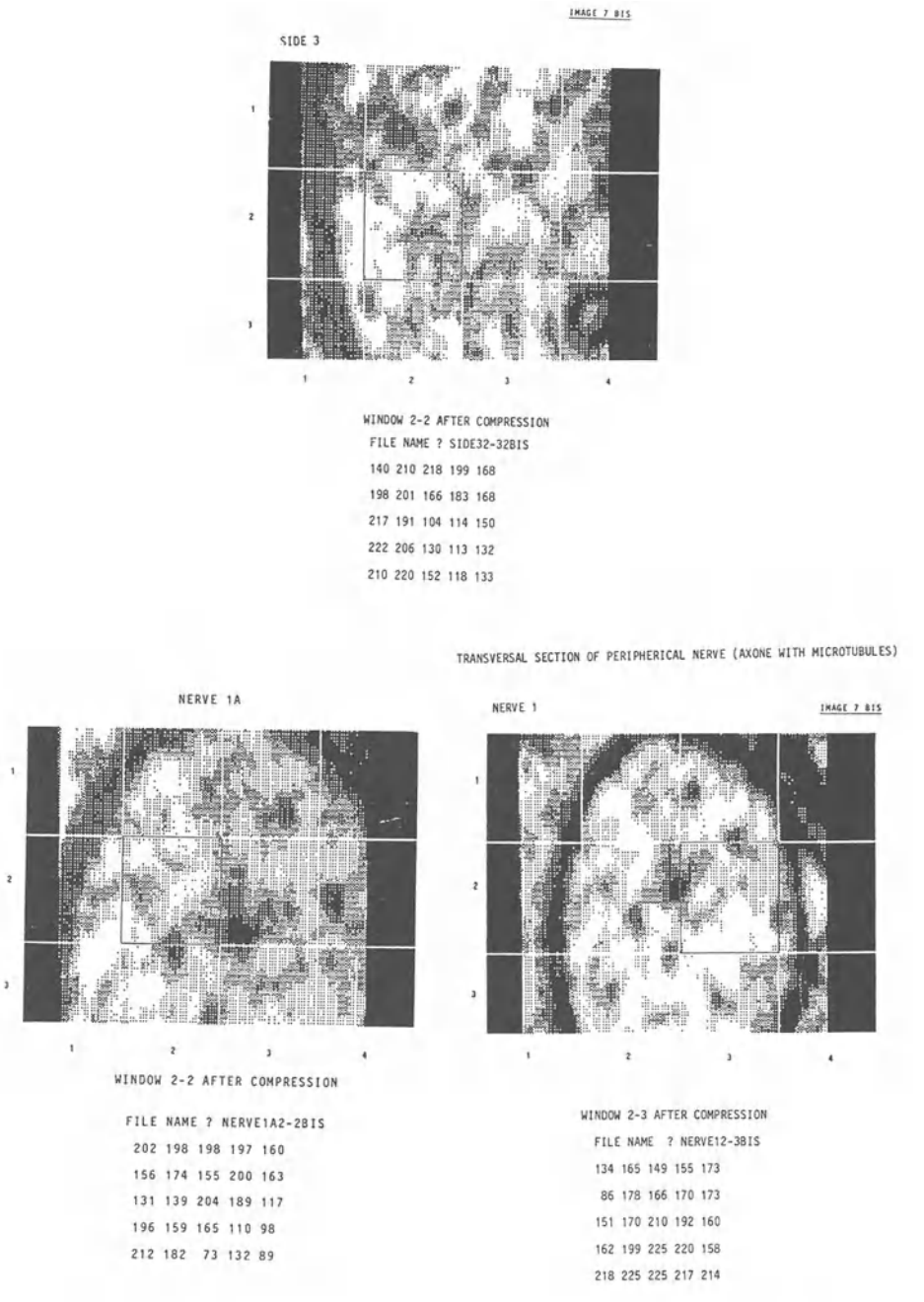


Fig 3 3 Different digital views of electron microscopy inside a nerve cell, where microtubule sections are apparent Segmentation windows have been chosen in order to characterize the cellular structures around the microtubules For analysis, the 70 × 70 pixels of the window have been compressed to a (5 × 5) coefficient matrix.

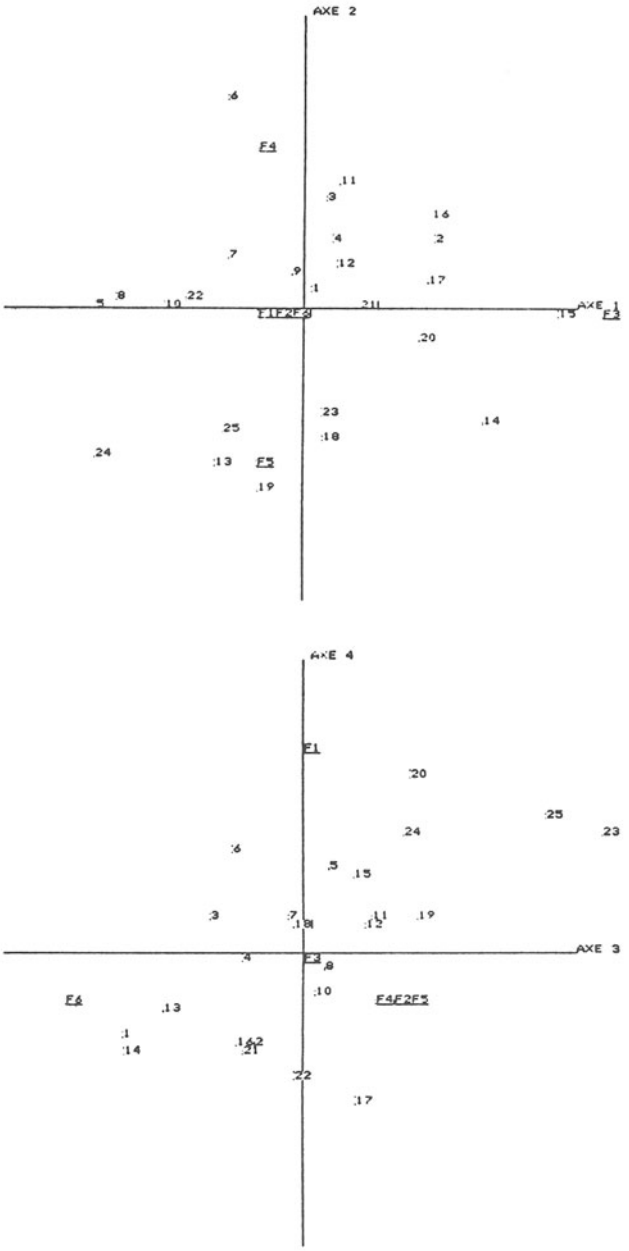


Fig. 3.4. Factor analysis of the correspondence of a series of 6 segmentations (dimension 70 × 70 pixels) after compression (dimension 5 × 5). (a) a plane built on two principal eigen vectors (1, 2). (b) a plane built on vectors (3, 4), which underline characters representing a position of the segmentation windows. Other numbers represent the coefficients obtained after compression. This example shows that the textures which surround the microtubules are different, following their position in the cell (axone, etc.). Texture analysis is an important instrument for analysis and reconstruction of complex molecular objects.

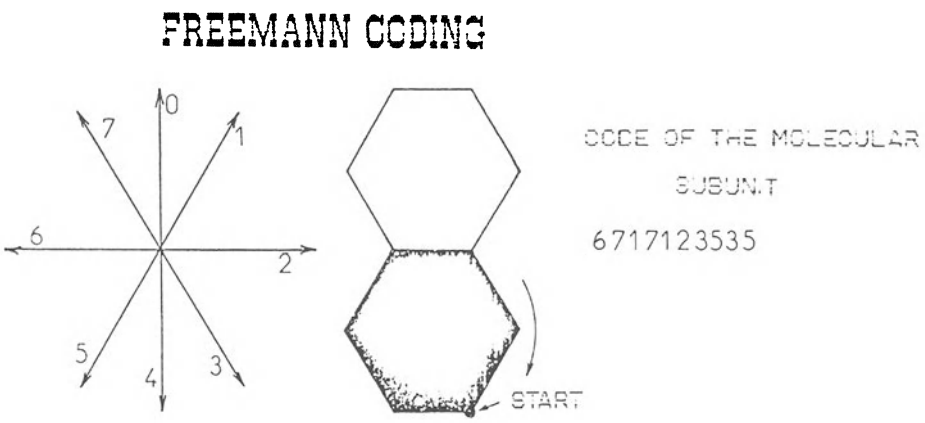


Fig. 4.1.

The coding method does not necessarily seek to produce an efficient code in the sense of minimizing the number of code bits required to describe a molecular boundary but rather provides some manipulatives properties.

Each chain is composed of line segments connecting adjacent pixels in a horizontal, vertical or diagonal direction.

When this boundary has been closed, a code group 0401 is inserted in the chain code and an "invisible line" connecting the two boundaries is encoded.

The dimer, which is described by a chain of Freemann-constituted by numbers, represents the contour of the elementary subunit to be integrated.

4.2. MOLECULAR ASSEMBLAGE STEPS

In the molecular polymerization simulation, different steps must be considered: rotations and translations, correlation, concatenation.

4.2.1. *Rotation*

Planar rotation of the dimers is necessary in order to simulate molecular collisions. Rotations are represented by additions or subtractions on character chains.

4.2.2. *Correlation*

Through correlation, it is possible for two dimers to be assembled together, and for one dimer to be connected to an already existing block [3].

The dimers are represented by numerical chains called octal chains, and we shall examine how to interpret a possible assemblage, i.e., how to detect a correlation between two chains.

We shall see that the Freemann code is very effective:

— we read the Freemann chain representing the constituted block, clockwise;
— we read the Freemann chain representing the dimer to be integrated into the block, counterclockwise.

By construction of the Freemann code, a correlation will appear between the two numerical chains only if a relation of congruence modulo 4 exists between two elements of the same order within each of the two chains to be compared.

Example:

Block: 4 5 7 5 7 0 1 3 1 3
Dimer: 0 7 5 7 5 4 3 1 3 1

From this example, we can see that a correlation between the elements of order 1 and order 6 appears between the two chains.

A correlation of order N is a correlation which exists between two sub-chains of length N. It is not possible to have a correlation of order N, i.e. between two sub-chains A and B, if each element of A (respectively, B) is congruent modulo 4 to each element of the same order of B (respectively, A).

4.2.3. *Concatenation*

Concatenation occurs after correlation. When we have two chains, we proceed to the assemblage between the block and the dimer; this assemblage is simulated by the procedure of concatenation, a procedure which obeys the rules illustrated by the following example.

Example:

A: 5 7 5 7 0
B: 0 5 1 3 4 Correlation of order 2

We have therefore two correlations of order 3. We have chosen the first one, called the window, the holding subchain. A chain representing the edifice block + dimer is generated.

- We read the chain representing the block, until it meets the window.
- Then, we read the chain representing the dimer — beginning with the element which is located immediately to the left of the window — until it meets element of order 1;
- then we read the chain representing the dimer — beginning with the element of order 10 — until it meets the element located immediately to the right of the window;
- then we read the chain representing the block — beginning with the element situated immediately to the right of the window — until it meets the element of largest order.

The software must generate this new chain from the knowledge we have obtained for the two correlated chains.

It must take into account some constraints for the assemblage. We have rejected certain types of assemblages, like the correlation by numerical chains (Figure 4.2).

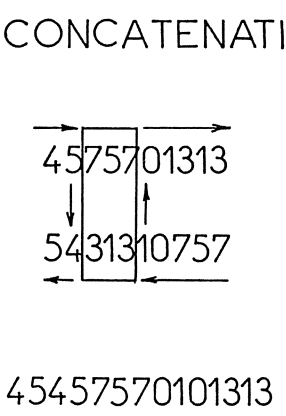


Fig. 4.2.

4.3. ASSEMBLAGE CONSTRAINTS

The complexity of the problem of the assemblage of dimers cannot be easily solved, due to the existence of various constraints, like the pH, colchicine, the temperature, etc. . . . . Many factors, once correctly quantified, desserve to be integrated into the software.

For several reasons, we have been able to conserve only the type of constraint described above. The software must take into account the fact that a certain number of constraints do not disrupt the structure of the existing programme.

4.4. MODELLING OF THE TUBULAR MOLECULES

4.4.1. *Subunit Structure*

Electrophoresis shows [2] that microtubule subunits are constituted with two types of spheroidal molecules which involve approximately 500 amino acids: the α and β tubulines (Figure 4.3).

These tubulines have two joining sites: guanidine nucleotides (guanosine triphosphate (GTP) and diphosphate (GDP)); alkaloids (colchicine, vinblastine, podophylline).

With the tubulines, another molecule plays a principal part: the dyneine (M = 500 000) which has an ATP-asic activity and which is associated with the tubulines in the presence of Mg.

In order to take account of environmental factors and constraints in the simulation, polymerization—depolymerization processes have been represented parallel to each other, although the two systems are in opposition to each other. The constraints, such as concentrations of Ca^{++}, Mg^{++} and colchicine, are determined and the environmental factors, such as temperature, are introduced.

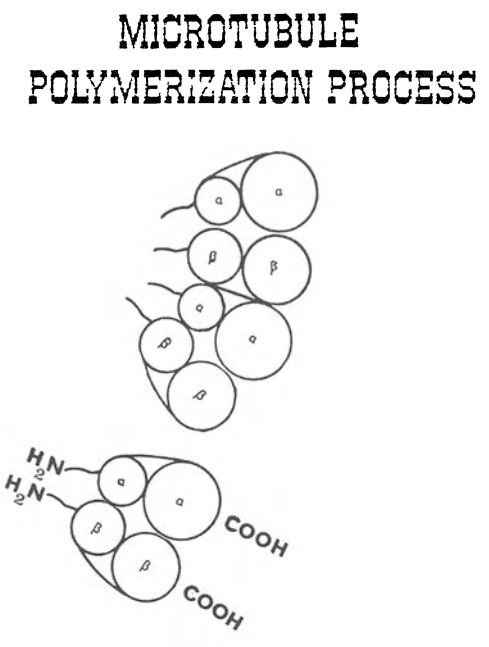


Fig. 4.3. Tubuline molecule assemblage forming microtubules. Each monomer is composed of an exposed tail on a COOH-terminal structural domain and NH2-terminal structural domain (left to right respectively). The size of the two structural domains is proportional to their molecular mass.

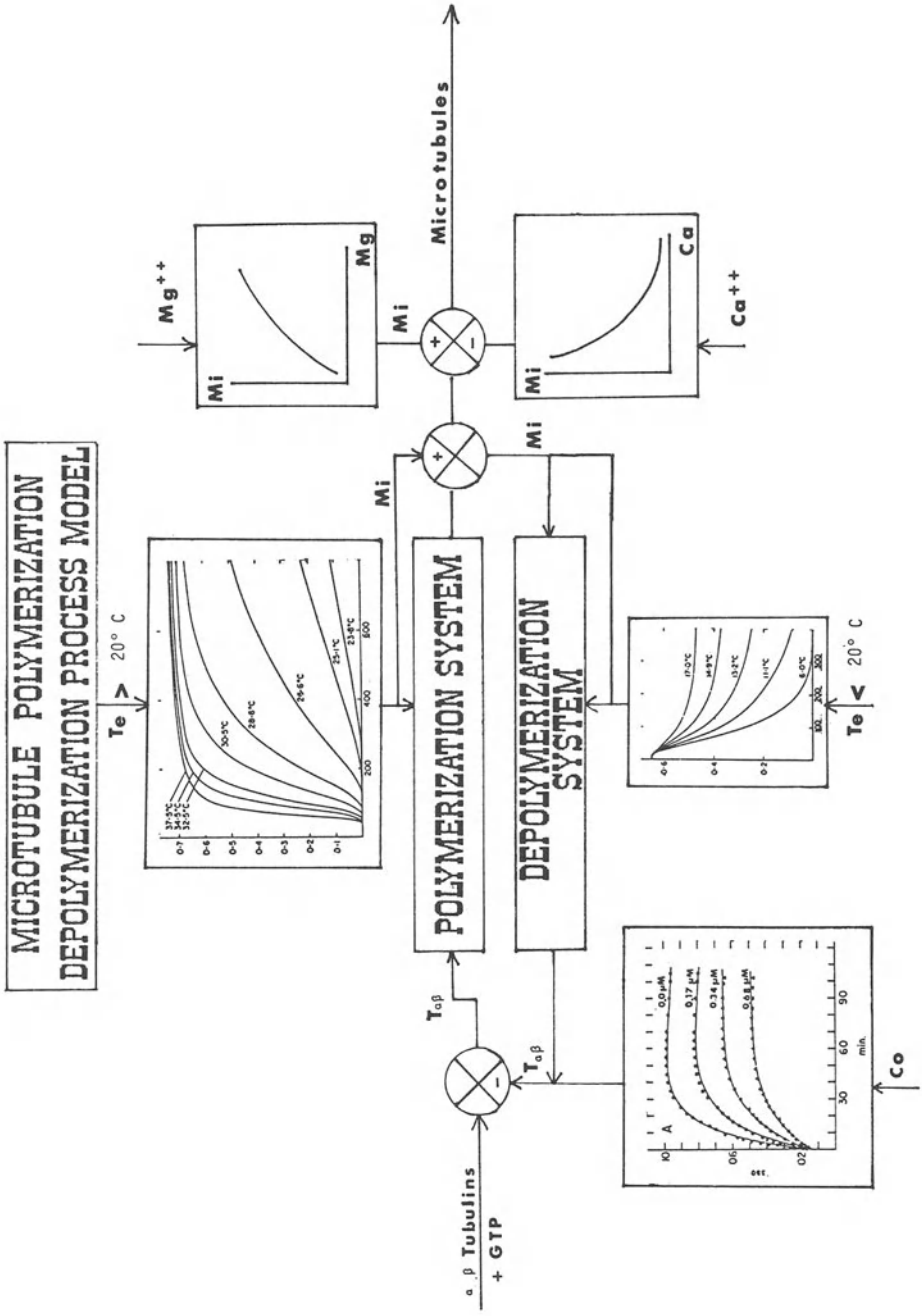
There are variations of concentration of α, β-tubuline at the input, and quantities of microtubules formed at the output. Although the system has multiple inobservable components, the confrontation between the model and the experiment can reduce the inobservable components (Figure 4.4).

4.4.2. *Simulation of an Inhibitor of the Microtubule Polymerization*

We have seen that the microtubule polymerization after a certain step, increase and proceed from a microtubule "germ" [2]. This part of the tube is formed and the dimers are fixed at the extremities of the microtubules (Figure 4.5).

Normally the microtubule is in dynamical equilibrium with the tubulin dimers in the hyaloplasm solution: the dimers, which leave the microtubule extremity, can be replaced by new ones principally during mitosis (Figures 4.6a, b). When an antimitotic molecule (colchicine, vinblastine, podophilline, etc.) enters the cellular hyaloplasm, it is fixed to each dimer and cannot replace a missing dimer at the end of the microtubule.

Thus the $\alpha\beta$ microtubules depolymerize progressively. This kind of process can be modelized with the manipulation of octal chains randomly associated with the microtubule extremity.


MICROTUBULE POLYMERIZATION
DEPOLYMERIZATION PROCESS MODEL
Te > 20° C
Te < 20° C
Mg++
Ca++
Mi
Mg
Ca
Microtubules
POLYMERIZATION SYSTEM
DEPOLYMERIZATION SYSTEM
Tαβ
Tαβ
α,β Tubulins
+ GTP
Co
min.

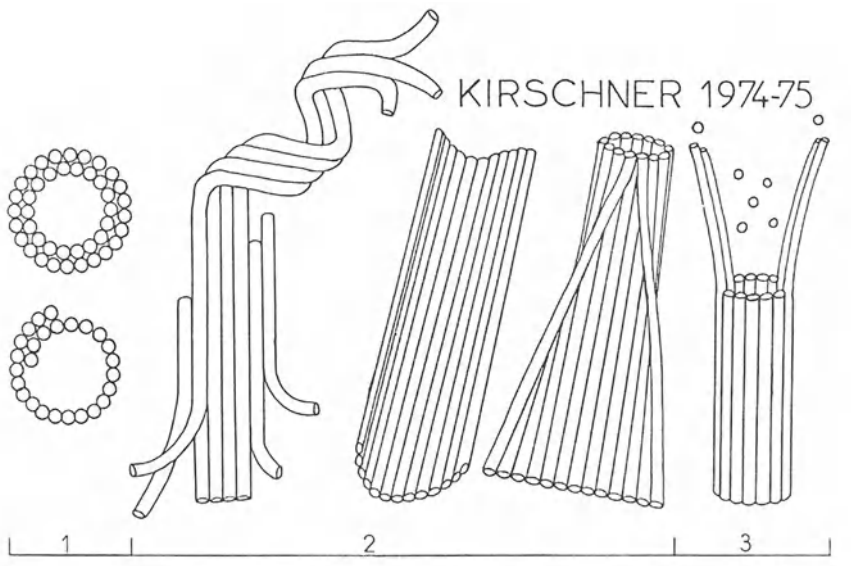


Fig. 4.5. After the ring step, the microtubule polymerization increases to form a tube.

The interactions between dimers are of hydrophobic nature. The depolymerization is in relation with the H_2O molecule and with the α and β tubuline peripheral molecules.

4.5. MODELLING OF THE HEMOCYANIN MOLECULES

To reconstruct the molecules, it is necessary to know the spatial structure of the subunits of the macromolecules and some chemical affinities or polarities.

On the basis of X-ray crystallography, subunits of *Pamulirus interruphis hemocyanin* and of *Limulus polyphemus* have a kidney shape appearance, that can be represented, using the Freemann code, by the corresponding character chain (Figure 4.7).

The hexamers of *Panulirus hemocyanin* occur as a triangular antiprism belonging to the point group S_6.

Following Gaykema, in each hexamer three domains are involved in the intersubunit contacts [8].

Helix 1.7 is in contact with a similar helix in a subunit related to its own by a 2-fold axis.

The 3 helix 2.3, from the 3 subunits interact with each other around the molecular 3-fold axis.

Helix 3.6 is close to helix 1.3 of another subunit.

Also, the 2-D correlation between the three dimers, which have produced the concatenation, do not correspond well to the chemical affinities of the known sites.

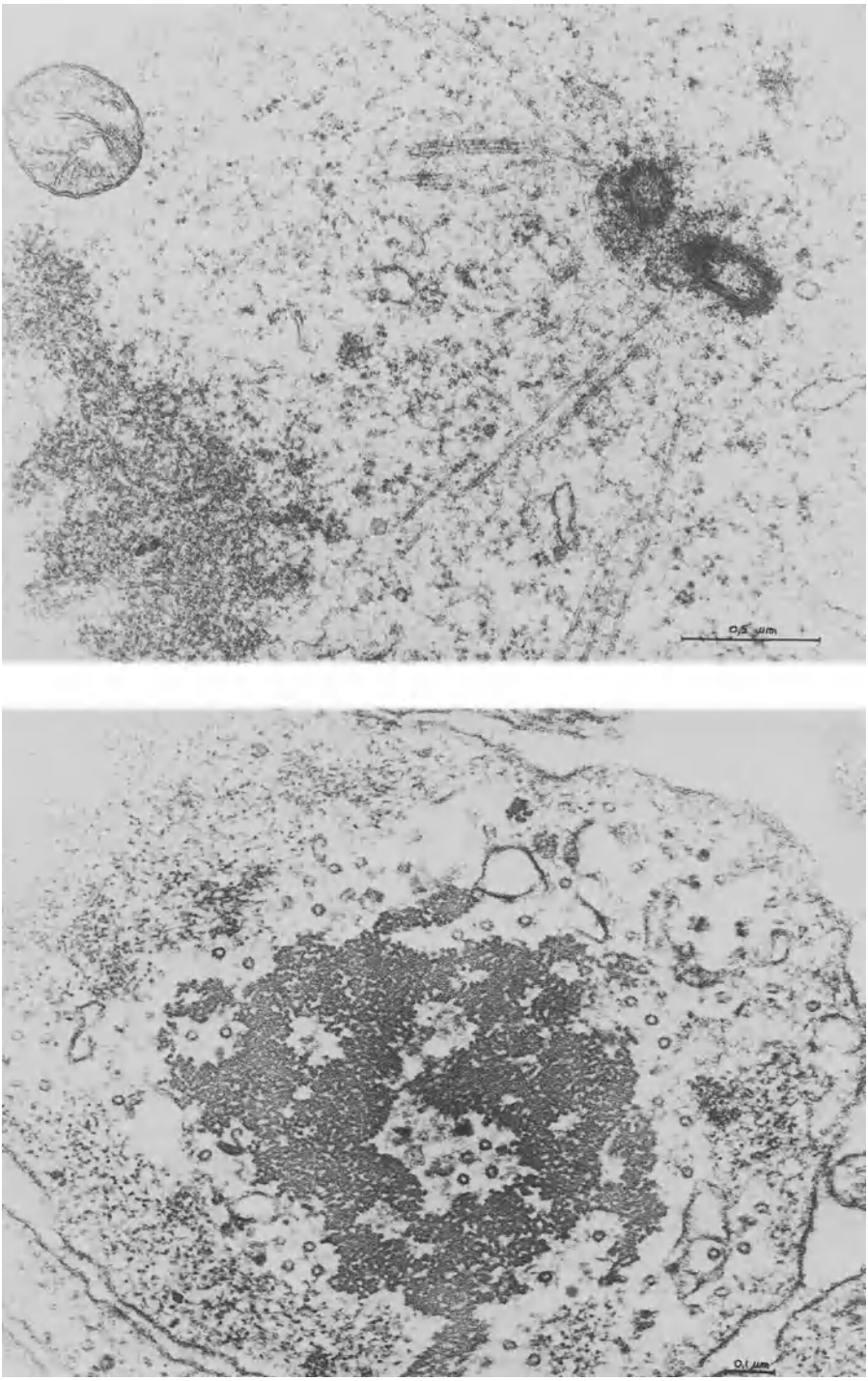


Fig 4 6 Pole of a mitotic spindle of a cell in a division period we can observe — asteriens centrioles — microtubules — chromosomes (a) Longitudinal view (b) transversal view Micro tubules polymerize and depolymerize continually during this period to separate the chromosomes (mouse femur B Arbeille Unite de Microscopie Electronique Tours)

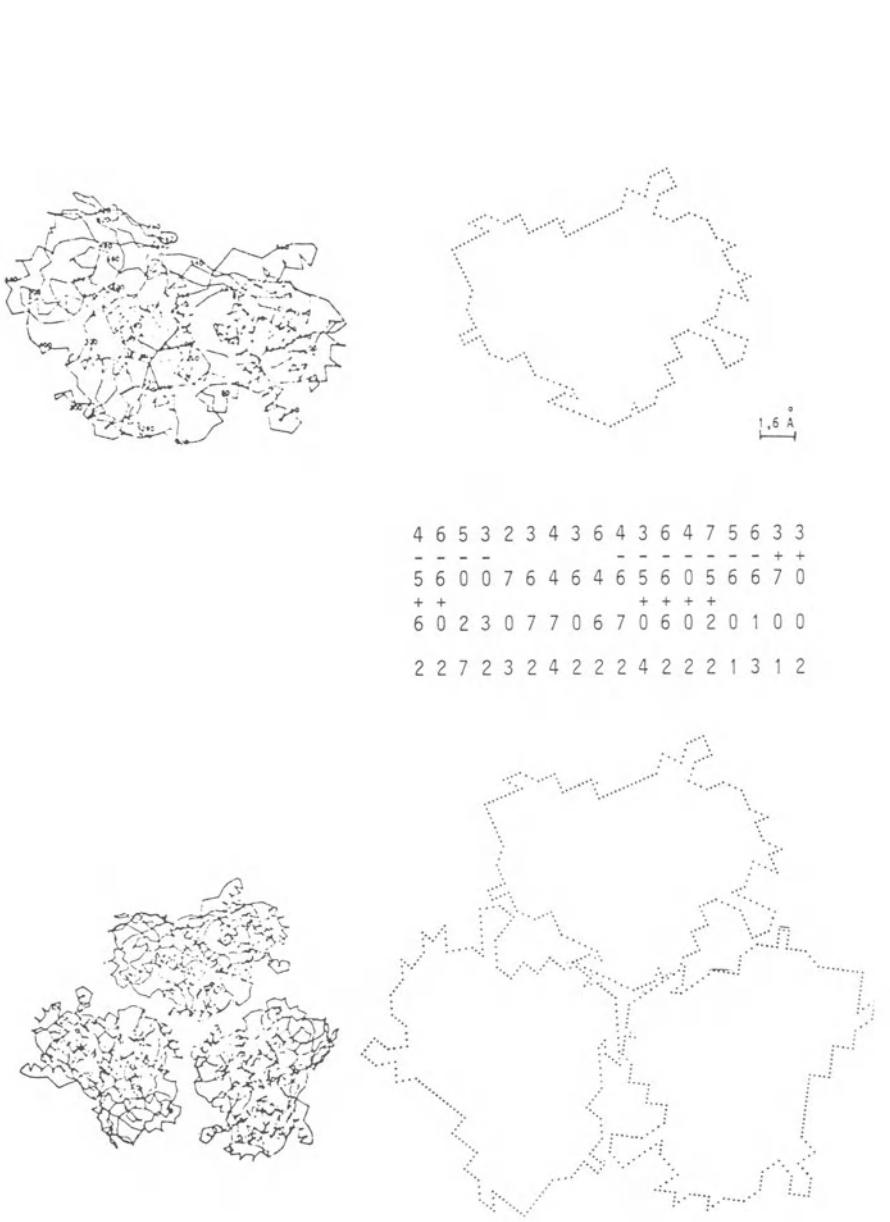


Fig 4 7

Other methods like the correlation of 3-D pseudo-Peano chains representing the 3-D development of peptidic chains of the dimer, could in the future, be a more precise approach

Fraction 1, constituting the subunits 3C and 5B, appears to play an important

role in the reassemblage portion of two dodecamers. This fraction is capable of *self-assemblage* and there are at least three building sites for another molecular fraction 1 which occurs on each molecule.

The dimer self-assemblage process can be represented by a formal system and a derivation tree of all possible assemblage combinations. The observation of the assembled dimers will be described by a "molecular context-dependent grammar" associated with the formal system (Figure 4.8).

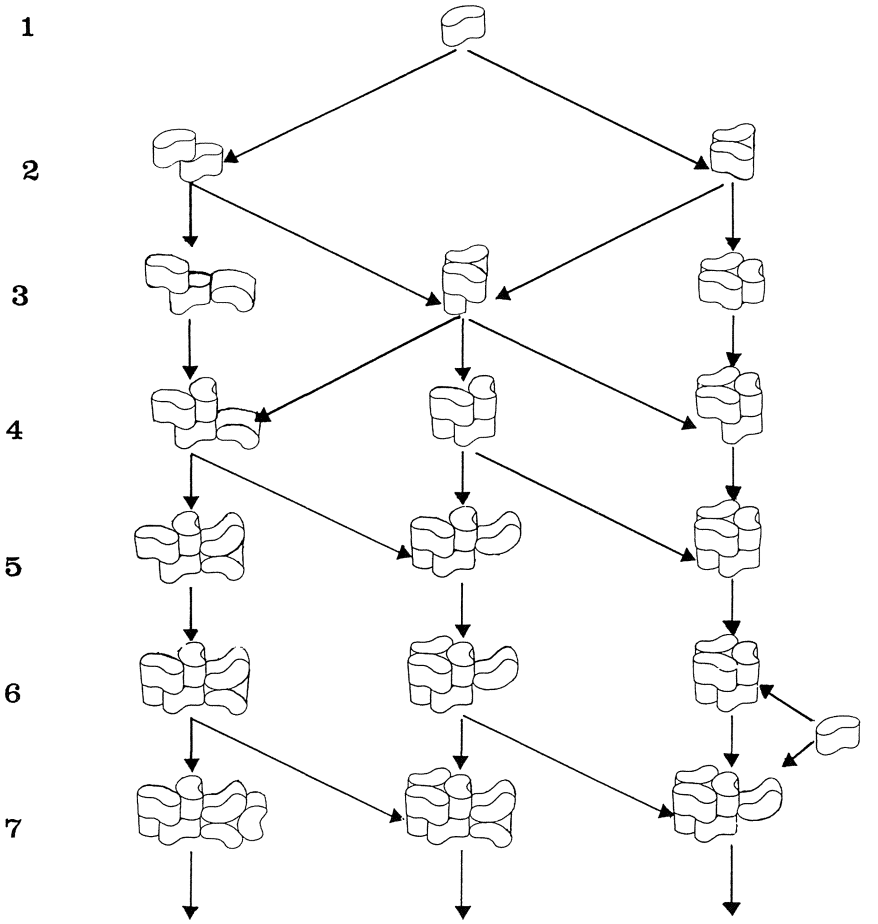


Fig. 4.8.

By using correlations with constraints, we simulate all forms of molecular collisions and interactions between the dimers.

These correlations can be described by the projection on a principal plane, calculated by the correspondence analysis previously described or the classification analysis (RECFORM method). Thus the correlations obtained can allow to define the structure of the molecular grammar.

5. Discussion and Conclusion

The 2-D assembly simulation must be extended to 3-D. This involves a good 3-D spatial representation of the dimer subunits and a good knowledge of the active assembly sites. Knowledge representation, for example of fractional structures, can be a useful tool when one takes into consideration the polypeptidic structure of the subunits in the polymerization process.

At present, it is difficult, from the observation of the isolated dimers (complex $\alpha\beta$ tubulin + GTP or hemocyanin subunit), to predict the structure of the whole molecule (the microtubule or the 24-mer hemocyanin molecule), by using the simulation programme of a self-organizing molecular system [23, 26].

Different possibilities can be considered in order to reduce the distance "object-model":

1 — the information concerning the intra and extra molecular interactions could be increased,
2 — the information concerning the native spatial structure of the whole molecule could be completed,
3 — the observations concerning the subunit structure and latent properties could also be completed.

This incompleteness is due to:

i — the biochemical methods or observation tools used (X-ray crystallography or electron microscopy);
ii — the fact that molecular potentialities which emerge after assemblage are not observable on the isolated subunits ($\alpha\beta$ tubulin or hemocyanin), Figure 5.1.

The theories of the emergence of potential systems, which have been developed by several authors [18, 23, 26], can be applied to molecular systems by using a decomposed algorithm in order to improve the knowledge of the properties of the subunits.

In the assembled molecule, the subunit structures and properties are not the same as in the isolated molecule. If new intramolecular interactions occur, this will allow us to redefine the model.

The application of automata theory has already been developed for molecular interactions by Atlan [27]. New developments of this theory can also allow one to

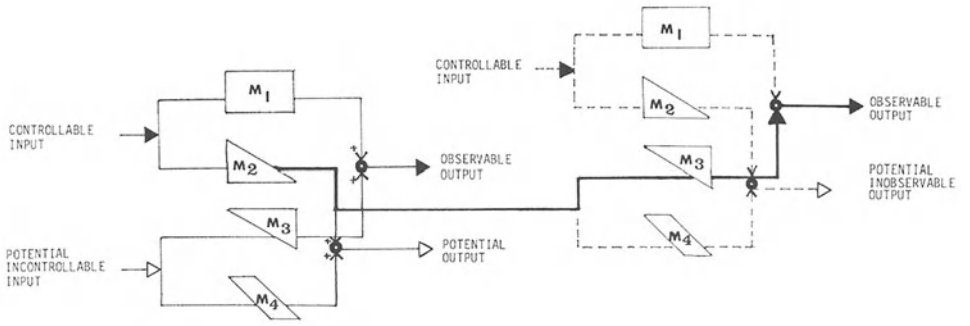


Fig 5 1 Simulation of the association of two molecular systems, broken down in fragments and connected in series, can be performed by application of fuzzy set theory in order to estimate the potentialities of each system (I Redon *Cell*, Pergamon Press, p 32, 1983)

describe the latent properties of molecular subunits, in particular the concept of fuzzy automata.

Acknowledgements

The authors wish to thank Professor Brigitte Arbeille, Laboratoire de Microscopie Electronique (Tours), for providing them with illustrations for the paper; Professor Joseph Thouvenot, Laboratoire de Physiologie (Tours), for stimulating discussions; Dr Denise Pignon, Hervé Le Pape, Isabelle Fontenas and Clotilde Saint-Requier, Société de Bio-Informatique et de Bio-Technologie (Tours), for technical assistance; Frederic Pathier, Société Automatisme et Avenir Informatique (Paris), for preparing the computer graphics, and Donna J. Chow for editing and translation help.

References

1 P Dustin *Microtubules*, Springer-Verlag, Berlin (1978)
2 A Berkaloff, J Bourguet, P Favard and J C Lacroix *Biology and Physiology, I Cellular*, Hermann, Paris, p 270 (1977)
3 C Gaudeau, E Michonnet and J F Touchard *Modélisation des Systemes Vivants Complexes et des Interactions avec leur Environnement*, Laboratoire de Bio-Informatique, p 112 (1982)
4 K Roberts and J S Hyams *Microtubules*, Academic Press (1979)
5 J Lamy, J Lamy, P Y Sizaret, P Billiald, P Jolles, J Jolles, R J Feldman and J Bonaventura *Biochemistry* **22**, 5573 (1983)
6 J Lamy, J Lamy, P Billiald, P Y Sizaret, G Cave, J Frank and G Motta *Biochemistry* **24**, 5532 (1985)
7 J Lamy, J Lamy, J Bonaventura and C Bonaventura *Biochemistry* **19**, 3034 (1980)
8 W P J Gaykema, W G J Hol, J M Vereyken, N M Soete, H J Bak and J J Beintema *Nature* **309**, 23 (1984)
9 C Gaudeau These de Biologie Humaine, Tours (1975)
10 D Dziedzina, C Gaudeau, M Lafargue, C Longevialle, P Morand and J Thouvenot *Analyse et traitement d'images bio-medicales*, Ed IRIES, Paris, p 53 (1980)

11. D. Dziedzina, C. Gaudeau, C. Longevialle and G. Routier: *Image Sequence Processing and Dynamic Scene Analysis*, Vol. 2, p. 625, Springer-Verlag, Berlin (1983).
12. M. Bouabdallah, C. Gaudeau, C. Kurc, C. Longevialle, C. Malet, P. Morand and J. Thouvenot: *Computers in Cardiology*, IEEE, p. 267 (1975).
13. M. Boiron, M. Bouabdallah, J. P. Crettez, C. Gaudeau, C. Longevialle and J. Thouvenot: *C. R. Journées d'Informatique Médicale*, Toulouse, 13—17 mars 1978, **1**, 259 (1979).
14. M. Boiron, D. Dziedzina, J. C. de la Souchère, C. Gaudeau and J. Thouvenot: 'Détection des contours par les automates dans la reconnaissance des formes', in *Analyse et traitement d'images bio-médicales*, Ed. IRIES, Paris, p. 1 (1980).
15. W. K. Pratt: *Digital Image Processing*, John Wiley, New York (1978).
16. H. C. Andrews, J. Kane and W. K. Pratt: *Proc. IEEE* **57**, 58 (1969).
17. J. L. Chermant and M. Coster: *Précis d'analyse d'images*, Editions du CNRS, Paris, p. 1 (1980).
18. G. Quaster: *Emergence in Biological Organization*, Yale University Press (1970).
19. E. W. Taylor: *J. Cell Biology* **25** (1, part 2) 145 (1965).
20. J. Lamy, S. Compin and J. Lamy: *Archives Biochem. Biophys.* **223**, Academic Press, p. 584 (1963).
21. T. Cavalier-Smith: *Bio-Systems* **10**, **93** (1978).
21.A. J. Travis, J. Kenealy and R. P. Allen: *J. Cell Biology* **97** (6), 1668 (1983).
21.B. R. P. Allen: *Annual Review of Biophysics and Biophysical Chemistry* **14**, 265 (1985).
21.C. R. P. Allen, D. G. Weiss, J. H. Hayden, D. T. Brown, H. Fujiwake and M. Simpson: *J. Cell Biology* **100** (5), 1736 (1985).
22. E. C. Raff: *International Review of Cytology* **10**, 1 (1979).
23. I. Redon, D. Morand and C. Gaudeau: *Fuzzy information, knowledge representation and decision analysis*, p. 31, Pergamon Press, New York (1984).
24. D. L. Sackett and J. Wolff: *J. Biological Chemistry* **261**, 9070 (1986).
25. M. van Hell: *Ultramicroscopy*, Vol. 13, p. 165, North-Holland, Amsterdam (1984).
26. H. von Foerster: *Self-organizing Systems*, pp. 31—50, Pergamon Press, New York (1960).
27. H. Atlan: *Ann. Phys. Biol. Med.* **1**, 15 (1970).

Modelling Molecular Displacements in Organic Solids

CLAUDE DÉCORET
Laboratoire de Chimie Industrielle, Université Claude-Bernard, 43, bd du 11 Novembre 1918, 69622 Villeurbanne Cedex, France.

1. Introduction

Modelling of molecular displacements in the organic solid state is devoted in this paper to large molecules in Van der Waals crystals and, in particular, to phase transition and reactivity. Such displacements could lead to drastic transformations or even the breaking down of the crystals. Modelling of small molecular motions such as vibrations and librations is quite a different question, since time-dependence is taken into account more or less explicitly [1, 2]. Attempts will be made throughout to consider the importance of spatial modelling and its direct representation, such as 3-D screen pictures, and derived ones, such as the packing analysis of Gavezzotti, to elucidate crystal phenomena. One major specificity of the solid state is the packing of molecules which must be reproduced by intermolecular interaction models. So energy calculation methods are the background of spatial representation. The approach to modelling displacements of molecules in the solid state could be derived both from:

— the various trends of crystallographers to visualize crystal arrangements (stereoscopic views, etc.),
— the developing area of drug design.

The need for spatial model management seems stronger than in usual quantum-chemical approaches where molecules are often considered by themselves in unorganized gaseous or liquid states. Figures 1 and 2 show such examples of spatial representations for the packing of the acetylanthranyl molecule in its crystalline state [3, 4]. No doubt these figures will be very useful for understanding the packing. Management would consist here, for instance, of labeling molecules deduced by a given symmetry. Distances and angles between given atoms could also be pointed out.

For a chemist, what is fascinating in the solid state is that the anisotropy of molecules is reflected, through their static interactions, in their macroscopic properties. Such anisotropic interactions are fixed to such an extent that they have a direct influence on a greater number of experimental observables than in any other state, where they are generally orientionally averaged. Observations are derived from X-ray or neutron diffraction measurements, and thermodynamic or

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, 159—171.

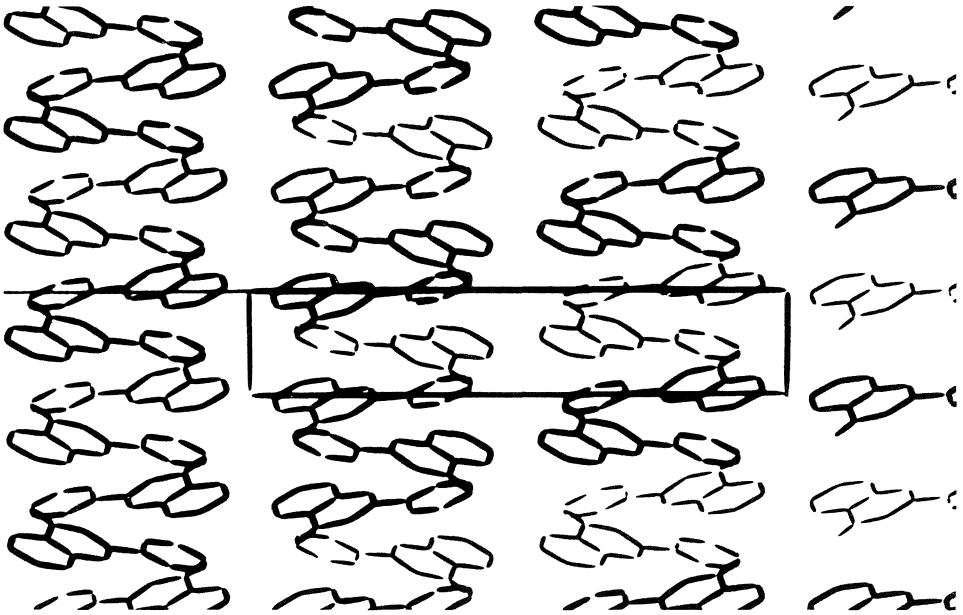


Fig. 1. Crystal packing of acetylanthranyl crystal. Bound view on PS300 graphic display.

optical and vibrational spectra measurements. The first task of a theoretician is to give a representation of intermolecular forces.

2. Methods of Calculation

2.1. ATOM–ATOM POTENTIAL

In this field of research, the pairwise atom–atom potential formula is the most widely used for the evaluation of intermolecular interaction [5]. For instance, the so called Buckingham or van der Waals type is given by:

$$E_{\text{cryst}} = \frac{1}{2} \sum_i \sum_j - \frac{A}{r_{ij}^6} + Be^{-\alpha \cdot r_{ij}} \tag{1}$$

r_{ij} being the distance between atoms i of molecule 1 and atom j of other molecules; A, B and α being parameters.

A Lennard-Jones formula will be given by:

$$E_{\text{cryst}} = \frac{1}{2} \sum_i \sum_j - \frac{A}{r_{ij}^6} + \frac{B}{r_{ij}^n} \qquad n = 9, 12 \tag{2}$$

Both types are often referred as six-exp. or six-9, six-12 respectively. This formula

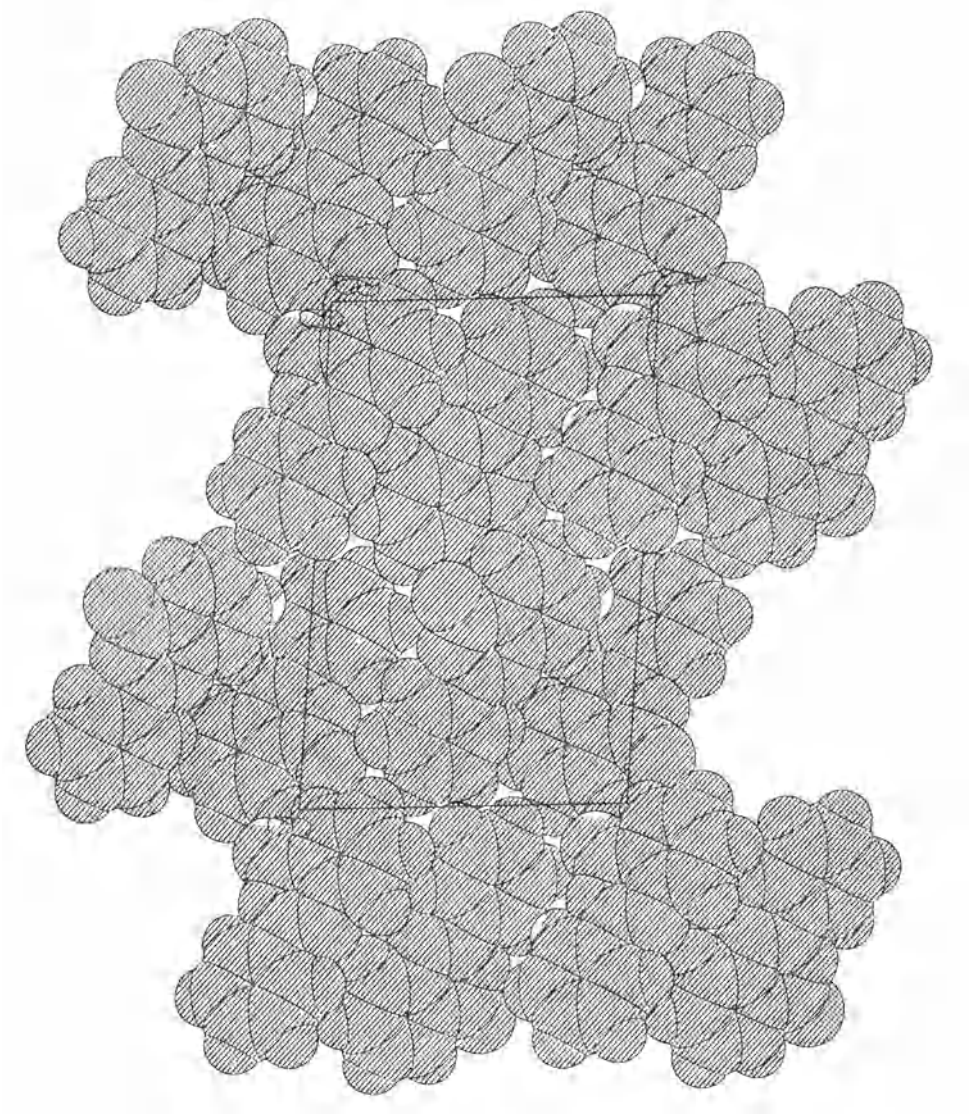


Fig. 2. Crystal packing of acetylanthranyl crystal. Van der Waals view on Benson plotter by PLUTO program.

can be made more sophisticated by adding induction terms which are necessary when dealing with polar molecules [6]. Multipole values are generally taken from *ab-initio* calculations or from experimental data [7]. The so-called dispersion-repulsion term (Eq. 1) can be formulated if we admit that it represents the dispersion and repulsion as usually defined more judiciously by adding terms such as $1/r^8$ and $1/r^{10}$ to the $1/r^6$ dispersion term. Possibly we also have to introduce a

damping term to the dispersion for short range interactions (Eq. 3) [8].

$$E_{\text{dispersion } ij} = -f(r_{ij})\,[C_6 r_{ij}^{-6} + C_8 r_{ij}^{-8} + C_{10} r_{ij}^{-10} + \ldots] \tag{3}$$

$$f(r_{ij}) = \exp -\alpha[(r_0 r_{ij}^{-1} - 1)^2] \quad \text{for} \quad r_{ij} \leqslant r_0$$

$$f(r_{ij}) = 1 \quad \text{for} \quad r_{ij} \geqslant r_0$$

The choice of parameter values is of vital importance. Many sets are available but comparisons are rarely made. Among the most commonly used, those by Mirsky [9] and Williams [10, 11] were determined for H and C atoms to fit the lattice energies of a certain number of crystals and equilibrium distances by taking into account up to 25 Å interaction distances. Those of Claverie [12] are to be used with formulae containing electrostatic terms. When dealing with lattice or molecular dynamics the choice of Dasheviskii [13, 14] is sometimes recommended.

The spheroidal molecule—molecule pairwise interaction model, developed by Aubert-Frecon [15], and the intermediate multicenter multipolar expansion model of Righini [1] can also be mentioned.

2.2. DYNAMIC CALCULATIONS

To treat phenomena such as phonon curves, thermodynamics, etc., theoreticians develop two kinds of calculations. Without entering into detailed discussions, let us remember that lattice dynamics is typically a treatment based on the harmonic oscillator model and is more or less related to absolute zero measurements. Corresponding calculations are relatively rapid but there are some difficulties in introducing the relaxation of lattice parameters and the temperature factor. Molecular dynamics is based on randomly distributed displacements of molecules or atoms and seems more appropriate for treating temperature dependent phenomena. Both models, lattice and molecular dynamics, as mentioned before, utilize a simple dispersion-repulsion term for atom—atom potential without induction terms which converge slowly. Computer improvements make these methods more and more available, either for small molecules with more precise calculations or for calculations on bigger molecules.

Among numerous recent works, the evidence of the "butterfly wing-like" motion of the two external rings of the anthracene molecule is worthy of notice [16, 17]. Calculation of thermodynamical properties of naphthalene fits also experimental results very well near melting point [18].

3. Spatial Representation

The principles of close packing introduced by Kitaigorodski [19] and the high repulsive short range interaction values in the atom—atom potential scheme have the following obvious spatial representation. Molecules are built up with atoms represented by their Van der Waals spheres. Molecules arrange themselves as

close as possible without sphere penetration. Management and handling of spatial data concerning such molecules in interaction is not trivial. In the field of organic solids, Gavezzotti developed the so-called packing analysis [20], which is a useful tool for spatial analysis of local interactions. On the one hand, the packing analysis can be considered as partitions of the lattice energy according to geometrical considerations. Gavezzotti wrote that it can be thought of as a good synergy between atom—atom potential method and the packing analysis he developed. This analysis is very useful for interpreting the packing. Thus, linear correlations have been found between volumes and surfaces of molecules and lattice energy of the crystal [21]. On the another hand, by analyzing the degree of occupancy of small volumes, that is, by calculating the ratio of points inside the VdW spheres relative to those outside (typically 1000 points per $Å^3$), one can obtain spatial features of empty space in a crystal in order to visualize holes, channels, etc. When studying displacements of a molecule in its crystalline environment, analysis of contact or penetration between VdW spheres may lead to a detailed and comprehensive description of the phenomenon, for instance, by visual inspection, aided by immediate sorting out of geometrical and energetic characteristics. Accurate determination of molecular volumes has also been proposed by Connolly and Richmond [20]. Such analysis are an obvious support for interpreting of potential hypersurfaces.

Let us take as an illustrative but imaginary example the rotation of the naphthalene molecule about its principal axis of inertia in the crystal. You can calculate the potential energy surface [1, 2], and look at the potential barrier, but you can also, with some interactive sysem display the displacement on a screen, the molecule being represented with its VdW spheres around atoms. You can detect the knocking between spheres and eventually trace a path without knocking that is more or less a free rotation. This, in fact, occurs with relaxation of the environment through increase of temperature [22, 23]. In the field of drug design, the hardening (non-penetrability) of such spheres are implicitly used by Cambillaud to calculate potential hypersurfaces. Here in a first approximation, energy calculations for atom—atom distances are eliminated below a given value when dealing with interactions between long side chains. This treatment is the translation of the non-penetrability of Van der Waals spheres in practical calculation [14].

Van der Waals spheres can be considered to a first approximation as electronic zones. Models for softening these spheres have to be proposed. In the field of quantum chemistry, ways have been explored [25] for spatial representation of deformation of electronic clouds without forgetting the numerous applications of the "Théorie des Loges" ideas [26] as, for instance, bond formation treatments [27].

4. Small Molecular Displacements

Concerning small molecular motion measurements, we have to keep in mind that

the recording measurements are typically 10^5—10^6 sec. for X-ray and neutron diffraction measurements, whereas time scaling for motion is of the order 10^{-3}—10^{-4} sec. Up until now we have access only to average motion phenomena. Nevertheless, small molecular displacements with time dependence, the so-called molecular motions, (e.g. vibrations or librations), are also relevant for spatial representation (see, for example, the film of the pulsation of DNA, [28]). These kinds of movements of molecules or fragments of molecules can be detected either by spectroscopy (Raman, IR, etc.) or crystallography (X-ray etc.). Numerous experimental data are available and are a true mine to be exploited by theoreticians. For example, by analyzing atomic MSDA (mean square displacement amplitude) from literature, Trueblhood [29, 30] studied the rotation of methyl in a large series of solids. The corresponding thermal ellipsoids are calculated and displayed by various programs such as ORTEP [31]. In our concern of large molecular displacements, they can serve as indicators through the directions of their principal axis. This was done for studies of configurational changes of spherand [32] and propellane molecules [33]. Evolution of a phenomenon in a series of similar molecules can be also studied as done by Gavezzotti on the annulene molecules [34]. The question is how a change in substituent borne by this molecule can undergo or hinder the rotation of the molecule in the crystal. It is treated by rotating the molecule in a fixed environment which can eventually be relaxed with calculation of the potential hypersurface. The rotation is found to be possible with certain substituents and impossible with others. Answers are suggested by the energetic curve profils and from geometrical consideration. The thermal ellipsoids around carbon atoms illustrate qualitatively the direction of the molecular rotation.

So far, only small motions were considered. Let us turn now towards displacements leading to rearrangements of molecules in different crystalline forms (polymorphism) and the break down of crystals by reactions between molecules.

5. Phase Transition and Polymorphism

The ability of molecules to arrange in different crystalline forms seems to be very common. Observed rearrangements are either irreversible (monotropic) or reversible, depending on the temperature range (enanthiotropic). In very few cases, single crystal to single crystal rearrangement has been observed [35, 36]. Most often, the transformation occurs through anamorphous phase, as that described in detail for paradichlorobenzene [37] or through break down of the monocrystal into microcrystallites. In fact, few experimental studies have been performed on the full change except for small molecules and some aromatic hydrocarbons, while the so-called plastic phase has been extensively studied [1]. Such plastic phases have also been the subject of many theoretical works dealing with various thermodynamical treatments [38, 39].

Since we are here generally not concerned with large molecular displacements,

reviews such as [1, 19, 40] can be consulted for those interested. Concerning large molecules, the existence of several crystalline forms for the same molecule or group of molecules is, in a certain number of cases, associated with different molecular conformations or configurations sometimes with different arrangement of hydrogen bonds, etc.

In past decades, theoreticians have focused on the relative stabilities of different polymorphic varieties in order to find evidence for the part played by intermolecular and intramolecular forces [41, 42]. These works are related to irreversible transformations. The question of modelling the transformation itself seems generally more complicated since it deals often with problems of symmetry changes and geometrical correspondence between polymorphs. As an example, let us examine the irreversible transformation of the diacetamide molecule (monotropic). This molecule has two crystalline forms, one for the *trans—trans* configuration, another for the *cis—trans* configuration. The *trans—trans* and *cis—trans* configurations are deduced by rotation of one acetyl group. In the polymorphic variety corresponding to the *trans—trans* conformation, molecules form long parallel chains in which molecules are linked by bifurcated hydrogen bonds, the plane of each molecule being almost perpendicular to its neighbour. In the *cis—trans* variety, molecules are arranged in dimers, whose planes are almost perpendicular to each other. Figure 3 shows an imaginary path from dislocation of chains to rearrangement in dimers with rotation of one acetyl group.

The theoretical study has been carried out by simultaneous symmetry constraint displacements of 3280 molecules without independent motion. The intermolecular energy is calculated with the atom—atom potential method developed by Claverie [12] and with *ab-initio* 6—31G for intramolecular energy. The potential energy hypersurface was calculated with minimization and a moderate potential barrier of 15 kcal/mole was found. Stabilities of various polymorphic varieties of amide group containing molecules have also been studied [43, 44]. All these studies are more or less explicitly related to the question of formation of different hydrogen bonds in peptide-like molecules.

Incorporation of thermodynamical treatments is to be expected for future studies on phase transition of molecules of chemical interest. Besides the interest of modelling of physical properties, such as motion-related phenomena, theoretical and experimental studies on organic molecule phase transition might afford, among others, important information on the behaviour of weak intermolecular bonds, and on intramolecular changes, such as configuration or conformation in the solid state.

6. Reactivity

Reactivity in organic solid state has been observed for a long time and among many notable experimental works those of Thomas [45], Paul and Curtin [46], Lahav [47], Dunitz [48] and Scheffer [49] are of particular interest. They often

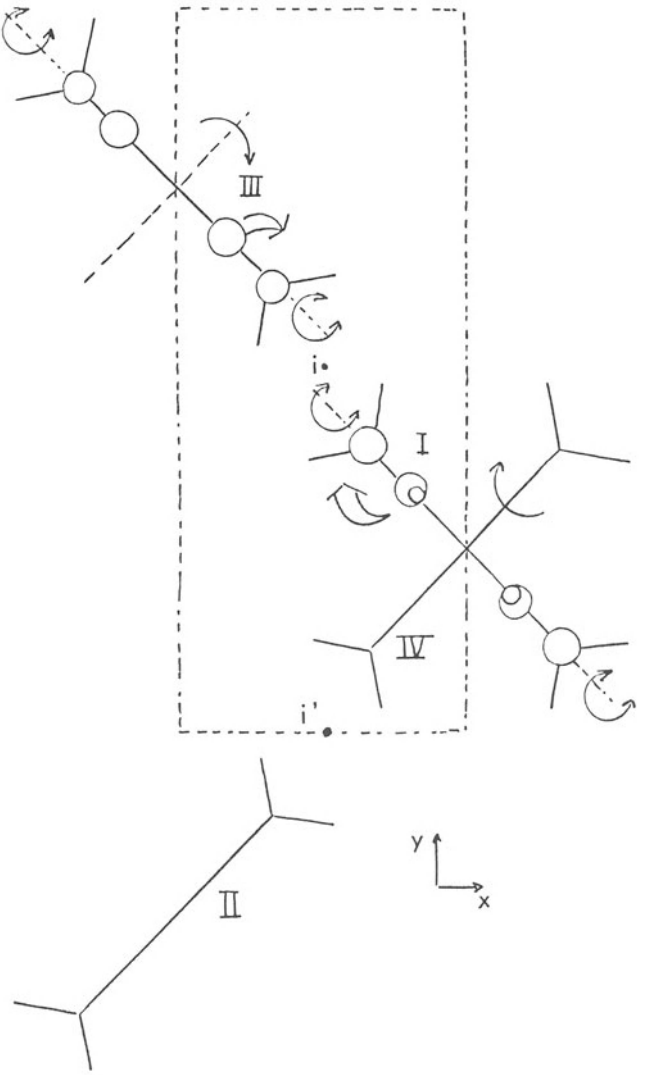


Fig. 3. Diacetamide path of transition from *trans—trans* to *cis—trans*. Beginning of the path: both two acetyl groups are rotating (double arrows) while the whole molecule is moving (single arrow).

refer to polymerization or dimerization. Among reviews those of Shklover [50], Simonetta [40], Kitaigorodski [19] are especially important.

Dunitz was one of the first to study reactivity mechanism through the analysis of crystal arrangement [51]. He compared crystallographic data of different molecules with the same reactive group and facing an antagonist group of another molecule in the crystal. He suggested that the deformations of such reacting groups under the influence of different antagonists or substituents could be thought of as frozen pictures of the beginning of transition state along reaction paths.

One of the most known theoretical works is that of Gavezzotti [52] on the transfer of methyl group from sulfonate to amino group in the *p*-dimethylamino-benzenesulfonate crystal to yield the corresponding trimethyllammonium ions. This reaction does not occur in solution. He modelled the displacement of a methyl group between sulfonate and amino group of two molecules, arranged head to tail as in the crystal. He calculated the potential energy surface by E.H.T. quantum-chemical method for the methyl transfer between only two fixed molecules. He also calculated the effect of environment of other molecules in the crystal by an atom—atom potential method. The reaction was shown to occur with reasonable barrier along chains of parallel head to tail molecules. Another very interesting work by the same author [53, 54] deals with the solid decomposition of peroxides with formation of carbon dioxide. Thanks to his packing analysis method, he was able to study in a detailed way the appearance of empty space in the crystal and the ability of the CO_2 molecule to slip through channels or not, depending on the kind of peroxide forming the crystal. Making full use of spatial consideration made possible by the packing analysis, he deduced many interesting features of this decomposition reaction without need of energetic calculations. We have to recall here that there is a close connection between penetration of VdW spheres and high peaks along energetic curves.

Another theoretical study [55] concerning reaction and rearrangement of molecules, consecutive to the breaking of a bond and the formation of a new one was done on the transfer of a chlorine in a crystal of 2,4,6,6-tetrachloro-3-methyl-5-isopropylcyclohexa-2,4-diene-1-one. This reaction between *ortho* and *para* positions lead to 2,4,4,6-tetrachloro-3-methyl-5-isopropylcyclohexa-2,5-diene-1-one crystal. Here, initial and final states were calculated by atom—atom potential intermolecular energy and, furthermore, the beginning of the reaction path was modelized by allowing the rearrangement of the initial crystal in which initial molecules were replaced by the same molecules containing a C_2—C_ω lengthened bond. Due partly to the increase of intermolecular electrostatic energy (the molecule is more polar), the rearrangement leads to such a positioning of molecules that the chlorine becomes closer to the para position of the antagonist molecule where it can react (see Figure 4). It must be pointed out that lattice parameters of the initial crystal were observed to change under illumination. Photo induced reaction in solid state can also be modelized by dynamical treatment of the rigid molecule. This has been recently described in a series of papers by Craigh [56]. Gas-solid reactions are an important field to be explored by means of theoretical studies, but until now few have been made [57]. Also very exciting are the problems concerning crystal growth [58]. A model based on the decomposition of the intermolecular energy according to directions of crystal faces has been proposed [59].

7. Perspectives

Displacements of organic molecules where solid state is concerned imply, for

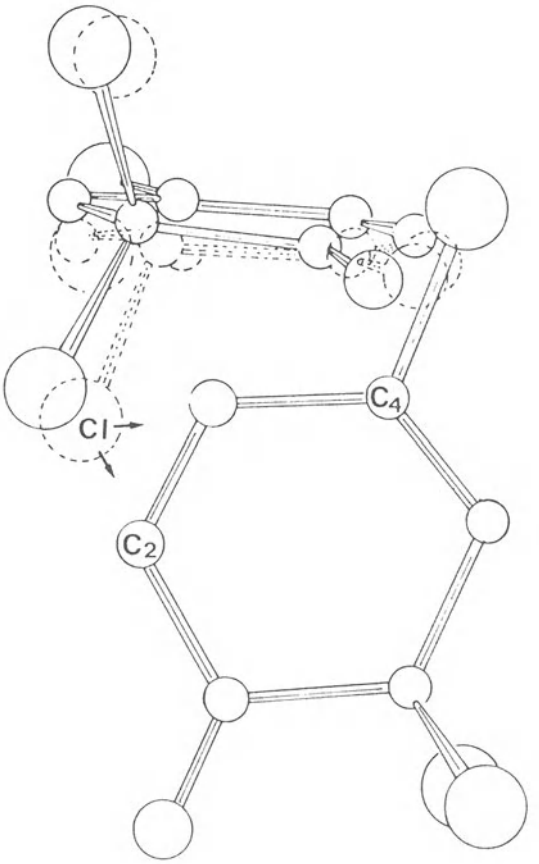


Fig. 4. Positions of two neighbouring molecules of 2,4,6,6-tetrachloro-3-methyl-5-isopropylcyclohexa-2,4-diene-1-one in the crystal from intramolecular packing calculations. The full lines represent the molecules in the initial crystal. The dashed lines represent one molecule after lengthening the C—Cl bond from 1.78 Å to 1.90 Å. Arrows mark the displacement of chlorine atom. Hydrogen atoms are not represented.

modelling purposes, a close relationship between energetic and spatial considerations. Relative to many other fields of theoretical studies, spatial informations are experimentally available at least for stable arrangement and so become important factors of comprehension. Because the amount of such information, geometrical representations (such as screen picture management) and treatments (such as packing analysis) are of vital importance and must be further developed. Molecular displacement modelling in solid state is expected to develop rapidly with progress and extensive use of molecular graphism and design. Thanks to the appearance of new devices such as the synchrotron [60] for measuring time-dependent location of atoms and nuclei, experimentators will be able to give more refined and more instantaneous pictures of displacement phenomena. Another trend leads to a better management of molecular displacement models with spatial representation

linked to energetic calculations. Let us survey briefly two typical examples of what can be done. Models can help crystallographers to choose between several plausible structures when they are dealing with very large molecules. With these techniques, the most stable conformation of mercerised cellulose compatible with crystallographic data was found by Pertsin [61]. Also, a camera film of the *z*-configuration of DNA was recently made [23]. In this film, Vergoten displays real displacements of molecules with time, as obtained from a dynamical treatment based on atom—atom potential calculations and IR, Raman, and X-ray data. On the other hand, displacement studies of molecules in organic solid state are closely related to drug design and, consequently, to devices and programs developed in this field. Based on X-ray data of biologically interesting molecules such as proteins (molecules of as much as 25 800 daltons have been crystallised [62]) and for pharmaceutical purposes in general, such studies deal with interactions among certain loci of proteins, for example, and various molecules to be tested regarding such or such properties to be fulfilled with regard to these loci [24]. Such problems, which are not exactly those of the solid state, show the possibilities of new devices and the direction in which phase transition and reactivity studies are going. Studies of displacements of organic molecules in solid state introduce a new approach to the question of representation. For hypersurfaces it is a problem to treat their many representative variables together, whereas spatial representations really contain the whole set of variables. The difficulties, but also the interest of such an approach could be compared to the functioning of our brain. Psychological science gives us more and more evidence for qualitatively different roles played by the right and the left parts of the brain. The right part seems to be devoted to the spatial treatment, "gestalt". Spatial agnosis is a very rare deficiency, due to the malfunctioning of the right side of the brain. People with such a deficiency seem to have no capacity for the representation of a map, for instance, in their head. Those people can memorize only sequential orders such as: I have to go 3 blocks, turn right, follow 5 blocks, turn left, count 5 entrances, and I arrive at my destination. They cannot recognize their way through a map representation in the head [63]. Some kind of intuition is probably lacking. For normal people, the question is worth putting: do we have a good balance between the two parts?

Acknowledgements

I would like to thank the American Chemical Society for facilities of access to the Chemical Abstracts databank through the Scientific and Technical Information Network.

References

1. R. Righini: *Physica* **131B**, 234 (1985).
2. G. Stuart Pawley: *J. Molecular Structure* **130**, 17 (1985).

3. J. Vicens, C. Decoret, C. Gaget, M. C. Etter, L. A. Errede: *Mol. Cryst. Liq. Cryst.* **96**, 39 (1983).
4. J. Vicens, C. Decoret, M. C. Etter and J. Royer: *Israel J. Chem.* **25**, 306 (1985).
5. A. J. Pertsin and A. I. Kitaigorodsky: *The Atom—Atom Potential Method in Physics and Chemistry of Organic Solids*, Springer Verlag (1986).
6. P. Claverie, N. Gresh and A. Pullman: *Int. J. Quantum Chem.*, Quantum Chemistry Symposium 13, 243 (1979).
7. F. L. Hirshfeld: *Theor. Chim. Acta* **44**, 129 (1977).
8. R. Ahlrichs, R. Penco and G. Scoles: *Chem. Phys.* **19**, 119 (1977).
9. K. Mirsky: *Proceedings of an International Summer School in Crystallographic Computing*, Delft University Press, 169—182 (1978).
10. D. E. Williams: *J. Chem. Phys.* **45**, 3770 (1966).
11. D. E. Williams: *J. Chem. Phys.* **47**, 4680 (1967).
12. J. Caillet and P. Claverie: *Acta Cryst.* **A31**, 448 (1975).
13. B. G. Dasheviskii: *Zh. Strukt. Khim.* **11**, 912 (1970).
14. J. Dumetz, G. Vergoten, J. P. Huvenne and G. Fleury: *J. Raman Spectroscopy* **11**(4), 221 (1981).
15. M. Aubert-Frecon, J. Royer, C. Decoret and B. Tinland: *Theor. Chem. Acta* **56**(2), 125 (1980).
16. G. S. Pawley and S. J. Cyvin: *J. Chem. Phys.* **52**, 4073 (1970).
17. S. L. Chaplot, G. S. Pawley, B. Dorner, V. K. Jindal, J. Kalus and I. Natkaniec: *Mol. Phys.* **39**, 251 (1980).
18. R. G. De La Vallee and G. S. Pawley: *Acta Cryst.* **A40**, 297 (1984).
19. A. I. Kitaigorodsky: *Chem. Soc. Rev.* **7**, 133 (1978).
20. A. Gavezzotti: *J. Am. Chem. Soc.* **105**, 5220 (1983); J. M. Connolly: *J. Am. Chem. Soc.* **107**, 1118 (1985); J. M. Connoly: *J. Mol. Graphics* **4**(2), 93 (1986); T. J. Richmond: *J. Mol. Biol.* **178**, 63 (1984).
21. A. Gavezzotti: *J. Am. Chem. Soc.* **107**, 962 (1985).
22. J. A. Ripmeester and R. K. Boyd: *J. Chem. Phys.* **71**, 5167 (1979).
23. A. H. Reddoch and N. S. Dalal: *J. Chem. Phys.* **74**, 1526 (1981).
24. C. Cambillau and E. Horjales: Proceedings of 'Journées systèmes informatiques et modélisation moléculaire', *U.C.B.L. Villeurbanne* (1986).
25. S. Berry, C. J. Roothaan and P. Rehmus: *Chem. Phys. Letters* **58**, 321 (1978).
26. R. Daudel: *Introduction to the Loge Theory* (Localization and Delocalization in Quantum Chemistry, vol. 1, O. Chalvet, R. Daudel, S. Diner and J. P. Malrieu Eds), pp. 3—39. Reidel (1975).
27. R. Constanciel, R. Daudel, O. Chalvet and L. Esnault: *Jerusalem Symposia on Quantum Chemistry and Biochemistry* **6**, 63 (1974).
28. G. Vergoten: Proceedings of 'Journées systèmes informatiques et modélisation moléculaire', *U.C.B.L. Villeurbanne* (1986).
29. K. N. Trueblood: *J. Mol. Struct.* **130**, 103 (1985).
30. K. N. Trueblood: *Acta Crystallogr.* **B39**, 120 (1983).
31. C. K. Johnson: ORTEP, *Report ORNL-3794*, Oak Ridge National Laboratory, Tennessee.
32. K. N. Trueblood, C. B. Knobler, E. Maverick, R. C. Helgeson, S. B. Brown and D. J. Cram: *J. Am. Chem. Soc.* **103**, 5594 (1981).
33. P. Chakrabarti, P. Seiler, J. D. Dunitz, A. D. Schluter and G. Szeimies: *J. Am. Chem. Soc.* **103**, 7378 (1981).
34. A. Gavezzotti, R. Bianchi and M. Simonetta: *Theochem* **28**, 391 (1986).
35. M. C. Etter and A. R. Siedle: *J. Am. Chem. Soc.* **105**, 641 (1983).
36. F. Bayard, C. Decoret, J. Royer, Y. G. Smeyers and J. J. Randez: *Int. J. of Comput. Chem.* **32**, 1 (1987).
37. S. R. Byrn: *Solid State Chemistry of Drugs*, Academic Press, p. 87 (1982).
38. A. J. Pertsin, Yu. P. Ivanov and A. I. Kitaigorodsky: *Acta Cryst.* **A37**, 908 (1981).
39. A. J. Pertsin: *J. Phys. Chem. Solids* **46**(9), 1019 (1985).
40. A. Gavezzotti and M. Simonetta: *Chemical Reviews* **82**(1), 1 (1982).

41. A. T. Hagler and J. Bernstein: *J. Am. Chem. Soc.* **100**(20), 6349 (1978).
42. J. Caillet, P. Claverie and B. Pullman: *Acta Cryst.* **A33**(6), 885 (1977).
43. J. Caillet, P. Claverie and B. Pullman: *Theor. Chim. Acta* **47**, 17 (1978).
44. Z. Berkovitch-Yellin and L. Leiserowitz: *J. Am. Chem. Soc.* **102**(26), 7677 (1980).
45. J. M. Thomas: *Pure Appl. Chem.* **51**(5), 1065 (1979).
46. R. S. Miller, I. C. Paul and D. Y. Curtin: *J. Am. Chem. Soc.* **96**(20), 6329 (1974); *J. Am. Chem. Soc.* **96**(20), 6334 (1974); *J. Am. Chem. Soc.* **96**(20), 6340 (1974).
47. M. Lahav, B. S. Green and D. Rabinovich: *Acc. Chem. Res.* **12**, 191 (1979).
48. H. B. Bürgi and J. D. Dunitz: *Acc. Chem. Res.* **16**, 153 (1983).
49. J. R. Scheffer: *Acc. Chem. Res.* **13**, 283 (1980).
50. V. E. Shklover and T. V. Timofeeva: *Usp. Khim* **54**(7), 1057 (1985).
51. J. D. Dunitz: *X-ray Analysis and the Structure of Organic Molecules*, Cornell University Press, Ithaca, N.Y., 312—318 (1979).
52. A. Gavezzotti and M. Simonetta: *Nouveau journal de Chimie* **2**(1), 69 (1977).
53. A. Gavezzotti and R. Bianchi: *Chem. Phys. Letters* **128**, 295 (1985).
54. A. Gavezzotti and R. Bianchi: *Tetrahedron* (1987) in press.
55. C. Decoret, J. Vicens and J. Royer: *Theochem* **22**, 13 (1985).
56. D. P. Graig, R. N. Lindsay and C. P. Mallet: *Chem. Phys.* **89**, 187 (1984).
57. G. Bertholon: *Quantum Theory of Chemical Reactions*, vol. 3, pp. 55—75, edited by R. Daudel, Reidel (1982).
58. A. A. Chernov: *Modern Crystallography III: Crystal Growth*, Springer-Verlag, Berlin (1984).
59. Z. Berkovitch-Yellin: *J. Am. Chem. Soc.* **107**, 8239 (1985).
60. D. Raoux: *La Recherche* **16**, 806 (1985).
61. A. J. Pertsin, O. K. Nugmanov, G. N. Marchenko and A. I. Kitaigorodski: *Polymer* **25**, 107 (1984).
62. C. Cambillau, B. Kerfelec, E. Foglizzo and C. Chapus: *J. Mol. Biol.* **189**(4), 709 (1986).
63. D. W. Zaidel: *La Recherche* **15**, 322 (1984).

Computer Molecular Modelling and Graphic Design

J. E. DUBOIS, J. P. DOUCET and S. Y. YUE
Institut de Topologie et de Dynamique des Systèmes, associé au C.N.R.S., Université Paris-7, 1, rue Guy de la Brosse, 75005 Paris, France.

0. Introduction

Man-machine communication is undergoing rapid change. After manual messages, i.e. via an alphanumerical text, the evolution, perhaps indeed the revolution, involves communicating by graphic interface. Already "image languages" are appearing, destined to gradually modify our ways of learning and communicating.

The association between image and idea is faster, often more direct, than that between words and ideas. In chemistry, the natural language is more that of formulas and molecular shapes than that of associated nomenclature. "A formula is worth many commentaries" might be a valid paraphrase of "a picture is worth a thousand words". However, in order to dialogue with images, one needs a material support which is at the same time very flexible and very interactive.

Molecular computer graphics had to wait for a computer generation endowed with graphic screens and efficient communication instruments (light pen or mouse).

Any *review of chemical graphics* is doomed to be quickly outdated because the graphic support is evolving swiftly (19 000 graphic stations in the world in 1986 and a predicted annual growth of 30% by 1992). This is not simply a matter of hardware. Also involved is a symbiosis in graphic stations, between the computer hardware functions and graphics, on the one hand, and the integration of perfected computer-aided design software in molecular imagery problems, on the other.

It is true that a molecular image can be conceived from structural primitives, but rendering the image then depends largely on the graphic station's tools (functions of lighting, color, display, zoom, texture). The first developments modeled molecules in the form of image-objects, more easily manipulated than material objects. Then graphic techniques and image synthesis encountered artificial intelligence and computer-assisted design. This meeting gave birth to "self-informed images", able to engender derived images and to react to constraints.

INTELLIGENT IMAGES

The image currently sought is an intelligent one whose information can be requested interactively. An intelligent image is often too heavy to be stored in the computer memory. Instead of memorizing it, one memorizes the steps leading to

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, 173–204.

its potential construction. When needed, it can be visually generated, step by step, by means of various programs, and this graphic construction (in calligraphic or raster versions) is often the most interesting, as it evokes filiations and analogies. From an informational standpoint, it is less cumbersome. Generative procedures have achieved a special status mainly in graph theory and for enumeration processes.

1. Simulation and Graphics

At first, chemical simulation was limited by computer hardware, as it depended on the association of large computers and graphic stations or terminals, whose performance speeds were very different. Now more and more powerful computers and graphic stations are making their appearance. Nevertheless, for *ab initio* calculation of large 3-D molecules, it is too soon to dream of such on-line imagery with a super-computer. On the other hand, on-line documentation exists for large files calling on common formulas.

In reality, thanks to the parallelism achieved between its computer hardware and its software as well as to the growth of its own power, 'infography' (a marriage of informatics and graphics) has become the communication tool for simulation. It provides molecular and biomolecular chemists with powerful interpretation tools developed in several important molecular fields (quantum chemistry, quantum biology, molecular mechanics, molecular topology, molecular pharmacology). Moreover, it gives scientists new techniques for the didactic presentation of theories of matter.

New application areas are opening before us and important fields of physical chemistry and theoretical chemistry, such as molecular dynamics, fluid dynamics and spectral methods, are more and more assisted by infography.

THE PRESENT SITUATION

In recent years the term 'molecular computer graphics' has become predominant, first as a logistics tool, then as a facet of artificial intelligence in its logic and linguistic aspects. We are probably witnessing the appearance of *chemical computer graphics*. This evolution takes place at the same time as the introduction of the general term *computational chemistry*, which is a return to the more general French expression of *chemical* informatics, broader because it includes all aspects of hardware and software.

1.1. AREAS OF MOLECULAR INFOGRAPHY

The first aim of molecular infography was the tridimensional representation of biological macromolecules [1, 2]. The idea was to create a computerised tool to visualise, manipulate and file chemical objects built from crystallographic data. Even building the usual conventional models (plastic stick and ball) for such

complex structures involves mechanical problems. The advantages of the graphic approach are then clear: capacity to detect and visualise non-standard situations (geometric distortions), to focus attention on local vision (seen from the internal areas of macromolecules), to treat these images in ways impossible with material models (comparisons, pattern recognition).

In fact, the first drawing conceived by C. Levinthal [1] dealt with a line segment for each bond and could be modified in real time. Such exploratory work aimed mainly at representing the biochemist's and the chemist's rigid structures (static modelling). Later the requirements of molecular biology gave particular importance to locating the *molecular surface*. Such a location is more important than that of the bond, for instance, in elucidating the interaction of an enzyme with a substrate or inhibitor [3].

This problem of molecular surface location is also essential in drug design strategies, i.e. interaction of a drug or hormone with a receptor site. These requirements provided the parallel impetus for progress in 3-D representation of molecules by stick and ball models, then by volume, calling upon spheric atoms coupled with various methods for visualising their bonds. Thus, different algorithms have been developed to represent surfaces either as unions of spheres or other smoothed surfaces.

Such representations have made it possible to simulate and represent atomic fluctuations around observed positions of balance in proteins [4] or concerted movements of fragments found in rearrangements or molecular interconversions [5]. Such examples constitute the application field of dynamic modelling. In order to interpret molecular behavior, in particular, to predict optimal structures capable of reacting with active sites or biosites, new tools have been developed to study molecular interactions. They broaden the scope of structural parameters: distances, surfaces, volume, accessible surface, hydrophilicity [6—9].

Molecular graphics uses the structural shape to support properties expressing the molecule's sensitivity to outside perturbations. In most cases, electronic indexes resulting from quantum calculation are used to characterise the behavior of interacting systems [10—11]. Localized electronic populations; total, π or frontier charges; super delocalizability define the spatial localization of electrons and their reactional capacity. The molecular skeleton becomes the "3-D molecular map", enhanced by localized information presented visually. For instance, a molecule creates an electrostatic potential in its neighborhood; an attacking reagent feels this first. We thus approach the molecule's behavior in relation to an external entity, and the consecutive dynamic interaction can be visualised either statically or dynamically. This example shows how simulation can gradually benefit from graphic tools. *Electron deformation densities* (the difference between electronic densities for the molecule and for the component atoms) play an important part insofar as they are directly comparable to crystallographic measurements [12]. We thus dispose experimentally of the real shape; this will be the support for graphically localizing finer electronic information.

The synergy of the computer and molecular graphics give us access to true simulation strategies for molecular interactions. The flexibility of the computer tool means that various types of representation will be available on demand, can be combined and can undergo, sequentially, pattern recognition operations based on related correlation treatments.

Often problems are formulated pragmatically. At the same time, image languages based on original or classical structural logic are developing. The DARC language, however, is both a metalanguage (topological codes) and a source of generic and/or specialized image languages [12, 13] that can be optimized for diverse documentation applications and for computer-assisted design (synthesis — structure elucidation — property prediction).

1.2. MOLECULAR INFOGRAPHY: NATURE AND SPECIFICITY

Molecular Computer Graphics is original in comparison with other areas of applied Computer Graphics because of the situations it deals with and the objects it manipulates. In these respects, we can distinguish three categories [14].

1. *Image Analysis* handles data arising from external probes (medical imaging, geographical descriptions from satellites). The images must first be scanned and sampled. Then digitalized data are structured (frequently via graphs) to facilitate handling of relevant information for subsequent treatment (pattern recognition, for instance) or various visualisations. Hierarchical ordering (tree building . . .) is often used to optimize traversing of model.
2. *Generative Graphics* creates pictorial representations of real or conceptual objects. First, an abstract description, understandable by the computer, is achieved as a model. This will be transformed into a picture on a display device.
3. *Cognitive Graphics* ("scene analysis") involves recognition of archetypes and relations between them. As such, it is related to the other categories.

Molecular Graphics is part of Generative Graphics. However, to synthesize computerized images representing usual objects or scenes, it is important first to search for a data structure or for simpler constituent parts ("primitives") to facilitate picture generation. This structuring can be achieved either on the object itself ("sketchpad") or on the space embodied (voxel subdivision).

For chemical objects this structure exists already. The problem is therefore to build mathematical images from known underlying algorithms. Atomic spheres constitute a set of primitives for the molecular body. For structural descriptors, generation and handling algorithms provide the organization laws. Electronic properties are derived from quantum mechanics programs (Figure 1).

These applications correspond to various complex levels. Representation of structural shapes relies mainly on geometric analytical handling. For more conceptual shapes associated with electronic localizations, the structuration algorithms must first generate the information to be represented in the 3-D space.

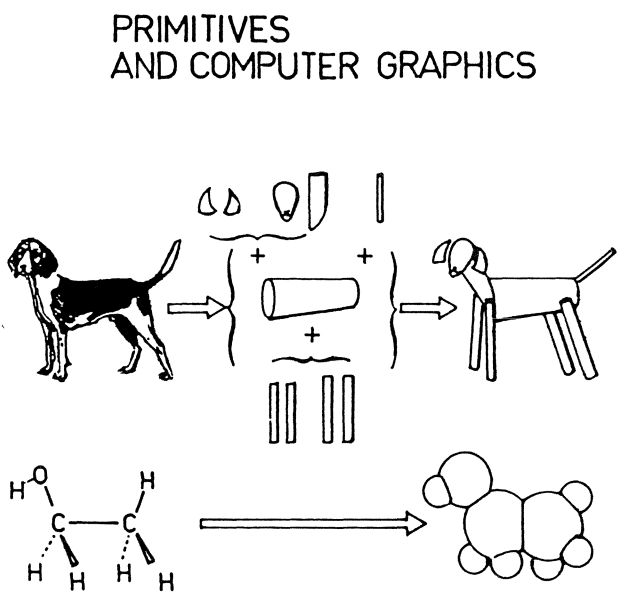


Fig. 1. Representation of real objects or beings first implies a search for primitives and knowledge of their combination laws (by trees or graphs). For chemical shapes primitives and/or structuration algorithms are already known.

The image is no longer a frozen visualisation detached from its physical context but an image informed about itself thanks to its correspondence with generation algorithms of a whole set of complementary information.

This specificity, apparently poorly perceived in the initial applications, leads us now to a clear distinction between the visualisation part, managed and assisted by classical graphic subsystems and the modelling part, controlled upstream by simulation programs (quantum mechanics, topological analyses). This distinction is beneficial on a practical level also, since it allows us to take advantage of interaction. Local handling of images, or limited perturbation of models are thus carried out graphically, and there is no need to return to the more cumbersome simulation programs.

1.2.1. *Display Devices and Basic Tools*

Without going into too many technical details about Graphic Systems, it is useful to observe that certain types of output favor certain representations. All systems are based on the cathode ray tube (CRT), but display devices use varied techniques.

In *calligraphic* (or vector) systems, line drawing is obtained by displacement of the electron beam of the CRT (random scan display). Good resolution is achieved with high interactivity. On-line animation allows for real-time dynamic modelling. However, the short duration of the phosphors requires the picture on the CRT to

be frequently refreshed, and this may cause "flickering" for complex drawings. Furthermore, solid areas ("images") cannot be displayed.

In *Raster* scan systems images are made by a matrix of discrete cells or pixels (picture elements), each of which can be made bright or not. The technology is close to that of television. The whole image is stored in a refresh frame buffer as a bit map. Color or grey levels can be incorporated by using additional bit planes. Such systems (cheaper than calligraphic ones) are very popular. They make it possible to represent both drawings, albeit with poorer resolution (at low definition, staircase effects may occur), and filled areas (solid images). However, interactivity is limited since the bit map must be updated after each object modification.

The direct view storage tube can be considered as a CRT with long lasting phosphors. The image is "written" as charges on a grid behind the screen. There is no need to refresh the image, but writing is a very slow process and since it is not possible to selectively erase parts of the screen, real time applications are limited.

1.2.2. *A Synthetic Camera*

The visualisation process has been compared to a synthetic camera of the chemical scene [18]. The view obtained depends on its position with regard to both the word and the observed objects. Local assistance is provided by the graphic system. It includes *geometrical transformations* (translation, rotation, scaling, zooming) and *rendering capabilities* to obtain more realistic images and particularly to suggest a 3-D vision from a 2-D screen.

Clipping (cutting slices of greater or lesser thickness) allows for views from within. *Depth cueing* (nearer points appearing brighter) and *perspective* both strengthen relief illusion.

Let us briefly mention some rendering operations, among those most frequently used. *Hidden-line or -surface* removal is one of the basic problems for both vector graphics and raster displays. Various algorithms have been proposed for working in the real object space (hidden line) or in the image space (hidden surface). For example, the popular z-buffer algorithm [15] has been used by Connolly [16] to represent Van der Waals surfaces. The z-buffer is separate from the main frame buffer storing the intensity of each image pixel: for each (x, y) value of the screen, it is updated with z values corresponding to the image elements closer to the observer, so as to select the visible parts. In the *3-D depth buffer*, hidden surfaces are not removed but rather graded along the axis according to their positions. Thus raster displays are given some of the capabilities of calligraphic screens, i.e. ability to clip images or to display a translucent surface through which the chemical structure may be seen from within.

In the scan line algorithm, the scene is sequentially explored in the image space (generally in the scan line order of the raster screen) to find the visible portions of the image.

For greater efficiency, z-buffer and scan line algorithms require some coherence

in the scene. The basic idea underlying Ray Tracing is that an observer views an object by means of light issuing from a source and reaching the observer by transmission, refraction or reflexion on the object surface. In fact, tracing rays from the observer to the object is computationally more efficient. Ray tracing was first used as a hidden or visible surface processor: rays terminate when they strike the surface of opaque objects. Special effects (reflexion of an object on the surface of another, transparency, shadows) have since been implemented. For solid models, intensity and colors to display the surface can be computed by *shading* software simulating the behavior of light on surfaces, taking into account reflectance, transparency of the model and illumination characteristics (diffuse or point source). For surfaces approximated by planar faced polyhedra, *smooth shading* (with Gouraud or Phong algorithms) restore the appearance of regularity [17—19].

1.2.3. *Creating 3-D Images*

A basic task of molecular infography is to create tridimensional views. The kind of image created and its construction mode will be influenced by the type of graphic station (vector or raster display) and the way of defining the shape to be handled, i.e. the modelling theory used to formulate a computer model of the object.

Once the chemical shapes have been defined and are considered as real objects, they can be represented directly as a volume; or represented by their surface envelope.

The efficiency of these methods is however strongly dependent on the upstream *data structure* created by analysis of the object and of the *model traversal* used during image creation.

For *solid modelling, constructive solid geometry* (CSG) describes the solid shape to be represented as a combination of other simpler shapes or "primitives". The scene is reconstructed through the basic operations of union and subtraction on simple objects (blocks, cyclinders). For easy welding and cutting, a convenient data structure can be provided by a directed acyclic graph [20] where each vertex represents a primitive or a combination of them.

Representation of chemical shapes by such solid models requires a previous object analysis and known primitives. Building molecular Van der Waals objects (e.g. CPK plastic models) needs only spherical primitives (representing the constitution atoms) combined by union operators. Richard's molecular surfaces [3] or re-entrant surfaces require both spherical and toroidal primitives.

For rendering purposes Fujimoto *et al.* recently proposed an accelerated Ray Tracing method to represent a DNA fragment of *ca.* 7011 atoms embedded in a spherical observation space [21].

Other treatments (*discrete solid modelling*) describe the solid object as a sum of small identical contiguous cells (operating at arbitrary but fixed resolution). Such systems are widely used in medicine (computer-assisted tomography) [22]. Rather than using identical cells, the *octree* method relies on recursive subdivisions of the cubic space investigated, into 8 smaller cubes which are encoded as "full", "empty"

or further divided till each final cube may be considered as homogeneous (Figure 2).

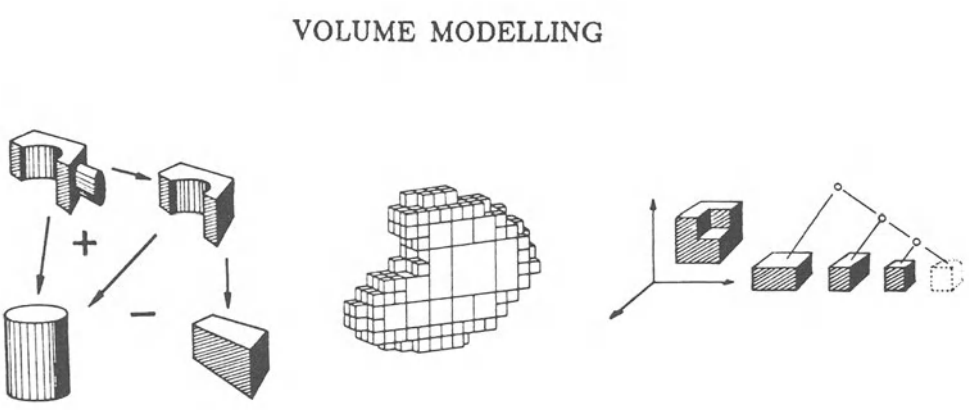


Fig. 2. Volume modelling via constructive solid geometry [20], discrete solid models [22] or bintree decomposition [22].

Surface modelling provides an alternative approach to 3-D body representations. Objects are described by shaping their envelope like a skin around them. Such representations — very popular in applications involving free form curves — are extensively used for representing electron properties (for which neither geometric primitives nor simple analytical expressions are available). Images can be presented by a set of 3-D *surface points*, a privileged representation in chemistry since it allows for maintaining the visibility of the molecular skeleton within (one way is to modify dot brightness), as a network of *polygon shaped planes* or as a *mesh of curved surface* patches [23] (Figures 3—6 [Color Plates I & II]).

1.2.4. *Serial Sectioning and Grid or Cube Methods*

As previously said, display of molecular surfaces or volumes can be approached through well-known geometrical primitives. More complex is the representation of property shapes (related to electron distributions) since they can be defined only through upstream application programs. No primitives can therefore be *a priori* determined and the preceeding methods of hierarchical trees are not of general interest. It should be noticed, however, that the octree technique has been recently used to determine the molecular surface on which electronic properties (MEP, electron densities) must be calculated. This CSG model is stored in a file to generate views of molecules where pixels are color encoded to give a 3-D map of electron properties. The hierarchical subdivision is mainly used here to reduce the computational space [10].

For the representation of electronic properties one must first generate a structured array of data and specify their organization (particularly the relations between the localization of the observation points and the values of the property

sought). In this field of molecular graphics, the modelling action generally uses a 3-D space division in a lattice "maintaining" the images, and manages data via either network or relational bases. In cube methods, computation and handling of numerical information are limited to the so-defined elementary cuboids ("voxels"). In grid methods they are defined by interpolating between the lattice nodes or sectioning voxels ("sub-voxel level").

These approaches rely on *serial sectioning* methods using cross section outlines (contour lines) in parallel planes (slices) usable in both surface and volume modelling. A combination of 2 packets of such contours in orthogonal planes provides a mesh of surface patches representing the 3-D surface, in chicken wire models [23, 26].

By connecting suitable data points of adjacent contours in a slice, triangulation algorithms give a mosaic of triangles defining the solid surfaces [24—28]. Although recently used in MCG for vector display (and well suited to plotters because of the simplicity of the drawing) [29], such triangulation algorithms have been mainly implemented on raster systems since they allow for color filling leading to solid representations by colored facets. Hidden surface treatment and Gouraud shading provide significant refinement in image rendering [24] and smooth surface. Serial sectioning can also be used in discrete solid modelling. The volume included between successive contour lines is divided into small cuboids considered as basic cells of the solid body.

At sub-voxel level, elementary cells can be spanned in smaller units (five tetrahedra are created by cutting off corners, plus the central part). Intersections of the contour (to be displayed) with the edges of these tetrahedra, determined by linear interpolation, define the vertices of the generated triangles [27].

Triangulation has also been used for quantitative treatment. The algorithm proposed by Connolly operates from analytically defined molecular surfaces [30]: triangles are created by recursive subdivision of the edges limiting the curved faces (part of spheres or tori) constituting the molecular surface (to use a unique algorithm, these faces are previously defined by similar geometric descriptors). Once coordinates of each elementary triangle are known, molecular parameters (areas, volume) can be easily computed.

2. Chemical Shapes and their Representations

Though normally described by 2-D chemical formulas with the molecular concepts of chemistry, compounds are currently considered as tridimensional objects.

The tendency at present is to develop "realistic images" for the purpose of elucidating the internal structure of the real (visible) world and of modelling it mathematically, in order to reproduce it with maximum precision. Chemical objects are invisible, but their volumes and properties can be represented through various mathematical tools. Such "objects" are discovered through quantum methods or molecular mechanics; thus, they correspond to their "realistic images".

By virtue of programs, one can derive, from these theoretical or semitheoretical models, various and realistic representative images, more or less stripped of certain details.

Various representation levels can be attained by computer graphics depending on the kind of information encoded or displayed. These images allow for model testing and simulation as well as for predictive ability. They can be perceived either as models of the real situation or partial visions of larger structures. Contrary to the magnet which is itself visible while its action is invisible but materialized by iron filings, for molecules synthetic images reveal two realities, invisible but known, thanks to theory and experimentation.

2.1. STRUCTURAL BODIES

The simplest modelizations show molecules according to the general organization laws of their constituting elements (atoms) endowed with a shape and with specific metrics in certain standard states.

Ball and stick or wire type representations allow for rapid comprehension of the nature of atoms and their geometric relations. Space filling models aim at a more realistic view of molecular volumes. Atoms viewed as spheres of prescribed radii are represented by computerized images similar to the popular plastic CPK (Corey, Pauling, Koltum) models [31] (Figure 7 [Color Plate II]).

For biological macromolecules too complex with all their component atoms to be individualized, larger primitives (prefabricated standard species) relieve the difficulty of complete representation by various simplifications. Descriptive structural primitives (amino acids coded symbolically by three-letter initials) or pictorial primitives (ribbon, cyclinder, spiral) condense the essential morphological characteristics. For instance, for on-line modelling the secondary structure of proteins, schematic representations can be obtained from cylinders or regular spirals to depict α helices, broad arrows for β strands ... randomly coiled polypeptide chains indicated only by their C_α—C_α bonds [32].

Ribbon models use the peptide sub-units as geometrical guide coordinates and *B*-spline technique to generate the set of nearly parallel threads running along the length of ribbon that they constitute [33].

In these complex situations, computer technique and graphic capabilities (windowing, zooming) are of prime interest in allowing for simultaneous displays of global qualitative features and local quantitative details (Figure 8 [Color Plate III]).

For such descriptions stressing the geometric characteristics, X-ray coordinates constituted the first source of input, particularly with extensive crystallographic data banks (Cambridge Crystallographic Databank, Brookhaven Protein Databank) [34, 35]. Theoretical methods (quantum mechanics, molecular mechanics) offer an alternative approach for geometric model building.

Recent computer techniques allow for interactively assembling fragments or interpolating between existing structural data to constitute complex systems, when the experimental information is not available.

For biological systems, increased interest focused on molecular internal and external surfaces and volumes. The external surface of a molecule gives a first image of the areas able to interact with suitable ligands or with the surrounding medium. The Van der Waals contour, union of the constituting atomic spheres, is accepted as a common representation of the molecular body [31].

In macromolecules, an important part of the molecular surface corresponds to the internal part of the edifice, and the boundary volume that can be occupied by the solvent without penetrating the molecule must also be specified. This is done by rolling onto the molecular shape a probe sphere simulating the solvent molecule or its size.

The *solvent accessible* area was first defined by Lee and Richards [3a] as the surface pushed outward from the Van der Waals surface by a distance equal to the radius of the rolling probe sphere (commonly 1.4 Å, to mimic a water molecule). Later the molecular surface was defined as the *contact surface* (part of the Van der Waals surface accessible to the probe sphere) and the *reentrant surface* (inward facing surface of the probe when it is simultaneously in contact with 2 or more atoms) [36] (Figure 9).

Various algorithms have been proposed to derive such surfaces and calculate the relevant parameters (surface, volume). One of the most popular is the dot

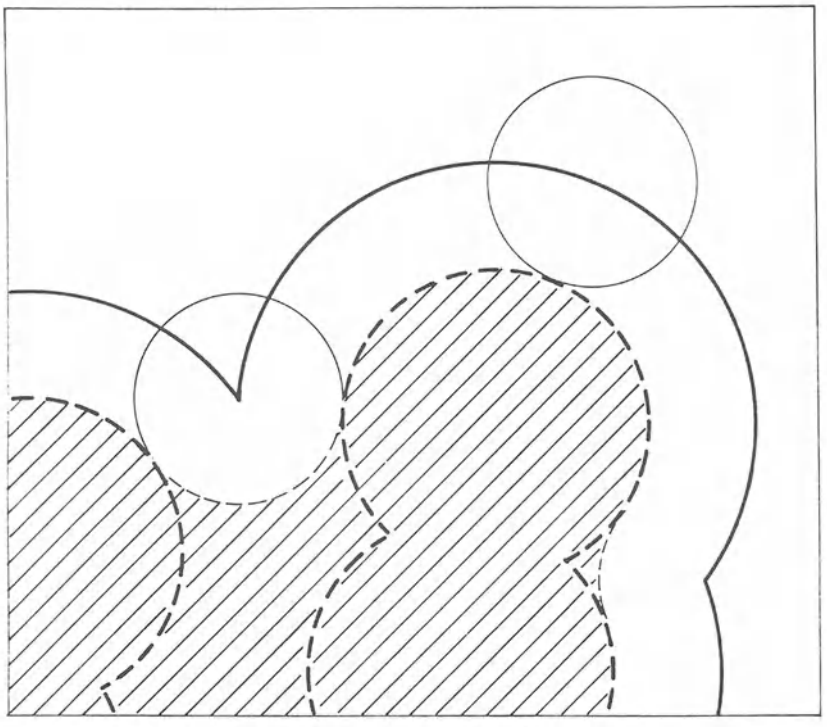


Fig. 9. Molecular body, molecular surface (dashed line) and solvent accessible area (straight line) [3, 36]. The small sphere mimics a solvent (or water molecule).

numerical algorithm of Connolly [37] where dots are scattered on the molecular surface at approximately constant density per unit area. Either numerical (subdividing the body into a large number of small similar cells through grid or cube methods) [38—40] or analytical software (using usual geometric formulas for the constituting parts: convex, saddle or concave faces from spheres or toruses) have been used [8]. In the cube method of Pavlov and Fedorov [39], the portion of space investigated is divided into small cubes (edge length 0.3 Å) encoded according to their location with respect to the molecular surface of the "hydrated molecule" (i.e. "within", "on its surface", "beyond it"). Similar treatments have been used by Pearl and Honneger [41] for the display of Richards accessible surface. Marsili *et al.* [42] proposed a simple method for Van der Waals surface calculation and Meyer *et al.* [43] derived molecular parameters (for structure/activity relationship application).

The *solvent excluded volume* has been defined as the sum of the molecular body plus those interstitial volumes too small to accept a probe molecule [36]. This solvent exclusion surface is composed of the points where probe spheres most closely approach protein atoms. Patches of the convex molecular surface are joined by concave parts of (water) probe spheres "blocked at the opening of the narrow molecular fissures". The accessible surface can be determined simply by starting from an (x, y) grid and dropping the probe along the z axis until it is in contact with the molecule [40].

2.2. PROPERTY SHAPES

One of the basic advantages of molecular graphics is its ability to superpose property shapes characterizing its behavior with regard to the surrounding environment (solvent, ligand, biological receptor) onto the symbolic molecular body. As to the preferred areas for solvent accessibility, the hydrophobic character is often sketched out (and color encoded) with only the nature of the atoms.

For more refined descriptions, one frequently refers to the display of electronic features derived from theoretical chemistry calculations.

Quantum parameters (such as electron population) have long since been accepted as quantitative reactivity indices to predict behavior in interacting systems. Rather than such numerical values, conventionally attributed to individualized sites, 3-D graphics displays the spatial localization of electron distributions and allows for a more global vision of the reactive areas embedded in their molecular environment. *Electron density* (sometimes Frontier Orbital densities in semiempirical methods) [44] and *Superdelocalizability* are more widely used. Charge distribution in the molecule creates an *electrostatic potential* (MEP) in its neighborhood. This is what an incoming electrophilic reagent or a binding site first perceives. This electrostatic potential corresponds to the main contribution to interaction energy (at least for large distances) and thereby gives an "electric image" of the molecule [11], more global than those attainable by localized

electron populations [45, 46]. *Electron deformation densities* (difference of electronic densities between the molecule and its supposedly isolated constituting atoms) can be directly compared to experimental crystallographic measurements and are very useful to define the electron reorganization occurring during bond formation [12].

Recently, the electrostatic field has been proposed in place of the electrostatic potential of which it is the gradient. This index is mainly governed by the electron distributions nearer to the observation point (the distance dependence is r^{-2} and not r^{-1}) [47]. Though a vectorial (and not scalar) quantity, its display can be approached by arrows. Similarly recent interest also focuses on gradients of electron density. This quantity (related for instance to the bond path defined by Bader [48]) can give new insights into the molecular structure particularly regarding the response to limited perturbations [27]. Shading isometric surfaces according to the magnitude of the gradients affords an efficient representation.

Quantum simulations often require lengthy calculations which weigh heavily on computer time requests and hamper interactivity. MEP, in particular, involves numerous integral calculations that can be alleviated by various approximations. In the semi empirical CNDO method, the VSS [49] treatment approximates the nucleus attraction integral by minus the repulsion integrals between *s* orbitals; "deorthogonalization" of the density matrix has been proposed to give better quantitative agreement with ab initio calculation.

Various treatments have been proposed to substitute faster analytical formula evaluations [50] for full integral calculations. In the crudest method, coulombic interactions are simply derived from the partial Del Re or CNDO electronic charges localized on atoms (monopolar approximation [51]). More refined treatments use multipolar expansion (associated with division of the investigated molecules in simple subunits).

For polypeptides and proteins a reparametrization of the empirical Hückel Del Re procedure was recently proposed for defining optimized monopole expansions which closely reproduce the electrostatic properties (potential, field) obtained from ab initio wave functions. The reduced computational time would be very useful when studying conformational influences on electronic properties [52]. Also relevant to this type of approach is the work of Clementi *et al.* [53], in which transferable atom pair potentials are defined by fitting ab initio results with simpler analytical expressions, in order to calculate interaction energies between Li^+ ions and simple organic substrates (ether, thioether, amide systems).

These various indices (the electrostatic field excepted) are defined as scalar functions of the observation point coordinates $P(x, y, z)$. We derive contour surfaces on which the property value $P(x, y, z)$ is constant ("isometric contours"). In such 3-D displays, in order not to overload the drawing, one can only represent a limited number of such surfaces since they often partially mask each other. Local graphic capabilities provided by the display strengthen the relief illusion allowing for a 3-D perception from a 2-D drawing on the display screen.

Heavy MEP visualisations can be limited to a few characteristic isometric contours. For instance, highly attractive values (−50, −100 Kcal/mole) can be used to identify privileged approach areas for an incoming electrophile. The neutral contour separates the attraction and repulsion areas around the molecule. The envelope associated with a strong repulsion threshold (100 Kcal for instance) might be related to the molecular body [54].

For such 3-D displays, various representations have been previously indicated (wire mesh models or colored facetted solid bodies). Reducing the representation space (by introducing geometric constraints) leads to simplified views allowing for simultaneous displays of a larger number of isometric contours.

Several options have been proposed:

(i) — map of isometric lines in a given plane: It has been recently proposed to use a set of such maps in equidistant parallel planes to obtain detailed 3-D information via a perspective view of all these maps [55];

(ii) — calculation of a given electronic property, especially MEP, on the Van der Waals surface (molecular body) or on various layer surfaces derived from it [10]. Such calculations can be made by using the dot algorithm (included, for instance, in the Gaussian QCPE package), or, more recently, through an octree technique: the molecular surface on which the MEP is to be calculated (and color encoded) is produced by recursive "spatial subdivision". The cubic space area investigated is divided into 8 smaller cubes. The empty ones are discarded as are the cubes situated entirely within the body. The remaining cubes contain some parts of the surface and are further subdivided into eight down to pixel resolution [10];

(iii) — 2-D view of these contours by projection onto a given plane — outside the molecular body [56].

With graphic raster displays of 2-D maps, one can color the areas between iso contours corresponding to preset values. For 3-D views (as previously seen) the envelopes can be presented as compact solids providing better insight into volumic features (Figures 10, 11 [Color Plate IV]).

2.3. BUILDING DIRECT AND COMPOSITE IMAGES

In order readily to obtain varied views of the molecular complexity, we recently developed a polymodelling system (POLYMOD). Flexible 2-D or 3-D displays of chemical shapes, whether structural or electronic, can be generated or combined into composite images [54, 57]. Important features involve:

— immersing the molecule or molecules studied in a unique representation space;
— structuring this space by a 3-D grid, on the nodes of which the property sought is evaluated;
— managing reversible correspondences (by coordinate changing procedures) between this unique grid stage and the initial coordinate systems in which individual molecules are defined (calculation space) (Figure 12).

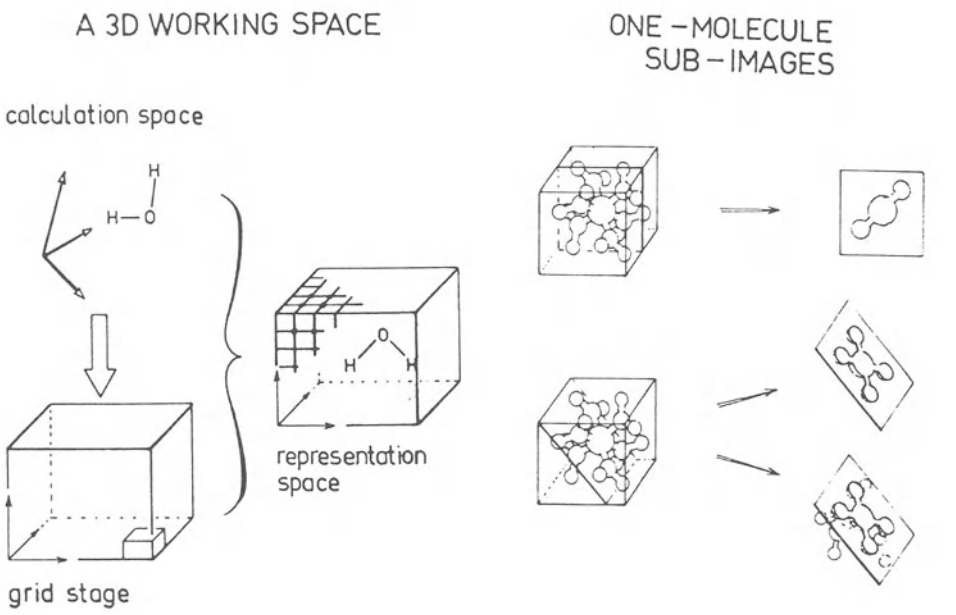


Fig. 12. In POLYMOD, by embedding the molecule into a grid stage acting as unique representation space, easy building of varied subimages is possible.

One useful feature in POLYMOD is the use of the grid *index* both to structure data associated with the property sought and to monitor the representation of isometric contours on a vector graphic device. Two files are used. One contains the description of the mesh. Each node is defined by a triplet index (i, j, k) which allows for unambiguous location in the representation space. The other file gathers the property values on the nodes ordered according to the sequence of their location indexes (i, j, k) (Figure 13).

This structure allows for easy handling of data, either for electronic properties defined as functions of space coordinates or for geometric conditions encoded. Multislice sections of the grid stage allow for determining the isometric contours associated with a given value of the property examined. Selection of the relevant nodes ("valid nodes") leads to dot-images at a first answer level. The representation can be refined by returning to numerical values associated with valid nodes and by using a search algorithm based on *voxel edge intersection*. This algorithm relies on the definition of frontier nodes (contiguous nodes respectively inside and outside the envelope sought for). Returning to initial numerical values, interpolations specify the limit points: intersection of the lattice edges with the envelope. For use with a calligraphic system, limit point extraction leads to a *linear linked list* indicating for each point the localization of its neighbors for sequential drawing of polygonal contours. Hidden line removal facilities are easily provided by the corresponding "valid nodes" or "limit point" algorithms for a more realistic presentation (Figures 14, 15).

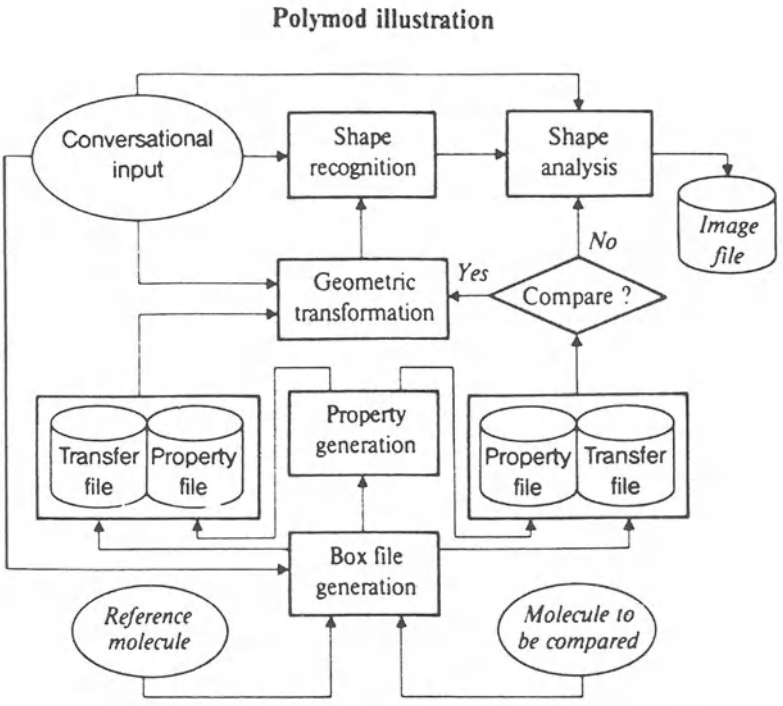


Fig. 13. Schematic diagram of the POLYMOD System [54].

This data management and representation approach makes it possible to perform Boolean operations on structural features and also (more originally) on electronic localization areas. (MEP, electron density). Such comparisons play an important role in shape recognition processes, intervening for elucidating interactions of systems in close spatial proximity or for studying the evolvement of properties in families of related structures. Local subprograms ensure management of coordinate changes, and geometrical transformations between the unique representation space and the various calculation spaces (this organization facilitates the computation of electronic properties particularly owing to the directional character of ab initio wave functions). Local assistance programs also handle geometric transformations and anchoring procedures required in shape comparisons (Figure 14).

2.4. CHEMICAL OBJECTS: STRUCTURAL AND IMAGE LANGUAGES

Certain applications can be handled through more symbolic representations, closer to the chemist's usual language; that of structural notation and of developed formulas.

Chemistry can be studied by classical informatics handling, but it can also lead to representations built on chemical vectorial representations.

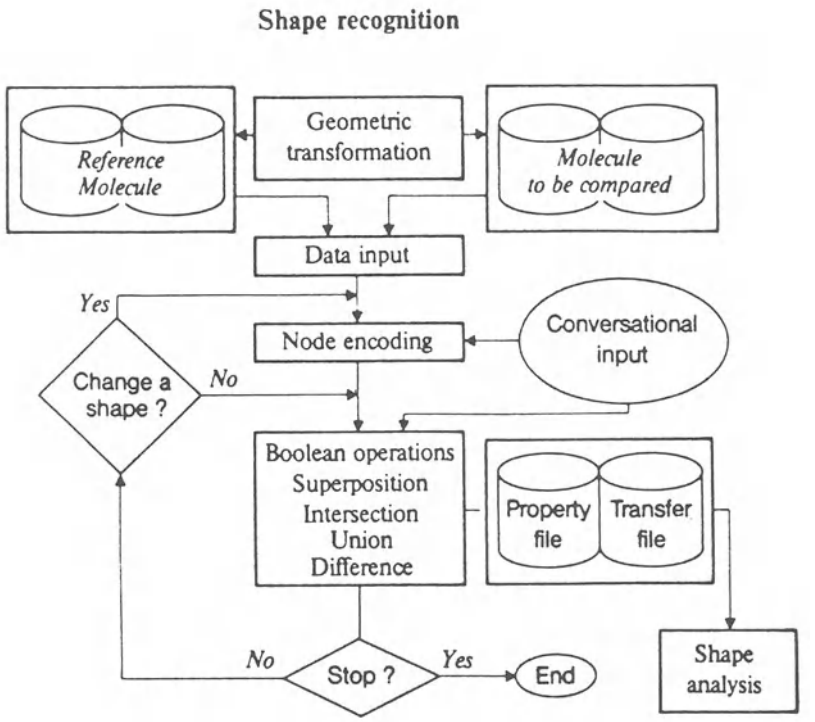


Fig. 14. Shape recognition in POLYMOD System. Comparison of two molecules via Boolean operations. The resulting data are stored in the Transfer and Property files and then passed to the "shape analysis" part for graphic display. After the "geometric transformation" step, files for the two molecules refer to the same nodes of the common representation space.

To benefit fully from computer possibilities, new languages had to be invented. This was an important part of the new field of "chemical informatics".

In standardizing chemical languages, the initial aim was to translate image-like formulas into alpha-numerical texts as is done, for instance, in systematic chemical names. Recent progress (DARC, CAS MORGAN) bypasses the text and ends up with a coding metalanguage of the images: graphs enriched by textual information constitute chemical pictures perceived as veritable ideograms [13, 58—60].

Extending the logic of structural chemistry, code families handle molecular information by structural formulas and by relationships.

Two important requirements are:

— to ensure efficient handling of large databases,
— to favor artificial intelligence treatment.

Indeed if the chemist's dialogue goes from the question-formula to the answer-formula or to combinations of these, we dispose of an image language, with primitive icones or substructures and with operations also translatable into image language (insertion, suppression).

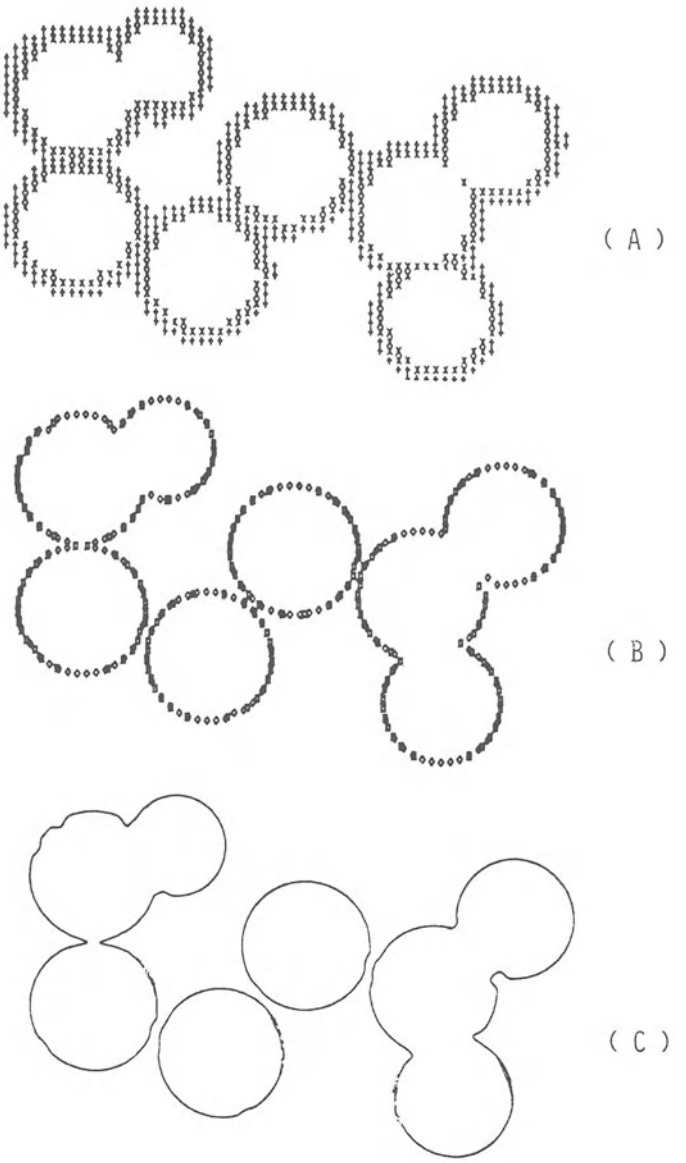


Fig. 15. Shape analysis through frontier nodes and limit points extraction: aminobutyric zwitterion (slice 0.2 Å above molecular plane). For illustration purposes, atomic sphere radii have been arbitrarily set to half Van der Waals values. (a) Frontier nodes (x inside, + outside); distance to atomic center = radius ± threshold. (b) Limit points extraction. (c) Polygonal contours.

The complex substructure thesaurus leads, in the DARC System, to an association of images and of metalanguages for handling them. This is made simpler by the graphic nature of chemical formulas which provides access to graph adjacency matrices useful for encoding purposes.

Among the areas of application of chemical informatics is that of documentation, including the management of chemical data bases comprising several million structures. The entity structure (reaction or compound), the foundation stone of information in such systems, is topological by nature and can be modelled by more or less colored chromatic graphs and by their connex matrices. Isomorphism and homomorphism operations in data handling allow for specific structure searching and for recognizing similarity classes for compound families with a common fragment.

Among the two systems internationally accepted as currently operational in this field, the CAS (Chemical Abstract Service) uses a complex language (Basic CAS) where the compounds are encoded into matrices via a dictionary of familiar fragments.

2.5. ORDERED STRUCTURES: DARC SYSTEM

In the DARC system, developed by J. E. Dubois since 1966, the interaction between chemist and computer uses the structural formula as a natural language via a numerical ordering based on topological environment concepts. All dialogues rely on interactive molecular imaging. From the structural formulas, the DARC codes are computer generated in a topological and chromatic representation relying on ordered chromatic graphs [13, 59, 61]. Basically, the DARC system orders any structural environment, either internal or external. Ordering is attained by progressive constructions controlled by algorithms.

2.5.1. *Geometric and Graph Segmentation and Organization*

The structuring of semantic spaces (molecules and populations) is based on a geometric organization of graphical structures into concentric subspaces and on an internal hierarchy of defined modules.

This approach makes it possible to replace the usual static fragmentations by a dynamic vision through which chemical entities can be perceived, either globally or with regard to their relevant fragments. The former notions of compound, function, family are replaced by those of structure, substructure, hyperstructure (S, SS, HS).

By this means, clear relationships are established on the local level (substructure SS), on the level of entities (structure S, where site ordering in achieved) and on that of populations (hyperstructure HS, where external hierarchy of entities is involved). This trilogy makes it possible to manage all applications with a single language and a single logic, thus increasing efficiency and access to new concepts.

Structure: In the DARC code, compounds are described by ordered chromatic graphs whose nodes represent atoms and whose edges represent bonds between these atoms. (Structural information, such as atom nature, bond nature, charge, . . . constitutes the chromatic elements). Adjunction and ablation operators, applied

sequentially, ensure generation and, through a modular concentric description, lead to an ordered linear descriptor (Figure 16).

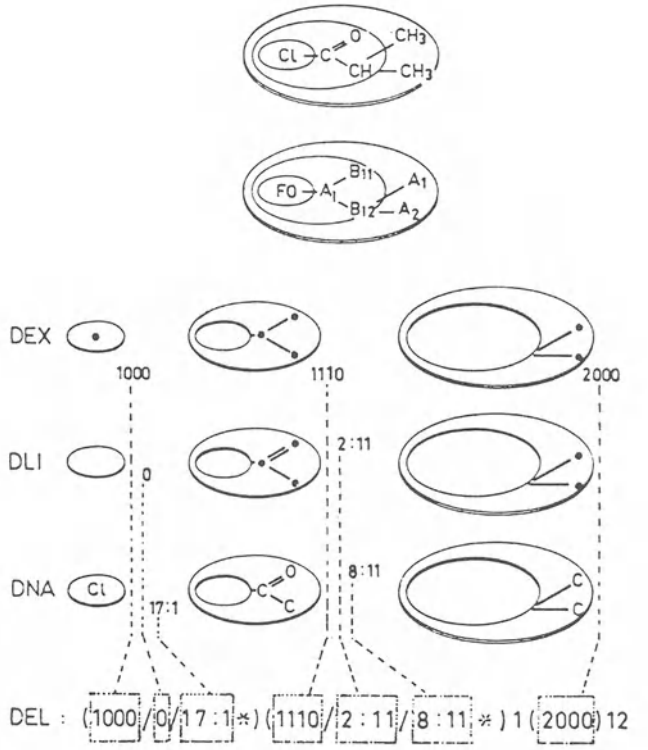


Fig 16 Concentric module description of a compound For each module, the Limited Environment Description (DEL) consists of three parts which describe respectively the topology (DEX), the bond multiplicity (DLI) and the atom nature (DNA) In the example the focus is described by the first three boxes topological description of the focus (first box), existing bonds in the focus (here 0), atomic number of the chlorine atom [17] The following three boxes rely on the first ELCO of the environment topology, bonds (double bond towards the B_{11} position), atoms (oxygen in B_{11}) and so on Chromatism of simple links and carbons are implicit

Substructure: Chemical information (reactivity, spectral or pharmacological properties) is very often linked to the presence of characteristic structural fragments. In the DARC system, these fragments or substructures are handled by concentric modules or FRELs (FRagments of Environment that is Limited). These FRELs, developed from a bond or from an atom, are associated with different information content according to the fineness of the structural description sought.

This latter is a factor of both depth (number of atom rows from the origin focus) and of precision (exact or fuzzy atom or bond chromatism). The ability to handle images with generic concepts is the basis for the DARC system's usefulness in artificial intelligence (Figure 17).

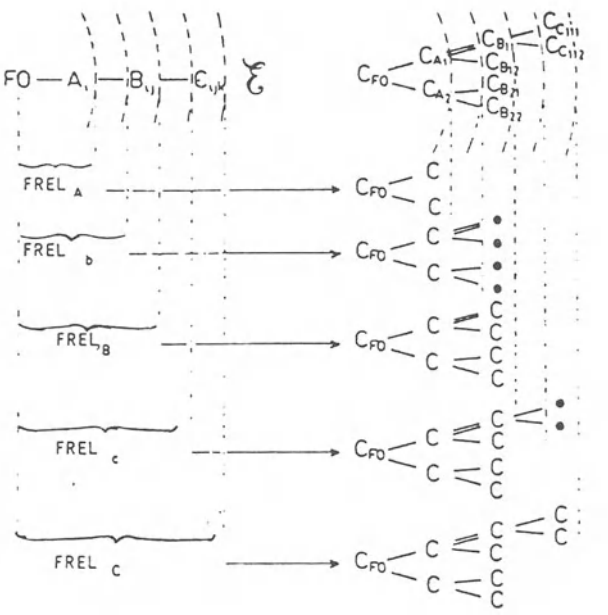


Fig. 17. Generation of Fragments Reduced by an Environment that is Limited (FRELs). The different types of substructures originating from a given atom are hierarchically ordered.

Hyperstructure: The set of compounds formally generated in this construction of a target structure is grouped in the generation hyperstructure. This hyperstructure, built in synchronism with the structures composing it, forms a representation and classification space wherein structural proximities and filiations within populations of chemical entities can be easily handled (Figure 18).

These three notions of structure, substructure and hyperstructure, in association with the theories of generation-description, topology-information and population-correlation, optimize the handling of chemical situations. They unite all the essential elements of major chemical representations: the topological one of structural chemistry and those of quantum mechanics and of molecular mechanics.

All three notions (S, SS, HS) are accessible by computer graphics and require color graphics to take full advantage of the topochromatic building of molecules, either through node by node construction (DUPEL) or through bigger fragments such as rings, functional substructures and superatoms. The graphical menus for such input of structural data usually reflect fragmentation expressed as "object-icones" and simple actions that can be "action-icones". The construction grammar allows for symmetrical action in query procedures. Multi-step graphical assistance, such as local flickering, local deleting, is usually considered (Figures 19 and 20 [Color Plate V]).

2.5.2. *CAD Structural Handling*

This type of recursive and association-based organization creates space states

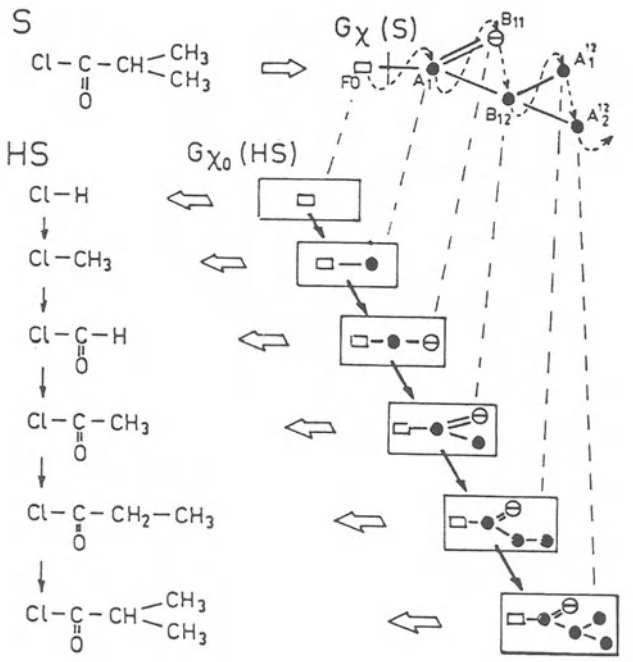
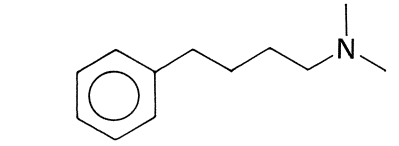


Fig. 18. Synchronous generation of a target compound and of the associated Hyperstructure HS.

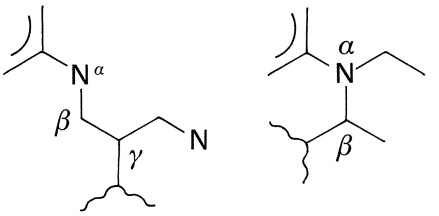
useful for localizing artificial intelligence and computer-assisted design problems. They also provide basic sets of tools for CAD. In this approach, largely oriented towards structural formulas, the use of graphic tools for input and output simplifies man-machine communication by employing the chemist's natural language (the codes are generated by the computer so as to be transparent for the user). Graphic representation is then ensured, not from atomic coordinates (such as crystallographic data) but through topochromatic elements generated by codes and corresponding to customary editing conventions. Such topological data bases are situated between the usual alphanumerical data bases and image data bases [12].

One example is the on-line querying of large databases: nearly 7 million compounds of the CAS file are available in the DARC-EURECAS file where structural questions and answers are handled through computer terminals. The EURECAS Data Bank Management System (DBMS) is based on structural matrices and codes, but input and output formulas are graphically generated by expert systems; graphic output is interactive. Computer-aided design has been extremely useful in property optimization and prediction. For example, we can optimize correlation procedures (physico-chemical properties or pharmacological activity) and efficiently manage systems of elucidation and spectra simulation (DARC-EPIOS system in ^{13}C NMR, DARC-SIRS in IR) [61—65], or propose new approaches for recognizing pharmacophores in drug design, all thanks to the manipulation of global or specific structural descriptors.

A recent pharmacological application relates to neuroleptic activity [66]. This was linked to the existence of Jensen (A) structures, but there are exceptions, and this interpretation required further investigation. Starting with a bank of psychotropic agents, by analyzing the FRELs constituting the molecules, we can refine the description and define those substructures participating in the activity (B) or with no influence (C).



(A)



(B) (C)

In the SIRS program (Simulation of Infra-Red Spectra), helpful for computer-aided elucidation, the IR spectrum of a molecule is reconstituted by extracting the FRELs centered on *different bonds* (substructures) and, for these structural primitives, searching for their associated subspectra in a bank of spectrum primitives [64].

3. Dynamic Modelling

The ability of graphic displays to fly rapidly from one type of representation to another is of prime interest in understanding the spatiotemporal behavior of chemical systems.

Chemical education movies have been proposed to present important typical processes: ammonia inversion, cyclohexane interconversion, cation rearrangements, electrocyclic reactions [5, 67—70].

One of the main advantages of such representations is the simultaneous display of structural modifications (changes in geometric or conformational parameters) and the corresponding energy and/or orbital aspects. The stereospecific approach to reagents, the orbital interactions, etc. can be illustrated in relation to the progressive evolvement of reaction intermediates. At the same time, the potential energy profile is followed as the reaction goes on. Interactivity allows the user to

stop the visualization during interesting events and to manipulate the scene displayed by rotation or zooming, etc. for better insight into relevant details.

In developing such animated sequences the modelling part often requires much more effort than static simulations. First, on-line visualization demands efficient graphic displays and application programs to generate successive pictures at a convenient frequency (i.e. 8—24 frames/s). (If such conditions cannot be fulfilled, frame by frame photography can be used to reconstruct animated sequences in films). Animation must therefore be supported and driven by an adequate data base. At several points along the reaction pathway, molecular geometry and energy of the evolving system must first be computed (generally through quantum mechanics programs). Then one has to define a strategy for continuous display and optimal illustration of the reaction process by exploring the corresponding data base.

A compromise between esthetic drawings and images illustrating the common chemical concepts seems to be a sensible guideline among the varied interpolation strategies available.

Another important application domain was recently opened with *Molecular Dynamics Simulations*, first initiated by Karplus *et al.* [71] with the study of time dependent fluctuations of macromolecules around their equilibrium positions. With such applications, as with the visualization of molecular vibrations, modelling action now brings the time dependence of microscopic phenomena down to our human time scale.

To understand dynamic properties on the molecular level, computer simulations (Molecular Mechanics, Molecular Dynamics, Monte Carlo Calculations) provide the more detailed theoretical approaches. However, such complex and lengthy calculations result in extensive output and large coordinate sets from which it is arduous to extract interesting data.

In this field where, up to now, no full theoretical descriptor can reproduce all features, computer graphics allows a selection of important events and acts as a guide in suggesting more quantitative treatment for a clearer understanding of these specific situations. Weiner et col. pointed out that, for biopolymers, molecular flexibility may have important implications for transmitting information from one part of a complex bioorganic system to another [72].

Relevant to this topic is the knowledge of internal correlated geometrical changes, the influence of strain in the motions of adjacent fragments, and the overall ability to relieve strain by small local adjustments.

From initial equilibrium positions, atomic motions can be visualized on a scale of picoseconds and their displacements can be color encoded for easier analysis.

Computational statistical mechanics complements the somewhat rigid view of energy minima provided by Molecular Mechanics; perturbations introduced as part of Molecular Dynamics enable us to investigate how the system can leap over small barriers to reach neighboring local energy minima [73].

The ability of recent efficient programs to handle calculated data and to build

up images at reasonable speed for systems up to 2000 atoms makes it possible to follow, step by step, a large number of the consecutive conformations generated, and to group these displays in animated sequences.

The new graphic tool seems very promising for the study of concerted motions or adaptations such as those occurring in dynamic intermolecular interactions (enzyme-ligand, receptor-inhibitor). In these situations, difficult to analyze from lengthy coordinate listings, it provides predictive capabilities.

Recent applications clearly point out the potentialities of that graphic tool in simulation strategies for selecting important features (otherwise hardly discernible) in the dynamic behavior of complex systems. Examples are:

- characterization of sugar puckering in DNA (sugars of guanine residues change conformation much more frequently than sugars of the cytosine residues) [72];
- study of orientationally disordered crystals [74]. In such media (intermediate between liquid or true crystals), spatial positions are ordered (as in a lattice) with more or less disordered orientations. Depending upon crystal field strength one can observe frequent jumps between a few discrete orientations or a whole range of continuously variable orientations;
- graphic displays of dynamic trajectories derived from extensive Brownian dynamics simulations have been used also to study the influence of the electric field of an enzyme steering substrate molecules towards its active site [75].

These dynamic applications take advantage of the eye's capacity for analysis and synthesis. The diverse component parts of a system and their temporal evolution are easily perceived when the starting point is a sequence of several images rather than a long commentary detailing each successive modification of the scene. Efficiency is increased by the interactivity that allows the user to freeze the image during a key situation deserving a detailed scrutiny, to backtrack, to modify observation conditions during limited perturbation treatments.

The privileged association of image and idea provides a holistic view of a complex reality and transfers a fine understanding of phenomena to the intuitive level. The animated image, valuable for didactic presentation, is also a source of innovation and prediction.

4. Modelling Chemical Behavior: Drug Design and Structure-Activity Relationships

Its capacity for synthesis, display and manipulation of computerized chemical shapes makes molecular graphics an invaluable tool for investigating molecular interactions, particularly in drug design. Besides contributing to a clearer understanding of observed behavior of structural models, MCG also acts as a guide for proposing predictive rules and even for designing new active compounds. The underlying hypothesis is that, in a set of structurally related drugs, common activities correspond to similar interactions with the receptor and reveal the

presence of a single bioactive structural fragment (the pharmacophore). If one also accepts that these interactions (biomolecules—ligand) imply adequation between the chemical shapes in contact, one then obtains a receptor mapping analysis, the pharmacophore providing a complementary shape of the receptor.

4.1. STRUCTURAL MATCHING

To spot these common fragments (substructures), responsible for the same activity, it is sometimes enough to examine the 2-D structural formulas. In the simplest cases, molecular topology (indicating the nature of the atoms and of the bonds linking them) reflects spatial organization and often permits a first level of investigation of molecular mimicry.

In more complex situations, pre-existing analogies or those resulting from a conformational adaptation are more easily perceived through three-dimensional representations.

Thus, the mere representation of the molecular skeletons of two analgesic neighbors suffices to highlight their conformational differences (rocking of a phenyl group), whereas a scrutiny of the 2-D planar formulas reveals only one difference, the apparently minor adjunction of a methyl group. Such treatment is particularly useful for characterizing features closely related to molecular topography, e.g. searching for the common denominator of a set of structures within a population (population focus) or determining its trace: imprint of the population set. Examples of immediate application are the study of steric factors in interaction processes or geometric deformation regarding a reference structure or an "average compound" in a family (Figure 21 [Color Plate VI]).

Various strategies have been developed to reveal structural or geometric similarities. To compare and superpose supposedly related molecules, anchoring is performed by choosing sets of comparable atoms and computing the rotation most able to achieve superposition [76]. If rigid active compounds are known, mapping onto their frozen skeletons provides an easier way to find the best molecular fitting. The comparison of different flexible structures among themselves creates a more complex situation. It no longer suffices to consider minimal energy geometries (experimental or calculated) for the isolated molecule. One must take into account local adaptation possibilities and the corresponding cost in conformational energy. In molecular matching, the progressive adaptation of a glove to a given hand is a more realistic image than the direct and rigid relationship between a key and a lock.

In the case of mescaline and LSD, a local shape identity between these two hallucinogenic entities can be obtained only if the mescaline molecule moves into one of its less stable conformations. In molecular fitting programs, recent improvements [77—79] allow for the internal rotational freedom of flexible molecules. The local fitting search is accelerated by using "branches" or "aggregates" (internally rigid sets of atoms) according to pre-defined constraints. The MAXIMIN program

[79] provides for a statistical best fit (minimization of conformational variance) allowing for both molecular adequation and conformational energy variations.

4.2. ELECTRONIC MATCHING

Geometric features constitute but one aspect of molecular matching. The nature of the molecular body plays an important role in interaction processes. Van der Waals surface displays make it possible to scrutinize conformational details and steric hindrances. Molecular surfaces or solvent accessible areas [3, 36] specify the conditions of solvent approach towards the molecular outer envelope. However, the main factors monitoring interaction processes often rely on electronic properties (such an electronic fit can even overshadow the more easily perceived geometric similarities). It is established that the similar behavior of metothrexate and dihydrofolate fragments is not directly linked to the evident resemblance between their skeletons but rather to a like localization of the MEP attraction zones (observed only after reversal of the pteridine cycles [80]).

Similarly, the diversified behavior of β-adrenergic drugs with various structures (arylethanolamines or aryloxypropanolamines) and with different action (β activating or β blocking) cannot be rationalized by a structure similarity search. Observations are, however, easily ordered in relation to MEP attraction and repulsion zones [81].

Color encoding of the polar or hydrophobic character of atoms provides initial insight into structural and energetic features. More refined treatment involves comparing the zones associated with the spatial localization of electronic properties. The Molecular Electrostatic Potential furnishes the molecule with an "electric volume" (not strictly related to the geometric body), more global than the vision obtained from electron populations localized on a few "interesting" atoms. This volume also indicates the most likely areas for the approach of an incoming electrophilic reagent.

4.3. BOOLEAN OPERATIONS AND COMPARED CONCEPTUAL SHAPES

Such comparisons can be made easily through elementary Boolean operations (union, intersection, difference) combined, when necessary, with more complex sequences. This area has invaluable possibilities for image synthesis.

On a display, where images can be manipulated, bleached, etc. on demand, such Boolean operations are simple. Despite the evident advantages, this possibility has been rarely exploited for electronic characteristics. The problem is more complex than for structural shapes (distance, volume, surface) since the properties to be compared are generally not defined or calculable by simple analytical expressions.

Our POLYMOD system proposes a three-dimensional ordered network for comparing structured entities with electronic data (MEP localization) [54].

The molecules under observation are immersed in a grid box: the nodes of the

network then constitute a unique set of control points defining the two shapes to be compared. Thanks to this organization of the data, instead of treatment by numerical values, we study conditions on the observation nodes and carry out logical operations on the resulting encoding. Thus, a common privileged approach zone is determined by intersection of 3-D areas where the attractive electrostatic potential is lower than a given negative threshold. On each network node, we examine the values of properties associated with the shapes to be compared. Thus, a set of *validated nodes* is defined. Point images are then obtained on a first level of precision.

The algorithm of *voxel edge intersection* (determination of limit points from frontier nodes) can also be used to define the shape resulting from the Boolean operation more precisely.

By maintaining data structuration (nodes anchored by triplets specifying their spatial localization), this approach is well adapted to logical operation sequences. It can be used downstream for creating and managing images in graphic display (Figure 22 [Color Plate VI]).

4.4. INTERPRETING REACTIVITY THROUGH MOLECULAR GRAPHICS

Until now molecular graphics was widely used in bioorganic chemistry (protein conformation, enzyme activity, drug design, etc.). With recent developments of quantum methods well suited for describing heavy atoms such as the X_α method, a new and promising field seems likely to open up in organometallics (with catalysis and complexation of biological substrates). Another application illustrates the importance and power of the image for interpreting molecular properties. Tetraperoxometallate-type complexes $(M(O_2)_4)^{n-}$, vary greatly in their reactivity and chemical properties [83, 84]. For example, the parent compound of molybdenum, $MoO(O_2)_2$ HMPT, for which $(Mo(O_2)_4)^{2-}$ can be considered an adequate model, is a very active catalyst for epoxydizing olefins [85], while both $(Cr(O_2)_4)^{3-}$ and $(Nb(O_2)_4)^{3-}$ are practically inert under similar conditions [86]. MEP can be used for a first-order description of the intermolecular interaction potential of a given compound. Applied to the case of tetraperoxometallates [87], it is seen that an important nucleophilic canal (i.e. open to attack by a soft nucleophile such as ethylene) is present in each symmetry plane of the compound $(Mo(O_2)_4)^{2-}$, whereas these conditions are not fulfilled for the Cr and Nb compounds. The appearance of MEP of $(Mo(O_2)_4)^{2-}$ indeed suggests that nucleophilic attack takes place on the metal, with the possibility of intramolecular rearrangement causing, in a later phase, an electrophilic attack on the 01 oxygen atom (the one on which we note the presence of a minimum of the MEP), as proposed by Mimoun [84] (Figure 10). These results, therefore, indicate that the MEP approach could lead to a rationalization and an interpretation of reactivity in coordination chemistry.

5. Molecular Graphics and Quantitative Data

One of the most often mentioned advantages of molecular graphics is its capacity to condense, in a single image, a vast amount of information and, thereby, to render perceptible in one instantaneous synthetic vision the significant elements of a complex situation.

Although essential, these aspects of representation, as well as the esthetics of certain images, should not overshadow the importance of molecular graphics as a tool for handling quantitative physical chemical data defined in a 3-D space.

Indeed, in pattern recognition or adequation procedures, tridimensional visualizations are essential for suggesting strategies and initiating new research directions.

Thus, in planning a rational drug design process [82], molecular graphics was very closely associated with quantum or molecular mechanics theoretical simulations for visualizing geometric elements or electronic localizations.

Moreover, *data structures* intervening in the very constitution of these synthetic images are often directly involved in elaborating or managing these quantitative data. Thus, in the POLYMOD software, in order to localize electronic properties, Boolean operations are carried out with, as a starting point, the grid nodes used to define images in the representation space.

Another significant example concerns the calculation of molecular surfaces and volumes. In models formed by uniting Van der Waals atomic spheres (sometimes linked by toroidal fragments), analytical formulations have been proposed using the classical geometric properties of these simple solids.

Several recent studies have established that precise quantitative evaluations of surfaces and volumes associated with molecular shapes could be obtained by graphics. For example, by counting the points of a grid situated within the molecular body one obtains a good idea of its volume [42, 43], in agreement with analytical calculations. The correlations obtained between values thus calculated on the microscopic level and several experimental properties measured on the macroscopic scale (molar refraction, b constant in the Van der Waals equation of state) confirm the validity and predictive capacity of such approaches.

Very recently Meyer and coworkers [43d] were thus able to determine the cross section areas of a series of alkanoic acids. These calculated molecular shape parameters are nearly identical to those independently derived through a fractal theory of adsorption (and resulting directly from experimental measurements) producing indirect molecular distance evaluation.

Two important conclusions can be drawn. One is valid in surface chemistry, and the other deals with atomic dimensions. On the molecular level, it has been proven that usual standard bond distances must be considered, and new values can be proposed.

Very often global and local shape parameters emerge instead of volume parameters. This is particularly true with the interpretation of the high selectivity

of drug-receptor interactions. "How should one mold the volume decided upon?" [43d] Such considerations are valid for small molecules and macromolecules (tertiary and quaternary structures in protein chemistry). To start from molecules and obtain the shape of biological receptors or, more precisely, of volumes or surfaces anchored on the active sites of receptors, is one of the basic problems that biology and biomedicine can now hope to solve. This illustrates the importance of the numerous modelling techniques, such as infography, that must be implemented if we are to progress in these essential fields.

6. Conclusions

In this review article, we have shown how Chemical Informatics and Molecular Computer Graphics have become complementary in their support of chemistry and biology. In our opinion, this synergy will shape the evolution of those sciences.

It is, moreover, likely that chemistry, which presided over the birth of graph theory, will play a significant role in defining the forthcoming fifth and sixth computer generations. These will need to integrate, among other things, symbolic and image languages with cognitive logic. Chemical problems provide a natural basis for such new ventures.

It is our feeling that, in such cognitive activities, chemistry, both within its own frontiers and in connection with other fields, will remain a pioneer and a leader.

References

1. C. Levinthal: *Sci. Am.* **6**, 42 (1966).
2. R. J. Feldmann, D. H. Bing, B. C. Furie and B. F. Furie: *Proc. Nat. Acad. Sci. U.S.A.* **75**, 5409 (1978).
3. (a) B. Lee and F. M. Richards: *J. Mol. Biol.* **55**, 379 (1971).
   (b) T. J. Richmond and F. M. Richards: *J. Mol. Biol.* **119**, 537 (1978).
4. J. A. McCammon, P. G. Wolynes and M. Karplus: *Biochem.* **18**, 927 (1979).
5. J. Weber, M. Roch, J. J. Combremont, P. Vogel and P. A. Carrupt: *J. Mol. Struct. Theochem.* **93**, 189 (1983).
6. T. Fujita, J. Iwasa and C. Hansch: *J. Am. Chem. Soc.* **86**, 5175 (1964).
7. M. L. Connolly: *Science* **221**, 709 (1983).
8. M. L. Connolly: *J. Am. Chem. Soc.* **107**, 1118 (1985).
9. R. J. Feldmann: *Computer Applications in Chemistry*, S. R. Heller, R. Potenzone Jr. (Eds.). Elsevier (1983).
10. P. Quarendon, C. B. Naylor and W. G. Richards: *J. Mol. Graph.* **2**, 4 (1984).
11. (a) E. Scrocco and J. Tomasi: *Advan. Quantum. Chem.* **11**, 115 (1978).
    (b) C. Petrongolo: *Gaz. Chim. Ital.* **1**, 445 (1978).
12. J. E. Dubois, D. Laurent and J. Weber: *The Visual Computer* **1**, 49 (1985).
13. J. E. Dubois: *Computer Representation and Manipulation of Chemical Information*, W. T. Wipke, S. Heller, R. Feldmann and E. Hyde (Eds.). John Wiley, New York (1974).
14. W. K. Giloi: *Interactive Computer Graphics*, Prentice Hall, New Jersey (1978).
15. E. Catmull: *Proc. IEEE. Conf. Comput. Graphics*, 11 (1975).
16. M. L. Connolly: *J. Mol. Graph.* **3**, 19 (1985).
17. W. M. Newman and R. F. Sproull: *Principles of Interactive Computer Graphics*, MacGraw Hill, New York (1979).

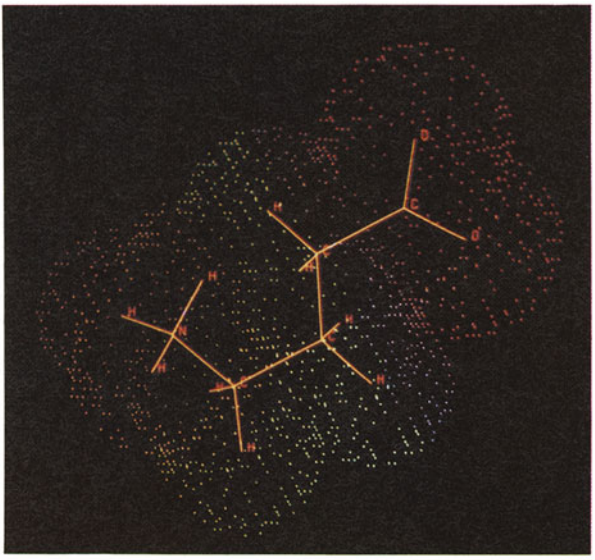


Fig. 3. Surface modelling and dot images – MEP drawn on a Van der Waals-type surface for γ aminobutyric acid (decreasing values from yellow to red).

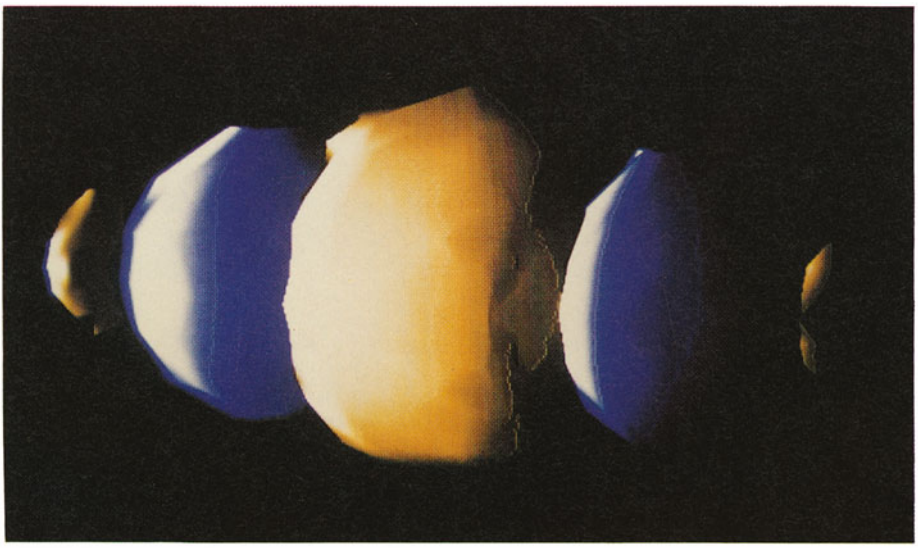


Fig. 4. Surface modelling by a mosaic of facets (raster display) – Electron Deformation Density of O_2 with positive (blue) and negative (yellow) surfaces (± 0.01 e/Å^3).

PLATE I

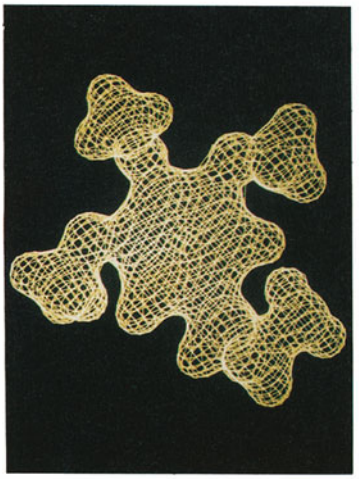


Fig. 5. 3D contour envelope for property shape – The 100 Kcal repulsive potential of mescaline, schematizing the molecular body.

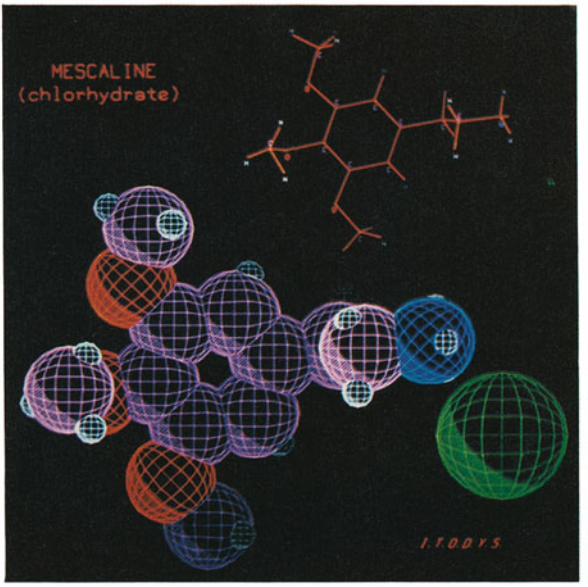


Fig. 7. Wire type and space filling models – Calligraphic display for mescaline chlorhydrate (with curved cross hatching and shading).

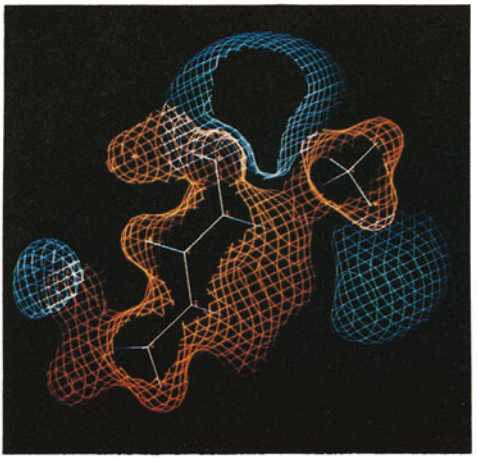


Fig. 6. Local Graphic Assistance – 3D representations of property shapes. (a) Clipping allows for viewing inside and reveals the molecular skeleton. MEP of mescaline. Orange repulsive potential; blue attractive areas. (b) Removing hidden lines strengthens relief impression. MEP of thiophene (STO-3G calculation). Yellow repulsive potential (100 Kcal) green attractive areas (– 7 Kcal).

PLATE II

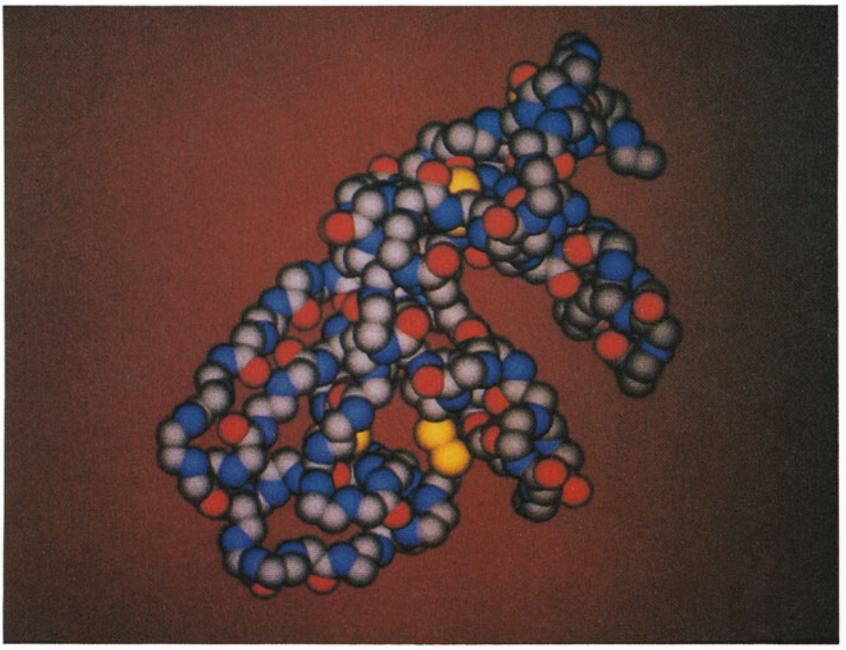


Fig. 8a.

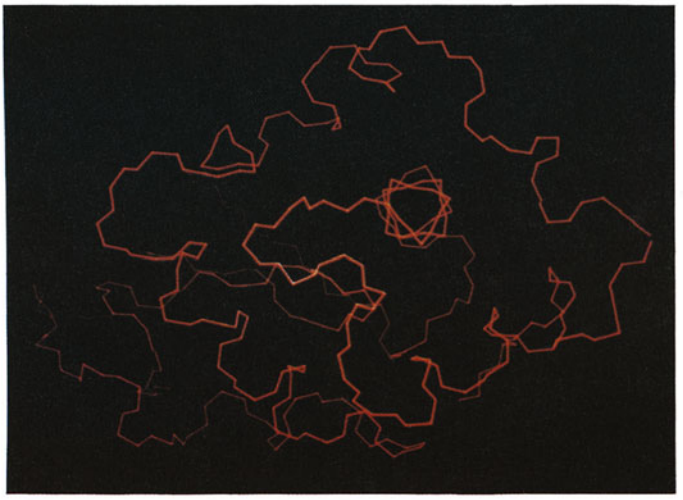


Fig. 8b.

Fig. 8. Representation of a small protein, lysozyme: (a) space filling model (raster display); (b) by simplifying the images one can grasp significant details. An α helix is visible in the center.

PLATE III

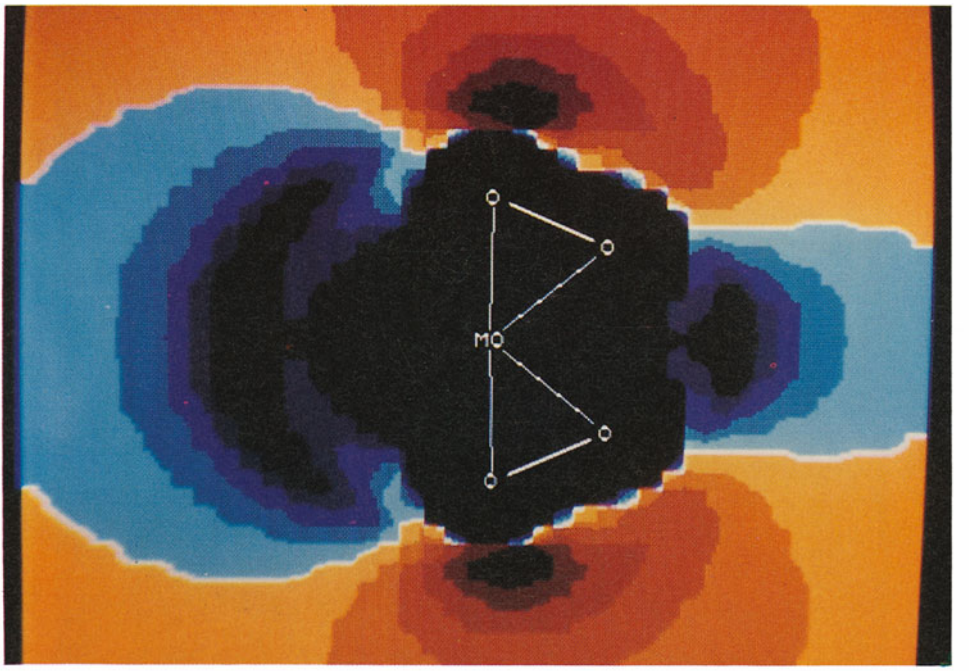


Fig. 10. Color Filled Contour Map of MEP (compensated for the anion negative charge) for $[Mo(O_2)_4]^{2-}$ in a symmetry plane. The horizontal blue channel corresponds to the favored area for nucleophilic attack.

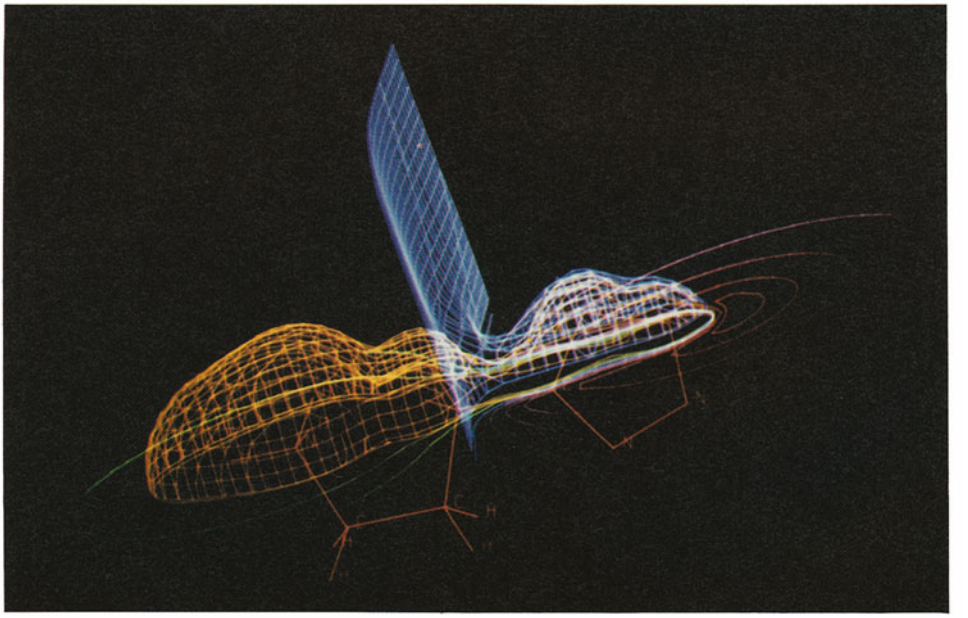


Fig. 11. Composite images for MEP representation – Display of a planar isopotential map with part of the isopotential surfaces above this plane for GABA (yellow envelope: positive 100 Kcal potential; blue: neutral zone). A fraction of the molecular skeleton appears in the lower part.

PLATE IV

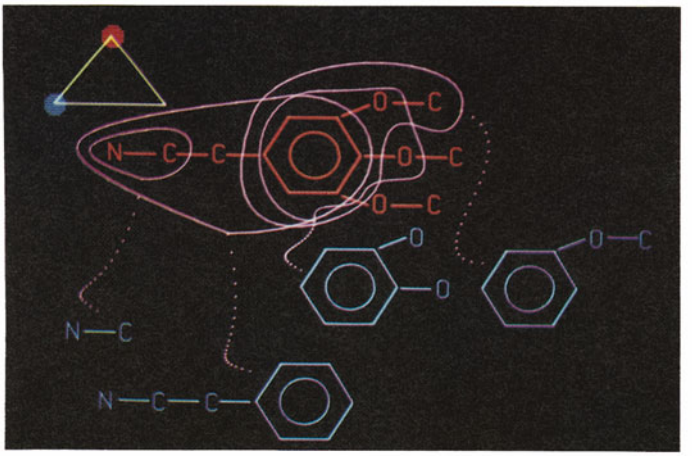


Fig. 19. Decomposition into fragments (icones). Graphic display illustrates various substructures (conceptual molecular fragments) issued from the mescaline molecule.

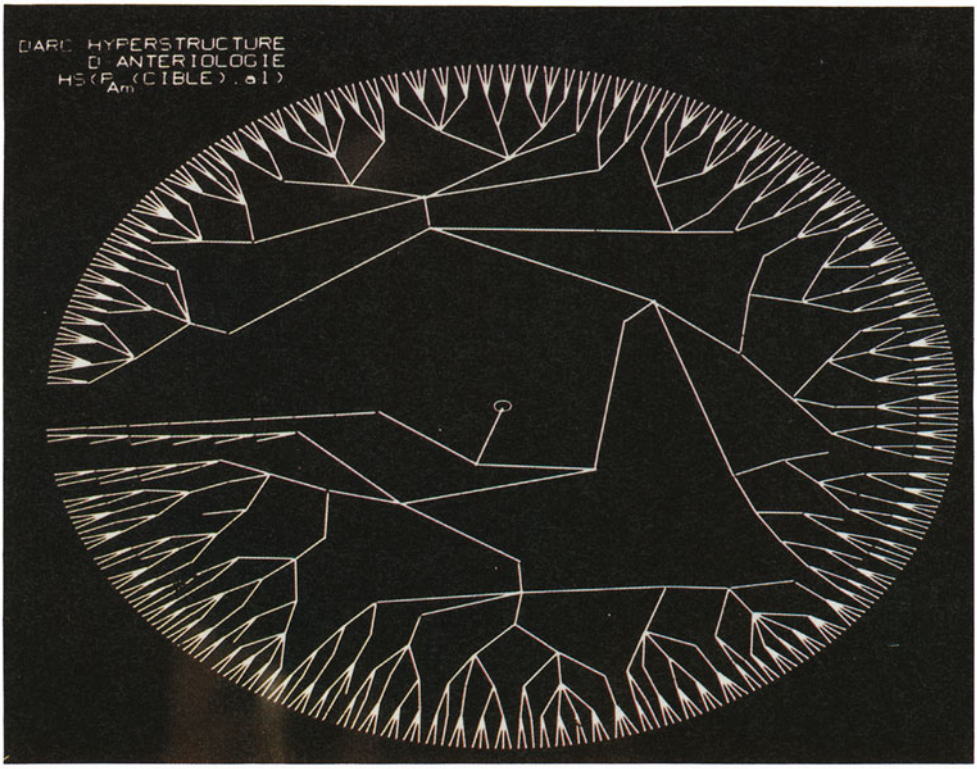


Fig. 20. Generative Hyperstructure of Alcanes (from methane to $C_{12}H_{26}$). Each node corresponds to a compound (355 isomers of duodecane and their 289 anteriomorphs). Such a hyperstructure, both classification and representation space, localizes each structure within a population set and ensures easy management of proximity or anteriology relationships.

PLATE V

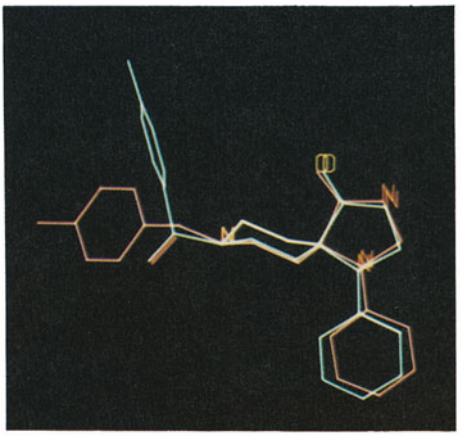


Fig. 21. Partial superposition of two neighboring analgesic drugs. A minor structural perturbation (adjunction of a methyl group) induces a large conformational change: rocking of the phenyl ring.

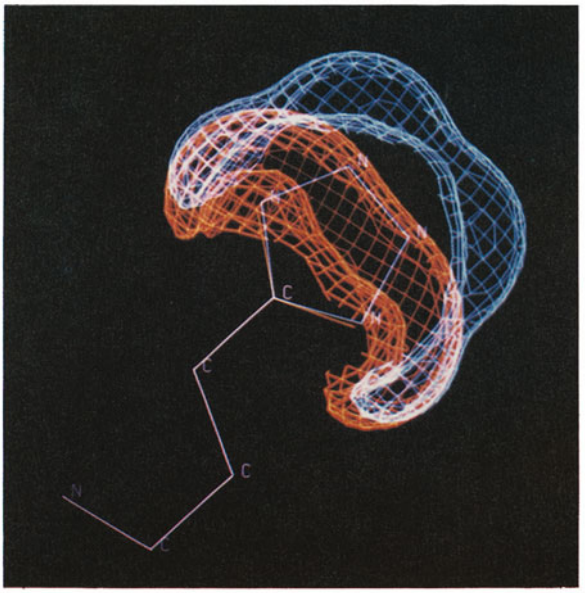
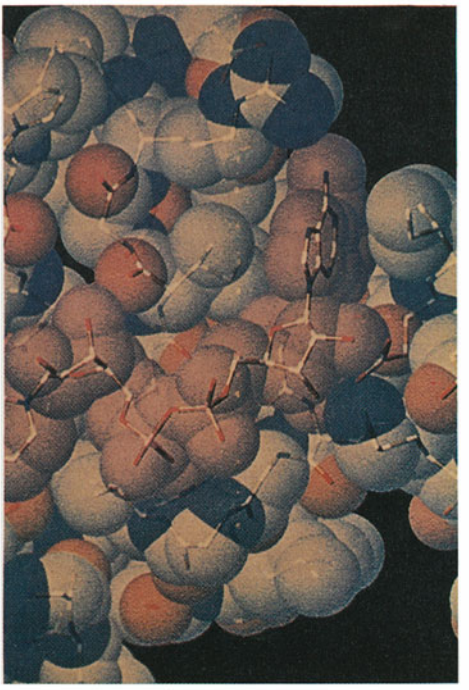
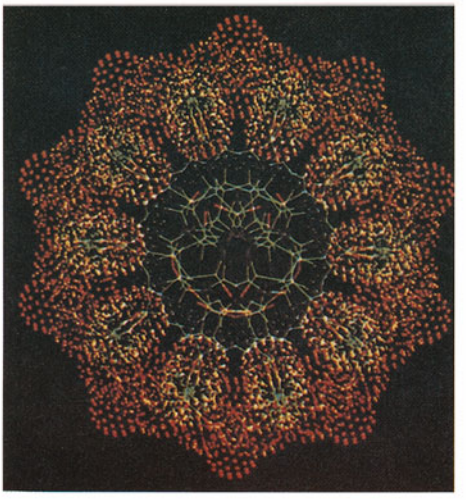


Fig. 22. Local comparison of property shapes – Superposition of STO-3G MEP surfaces (attractive – 100 Kcal/mole) for amino butyric acid (red) and a tetrazole analog (blue) (the calculations being made with standard geometry, the skeletons are strictly coincident in the acyclic part).

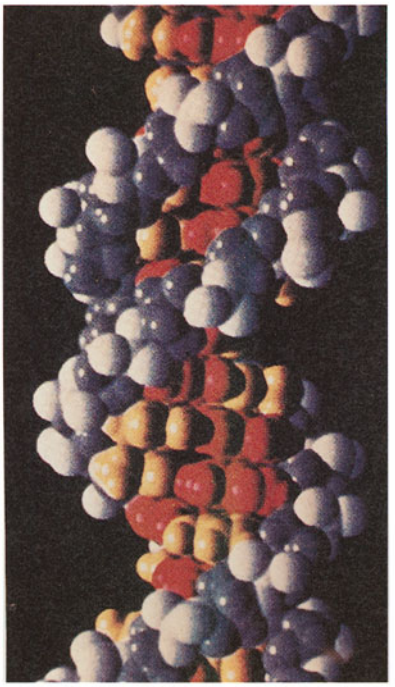
PLATE VI

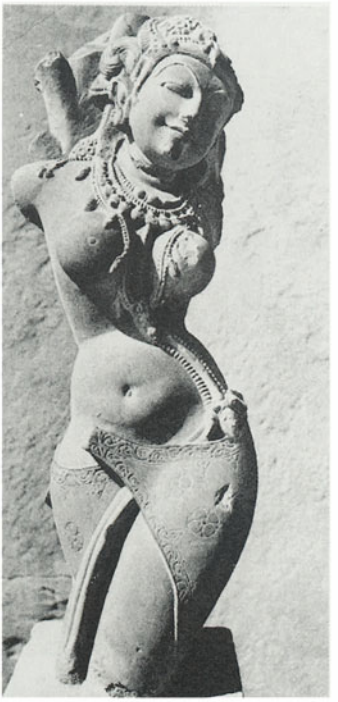


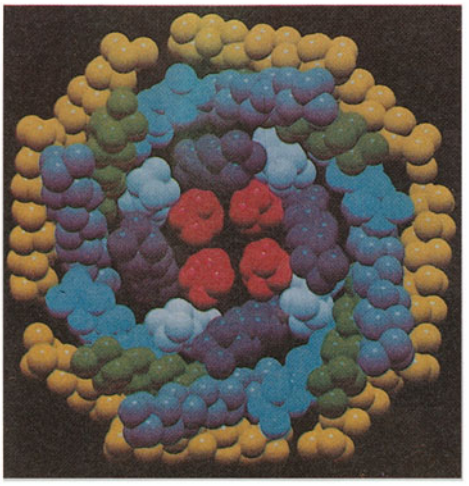


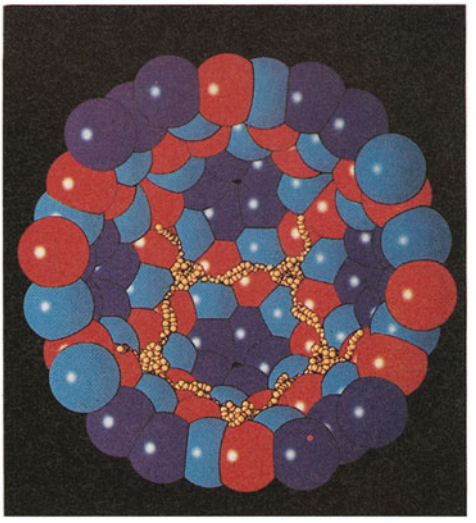
Binding site of liver alcohol dehydrogenase with a bound NAD molecule (upper left, from Dr J. M. Burridge, IBM UK Scientific Centre, by permission), botryoidal growth of smithsonite crystals (upper right), B-DNA in a plane perpendicular to the helix axis (lower left, from Prof. R. Langridge, copyright regents, Computer Graphics Laboratory, University of California, by permission), Kazak carpet from Caucasus (lower right).

PLATE VII









B-DNA in space-filling representation (upper left, from Dr R J Feldmann, NIH Computer Research, by permission), Vrikshaka Tree Goddess, 10th century, India (upper right), sodium ion channel protein (lower left, from Dr R J Feldmann, NIH Computer Research, by permission), The Cosmic Laser Concert, designed by Ivan Dryer (lower right)

PLATE VIII

18 J Foley and A Van Dam *Fundamentals of Interactive Computer Graphics*, Addison Wesley, Reading (1981)
19 H Gouraud *IEEE Trans Comput* **20**, 623 (1971)
20 G Wyvill and T L Kunii *The Visual Computer* **1**, 3 (1985)
21 (a) A Fujimoto, T Tanaka and K Iwata *IEEE Computer Graphics and Applications*, 6 (1986)
(b) A Fujimoto, C G Perrott and K Iwata in *Advanced Computer Graphics*, T L Kunii (Ed) Springer (1986)
22 K Kronlof and M Tamminen *The Visual Computer* **1**, 24 (1985)
23 J E Scott *Introduction to Interactive Computer Graphics*, John Wiley, New York (1982)
24 M Roch Ph D Thesis, Geneva (1986)
25 H N Christiansen and T W Sedenberg *Comput Graph* **12**, 187 (1978)
26 W L Jorgensen *Q C P E Prog* 340, Indiana University
27 G D Purvis and C Culberson *J Mol Graph* **4**, 88 (1986)
28 (a) S Granapathy and T G Dennely *Comput Graph* **16**, 69 (1982)
(b) J H Boissonat *A C M Trans Graph* **3**, 266 (1984)
29 A Koide, A Doi and K Kajioka *J Mol Graph* **4**, 149 (1986)
30 M L Connolly *J Appl Crystallogr* **18**, 499 (1985)
31 (a) N L Max *J Mol Graph* **2**, 8 (1984)
(b) N L Max *J Med Sci* **6**, 485 (1982)
32 J M Burridge and S J P Todd *J Mol Graph* **4**, 220 (1986)
33 M Carson and C E Bugg *J Mol Graph* **4**, 121 (1986)
34 F H Allen, S Bellard, M D Brice, B A Cartwright, A Doubleday, H Higgs, T Hummelink, B G Hummelink-Peters, O Kennard, W D S Motherwell, J R Rogers and D G Watson *Acta Crystallogr Sect B* **35**, 2331 (1979)
35 O Kennard, D G Watson and W G Town *J Chem Doc* **12**, 14 (1972)
36 F M Richards *An Rev Biophys Bioeng* **6**, 151 (1977)
37 M L Connolly *J Appl Crystallogr* **16**, 548 (1983)
38 J J Muller *J Appl Crystallogr* **16**, 74 (1983)
39 M Y Pavlov and B A Fedorov *Biopolymers* **22**, 1507 (1983)
40 J Greer and B L Bush *Pr Nat Acad Sci , U S A* **75**, 303 (1978)
41 L H Pearl and A Honegger *J Mol Graph* **1**, 9 (1983)
42 M Marsili, P Floersheim and A S Dreiding *Comput Chem* **7**, 175 (1983)
43 (a) A Y Meyer *J Chem Soc Perkin* **2**, 1161 (1985)
(b) A Y Meyer *J Comput Chem* **7**, 144 (1986)
(c) A Y Meyer *J Mol Struct* **124**, 93 (1985)
(d) A Y Meyer, D Farin and D Avnir *J Am Chem Soc* **108**, 7897 (1986)
44 (a) K Fukui in *Molecular Orbitals in Chemistry, Physics, and Biology*, P O Lowdin and B Pullman (Eds) Academic Press, 513 (1964)
(b) H Fujimoto and K Fukui in *Molecular Interactions*, H Ratajzack and W J Orville Thomas (Eds) J Wiley and Sons, 89 (1980)
45 (a) P Politzei, L Abrahmsen and P Sjoberg *J Am Chem Soc* **106**, 855 (1984)
(b) J S Murray, B A Zilles, K Jayasuriya and P Politzer *J Am Chem Soc* **108**, 915 (1986)
(c) P Politzer and D G Truhlar *Chemical Applications of Atomic and Molecular Electrostatic Potentials* Plenum Press, New York (1981)
46 G Grassy, Y Rival, M Bonnafous, Y Adam, J C Teulade and J P Chapat *Eur J Med Chem Chim Ther* **20**, 501 (1985)
47 R Lavery, A Pullman and B Pullman *Theor Chim Acta* **62**, 93 (1982)
48 R F W Bader, T H Tang, Y Tal and F W Biegler Konig *J Am Chem Soc* **104**, 940 (1982)
49 C Giessner-Prettre and A Pullman *Theor Chim Acta* **25**, 83 (1972)
50 A Goldblum, D Perahia and A Pullman *Int J Quantum Chem* **XV**, 121 (1979)
51 J T Egan and R D Marc Elroy *J Comput Chem* **5**, 72 (1984)
52 K Zakrzewska and A Pullman *J Comput Chem* **6**, 266 (1986)
53 C Corongiu, E Clementi, E Pretsch and W Simon *J Chem Phys* **72**, 3096 (1980)
54 J E Dubois, S Y Yue and J P Doucet *The Visual Computer* **2**, 367 (1986)

55. M. M. Gilbert, J. J. Donn, M. Peirce, K. R. Sundberg and K. Ruedenberg: *J. Comput. Chem.* **6**, 209 (1985).
56. R. Lavery and B. Pullman: *Int. J. Quantum Chem.* **20**, 259 (1981).
57. J. E. Dubois, S. Y. Yue and J. P. Doucet: in *Advanced Computer Graphics*, T. Kunii (Ed.). Springer Verlag, Tokyo (1986).
58. J. E. Dubois: 'Ordered Chromatic Graph and Limited Environment Concept' in *Chemical Applications of Graph Theory*, A. T. Balaban (Ed.). Academic Press, London (1976).
59. J. E. Dubois, D. Laurent and H. Viellard: *C. R. Acad. Sci. Paris* **263C**, 764 (1966).
60. H. L. Morgan: *J. Chem. Doc.* **5**, 107 (1965).
61. R. Attias: *J. Chem. Inf. Comput. Sci.* **23**, 102 (1983).
62. J. E. Dubois, M. Carabédian and I. Dagane: *Anal. Chem.* **158**, 217 (1984).
63. Y. Sobel, I. Dagane, M. Carabédian and J. E. Dubois: *Proceedings of the 9th International CODATA Conference*, pp. 395—399. P. Glaeser (Ed.). Elsevier Science Publishers B.V. (North-Holland) (1985).
64. A. Panaye, J. P. Doucet, P. Peguet and J. E. Dubois: *10th International CODATA Conference*, pp. 32—37. P. Glaeser (Ed.). Elsevier Science Publishers B.V. (North-Holland) (1986).
65. J. E. Dubois, C. Mercier and A. Panaye: *Acta Pharm. Jugosl.* **36**, 135 (1986).
66. J. R. Chrétien, J. Szymoniak, J. E. Dubois, M. F. Poirier, M. Garreau and P. Deniker: *Eur. J. Med. Chem. Chim. Ther.* **20**, 315 (1985).
67. J. Weber, G. Bernardinelli and J. J. Combremont: *Proceedings of Eurographics '80.* C. E. Vandoni (Ed.). North-Holland, Amsterdam, p. 139 (1980).
68. J. J. Combremont, M. Roch and J. Weber: *Computer Graphics Forum* **2**, 89 (1983).
69. (a) C. W. Jefford, J. Mareda, J. J. Combremont and J. Weber: *Chimia* **34**, 354 (1984).
(b) D. Mottier, J. Weber, P. A. Carrupt and P. Vogel: *J. Chim. Phys.* in press.
70. (a) J. Weber, M. Roch, J. P. Doucet and J. E. Dubois: *Imagerie et Animation Graphique en Chimie.* Film 16 mm — Réalisation distribution: UAV ENS, Saint-Cloud, France (1986).
(b) J. E. Dubois and S. Stanojevic: *Les molécules attendaient l'ordinateur depuis la nuit des temps.* Film 16 mm — Réalisation CNRS, Paris, France (1986).
71. M. Karplus and J. A. MacCommon: *Crit. Rev. Biochem.* **9**, 293 (1981).
72. P. K. Weiner, S. L. Gallion, F. Westhof and R. M. Levy: *J. Mol. Graph.* **4**, 203 (1986).
73. T. Fujita, S. M. Swanson and E. J. Meyer Jr.: *J. Mol. Graph.* **4**, 208 (1986).
74. M. T. Dove, D. Fincham and R. E. Hubbard: *J. Mol. Graph.* **4**, 79 (1986).
75. G. Ganti and J. A. McCammon: *J. Mol. Graph.* **4**, 200 (1986).
76. S. C. Nyburg: *Acta Crystallogr. B* **30**, 251 (1974).
77. L. Barino: *Comp. Chem.* **5**, 85 (1981).
78. J. Lejeune, A. G. Michel and D. P. Vercauteren: *J. Mol. Graph.* **4**, 194 (1986).
79. J. Labanovski, I. Motoc, C. S. Naylor, D. Mayer, and R. A. Dammkochler: *Quant. Struct. Act. Relat.* **4**, 138 (1986).
80. P. Gund, J. D. Andose, J. B. Rhodes and G. M. Smith: *Science* **208**, 1425 (1980).
81. B. Macchia, F. Macchia and A. Martinelli: *Eur. J. Med. Chem. Chim. Ther.* **18**, 85 (1983).
82. A. J. Hopfinger: *J. Med. Chem.* **28**, 1133 (1985).
83. M. Roch, J. Weber and A. F. Williams: *Inorg. Chem.* **23**, 4571 (1984).
84. H. Mimoun: *Angew. Chem. Int. Ed.* **21**, 734 (1982).
85. H. Mimoun: I. Serée de Roch and L. Sajus; *Tetrahedron* **26**, 37 (1960).
86. H. Mimoun, M. Postel, F. Casabianca, J. Fischer and A. Mitschler: *Inorg. Chem.* **21**, 1303 (1982).
87. J. Weber, M. Roch and A. F. Williams: *Chem. Phys. Lett.* **123**, 246 (1986).

The Harmony of Molecules

GABOR NARAY-SZABO
CHINOIN Pharmaceutical and Chemical Works, P.O.B. 110, H-1325 Budapest, Hungary.

1. Introduction

The meaning of the word 'harmony' is 'agreement', 'congruity', 'combination or arrangement of parts to form consistent and orderly whole'. Furthermore, 'harmonize' is equivalent to 'be agreable in artistic effect' [1]. These everyday concepts are also very important in the world of molecules though they are described by other words. In the following we shall try to show that the various meanings are related in the scientific sense too, and even harmonize to a certain extent.

Analysis of molecular harmony may lead to simple and general rules that are qualitative but help very much in understanding important aspects of non-covalent bonding in molecular associations. Present-day science has achieved enormous advancement in quantifying the description of natural phenomena; therefore qualitative perceptions are often a bit suspect to a considerable fraction of the scientific community. However, we need qualitative descriptions at least for two reasons. First, as Heisenberg stated [2], mathematical formulation alone is not sufficient for the complete understanding of a natural phenomenon. Second, qualitative rules may become an integral part of the scientist's knowledge, influencing the way of his/her thinking, and may therefore have an important impact on the elaboration of working hypotheses.

Let us now treat various meanings of molecular harmony. Agreement or congruity between species of a given molecular family means, in other terms, *similarity*. This is an extremely important concept allowing us to classify certain groups of molecules that behave similarly in physical, chemical or biological processes. Up to now several dozens of definitions of molecular similarity have been proposed; however, none has gained an overwhelming popularity. A survey on recent results and propositions for a consistent formulation will be given in Section 2.

The second meaning of harmony, namely 'combination or arrangement of parts to form a consistent and orderly whole', closely corresponds to *molecular recognition* of certain species by others to achieve optimum binding conditions. Recognition is of primary importance in supramolecular associations, like host—guest complexes, crystals and biomolecular aggregates. It is one of the most important processes in molecular biology and is the target of numerous current theoretical and experimental studies. Molecular recognition is associated with non-covalent forces between more or less isolated entities in condensed phases; therefore a

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, 205—231.

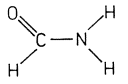
thorough analysis of the nature of intermolecular interactions is prerequisite for an adequate description. This topic will be discussed in Section 3.

The scientist cannot disregard the esthetic impact, variety and wealth of mechanical and computer models of various molecules. Owing to the widespread application of computer graphics in visualizing complicated molecules and their associations, colored illustration introduces a unique sense of beauty into everyday laboratory work. This idea will be evaluated in Section 4.

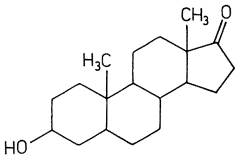
In the last section we deal with the human aspects of harmony that are needed to accomplish and maintain tolerance in the large family of molecular scientists. Science is nowadays split into tiny subfields and experts of certain phenomena incline to disapprove of experimental or theoretical techniques different from their own. This contradiction has to be resolved by interdisciplinarity and the ambition to inquire about related areas of molecular sciences.

2. Harmony in Molecular Families: Similarity

The definition of molecular similarity depends strongly on the property it is related to. For example, dipole moments of formamide (**I**) and androsterone (**II**) are almost the same, they are equal to 12.4 and 12.3 C.m, respectively, but it is quite clear that these molecules behave completely differently in chemical reactions and biological processes. Accordingly, similarity should be defined differently if a certain physical property, the chemical behaviour or the biological activity, is concerned.



I



II

Molecular properties, *e.g.* bond lengths, bond angles, dipole moments, electric polarizabilities or magnetic susceptibilities are similar for members of homologous series, like alkanes, aliphatic alcohols or ketones. This is closely related to the transferability of bond properties [3]. Similarity of molecules in connection with their physical constants can thus be formulated in terms of bond transferability. However, transferability is only a necessary but not a sufficient condition since additive properties like size, molecular weight or diamagnetic susceptibility may considerably differ for certain species belonging to the same family. So, concerning their physical properties, molecules can be stated to be similar if they are built up of the same bonds and have molecular weights relatively close to each other. According to this definition, *n*-propanol, *i*-propanol, *n*-butanol, *i*-butanol, 1-hydroxy-2-methylpropane and 2-hydroxy-2-methylpropane are similar, but as

we go further in the homologous series $C_nH_{2n+1}OH$, similarity with the first member diminishes increasingly.

Chemical similarity is related to the transferability of functional groups [4]. Certain units such as hydroxyl, carbonyl, carboxyl, phenyl, amide or sulfonyl behave chemically similarly even if they are linked to different molecular skeletons. Accordingly, molecules having the same functional groups like, for instance, substituted benzoic acids, may be considered to be chemically similar. This concept is rationalized in terms of the transferability of group properties, especially electrostatic potentials [5] and is very useful in organic chemistry since the terrifying amount of information can be nicely organized on this basis [6].

Maybe the most important aspect of molecular similarity is related to biological activity. It is a complicated task to define similarity in this case since biological activity itself is hard to define. Therefore it seems to be appropriate to restrict ourselves to the molecular level and to study the similarity of relatively small molecules only in connection with their binding properties to biopolymers. As is known, biopolymer—ligand binding is the molecular event that in most cases triggers some biological response. Since biopolymer—ligand binding is non-covalent in nature we try to define similarity for such interactions between molecules. Non-covalent binding plays an important role not only in biological processes but also in host—guest complexation, crystal packing and solute—solvent interactions. Accordingly, we call two molecules similar if they are interchangeable in either of the aforementioned processes. The similarity concept seems to be very useful in the case of these phenomena since, in general, they are too complicated to be handled at present at a rigorous, quantitative level.

The oldest approach to molecular similarity in biological processes is *bioisosterism* [7]. Bioisosteres are groups or molecules possessing chemical and physical similarities and producing broadly similar biological properties. It is possible to define atomic groups which are interchangeable when designing molecules that are intended to bind similarly to biopolymers. In making a bioisosteric replacement the following parameters of the group being changed could be considered [8]:

(a) size,
(b) shape,
(c) electronic distribution,
(d) lipid solubility,
(e) water solubility,
(f) pK_a,
(g) chemical reactivity,
(h) hydrogen bonding capacity.

Some classical examples for isosteric groups are given in Figure 1 [8]. Since the groups contain only a few atoms the similarity is easily recognizable.

The concept of biological isosterism is very useful in understanding the similar

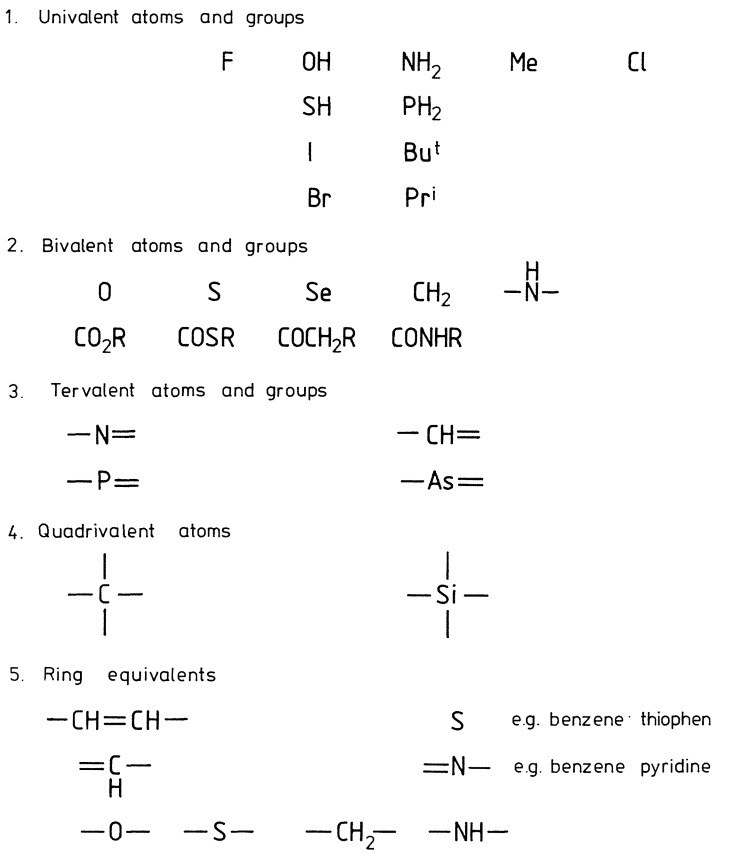


Fig. 1. Classical isosteres.

biological activity of molecules possessing similar structures. However, there are molecules, *e.g.* β-adrenergic stimulants depicted in Figure 2, where similar activities are observed for seemingly very different structures [9—15]. To be able to give a broader definition of biological similarity of molecules we have to go into greater detail.

Several attempts have been made to define similarity in mathematical terms: it is the theory of pattern recognition [16] that is best suited for a mathematical description. In pattern recognition no exact functional form is fitted to the data set of molecular properties; rather, relationships are sought that provide a definition of similarity among diverse groups. In essence, pattern recognition techniques can be thought of as providing relations that uncover common properties. Once such relations are developed they may be used to predict the properties of members that were no part of the original data set.

In order to apply the pattern recognition technique, first we have to scale and normalize the source data in order to convert them into a compatible form.

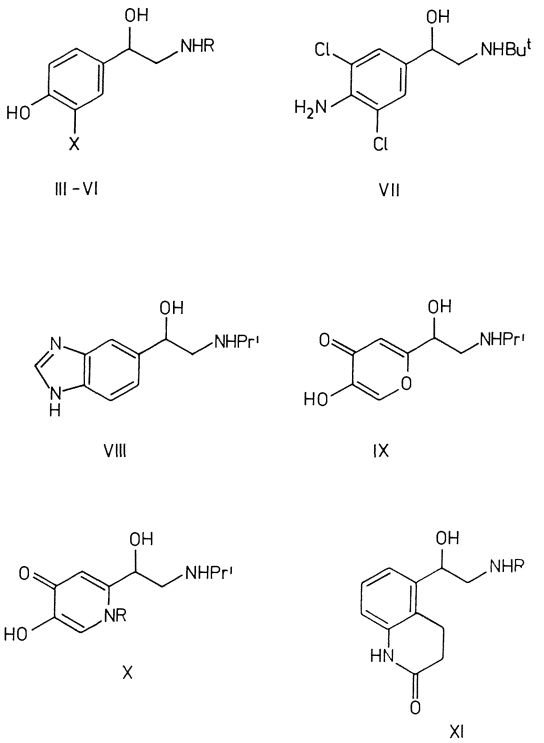


Fig. 2. Some examples of bioisosterism in the family of β-adrenergic stimulants. III: R = Me, X = OH (Adrenaline), IV: R = *t*-Bu, X = CH_2OH (Salbutamol), V: R = *t*-Bu, X = $NHCONH_2$ (Carbuterol), VI: R = *i*-Pr, X = $NHSO_2Me$ (Soterenol), VII: Clenbuterol.

Afterwards, we have to select one or more molecular descriptors that are suitable to characterize similarity and we should attempt to encode them into a numerical form. There are several algorithms available that divide such preprocessed data sets into a number of clusters. Most of these depend on the use of various sorts of distance metrics. Distance measures, however, are only one of several ways in which clusters can be defined. Other techniques are described *e.g.* by Hartigan [17].

Let us now discuss some practical descriptors of molecular similarity. A very popular way of describing a molecule is the *Wiswesser Line Notation* [18]. It is constructed by applying a set of rules to the structure to generate a string of symbols, each corresponding to a structural fragment of the molecule. The notation generated is compact, since most bonds are cited implicitly, and large structural fragments such as rings are often coded with one or only a few symbols. Linear notations are advantageous since no special equipment is necessary for the computer input, and a unique structural representation is obtained.

Another possibility to represent molecular structures is the use of *connection tables.* A connection table is a square matrix where the atom identifiers are stored

in the diagonal while bonding information lies in the off-diagonal elements. Off-diagonal elements are non-zero for bonding pairs, zero otherwise. The value assigned to each element is chosen to represent the different atom and bond types (*e.g.* carbon: 1, oxygen: 2, *etc.*, single bond: 1, double bond: 2, *etc.*). Connection table representations can be stored and used as *e.g.* substructure descriptors allowing us to handle molecular similarity on a straightforward manner.

A great number of molecular descriptors are topological in nature. They include fragment descriptors [19, 20], which code the atom and bond types; substructure descriptors [21, 22] which code the presence or absence of explicitly defined substructures; environment descriptors, which code the immediate surroundings of a substructure; and connectivity descriptors [23, 24] which are indices of the extent of branching in the molecular structure. The use of topology in defining molecular similarity emerged from the old perception in chemistry, namely that similar structural formulae represent molecules that behave similarly.

A useful topological descriptor is the *branching index* or its generalized forms, the various *molecular connectivity indices* [23, 24]. The first-order molecular connectivity index ($^1\chi$) is calculated from the non-hydrogen part of the molecular skeleton. Each non-hydrogen atom is described by its δ value, which is equal to the number of adjacent non-hydrogen atoms. This index is then calculated from the atomic δ values using the following expression:

$$^1\chi = \sum_{s=1}^{n} (\delta_i \delta_j)_s^{-1/2} \tag{1}$$

where i and j correspond to the pairs of adjacent non-hydrogen atoms and the summation is over all bonds between non-hydrogen atoms. Applications of the branching index to several molecular properties and to studies of structure-activity relationships have been carried out by Kier and coworkers [24].

An appealing algorithm for encoding molecular structures is the use of *molecular transforms* that can be obtained experimentally and represent all important informations on the molecular geometry. Soltzberg and Wilkins used a simplified form as a tool for the distinction between drug molecules possessing closely related biological activities [25]. Their formula is

$$I(s) = \sum_{i<j=2}^{N} Z_i Z_j \frac{\sin x}{x} \tag{2}$$

where Z_i and Z_j are atomic numbers, $x = sr_{ij}$, r_{ij} is the distance between atoms i and j. s measures the scattering angle of the electron beam. $I(s)$ is then a relative measure of the intensity of the scattered radiation observed in the electron diffraction measurement. Soltzberg and Wilkins used the zero crossings of the function (2) for the classification and applied mathematical techniques of pattern recognition. As a result, they were able to define two distinct groups that can be considered as tranquilizers and sedatives.

Since it is often quite laborious to construct molecular geometries and to find zeros of molecular transforms we attempted to define a topological form of $I(s)$ [26, 27]. Our definition is as follows:

$$I_t(s) = C^{-1} \sum Z_i Z_j \frac{\sin x'}{x'}, \tag{3}$$

where $x' = sD_{ij}$, D_{ij} is a topological distance, *i.e.* the minimal number of bonds connecting atoms i and j. C is a normalization constant. A measure of distance, t_{ab}, can be defined and molecules a and b are considered to be the more similar the smaller the value of t_{ab}.

The topological molecular transform method has been applied to a set of substituted benzamidines, known to be potent inhibitors of β-trypsin [27]. On the basis of our cited work, we used two descriptors to divide these molecules in two groups. One is their abstract distance, t_{ab}, another is their inhibitory potency. As can be seen in Figure 3, the classification is quite successful if the topological distance is used. Our molecules form two classes, A and B. Within each class $t_{ab} < 0.5$ and $t_{ab} > 0.5$ if a belongs to A and b to B. There is only one exception for 21 pairs and 10 exceptions for 66 pairs in groups A and B, respectively. The distinction between groups A and B fails for 15 cases of 84 comparisons. This approximate grouping corresponds nicely to that obtained on the basis of inhibi-

	4-Me	4-NH_2	4-OH	4-OMe	3-Me	3-OH	3-OMe	4-OEt	4-NO_2	4-COOMe	4-COOEt	4-COMe	4-CONHMe	3-OEt	3-NO_2	3-COOMe	3-COOEt	3-COMe	3-CONHMe
4-Me																			
4-NH_2	8																		
4-OH	16	10																	
4-OMe	47	39	31																
3-Me	62	40	21	50															
3-OH	17	15	13	36	16														
3-OMe	45	38	30	15	46	31													
4-OEt	85	77	70	40	89	75	47												
4-NO_2	72	64	56	28	75	62	36	23											
4-COOMe	101	93	86	56	105	90	62	20	31										
4-COOEt	132	124	117	87	136	122	94	49	63	33									
4-COMe	65	57	50	23	69	55	31	28	71	37	70								
4-CONHMe	104	100	93	58	107	93	65	22	33	3	30	49							
3-OEt	82	74	67	37	85	70	40	15	23	27	57	27	29						
3-NO_2	69	61	54	29	70	55	27	32	20	41	72	41	43	23					
3-COOMe	98	91	82	53	100	85	56	25	31	19	44	19	19	20	32				
3-COOEt	128	121	113	83	131	117	88	45	59	30	17	30	28	48	64	33			
3-COMe	62	55	48	25	64	49	22	37	22	46	78	46	49	28	68	38	71		
3-CONHMe	100	92	85	56	103	88	58	26	33	18	41	18	18	22	34	3	31	40	

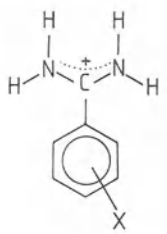


Fig. 3. Similarity matrix (values of $1000t_{ab}$, t_{ab} is the abstract distance between topological molecular transforms of species a and b) for 3 and 4-substituted benzamidines. Bold face figures denote values larger than 50. Me = methyl, Et = ethyl.

tory potencies (*cf.* Figure 4); there are only two outliers, 3-OEt and 3-COMe benzamidine, that have a pK_i larger than 4.2, yet fall into group *B* of Figure 3.

4-Me	4 523	4-OEt	4 000
4-NH_2	4 990	4-NO_2	3 482
4-OH	4 301	4-COOMe	3 523
4-OMe	4 495	4-COOEt	3 699
3-Me	4 523	4-COMe	3 495
3-OH	4 301	4-CONHMe	3 854
3-OMe	4 602	3-NO_2	3 887
3-OEt	4 569	3-COOMe	4 056
3 COMe	4 620	3-COOEt	4 187
		3-CONHMe	3 824
pK_i	>4 2	pK_i	<4 2

Fig 4 pK_i values for the inhibition of trypsin by substituted benzamidines (K_i in mol/l) Substituents in full boxes fall in group *B* of Figure 3, yet have a pK_i larger than 4 2 Substituents in dashed boxes do not obey Eq (4)

Mathematical measures of similarity may be quite successful but they do not give a full account on the structural basis of the problem. If one desires to study structural aspects in more detail, geometric comparisons should be made first, in order to find the main features determining similarity. A convenient way is offered by molecular graphics which allows superimposition of two or more structures and thus analyse similarity. An example is given in the work of Cohen [29]. He analyzed three-dimensional features of a set of representative active and inactive β-lactam structures and found that geometry may play a key role in the recognition of the antibiotic by enzymes in the biological processes. Some representative molecules chosen for 3-D comparison are shown in Figure 5. It becomes clear from the superimposed geometrical structures that active molecules have similar distances between the oxygen atom of the β-lactam amide group and the carbon atom of the carboxy group (*cf.* Figure 6). The distance is less than 390 pm for active and more than 410 pm for inactive compounds.

Geometrical similarity of various molecules can be studied using more advanced techniques of computer graphics, too. It is possible to superimpose van der Waals surfaces allowing us to judge the similarity of space-filling properties of a series of molecules [30].

The concept of geometrical similarity is not necessarily restricted to single molecules. An X-ray study of the crystal structure of the associate of 1,1′-

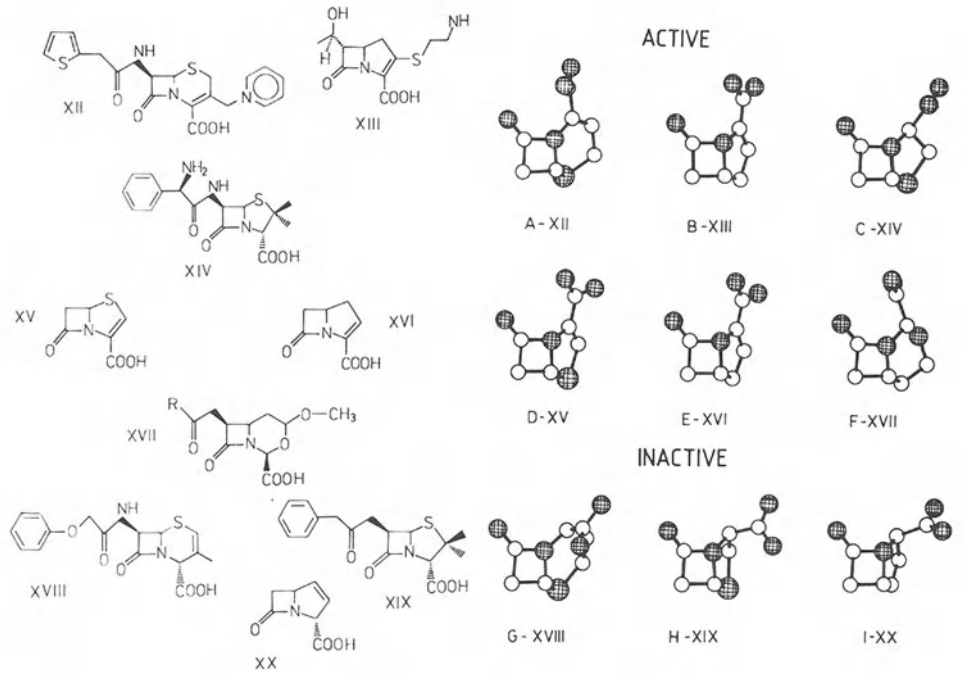


Fig. 5. Representative β-lactam antibiotics and their three-dimensional models (from [29] by permission).

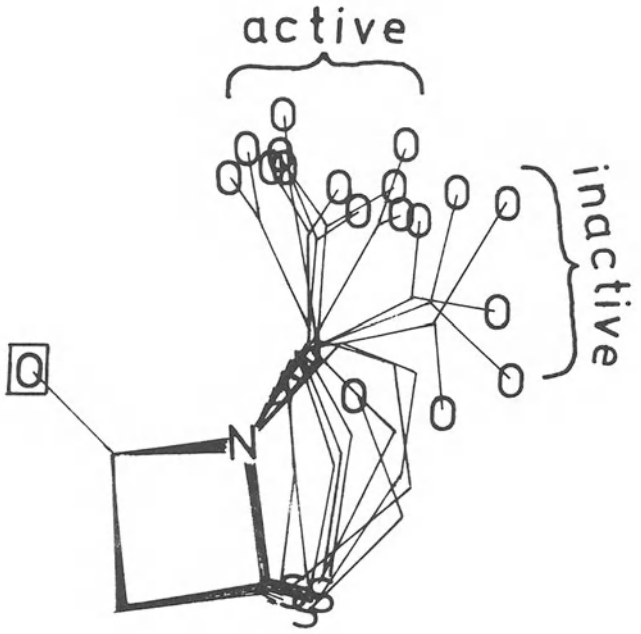
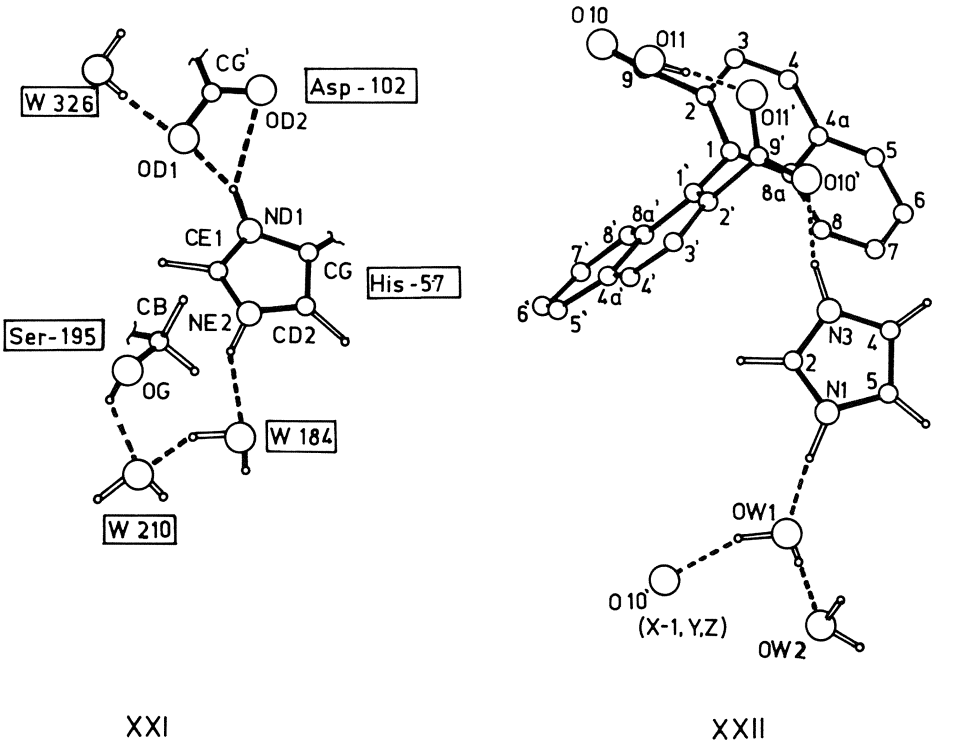


Fig. 6. Geometrical separation between active and inactive structures depicted in Figure 5 (from [29] by permission).

binaphthyl-2,2′-dicarboxylic acid with imidazole and water revealed striking congruence with the active site of an enzyme: *Streptomyces Griseus protease A* [31] (Figure 7). This similarity is not restricted to the geometry; charge distributions and electrostatic potential patterns coincide, too. Accordingly, the above associate, that can be very simple, may be used in certain aspects as an adequate model of the active site of the complicated enzyme.



XXI XXII

Fig. 7. Geometric arrangement of the carboxylate—imidazolium couple in *Streptomyces Griseus protease A* (**XXI**) and the 1,1′-binaphthyl-2,2′-dicarboxylic acid : imidazole : water adduct (**XXII**). Hydrogen bonds are indicated by dashed lines.

Of the eight requirements for an adequate isosteric replacement to produce a molecule similar to the parent one (*vide infra*), geometry represents only size and shape. Others, such as electronic distribution, pK_a, chemical reactivity and hydrogen bonding capacity may be accounted for by comparing molecular electronic densities or electrostatic potentials (MEP). Some works define molecular similarity by comparing electronic densities [32] but the use of MEP is more popular. As has been shown by Scrocco and Tomasi [5] the MEP is obtained from the molecular electronic density by integration and gives useful information on possible sites and

strength of protonation, reactivity in charge-controlled chemical reactions and hydrogen-bonding capacity. This is why several authors have used MEP maps to define molecular similarity.

Most studies concentrate on biological aspects, *i.e.* they analyze common features in the MEP of a series of selected molecules in order to understand details of enzyme or receptor binding. Maybe the first results on this field were published by Loew *et al.* who compared MEP maps of various morphine analogs [33]. On the basis of differences in the electrostatic potential they predicted that the interaction with a receptor, *i.e.* the biological activity, decreases in the following order:

6-monoacetylmorphine > morphine > *cis*-heroin > aminomorphine > > *trans*-heroin > *cis*-codeine > *trans*-codeine.

Their predictions are in complete agreement with, and may be considered to explain, the experimental results for these compounds. Following this work, the use of MEP became very popular in interpreting any similarity of drug molecules with respect to their action [34—41].

Similarity studies of whole MEP maps may become quite complicated since they contain too much information. It is therefore better to define a pattern of reference points around the molecules to be compared and analyze similarity in terms of the electrostatic potential values in these reference points [27, 42, 43]. Reference points are located outside the molecule near potential hydrogen-bonding groups and hydrophobic CH bonds. They are located in hypothetical lone-pair centres 100 or 150 pm apart from N, O and F or S and Cl atoms, respectively, and in the direction of CH, NH, OH or SH bonds at a distance of 100 pm from the hydrogen atom. These points lie inside the van der Waals envelope, in order to define potential hydrogen-bonding sites that are closer to proton donor and acceptor atoms than their van der Waals radii. It is important to notice that host and guest reference points do not necessarily coincide in the complex though they are located close to each other. Since, on the contact surface, the potential changes quite slowly, we may state that electrostatic complementarity is fulfilled in a given pair of reference points if the potentials in these points have the opposite sign, *i.e.* if the interaction is attractive in this region.

Electrostatic patterns, as defined above, are depicted in Figure 8 for two of the compounds given in Figure 4. As can be seen, the benzamidine parts show perfect similarity but the potential pattern is completely different around the substituents. While it is positive around the 4-OH groups, negative potentials represent the environment of the 4-NO_2 group. Since the potential in the active site of β-trypsin, where substituted benzamidine molecules bind, is everywhere negative it is expected that the hydroxy-substituted derivative will bind better than the nitro-substituted one. This finding can be generalized and, according to the above

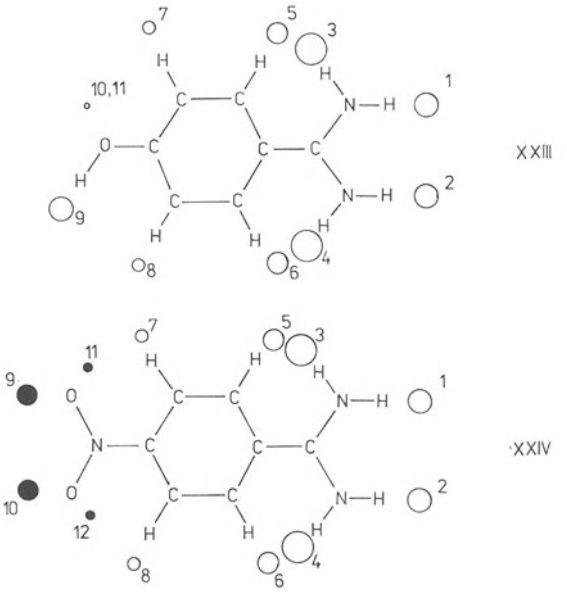


Fig. 8. Electrostatic potential pattern of 4-OH (**XXIII**) and 4-NO_2 (**XXIV**) substituted benzamidines. Empty and full circles stand for positive and negative potentials, respectively. Potential values (in kJ/mol) for **XXIII** and **XXIV** (in parentheses) are: 1: 564(619), 2: 566(619), 3: 547(623), 4: 559(623), 5: 479(561), 6: 496(561), 7: 280(338), 8: 371(338), 9: 465(−410), 10: 85(−410), 11: 85(−283), 12: (−283).

criterion, substituents can be divided in two groups showing similar electrostatic behaviours. These are:

S_I: 4-NH_2, 4-Me, 4-OH, 4-OMe, 4-OEt, 3-NH_2, 3-Me, 3-OH, 3-OMe, 3-OEt

S_{II}: 4-NO_2, 4-COOMe, 4-COOEt, 4-COMe, 4-CONHMe, 3-NO_2, 3-COOMe, 3-COOEt, 3-COMe, 3-CONHMe

It is postulated that for substituents in S_I and S_{II} the following inequality holds:

$$pK_i(S_I) > pK_i(S_{II}) \tag{4}$$

Comparing Eq. (4) with the data in Figure 4, it is found that there are only two outliers, the 3-COMe and 4-OEt derivatives, for which it does not hold. Generalizing Eq. (4) it was possible to define 'electrostatic isosteres' which are interchangeable without significant change in the binding strength of the molecule (*cf.* Table I).

Though the above studies are all focused on biomolecular interactions the concept of electrostatic similarity can be generated automatically for other problems like host—guest complexation [44] and crystal packing [45].

Discussing molecular similarity we should compare, finally, lipid and water solubility. As is known, solubility of solutes possessing sizeable saturated or

Table I. Classification of molecular fragments present in low molecular weight inhibitors of trypsin with the general formula R_c—R_h—S.

	R_c	R_h	S
good inhibitors	amidinium (Am^+)	phenyl	=NH
	guanidinium (Gu^+)	cyclohexyl	NH_2
	methylammonium (Ma^+)	*n*-propyl	OH
			CH_3(H)
weak inhibitors	H_2N^+=C=	alkyl $n = 3$	NO_2
		benzyl	COOR
			CHO

unsaturated hydrocarbon moieties, is much larger in lipid solvents than in polar media. It is the *hydrophobic effect* which is responsible for this phenomenon [46]. Owing to the decrease of solvent entropy around the apolar solute in an aqueous or other polar medium, the overall gain in the Gibbs free energy upon solvation diminishes and the solubility decreases. This unfavourable effect can be minimized if apolar groups of the solute molecules associate and exclude water from the contact region. This is why *e.g.* the so-called hydrophobic pocket of α-chymotrypsin represented by the apolar side chains Cys-191, Met-192, Val-213, Trp-215 and Cys-220 preferentially binds aromatic substrates. Hydrophobic groups tend to associate in crystals, too, though entropy effects play a minor role here.

A very popular measure of hydrophobic similarity is the Hansch π-value [47]. It is defined as $\log(P/P_0)$ where P is the partition coefficient between octanol and water for a substituted compound and P_0 is the coefficient for the unsubstituted compound. The π-values are additive so that the value of an unknown substituent can be approximated by summing up the values for fragments of the substituent. In a given series of compounds they may be proportional to molecular surface area.

While the Hansch approach is extremely successful in a great variety of cases, it does not lead to a structural picture of what is happening at the molecular association for which similarity has to be defined. Therefore we proposed to use molecular electrostatic fields, calculated at certain reference points around a molecule, for the characterization of molecular similarity [42]. At a given site the molecular electrostatic field orients and interacts with point dipoles, modelling water molecules participating in hydration. The stronger the interaction the larger is the absolute value of the electrostatic field [48—51]. Accordingly, hydrophobic regions around a molecule are indicated by small fields. Hydrophobic similarity between two molecules means matching of small field regions.

3. Harmony in Molecular Aggregates: Complementarity

The concept of complementarity between interacting molecular species has its

origin in the classical lock-and-key analogy of Emil Fischer. According to this, enzymes are modeled by locks and the successful interaction with a substrate can be considered as fitting a key into this lock. It is clear that this analogy can be generalized to other non-covalent molecular complexes, too. Since the same lock may accommodate several similar keys, there exists a close relationship between similarity and complementarity if the former is restricted to molecular interactions as in Section 2.

Complementarity is a prerequisite for successful *recognition* of certain molecules by others. Recognition is one of the fundamental events in molecular biology [52, 53], in host—guest complexation [54] and in optical resolution of chiral compounds [55, 56]. Molecular recognition requires that the interacting molecules be in contact over a relatively large area. Thus the lock (biopolymer, host, *etc.*) should contain intramolecular cavities sufficiently large to allow inclusion of the key (ligand, guest, *etc.*). Proper knowledge of the factors determining successful binding between complementary molecular systems allows to construct hollow molecules of defined architecture that can bind *i.e.* recognize specific ligands. Accordingly, the concept of molecular recognition is especially important in the design of drugs and various supramolecular devices [54].

Going further in the definition of molecular complementarity it has to be stressed that it is not completely equivalent to energetics. Two molecular species interact in order to minimize their total Gibbs free energy, so association of complementary structures involves a decrease of this quantity. However, this is not enough, since to achieve a perfect fit all regions of the associating molecules have to interact favourably. Accordingly, if the repulsion between certain regions is overbalanced by a strong attraction between others, we do not call complementarity necessarily better than in a case where the total Gibbs free energy of interaction is not as large as in the above case but all regions of the associating molecules interact favourably.

The total Gibbs free energy of association between the lock and the key can be partitioned as follows:

(a) exchange repulsion,
(b) dispersion interaction,
(c) electrostatic interaction,
(d) inductive interaction,
(e) solvent effects,
(f) entropy effects.

As in the case of similarity, we can trace back these effects to three major factors. These are geometric and electrostatic complementarities and matching of nonpolar regions.

Geometric complementarity involves exchange repulsion and dispersion interactions. This means that nonbonding atoms of the interacting species cannot get closer to each other than a certain limit. On the other hand, the ligand tends to fill

the crevice of the host as perfectly as possible in order to maximize non-specific (van der Waals) interactions.

If the interacting molecules possess polar groups the electrostatic and inductive terms become important. Roughly speaking, positively and negatively charged or inversely polarized regions should get as close to each other as possible.

One of the major components of solvent effects, the hydrophobic effect can be brought into connection with matching of non-polar regions of the associating molecules. If the interacting molecules are in an aqueous medium, unfavorable contacts with water should be avoided, *i.e.* apolar regions should get close to each other and extrude water from the contact surface. Thus the net hydration energy of the system increases since the unfavorable hydrophobic effect is reduced to a minimum.

Finally, pure entropy effects, like freezing of rotational degrees of freedom upon association, are not discussed here since they are not related to molecular complementarity. The only thing that we can say is that rigid structures ensure better complementarity since in this case unfavorable entropy effects do not diminish the probability (*i.e.* the total Gibbs free energy) of association.

Summarizing the above considerations, we can formulate three requirements for complementarity between two interacting molecules [57, 58]:

(1) geometric complementarity,
(2) electrostatic complementarity,
(3) matching of nonpolar regions.

Let us now discuss various aspects in detail.

First of all we define a useful notion. There are certain crucial distances, angles and cases of optical handedness that are recognized by the receptor. This combination of geometric features is reduced to the lowest common denominator of a class of molecules with similar biological activities and may be called a *pharmacophore* [59]. The pharmacophore defines the complementary crevice of the lock, too; therefore it is an especially useful concept if only one participant of the non-bonded association (drug, inhibitor, host, *etc.*) is known. Geometric aspects of complementarity can best be studied on three-dimensional models that can be built by molecular graphics or determined experimentally by X-ray diffraction [60].

An interesting example is depicted in Figure 9 [54]. The $^{+}H_3N—(CH_2)_n—NH_3^{+}$ substrate, that is preferentially bound to the host, has a length complementary to the length of the molecular cavity with $n = 4$, 5 and 7 for **XXV**, **XXVI** and **XXVII**, respectively. Thus the crown ether **XXVI** discriminates between cadaverine ($n = 5$) and putrescine ($n = 4$) cations. These phenomena are evidence for geometric complementarity between host and guest being the basis of molecular recognition.

While it is clear that several similar keys may fit into the same lock, a perfect fit is possible only for a few keys, or just for one. If geometric complementarity only is considered, a perfect fit means that the binding ligand fills the crevice of the host

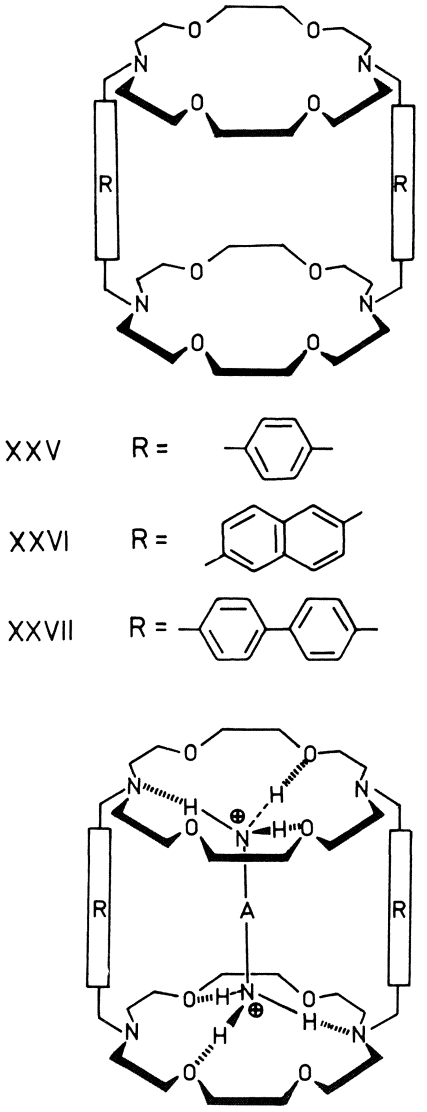


Fig. 9. An example of size-selective host—guest interaction. A = $(CH_2)_n$ (from [54] by permission).

as much as possible. Such a perfect space-filling maximizes nonspecific interactions and results in stronger binding [61]. The principle has been exploited by Blaney *et al.* to design potent thyroid hormone analogues [62].

Let us now turn to electrostatic complementarity. An important series of papers treating this problem has been published by Weinstein and coworkers [35, 63, 64]. They investigated a group of serotonin (5-hydroxytryptamine) congeners: tryptamine and various hydroxy-substituted derivatives, which incorporate high-affinity, low-

affinity and intermediate compounds. The group of molecules bear a strong structural similarity to LSD. They used electrostatic potential maps to provide an interaction pattern for the agonist with the receptor which must be matched by other agonists if they are to possess similar affinities. This pattern has been called the 'interaction pharmacophore'.

It has been postulated that the vector connecting the minima of the molecular electrostatic potential map serves as an indicator of the preferred alignment of the molecule towards the field generated by the receptor. The orientation vectors for other agonists may considerably differ from that of serotonin, *e.g.* for 6-hydroxytryptamine it is almost perpendicular to the former. This difference may be postulated as one reason for the differing activities. Further it is possible to set up a hypothesis concerning recognition at the serotonin receptor sites:

— the positively charged ethylamine side chain interacts with an anionic group of the receptor, preparing the indole position to form a stacking complex;
— the electrostatic orientation vector determines the position of the indole ring system.

It was further suggested that the other hydroxytryptamines anchor themselves in a similar manner but have to adjust their conformation to orientate the electrostatic vector. The measured difference between the affinity of serotonin and any other member of the series for the LSD—serotonin receptor results from the discrepancy between the orientation vectors. This hypothesis has been checked by experimental measurements and was successfully applied to the design of new, potent LSD analogs [65].

While it is obvious that molecular electrostatic potential maps are especially useful in the study of complementarity, they extend over an infinitely large space therefore contain too many informations. It is better to consider a reasonably selected fraction of the map, *e.g.* the van der Waals envelope where equipotential regions may be indicated by different colors [66]. Recently we have suggested the use of a few reference points around the associating molecules that are characteristic for the interaction and allow us to analyze complementarity in the necessary detail (*cf.* Section 2).

Let us first study the electrostatic complementarity between trypsin and basic pancreatic trypsin inhibitor, BPTI [67]. Analysis of the electrostatic potential patterns around the interacting molecules has shown that there is no complementarity in the peptide carbonyl regions of the Pro-13,Cys-14,Lys-15,Ala-16,Arg-17 ($P_3P_2P_1P'_1P'_2$) fragment of BPTI which is bound to the active site of the enzyme. Both enzyme and inhibitor potentials are negative in the corresponding reference points. This is due to two factors. First, the positive charge located on Lys-15 in BPTI is insufficient to overbalance the negative potential in the lone-pair region around backbone carbonyl atoms of the inhibitor. Second, the relative arrangement of proton donor and acceptor groups in the interacting molecules is not suitable for an optimal hydrogen bonding. The situation for the Pro-13,Cys-14

region is illustrated in Figure 10. Note *e.g.* that the geometry of the NH(Glu-192) . . . OC(Cys-14) hydrogen bond is considerably distorted.

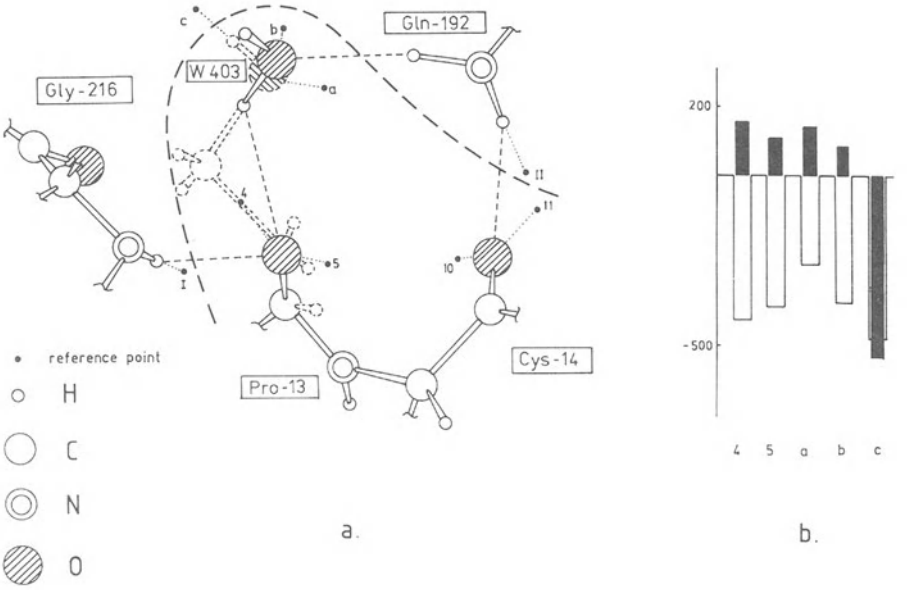
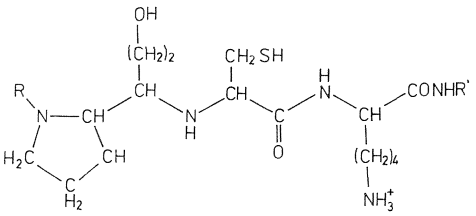


Fig. 10. Geometric arrangement of the Pro-13 region of BPTI in the enzyme-inhibitor complex (a). Hydrogen bonds and atoms of the hypothetical isostere, $CHCH_2CH_2OH$, are indicated by dashed lines. The heavy dashed line is the border between trypsin and BPTI. Electrostatic complementarity around the carbonyl group and the isostere is also illustrated (b).

The somewhat astonishing imperfection in the electrostatic fit between trypsin and the otherwise very strongly bound BPTI in the lone-pair regions is partly overbalanced by a structural water molecule, W403, bound at the contact surface between the enzyme and the inhibitor. The overall electrostatic fit in the new reference points *a, b* and *c*, determined by W403, is considerably better, as can seen from Figure 10. The better electrostatic fit suggests that W403 should be built into the inhibitor, *i.e.* the =C=O. . .HOH structural unit should be replaced by a $=CHCH_2CH_2OH$ group. This means, that an appropriately modified tripeptide, **XXVIII**, should ensure better electrostatic matching, *i.e.* stronger



XXVIII

binding and larger inhibitory power. Beside the gain in electrostatic binding energy the $=CHCH_2CH_2OH$ group is entropically also favored over the carbonyl since it expels W403 from the contact surface to bulk water, increasing the total entropy of the system. Incorporation of a structural water into the inhibitor is thought to be responsible for the especially strong binding of pepstatin to pepsin and penicillopepsin [68].

We gave here an example of a novel type of non-classical bioisosterism discussed in the previous section. Further possible isosteric replacements are summarized in Table II. The gain in the Gibbs free energy of binding is estimated after Andrews *et al.* [69] who derived averaged binding energies of common bifunctional groups to determine the goodness of fit of a drug to its receptor. The large increase in ΔG for the substitution of $—NH_3^+$ by $—CH_2CH_2NH_3^+$ or $—CH(CH_2NH_2)CH_2NH_3^+$ in Lys-15 is due to the displacement of W414 and W416 from the contact surface to bulk water.

Table II. Possible isosteric replacements in tripeptide analog inhibitors of trypsin. ΔG is the lower bound for the increase in Gibbs free energy of binding (kJ/mol).

residue	group	isostere	ΔG
Pro-13 (P_3)	backbone $=CO$	$=CHCH_2CH_2OH$	+6
Cys-14 (P_2)	backbone $=CO$	$=CHOH$	−7
Lys-15 (P_1)	backbone $=CO$	$=CHOH$	−7
	side chain	$—CH_2CH_2NH_3$	+13
	$—NH_3$	$—CH(CH_2NH_2)CH_2NH_3$	+28

Matching of non-polar regions in host—guest complexes is more difficult to handle by simple theoretical methods than is electrostatic complementarity. Electrostatic fields around molecules are much more sensitive to the location of reference points than potentials, therefore it is not easy to define the term 'matching'. Hydrophobic regions of a host can be characterized by small fields in their vicinity. While it is true that hydrophobic groups of the guest producing strong fields should avoid these regions, the mutual arrangement of host and guest hydrophobic regions may vary considerably, *i.e.* they can get closer to or further away from each other.

The hydrophobic complementarity can be studied on the fit between the Pro-13 side chain of the Pro-13,Cys-14,Lys-15 tripeptide part of BPTI and its complementary region in trypsin [67]. The enzyme fields in reference points near the $—CH_2CH_2—$ moiety are 0.2 and 0.4 V/nm, while the fields emerging from the tripeptide are 1.1 and 1.2 V/nm, respectively. This indicates that this hydrophobic region of the enzyme finds a fragment of the key that produces also a relatively small field and is therefore unsuitable for hydration. On the other hand, the improper fit in the Pro-13,Cys-14 carbonyl regions (*cf.* reference points 4, 5, 10

and 11 in Figure 10) is indicated also by the difference between enzyme and inhibitor electrostatic fields. The corresponding values in V/nm are 0.7(3.6), 0.2(4.0), 0.4(4.0) and 1.2(4.2), tripeptide values are in parentheses. Owing to the small enzyme fields around the inhibitor hydrogen bonding is not ensured and the fit is imperfect.

While the pharmacophore is the least common multiple of various ligands, *i.e.* of different keys unlocking the same receptor, it is also possible to define the complementary: different locks, opened by the same 'master key'. The three requirements for successful binding, as given above, are the same for different types of host—guest complexes therefore it can be expected that a host should find very similar local environments in guests otherwise considerably differing from each other. Let us now look at two examples.

Liebman used crystallographic data to analyze the binding pocket of the receptor interacting with serotonin and its methylated congener, bufotenine [70]. He examined the 1:1:1 stoichiometric complex of serotonin:picrate:water; he also examined bufotenine, having no solvent present in the crystal but two independent molecules in the asymmetric unit. The molecules under study were considered to be the centers of a local environment formed by their nearest neighbors in the crystal. Comparison of the geometric features of the local environments and the electrostatic potentials produced by them gave important information on the serotonin receptor. Liebman's approach has been used to examine a mitotic spindle poison that binds reversibly to the colchinine binding site of tubulin [71].

Another example is the carboxylate—imidazolium diad embedded either in a crystal of the 1,1′-binaphthyl-2,2′-dicarboxylic acid dihydrate—imidazole adduct or in an enzyme, *Streptomyces Griseus protease A* (SGPA, *vide supra*) [31, 32]. The diad, depicted in Figure 7, sees similar electrostatic potentials in both, otherwise very different, systems. On Figure 11 the electrostatic complementarity between atomic charges of the carboxylate—imidazolium ion pair and its environment in the crystal and in the enzyme is displayed. It is seen that crystal and enzyme potentials show a very similar variation along atoms of the diad. The minimum character is somewhat more pronounced in SGPA indicating the stronger stabilization of the diad by the enzyme than by the crystal environment.

The above two studies may serve as first examples for the rational design of crystal architectures that mimic receptor/protein binding sites. Another approach is the synthesis of artificial enzymes that are mostly based on crown ether or cyclodextrin structures [54]. These methods may complement each other. A considerable advantage of the crystal analogs is that they can be produced relatively easily and their X-ray structure determination is almost routine, in general.

A potentially powerful approach for modelling receptor/enzyme active sites is the methodology of distance geometry [73], which allows the construction of binding modes between a receptor and its substrate that meet selected distance

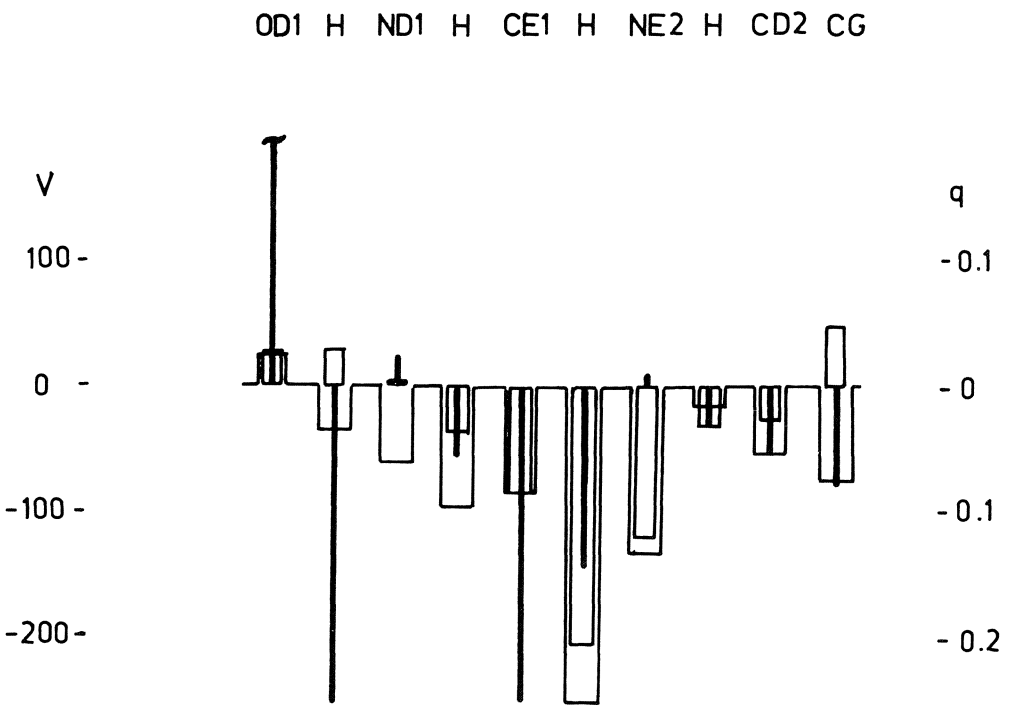


Fig. 11. Electrostatic complementarity between atomic charges (q) of the carboxyl—imidazolium couple (solid bars, in electrons), the electrostatic potential of *Streptomyces Griseus protease A* protein environment (narrow empty bars, in kJ/mol) and the 1,1′-binaphthyl-2,2′-dicarboxylic acid crystal environment (broad empty bars, in kJ/mol).

constraints. These can then be evaluated according to energetic considerations and other geometric criteria. The distance geometry manipulation can superimpose two bodies without explicitly calculating the necessary rigid rotation and translation. Thus, comparison of a large number of allowed conformations of various ligands becomes possible: the speed of the algorithm is considerably higher than in case of three-dimensional coordinate manipulation.

4. Harmony of Molecules as Objects of Nature: Beauty

As we mentioned in the introduction, 'harmonize' means also 'be agreeable in artistic effect'. Therefore, the purely scientific aspects of the term 'harmony' discussed in the two previous sections may be complemented by another one which is related to the fine arts. Molecules are parts of Nature, too; therefore, like other macroscopic objects as landscapes, plants, animals or human beings, they have a certain intrinsic beauty. An aspect of molecular beauty is symmetry, which has been extensively discussed in a recent book edited by Hargittai [74].

Molecular beauty is, however, hidden from most people. Even the concept of

molecules is not much older than a hundred years and the first detailed experimental information became available only in the 'twenties. To develop the conception of a molecule needed, until recently, years of training in some molecular science, preferably X-ray diffraction, spectroscopy or theory. It is therefore understandable that the overwhelming majority of people, including most scientists, are a bit anxious when hearing about molecules. This is the usual reaction when encountering completely unknown persons or events. The attitude is similar to the relation between most people and the forests in the Middle Ages. Even in Shakespeare's Europe the forest was thought to be filled with more or less harmful fairies, elves, goblins and other living or ethereal beings [75]. As civilization spread and more and more individuals discovered the attractive features of the wilderness, it became the object of arts and poetry.

The continuously increasing performance of instruments, especially photography, displaying physical, chemical and biological events had an impact on the fine arts, too. Kepes was one of the first who recognized the importance of new techniques in artistic photography and displayed *e.g.* a flying bullet with Mach waves around it or the eye of a fly in 200-fold enlargement [76]. The unbiased spectator may discover harmony in and feel attracted to these, otherwise scientifically also particularly interesting, pictures, and the concept of beauty will be extended.

Beyond photography, almost inconceivable perspectives have been opened in the visualization of objects of the macro and micro-world by computer graphics [77]. Its application to molecules, molecular graphics, has become a new branch of theoretical chemistry [78, 79]. Visualization of molecules allows us to understand structural features much more easily, especially if colors are also used. Accordingly, both specialized experts and laymen get considerably more insight in the world of molecules than would be possible without computers. Molecular graphics may help very much in the popularization of the molecular sciences, too. While most scientists do not meditate on such problems, popularization is very important, especially when the deteriorating image of the natural sciences, in particular chemistry, is considered.

Anybody who is viewing a complicated molecule, displayed on a computer screen, gets impressed by the turbulence of shapes and colors that is inherently governed by a functional order. It is therefore not absurd to speak about the beauty of molecules represented by computer or mechanical models. In his paper, Lehn compares several slides depicting sculptures by Henry Moore and CPK models of various cage molecules, synthesized from crown ether analogs [80]. The striking analogy may serve as a further manifestation of molecular beauty.

On the following pages (Figure 12, Color Plates VII and VIII) we present some pairs of photographs where molecules, displayed by the computer, are compared to an object of nature or fine art. We feel that the analogy can be recognized within each pair and the reader (or better: spectator) may get some feeling about the artistic effect brought about by sophisticated molecular models.

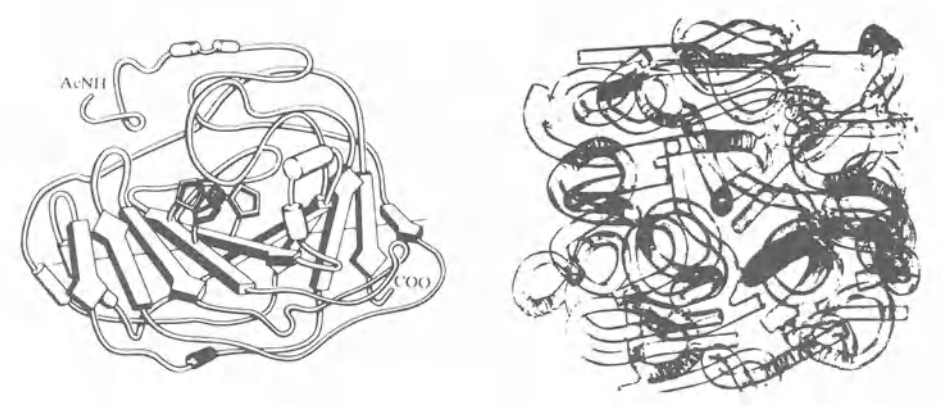


Fig. 12. Schematic model of human carbonic anhydrase (left); V. Liskova: *Glass Plastics*, 1982 (right).

If molecules have an esthetic value, they may become objects of fine art, as well. Indeed, as Professor Daudel shows in his paper [81], a renowned artist, Madame D'Agaggio, has chosen molecules to express her sentiments towards life and death, love and disillusion. It does not seem impossible that molecular art will become an integral part of modern spiritual life, like expressionism or pop art. It is not beyond the bounds of reality to suggest that some molecular scientists would get involved in such an activity with pleasure.

5. Harmony Among Molecular Scientists: Interdisciplinarity

At last we have arrived at a point that is of primary importance in scientific work: the human factor. Scientists, especially molecular scientists, belong to the type of intellectuals who know 'everything about nothing'. The highly specialized experimental techniques and theoretical methods need years of training; therefore it is understandable that *e.g.* an NMR spectroscopist is not necessarily familiar with X-ray diffraction or quantum chemistry. The diversity of branches and subbranches in molecular sciences has lead to the development of scientific ghettos and, in general, experts applying different techniques do not have very much to do with each other. Interdisciplinary studies are rarely, or not as frequently, carried out as is desirable.

A prerequisite for interdisciplinarity is tolerance. Experts in a given field should give credit to techniques different from their own, too. An example of intolerance has been the reaction of organic chemists against the applications of X-ray diffraction in the elucidation of mechanistic problems in solution. The objection raised very often has been that the crystal structure determination refers to the solid phase while organic reactions take place mostly in solution. However, in the past decade appropriate use of the coordinate holdings of the Cambridge Crystallographic Data Files has yielded a lot of examples where structural information

referring to the solid state gave important hints on reaction mechanisms in solution [82].

A further example may be the long debate between theoretical chemists using more or less qualitative concepts like hybridization, aromaticity or frontier orbitals, and sophisticated computer hardware and software, respectively, to elucidate a given molecular phenomenon. Some computational chemists, mostly possessing powerful hardware, deem qualitative concepts more or less superfluous; they should be applied in the education process, if at all. On the other hand, several theoreticians follow the line of Pauling and use simple models to understand certain molecular properties or events [83]. It seems that both approaches should have about an equal weight in theoretical chemistry. While smaller molecules can be handled quantitatively at the *ab initio* level, sometimes with an accuracy exceeding that of experiments, for larger systems simpler methods have to be applied. On the other hand, quantitative figures and qualitative concepts should be used together to get a deeper insight into the nature, properties and interactions of molecules.

The ambition to achieve interdisciplinarity may not mean that the rigorous treatment of a given problem should be abandoned. Clearly, careless work, improper argumentation or logical errors cannot be tolerated in a scientific study. Interdisciplinarity therefore involves profound studies in the related areas that are not always practicable for a single person. Team-work is necessary, coordinated by an emphatic personality who should be vigorous enough and possess sufficient ability to gain an overall view of various disciplines. To work in such a team likewise presumes tolerance on the human and scientific levels, as well. Fortunately, several interdisciplinary teams are working worldwide and they often achieve spectacular success in *e.g.* molecular biophysics, supramolecular chemistry or solid-state physics. The number of interdisciplinary symposia is continuously growing, a spectacular example has been that on the role of molecules in physics, chemistry and biology organized by Professors Csizmadia and Maruani and serving as a basis for the chapters of this book. Though we have pleaded for more tolerance in the community of molecular scientists, it is far better here than it is in everyday life. It would be nice if politicians had only as much of a problem concerning tolerance as we do.

6. Conclusions

In this chapter we have attempted to discuss the term 'harmony' as related to molecules. The scientific aspects are similarity and complementarity, both related to noncovalent bonding between host and guest molecules. Three requirements for similarity and complementarity are formulated in a decreasing order of importance: geometric, electrostatic and hydrophobic. Prerequisite for binding is the geometric fit between the interacting species, while electrostatic complementarity is important, but may be violated in some cases. The smallest effect on the

interaction energy, emerging from the matching of nonpolar regions through the related hydrophobic effect, may be definitive in some cases. Based on the above requirements, the concepts of the lock (host or pharmacophore) and key (guest or ligand) could be defined a bit more clearly. Some examples have been given for cases where different keys fit into the same lock or, on the other hand, the same key opens different locks.

Molecular harmony also has artistic and human aspects that are linked to the beauty of molecular models and interdisciplinarity in molecular sciences. The colored illustrations may give some hints on the former while the Paris Conference dedicated to Professor Daudel has been an excellent example of the latter.

This paper is dedicated to Professor Daudel, a man who knows how to integrate the harmony of molecules and the harmony of the arts in his own personality.

References

1. J. Coulson, C. T. Carr, L. Hutchinson and D. Eagle (Eds.): *The Oxford Illustrated Dictionary*, 2nd Edn., Clarendon, Oxford (1982).
2. W. Heisenberg: *Physics and Beyond*, Harper and Row, New York (1971).
3. L. Pauling: *The Nature of the Chemical Bond*, 3rd Edn., Cornell University Press, Ithaca (1960).
4. J. March: *Advanced Organic Chemistry, Reactions, Mechanisms and Structure*, 3rd Edn., Wiley, New York (1985).
5. E. Scrocco and J. Tomasi: *Adv. Quant. Chem.* **11**, 115 (1978).
6. S. Patai (Ed.): *The Chemistry of Functional Groups*, Wiley, London (1964 onwards).
7. A. Burger: *Introduction* (Medicinal Chemistry, 3rd Ed., Ed. A. Burger), p. 1. Wiley (1970).
8. C. W. Thornber: *Chem. Soc. Rev.* **8**, 563 (1979).
9. D. T. Collin, D. Hartley, D. Jack, L. H. C. Hunts, J. C. Press, A. C. Ritchie and P. Toon: *J. Med. Chem.* **13**, 674 (1970).
10. C. Kaiser: *J. Med. Chem.* **17**, 49 (1974).
11. A. A. Larsen, W. A. Gould, H. R. Roth, W. T. Comer, R. H. Uloth, K. W. Dungan and P. M. Lish: *J. Med. Chem.* **10**, 462 (1967).
12. J. Keck, G. Kruger, K. Noll and H. Machleidt: *Arzneimitt. Forsch.* **22**, 861 (1972).
13. C. D. Arnett, J. Wright and N. Zenker: *J. Med. Chem.* **21**, 72 (1978).
14. H. W. R. Williams: *Can. J. Chem.* **54**, 3377 (1976).
15. S. Yoshizaki, K. Tasimura, S. Tamada, Y. Yabuuchi and K. Nakagawa: *J. Med. Chem.* **19**, 1138 (1976).
16. A. J. Stuper, W. E. Brugger and P. C. Jurs: *Computer Assisted Studies of Chemical Structure and Biological Function*, Wiley, New York (1979).
17. J. A. Hartigan: *Clustering Algorithms*, Wiley, New York (1975).
18. W. J. Wiswesser: *A Line-Formula Chemical Notation*, Crowell, New York (1954).
19. G. Klopman: *J. Am. Chem. Soc.* **106**, 7315 (1984).
20. G. Klopman, O. T. Macina, E. J. Simon and J. M. Hiller: *J. Mol. Struct. THEOCHEM* **134**, 299 (1986).
21. G. W. Adamson and J. A. Bush: *J. Chem. Inf. Comput. Sci.* **15**, 55 (1975).
22. P. Willett and V. Winterman: *Quant. Struct.-Act. Relat.* **5**, 18 (1986).
23. M. Randic: *J. Am. Chem. Soc.* **97**, 6609 (1975).
24. L. B. Kier and L. H. Hall: *Molecular Connectivity in Chemistry and Drug Research*, Academic, New York (1976).
25. L. J. Soltzberg and C. L. Wilkins: *J. Am. Chem. Soc.* **98**, 4006 (1976); **99**, 439 (1977).
26. Z. Gabanyi, P. Surjan and G. Naray-Szabo: *Eur. J. Med. Chem.* **17**, 307 (1982).
27. G. Naray-Szabo: *J. Mol. Struct. THEOCHEM* **134**, 401 (1986).

28 G Moreau and P Broto *Nouv J Chim* **4**, 359 (1980)
29 N C Cohen *J Med Chem* **26**, 259 (1983)
30 A J Hopfinger *J Am Chem Soc* **102**, 7196 (1980)
31 M Czugler, J G Angyan, G Naray-Szabo and E Weber *J Am Chem Soc* **108**, 1275 (1986), P Nagy, J G Angyan and G Naray-Szabo *J Mol Struct THEOCHEM* **149**, 169 (1987)
32 R Carbo, L Leyda and M Arnau *Int J Quant Chem* **17**, 1185 (1980), E E Hodgkin and W G Richards *JCS Chem Commun*, 1342 (1986)
33 G H Loew, D Berkowitz, H Weinstein and S Srebrenik *Quantum Chemical Studies of Morphine-Like Opiate Narcotics Effect of Polar Group Variations* (Molecular and Quantum Pharmacology, E D Bergman and B Pullman Eds), p 355 Reidel (1974)
34 G H Loew and D S Berkowitz *J Med Chem* **18**, 656 (1975)
35 H Weinstein, D Chou, S Kang, C L Johnson and J P Green *Int J Quant Chem Quantum Biol Symp* **3**, 135 (1976)
36 C Petrongolo, B Macchia, F Macchia and A Martinelli *J Med Chem* **20**, 1645 (1977)
37 C Petrongolo, H J T Preston and J J Kaufman *Int J Quant Chem* **13**, 457 (1978)
38 M Martin, F Sanz, M Campillo, I Pardo, J Perez and J Turmo *Int J Quant Chem* **23**, 1627 (1983)
39 S Hirono, H Umeyama and I Moriguchi *Chem Pharm Bull* **32**, 3061 (1984)
40 H van de Waterbeemd, P A Carrupt and B Testa *Helv Chim Acta* **68**, 725 (1985)
41 D Kocjan, M Hodoscek and D Hadzi *J Med Chem* **29**, 1418 (1986)
42 P Nagy and G Naray-Szabo *Can J Chem* **63**, 1694 (1985)
43 G Naray-Szabo *J Mol Struct THEOCHEM* **138**, 197 (1986)
44 G Naray-Szabo, G Kramer and P Nagy *J Indian Chem Soc* **62**, 860 (1985)
45 P Laczik, G Naray-Szabo and K Simon *Magy Kem Folyoirat*, in press
46 G Ravishanker and D L Beveridge *Theoretical Studies of the Hydrophobic Effect* (Theoretical Chemistry of Biological Systems, G Naray-Szabo Ed), p 423 Elsevier (1986)
47 C Hansch *Acc Chem Res* **3**, 232 (1969)
48 G Peinel, H Frischleder and F Birnstock *Theoret Chim Acta* **57**, 245 (1980)
49 A Pullman, B Pullman and R Lavery *J Mol Struct THEOCHEM* **93**, 85 (1983)
50 J A Tainer, E D Getzoff, J Sayre and A J Olson *J Mol Graph* **3**, 303 (1985)
51 P Nagy, J G Angyan, G Naray-Szabo and G Peinel *Int J Quant Chem* **31**, 927 (1987)
52 G Naray-Szabo and K Simon (Eds) *Steric Aspects of Biomolecular Interactions*, CRC Press, Boca Raton (1987)
53 A S Horn and C J De Ranter (Eds) *X-Ray Crystallography and Drug Action*, Clarendon, Oxford (1984)
54 J M Lehn *Science* **227**, 849 (1985)
55 R O Gould, R Kelly and M D Walkinshaw *JCS Perkin Trans II*, 847 (1985)
56 E Fogassy, M Acs, F Faigl, K Simon, J Rohonczy and Z Ecsery *JCS Perkin Trans II*, 1881 (1986)
57 G Naray-Szabo *J Am Chem Soc* **106**, 4584 (1984)
58 G Naray-Szabo and P Nagy *Electrostatic Recognition in Proteins* (Molecular Basis of Cancer, Part B Macromolecular Recognition, Chemotherapy and Immunology, R Rein Ed), p 105 Alan Liss (1985)
59 P Gund *Progr Cell Mol Biol* **117** (1977)
60 P J Goodford *J Med Chem* **27**, 557 (1984)
61 P A Kollman *Drug-Receptor Binding Forces* (X-Ray Crystallography and Drug Action, A S Horn and C J De Ranter Eds), p 63 Clarendon (1984)
62 J M Blaney, E C Jorgensen, M L Connolly, T E Ferrin, R Langridge, S J Oatley, J M Burridge and C C F Blake *J Med Chem* **25**, 785 (1982)
63 H Weinstein, R Osman, W D Edwards and J P Green *Int J Quant Chem Quantum Biol Symp* **5**, 449 (1978)
64 H Weinstein, R Osman, S Topiol and J P Green *Ann N Y Acad Sci* **367**, 434 (1981)
65 H Weinstein, R Osman and P Mazurek *Simulation of Molecular Stereoelectronic Mechanism for the Interaction of Hallucinogen and Indole Derivatives at 5-HT Receptors* (Steric Aspects of Biomolecular Interactions, G Naray-Szabo and K Simon Eds), CRC Press (1987)

66 H Nakamura, K Komatsu, S Nakagawa and H Umeyama *J Mol Graph* **3**, 2 (1985)
67 G Naray-Szabo and P Nagy *Enzyme* **36**, 44 (1986)
68 D H Rich *J Med Chem* **28**, 263 (1985)
69 P R Andrews, D J Craik and J I Martin *J Med Chem* **27**, 1648 (1984)
70 M N Liebman *J Mol Graph* **4**, 61 (1986)
71 M N Liebman *Topological Analysis of Chemotherapeutic Systems* (Molecular Basis of Cancer, Part B Macromolecular Recognition, Chemotherapy and Immunology, R Rein Ed), p 285 Alan Liss (1985)
72 J Szejtli (Ed) *Cyclodextrins*, Reidel, Dordrecht, 1982
73 A K Ghose and G M Crippen *J Med Chem* **6**, 350 (1985)
74 I Hargittai (Ed) *Symmetry Unifying Human Understanding*, Pergamon, New York (1986)
75 W Shakespeare *The Tempest* (Shakespeare Complete Works, W J Craig Ed), Oxford University Press (1969)
76 G Kepes *The New Landscape in Art and Science*, Paul Theobald, Chicago (1956)
77 J Foley and A Van Dam *Fundamentals of Interactive Computer Graphics*, Addison-Wesley, Reading (1981)
78 R Langridge, T E Ferrin, I D Kuntz and M L Connolly *Science* **211**, 661 (1981)
79 W G Richards (Ed) *Journal of Molecular Graphics*, Butterworths, Guildford, quarterly since 1983
80 J M Lehn *Perspectives in Supermolecular Chemistry Molecules, Supermolecules, Molecular Devices*, lecture held at the International Symposium on Molecules in Physics, Chemistry and Biology, Paris, 1986
81 R Daudel *De la Science Moleculaire a l'Art Moléculaire*, lecture held at the International Symposium on Molecules in Physics, Chemistry and Biology, Paris, 1986
82 H B Burgi and J D Dunitz *Acc Chem Res* **16**, 153 (1983)
83 Z B Maksic (Ed) *Conceptual Approach in Quantum Chemistry* Special Issue of *Croatica Chemica Acta*, Zagreb (1984)

From Molecular Science to Molecular Art

RAYMOND DAUDEL
European Academy of Arts, Sciences and Humanities, 60, rue Monsieur le Prince, 75006 Paris, France.

1. Introduction

Nowadays four kinds of relations between sciences and plastic arts are particularly important [1, 2].

(1) Many artists integrate, in their works, products of sciences and technologies, like neon tubes, lasers, holograms, and so on.

Milton Komisar's computer-programmed illuminated sculptures [3] represent a remarkable example of such artworks. The light sculpture NISUS is made of five differently designed spherical satellites. They are colored by means of clear acrylic rods which can be lighted as neon tubes. The sculpture as a whole rotates on its axis every seven minutes and has a fifty-minute electronic cycle of light and sound which, due to its intricacy, makes the program's repetition almost imperceptible.

(2) Various technologies deriving from recent scientific strides have led to completely new means of expression, such as infography, electrography, video-art, etc.

Along that line Bonacic and his BCD group have introduced the concept of dynamic object, in which impregnable unity is established between the computer system and a work of art. The BCD group realized a cybernetic sculpture generating new visual and tone patterns of a Galois field. Its audio-visual pattern is repeated after the lapse of 24 days, 6 hours, 32 minutes, 32 seconds, with a rhythmic change occurring about every 2 seconds. This sculpture has been introduced by the Franciscans in their preparation for meditation at St. Kilian's Church, Wiesbaden, German Federal Republic.

(3) The aesthetic dimension of science is also an interesting aspect of the relations between sciences and arts. Marcel Bessis [4] has made photographs of blood red cells which can compete with good abstract paintings.

(4) Some artists start from scientific theories or scientific concepts to produce artistic works. In other words sciences can inspire artists. Art inspired by microphysics and molecular art are good examples of such a behaviour.

2. Art Inspired by Microphysics

Bettina Brendel [5] realized six joint paintings on the theme "Particles or Waves?" On these paintings there are many thin short lines. Bettina Brendel says that these

Jean Maruani (ed.), Molecules in Physics, Chemistry, and Biology, Vol. I, 233–236.

lines represent a fundamental constant in nature: "the energy radius of a moving particule". The paintings also evoke components of apparatus that one might find in an experiment using optical or particle beams. Thus the title "Particles or Waves?" may be taken to refer not only to the formation of wave-like patterns from particle-like components but also to the similarity of the tools that science uses to manipulate both types of matter.

Nicole D'Agaggio [6] realized a large fresco (220 × 700 cm) evoking the three main principles of Wave Mechanics. In the background there appear the columns of a temple erected to the glory of Wave Mechanics. The painting is divided into three parts and therefore constitutes a triptych. The first panel evokes the first principle, which associates an operator to each particle property. Each operator is represented by the shadow of the corresponding property. In the central panel the famous wave equation is enthroned in the shadow of a Neptunian God. The third panel evokes the third principle of Quantum Mechanics. All the painting is crossed by light beams, like in the classical Young experiment, to recall that Wave Mechanics was born out of a deep meditation on the evolution of the theory of light.

3. Molecular Art

In certain cases fortuitous relations appear between the concept of molecule and works of art.

Our colleague Professor Lehn has shown that the shapes of some of the sculptures made by Arp or by Moore present remarkable analogies with the shapes of some supermolecules.

The philosopher Bernard Dahhan [7] considers that the artistical production of Vasarely contains a molecular art. There is effectively, for example, a certain analogy between a painting like Tridim A and the notation used by Perutz to represent a macromolecule like haemoglobin.

But in other cases artists effectively use molecular concepts to build their artworks.

The work of Salvador Dali inspired by the DNA molecule is well known.

Nicole D'Agaggio is responsible for the most beautiful illustrations of my book: Molecular Vision of the World [8, 9]. Many of them are concerned with molecules. One is dedicated to the origin of Life on the Earth. It evokes the formation of the first "living" molecules during the "marriage" of the air and sea under the effect of solar radiations and lightning. Her work "Molecular Coupling" describes the bonding between the molecules which form the letters of the genetic message contained in every one of our cells (Figure 1). The transparencies evoke the vacuity of the molecular world and hint at a prelude to Life [10]. The last plate of the book "Molecules, Supports for Thought" allows a glimpse at a strange swarming of molecules transmitting the nervous impulse between two neurons.

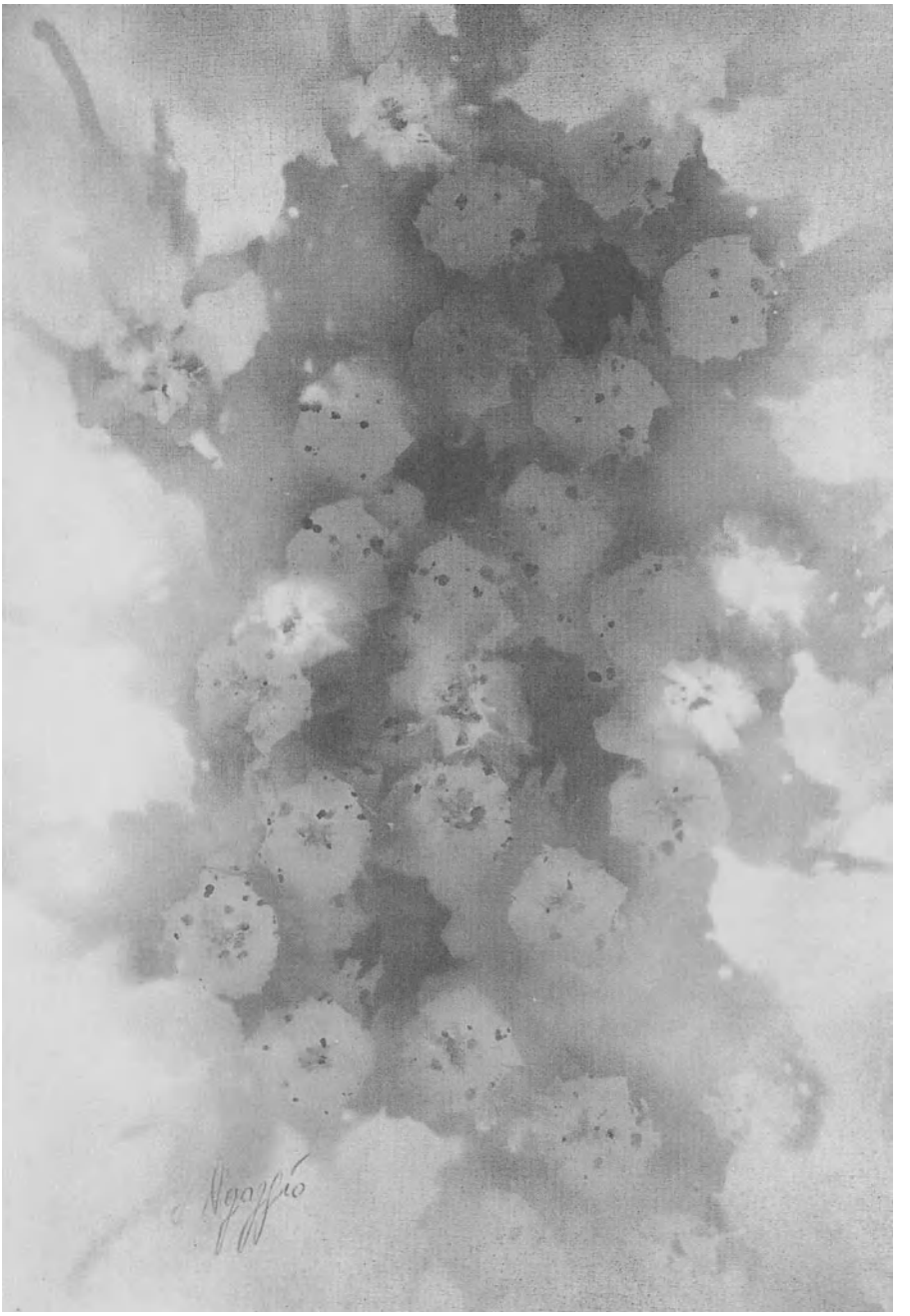


Fig. 1. Molecular Coupling (A painting by D'Agaggio).

From this work surges forth a distressing, diabolical, ambiguous, Manichean atmosphere, as if these molecules were already transporting good and evil.

By painting these canvases the artist dared to take some distance from the scientist ideal in order to make a place for mystery. She has constructed an aesthetic discourse parallel but not identical to that of science.

Nicole D'Agaggio has also painted a vast canvas dedicated to the chemistry of love. It is a triptych. Each panel evokes one of the phases of love: attraction, attachment and ecstasy, disenchantment. Each of these phases must be read on three levels. The backgrounds express the different moods of the soul. On each panel there appear the various molecules causing feeling by linking neurons. For example, during ecstasy there appear the molecules causing pleasure, represented by golden spots. And the famous circuits of pleasure and anguish are evoked on the frontispiece of this remarkable work.

Various artists have used molecular electronic isodensity curves in their artworks.

Hager has made a painting based on the isoelectronic curves of insulin.

Iera found a certain voluptuousness in the electronic isodensity maps concerning small molecules. He used such maps to underline the sensuality of the body of a woman.

References

1. R. Daudel: *Revues des Questions Scientifiques* **154**, 3 (1983).
2. N. Lemaire D'Agaggio: Thesis, Sorbonne, Paris (1986).
3. S. Boettger: *Vital Signs: The Sculpture of Milton Komisar*, San José Museum of Art Pub. (1983).
4. M. Bessis: *Corpuscules*, Springer International Pub. (1983).
5. B. Brendel: *Paintings*, Printed by S. and P. Printing (1983).
6. R. Daudel, J. A. França, A. Glibota, R. Huyghe, C. Langlois, A. Lanoux, P. Roberts-Jones and S. J. Wang: *D'Agaggio, La Volupté dans l'Hyperabstraction*, Paris Art Center Pub. (1985).
7. B. Dahhan: *Vaserely, Connaissance d'un Art Moléculaire*, Denoël Gonthier Pub. (1979).
8. R. Daudel and N. D'Agaggio: *Vision Moléculaire du Monde*, C.N.R.S. et Hachette Pub. (1981).
9. N. Lemaire D'Agaggio: *Leonardo* **15** (1982).
10. See also M. Alfert: *Leonardo* **19** (1986).

Index of Names

Index of Subjects