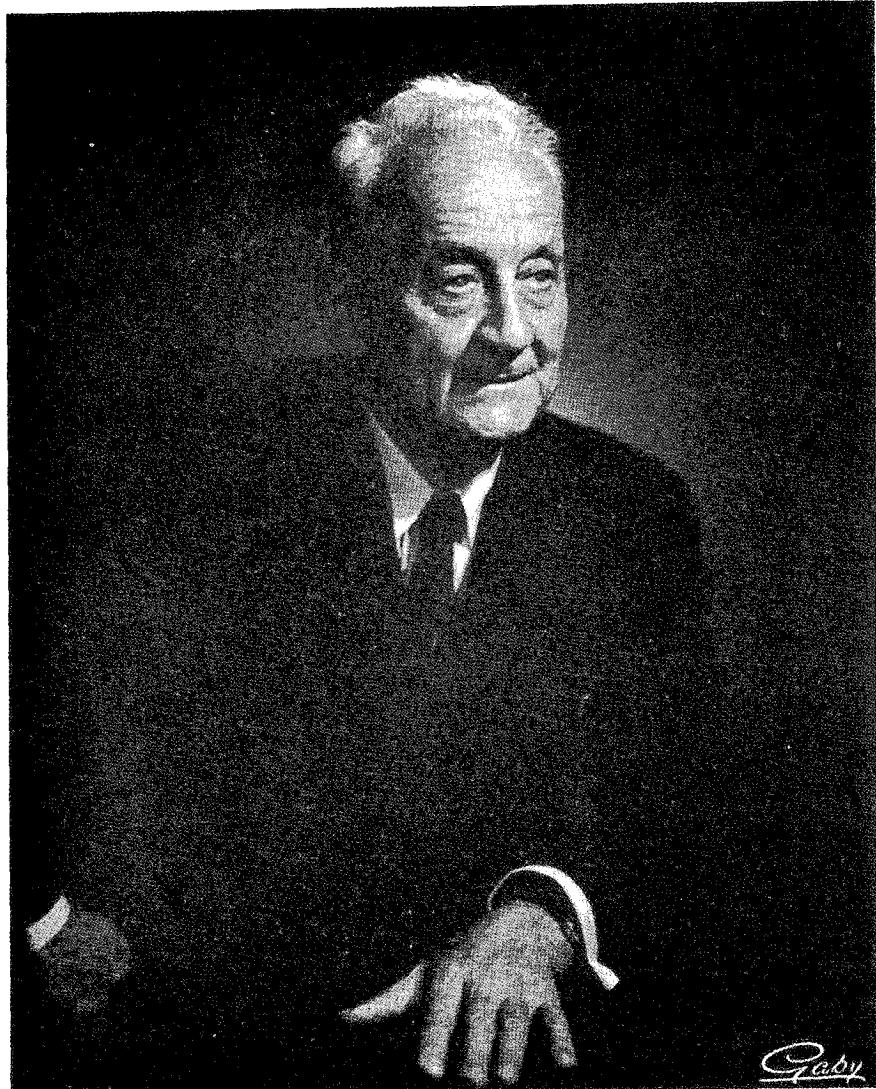


HORIZONS IN BIOCHEMISTRY

Albert Szent-Györgyi Dedicatory Volume

WITHDRAWN



Copyright by GABY-Montreal, Canada

Albert Szent-Györgyi

HORIZONS IN BIOCHEMISTRY

Albert Szent-Györgyi Dedicatory Volume

Edited by

Michael Kasha
*Institute of Molecular Biophysics
Department of Chemistry
Florida State University
Tallahassee, Florida*

Bernard Pullman
*Institut de Biologie
Physico-Chimique
Université de Paris
Paris, France*



Academic Press · New York · London · 1962

COPYRIGHT © 1962, BY ACADEMIC PRESS INC.

ALL RIGHTS RESERVED

NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM
BY PHOTOSTAT, MICROFILM, OR ANY OTHER MEANS,
WITHOUT WRITTEN PERMISSION FROM THE PUBLISHERS.

ACADEMIC PRESS INC.
111 FIFTH AVENUE
NEW YORK 3, N. Y.

United Kingdom Edition
Published by
ACADEMIC PRESS INC. (LONDON) LTD.
Berkeley Square House, London, W.1

Library of Congress Catalog Card Number 62-18226

First Printing, 1962
Second Printing, 1963

PRINTED IN THE UNITED STATES OF AMERICA

Contributors

J. D. BERNAL, *Birkbeck College, University of London, London, England*

L. BRILLOUIN, *Columbia University, New York, New York*

MELVIN CALVIN, *Department of Chemistry and Lawrence Radiation Laboratory, University of California, Berkeley, California*

BARRY COMMONER, *The Henry Shaw School of Botany and The Adolphus Busch III Laboratory of Molecular Biology, Washington University, St. Louis, Missouri*

JULES DUCHESNE, *University of Liège, Liège, Belgium*

D. D. ELEY, *Department of Chemistry, Nottingham University, Nottingham, England*

HENRY EYRING, *Departments of Chemistry and Experimental Biology, University of Utah, Salt Lake City, Utah*

S. FLEISCHER, *Institute for Enzyme Research, University of Wisconsin, Madison, Wisconsin*

EARL FRIEDEN, *Department of Chemistry, Florida State University, Tallahassee, Florida*

HANS GAFFRON, *Fels Fund, Department of Biological Sciences and the Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida*

D. E. GREEN, *Institute for Enzyme Research, University of Wisconsin, Madison, Wisconsin*

CHARLES HUGGINS, *The Ben May Laboratory for Cancer Research, University of Chicago, Chicago, Illinois*

VERNON M. INGRAM, *Division of Biochemistry, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts*

LEON L. JONES,¹ *Departments of Chemistry and Experimental Biology, University of Utah, Salt Lake City, Utah*

¹ Present Address: Department of Chemistry, University of Wisconsin, Madison, Wisconsin.

CONTRIBUTORS

HERMAN M. KALCKAR,¹ *McCollum-Pratt Institute and the Department of Biology, The Johns Hopkins University, Baltimore, Maryland*

MICHAEL KASHA, *Institute of Molecular Biophysics, Department of Chemistry, Florida State University, Tallahassee, Florida*

IRVING M. KLOTZ, *Department of Chemistry, Northwestern University, Evanston, Illinois*

ARTHUR KORNBERG, *Department of Biochemistry, Stanford University School of Medicine, Palo Alto, California*

D. E. KOSHLAND, JR., *Biology Department, Brookhaven National Laboratory, Upton, New York*

H. A. KREBS, *Medical Research Council Unit for Research in Cell Metabolism, Department of Biochemistry, University of Oxford, Oxford, England*

ALBERT L. LEHNINGER, *Department of Physiological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, Maryland*

W. D. McELROY, *Department of Biology and McCollum-Pratt Institute, Johns Hopkins University, Baltimore, Maryland*

SEVERO OCHOA, *Department of Biochemistry, New York University School of Medicine, New York, New York*

LINUS PAULING, *Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California*

JOHN R. PLATT, *Department of Physics, University of Chicago, Chicago, Illinois*

ALBERTE PULLMAN, *Institut de Biologie Physico-chimique, Paris, France*

BERNARD PULLMAN, *Institut de Biologie Physico-chimique, Paris, France*

ALEXANDER RICH, *Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts*

FRANCIS O. SCHMITT, *Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts*

H. H. SELIGER, *Department of Biology and McCollum-Pratt Institute, Johns Hopkins University, Baltimore, Maryland*

¹ Present Address: Biochemical Research Laboratory, Harvard Medical School and Massachusetts General Hospital, Boston, Massachusetts.

CONTRIBUTORS

JOHN D. SPIKES, *Departments of Chemistry and Experimental Biology, University of Utah, Salt Lake City, Utah*

GEORGE WALD, *The Biological Laboratories, Harvard University, Cambridge, Massachusetts*

RENÉ WURMSER, *Institut de Biologie Physico-chimique, Paris, France*

EMILE ZUCKERKANDL, *Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California*

Preface

When invitations were sent out to various friends and colleagues of Albert Szent-Györgyi for contributions to a Dedicatory Volume, the responses were enthusiastically and wholeheartedly in favor of the project, for all of us have felt the stimulating influence of his fruitful research endeavors.

Since Albert Szent-Györgyi has himself projected so much interest in searching for new biochemical mechanisms, we felt it appropriate that a collection of papers in such a volume should not consist merely of reviews and summaries of past ideas and results. Therefore, we particularly stressed in our invitation that each contributor should take this opportunity to express new viewpoints, new analyses, or philosophical viewpoints projected toward the future of modern biochemistry and molecular biology. Our efforts were fully rewarded by the responsive cooperation of the contributors.

The Editors are proud to present this volume as a dedication to Albert Szent-Györgyi. This unique collection of original essays by 28 contributors is a fitting token of our friendship and esteem for him. As Editors, we have enjoyed the double pleasure of presenting this collection to Albert Szent-Györgyi, and also preparing a delightful feast for interested researchers and students of the future of this science.

*Tallahassee, Florida
April 12, 1962*

M. KASHA
B. PULLMAN

Contents

CONTRIBUTORS	v
PREFACE	vii
Albert Szent-Györgyi and Modern Biochemistry	1
RENÉ WURMSER	
<i>Biochemical Evolution</i>	
Biochemical Evolution	11
J. D. BERNAL	
Evolutionary Possibilities for Photosynthesis and Quantum Conversion	23
MELVIN CALVIN	
On Dating Stages in Photochemical Evolution	59
HANS GAFFRON	
Origin and Evolution of Bioluminescence	91
W. D. McELROY AND H. H. SELIGER	

On the Problems of Evolution and Biochemical Information Transfer ALEXANDER RICH	103
Life in the Second and Third Periods; or Why Phosphorous and Sulfur for High-Energy Bonds? GEORGE WALD	127
 <i>Molecular Genetics</i>	
A Biochemical Approach to Genetics VERNON M. INGRAM	145
Enzymatic Mechanisms in the Transmission of Genetic Information SEVERO OCHOA	153
A "Book Model" of Genetic Information—Transfer in Cells and Tissues JOHN R. PLATT	167
Molecular Disease, Evolution, and Genic Heterogeneity EMILE ZUCKERKANDL AND LINUS PAULING	189
 <i>Biochemical Catalysis</i>	
The Significance of Absolute Configuration in Optical Rotation and in Catalysis HENRY EYRING, LEON L. JONES, AND JOHN D. SPIKES	229

On the Metabolic Significance of Phosphorolytic and Pyrophosphorolytic Reactions ARTHUR KORNBERG	251
Catalysis in Life and in the Test Tube D. E. KOSHLAND, JR.	265
Enzyme Activity and Cellular Structure H. A. KREBS	285
 <i>Molecular Organization</i>	
Giant Molecules and Semiconductors L. BRILLOUIN	295
Is DNA a Self-Duplicating Molecule? BARRY COMMONER	319
Remarks on Some Physical Properties of Nucleic Acids JULES DUCHESNE	335
Semiconductivity in Biological Molecules D. D. ELEY	341
On the Molecular Organization of Biological Transducing Systems D. E. GREEN AND S. FLEISCHER	381
Respiration-Linked Mechanochemical Changes in Mitochondria ALBERT L. LEHNINGER	421

Psychophysics Considered at the Molecular and Submolecular Levels FRANCIS O. SCHMITT	437
 <i>Biochemical Molecular Structure</i>	
The Complex Copper of Nature EARL FRIEDEN	461
Cancer and Necrosis Induced Selectively by Hydrocarbons CHARLES HUGGINS	497
Some Perspectives of Antispace Research in Biology HERMAN M. KALCKAR	513
Water IRVING M. KLOTZ	523
 <i>Quantum Biochemistry</i>	
From Quantum Chemistry to Quantum Biochemistry ALBERTE PULLMAN AND BERNARD PULLMAN	553
Quantum Chemistry in Molecular Biology MICHAEL KASHA	583
INDEX	601

Albert Szent-Györgyi and Modern Biochemistry

RENÉ WURMSER

Institut de Biologie Physico-chimique, Paris, France

As a contemporary of Albert Szent-Györgyi, it was possible for me to follow, in the atmosphere that reigned at the moment, so to speak from day to day, his so important contribution to biochemistry between the two wars and subsequently at the advent of modern biochemistry.

Szent-Györgyi has entered research, as a student of medicine, in the laboratory of his uncle, the histologist Michael Lenhossec. This was but a family tradition, and he was soon "dissatisfied with what dead tissues could tell him." World War I and events which followed in Hungary did not, however, permit him for several years to work as he liked in his chosen field. It was only after a wandering life spent between universities in Poland, Germany, and Holland that he could at last accomplish his first researches in biochemistry in the laboratory of Hamburger, at Groning.

This was the period when, following on the researches of Hill, Meyerhof had begun his work on muscle chemistry, studies which were to lead subsequently to a thorough understanding of glycolysis and the discovery of the role of phosphate bonds. But the most disputed question at that time concerned the mechanism of tissue oxidations. Two theories were in opposition, one of Warburg and the other of Wieland. Was the catalytic process determined essentially by the activation of hydrogen by Thunberg's dehydrogenases, or by the activation of oxygen by Warburg's *Atmungsferment*? The question seems strange today, habituated as we are to think that the two activations are both necessary. It was Szent-Györgyi who established this in a most convincing manner (Banga and Szent-Györgyi, 1925). The essential features of biological oxidation were shortly afterwards elucidated. While Keilin showed what was the *Atmungsferment*, namely a catalyst acting on cytochromes, Szent-Györgyi turned his attention to the intermediary transporters of hydrogen. After having rediscovered von Euler's cozymase quite independ-

ently, he discovered the "cytoflav" (Göszy and Szent-Györgyi, 1934), whose constitution was to be worked out later by Kuhn. His researches on vegetable tissues led Szent-Györgyi finally to the isolation and characterization of hexuronic acid—later ascorbic acid (Szent-Györgyi, 1924; Svirbely and Szent-Györgyi, 1932)—which proved to be none other than vitamin C (Szent-Györgyi, 1927).

This last work, which earned for Szent-Györgyi the Nobel Prize, was begun in a small basement of Groning and continued afterwards at Cambridge where Hopkins had invited him to his laboratory, a veritable Mecca for biochemists at that time. Szent-Györgyi returned to Hungary in 1930, and it was then, at the Institute of Medical Chemistry of the University of Szeged, that, while completing his studies on ascorbic acid, thanks to the paprika of his country, he discovered, with a brilliant team of students, the catalytic effect of fumaric acid on tissue respiration (Szent-Györgyi, 1928, 1937a).

The respiration of minced pigeon-breast muscle, initially very active, decreases with time, but it can be restored to its original level by addition of minimal amounts of succinate or fumarate. This is no doubt a catalytic process, since there exists no stoichiometric relation between the added dicarboxylic acid and the resulting oxygen consumption. This was indeed a new idea. The catalyst was not an enzyme but an enzyme-metabolite pair. In the proposed scheme, the hydrogen from a donor, for example a carbohydrate, reduced a first dicarboxylic acid, the oxaloacetic acid; the resulting malic acid reduced fumaric acid, the succinic acid thus produced transferred in its turn its hydrogen to cytochromes. This fundamental concept has been modified and completed since the discovery by Krebs of the role of citric acid. But we can say, like Baldwin in his "Dynamic Aspects of Biochemistry" (1948) that what we know about aerobic metabolism is founded on this study by Szent-Györgyi of the minced pigeon-breast muscle.

Respiration is a very spectacular phenomenon. Processes taking place in the absence of free oxygen are nevertheless of the utmost importance, even in the case of aerobic beings. Syntheses are nothing but oxidoreductions interwoven with condensations and hydrolyses. It was in order to understand their coupling, be it mentioned in passing, that I devoted so many years to determining the redox-reduction potentials of metabolite systems. Free oxygen plays a role

only at the last stage in the journey of hydrogen across the successive transporter systems. In all the preceding steps free energy is available to compensate for the endergonic process. Szent-Györgyi has insisted on the economy of this progressive oxidation. The cell would have been at a loss to utilize the large quantity of energy corresponding to the combustion of hydrogen in foodstuffs, had the latter been immediate. It pays for its functioning in small coin and thus suffers no loss (Szent-Györgyi, 1937b).

The direct demonstration of rapid hydrogen exchange between different metabolites is interesting from another point of view. This hydrogen transfer proceeding by random exchange between many donors and acceptors corresponds to one of those levels of intracellular oxidoreduction to which, for my part, I have attributed a definite role in the chemical organization of the cell.

Biochemistry, between the two wars, has acquired much information on the interdependence of the manifold metabolic reactions, on the enzymes and coenzymes which intervene and on the structure of these catalysts. The role played by phosphorylated compounds in the energetic coupling of reactions has, in particular, been studied in detail; much light has been shed on biosynthesis at the stage of combination between small molecules. A domain left practically unexplored so far was going to be the favored field of modern biochemistry.

Toward 1938 Szent-Györgyi undertook a series of researches best described, in my opinion, by the term molecular physiology. These researches soon led him to what he himself called submolecular biology.

The expression molecular physiology is used since the aim was to find out, in the same molecule, in this case the molecule of a muscle protein, a physiological effect and at the same time its origin: in other words, to elucidate certain typical phenomenon of life, such as muscle contraction, by the structure of certain specific molecules. Whether this is at all possible is by no means certain. In the words of Szent-Györgyi (1960) "one of the basic principles of life is organization by which we mean that if two things are put together something new is born, the qualities of which are not additive and cannot be expressed in terms of the quality of the constituents." It is quite likely, however, that if a certain kind of molecule possesses a rudiment of function prefiguring a physiological act, a part of the explanation should indeed be residing in the structure of the molecule.

This is what Szent-Györgyi (1947) had shown through his studies on muscle proteins.

Since the discovery of Engelhardt, Ljubimova, and Meitina, filaments of what was then called myosin were known to be extensible in the presence of adenosine triphosphate (ATP). Szent-Györgyi and his group succeeded in separating this substance into two proteins, myosin properly speaking and actin, the union of the two resulting in actomyosin. It is this complex which shows contractile properties. In solution it responds to ATP by reducing its viscosity. In the form of a continuous gel, ATP produces a shrinking. This modification prefigures the contraction of living muscle.

Differences no doubt exist between the two phenomena. Extraction procedures are bound more or less to alter the proteins. In any case, the conditions in the filament are not the same as in muscle fibrils. Short of believing, as do the partisans of the principle of generalized complementarity, that biological complexities are inextricable, we can foresee that they should be understood one day in terms of intermolecular forces. That day has not yet arrived. But one can already begin to attack the problem of relation between the individual properties of a molecule and its structure. This was the path that Szent-Györgyi engaged upon.

Quantitative data show that the effect produced by the breakdown of a small molecule of ATP on the whole of a large molecule of myosin raises a problem of transmission: the reaction has a too large a sphere of influence to be explained in terms of Coulomb forces; whence the idea that there exist electronic conduction bands in macromolecules. Szent-Györgyi extended this concept of semiconductivity to other biological problems. This could explain those enzymatic reactions for which one is led to exclude a simple collision mechanism and where a certain organization in space seems likely (Szent-Györgyi, 1941).

The moment was ripe. In science, in order that something tangible be accomplished, going beyond speculation, creative imagination must find the necessary support in the general movement of ideas and techniques. This is what was to happen in this case. A considerable amount of work based on the hypothesis of semiconductivity is coming out. True as it is that conduction bands have until now been demonstrated only in certain special systems excited by high energy quanta, the study of biochemistry on a submolecular level nevertheless received an impetus. Szent-Györgyi himself has turned

his attention to another type of electron transfer. When two suitable molecules come into intimate contact, the electronic orbitals of each of them may overlap and a charge-transfer complex can result therefrom. An electron passes from the highest filled orbital of the donor to the lowest empty orbital of the acceptor. The donor turned acceptor and the acceptor turned donor can constitute a chain for the transfer of electrons between molecules separated from each other. In any event, the study of these charge-transfer complexes in enzymatic catalysis is rich in promise since there is intimate contact between the enzyme and the substrate.

It is doubtful if enzyme action could be entirely explained without taking into account rearrangements in the disposition of atoms. The idea that a kind of reversible denaturation is concerned in the catalytic process is supported by data on entropy change corresponding to the formation of an enzyme-substrate complex and its activation. But these structural changes are as much dependent on electronic events as are the changes of reactivity by charge transfer. Anyway among the most significant consequences of the present researches of the Szent-Györgyi school will be their deliberate insistence on treating problems, theoretically and experimentally, on the electronic level. How is one to conceive the future of biology if not as a theoretical biochemistry, which, based on quantum mechanics, will seek to elucidate the properties of life from the intra- and intermolecular infrastructure?

Demonstration of charge-transfer through spectral evidence and through electron-spin resonance and their interpretation based on the calculations of the Pullmans have already led to the explanation of a number of biological properties.

In the fortys, during the period when Szent-Györgyi undertook his direct studies on submolecular physiology, was born what has been called molecular biology. Right or wrong, tradition makes a distinction between the study of the functioning of organisms, physiology, and the study of the formation of the same organisms, which appertains to biology. In view of the progress in the knowledge of protein structure and the obtaining of paracrystals of virus which could perhaps be considered as the simplest form of life, some authors have attempted mechanistic explanation of the formation of specific constituents of organisms, if not of the organisms themselves. The main difficulty in a protein biosynthesis scheme is the improbability of a given sequence of amino acids, not because of

the resulting decrease in entropy, metabolism can take care of it, but since it is necessary to imagine a process of selection of these amino acids. This therefore is a problem of intermolecular attractions. Pauling has, since 1940, discussed the nature of the forces in action. He concluded that these were forces brought into play between complementarity surfaces by virtue of their forms and their charges.

Molecular biology has truly derived benefit from decisive progress in genetics. On the one hand, there has been an explanation of the effect of the gene by its capacity to synthesize a particular chemical species, and on the other hand the proof that the support of information controlling the biosynthesis of specific substances resides in nucleic acids. The current concept is that a particular amino acid associates itself to a complementary sequence of nucleotides in the nucleic acid chain.

In the case of antigen-antibody and enzyme-substrate combinations, it has been possible to calculate the magnitude of intermolecular forces and to verify the results experimentally. A similar treatment may conceivably be applied to a system amino acid-nucleotide. This evidently implies the knowledge of charge distribution, of polarizability, and of ionization potential.

Two currents of research, one emanating from the study of muscular contraction and enzymatic activation, the other from the study of specific protein biosynthesis, thus led to common problems and techniques considered not long ago as being the domain of physicists alone. It is the beginning of a story that we foresee should be long. As Szent-Györgyi (1960) says in his last book, "in order to approach the central problems of biology, we have to extend our thinking in two specific directions, into both the sub- and the supramolecular." Some begin to look to solid state physics. But the extremely simple arrangement of systems treated therein contrasts sharply with the strange constructions of life.

Szent-Györgyi has insisted on this gap which remains to be filled up. But his enthusiasm and faith in the future kept him in an optimistic spirit. His last book, so typical of this viewpoint, exerts an incontestable influence on young biochemists, whose number increases from day to day on the new borderlines of biology and physics.

REFERENCES

- Baldwin, E. (1948). "Dynamic Aspects of Biochemistry." Cambridge Univ. Press, London and New York.

- Banga, I., and Szent-Györgyi, A. (1925). *Biochem. Z.* **157**, 50.
 Gözsy, B., and Szent-Györgyi, A. (1934). *Biochem. Z.* **224**, 1.
 Svirbely, J. L., and Szent-Györgyi, A. (1932). *Biochem. J.* **26**, 865.
 Szent-Györgyi, A. (1924). *Biochem. Z.* **150**, 195.
 Szent-Györgyi, A. (1927). *Nature* **129**, 782.
 Szent-Györgyi, A. (1928). *Biochem. J.* **22**, 1387.
 Szent-Györgyi, A. (1937a). "Studies on Biological Oxidations and Some of Its Catalysts." J. A. Barth, Leipzig.
 Szent-Györgyi, A. (1937b). "Über den Mechanismus der biologischen Verbrennungen." Nobel Vertrag.
 Szent-Györgyi, A. (1941). *Science* **93**, 609.
 Szent-Györgyi, A. (1947). "Chemistry of Muscular Contraction." Academic Press, New York.
 Szent-Györgyi, A. (1960). "Introduction to a Submolecular Biology." Academic Press, New York.

Biochemical Evolution

Biochemical Evolution

J. D. BERNAL

Birkbeck College, University of London, London, England

It was in the early nineteenth century that the world of chemistry was divided into two parts, the mineral chemistry which was amenable to the ordinary operations of the laboratory, of furnaces, oxidations and reductions, and the chemistry of animal and vegetable products. The origin of these *organic* chemicals was deemed to imply the action of some *vital principle* whose main characteristics were the complexity and instability of compounds which it produced by means of fermentations and putrefactions.

Since Wöhler's discovery of the synthesis of urea in 1828, this view of the action of a vital principle gradually faded. There were, it is true, partial setbacks. Pasteur made the most remarkable and significant of these in establishing the asymmetric nature of certain molecules which he maintained could only be produced by vital processes. This optical activity has remained a characteristic of biochemical molecules almost to this day for, although Pasteur himself had separated right- and left-handed molecules in the laboratory, he recognized that he was an organism himself and, therefore, this resolution might be attributed to vital forces as well.

Nevertheless, the general view was that there was nothing in organic chemistry than a more complicated version of inorganic chemistry and if we pursued the methods of inorganic chemistry long enough into the field of the carbon and nitrogen compounds, the whole matter would become plain.

Now we are beginning to go back to some extent to the older view when we consider the actual processes of biochemistry. Some of the compounds have extraordinarily complicated molecules but now that they can be recognized and purified, it is evident that in life they are being synthesized in a perfectly regular and ordered way. Moreover, these syntheses are carried out rapidly in a tiny space at ordinary temperatures without any of the elaborations of apparatus which the chemists of today or yesterday found necessary. But it is above all the discoveries of the last 5 years, particularly that of the role of

nucleic acid in protein and especially in enzyme synthesis, and those arising from the ultra histological observations of the interior of cells, that have altered our attitude toward biological chemistry.

We realize now that at the start of the subject, when Hopkins was investigating the essential metabolites, the new biochemists were only touching the fringe of the subject. They had, indeed, first to recognize the extremely simple molecules which were operated as the components for assembly in this elaborate production chain. And when we could see something of the process in its entirety, the cyclically linked reactions, for instance, in fermentation or oxidation or photosynthesis, we came to appreciate that the whole metabolism was a construct of a dynamical kind and that it was of an altogether different order of complexity from anything man had ever conceived of in his chemical operations. What, in a sense, was even more surprising, was that taking biochemistry as a whole, covering all species of animals and plants and extending from bacteria to viruses, there was an extraordinary unity and economy. The same chemical reactions, the same structures down to atomic details, occur over and over again. Even where there is variation, it is variation on a theme: the theme, for instance, of the porphyrins is utilized in respiratory ferments, in photosynthesis, and in the oxygen carrying of the higher animals.

All this is leading us to a belief which is rapidly taking form—indeed, in some sense the main new idea of the Fifth Biochemical Congress in Moscow—that there is an underlying *unity* in biochemistry which implies a *biochemical evolution* much more elaborate and much earlier than the biological evolution that gave us all the various forms, performances, and behaviors of the plants and animals of today.

The study of the origin of life, which began a quarter of a century ago as an interesting speculation needed to link up the development of the earth with that of the living forms on it, has now been recognized almost officially as an essential aspect of biochemistry itself, one introducing a temporal and historical element into what had hitherto been a description of contemporary structures and reactions.

The earlier studies of the origin of life concentrated on establishing a case for it, that is, on showing that life could have originated from inorganic substances on the surface of the earth. We now see that the case no longer needs to be made: it can be accepted. Moreover, we see that there are many ways in which the process could

have started. What interests us now is not *that* it could happen but precisely *how* it happened.

Here the evidence from speculation or even from the reasonable use of geochemistry and astrophysics proves quite insufficient. The actual track of biochemical evolution can only be discovered by following it out from current biochemistry. This necessarily implies an assumption for which, however, there is very considerable circumstantial evidence. We must assume that by and large the chemical substances and the chemical reactions that are found today widespread among all living organisms, plants as well as animals, do repeat—if with variations—the reactions and substances that existed at the very beginning of life or, strictly, before any organisms as such existed.

The implication is that there has been, all the way through, a definite process of reproduction, not of an organism or even a cell, but what seems at first sight a much simpler thing, let us say a glucose molecule. There is no doubt that this reproduction is taking place now and has been as far back as we can trace in geological time. Wherever the appropriate modern analytic methods are used—as, for instance, by Abelson—we find in the rocks the very same molecules that are now the cornerstones of modern biochemistry, particularly amino acids, purines, and porphyrins. But when we come to consider the present-day biosynthesis of these molecules, how they are actually formed in organisms, we see that we are up against an apparent contradiction. The actual formation in an organism of such simple molecules, or perhaps even more of the simplest molecules of all, like hydrogen or carbon dioxide, requires complete interlinked cycles of reactions. Moreover, for this production, numerous enzymes, nucleic acids, and other far more complicated molecules have to be involved.

If we did not know better, we would postulate from this the eternity of life, not necessarily life as we know it, but life as the biochemists know it: life with all its enzymes, nucleic acids, and interwoven cycles. But, logically, this can only be true back to a certain point in time and what that point is, is what we have to determine. It may be one thousand million, it may be even three thousand million years ago. From then on, the whole mechanism—as I have said, with variations—has continued to churn out the same kinds of molecules. But before that it could not have been done because the mechanism itself did not exist. If it were really true, as the tendency

of modern biochemistry seems to indicate, that the *reproduction* of vital processes is through the mechanism of nucleic acid and protein synthesis, then we must admit that the nucleic acid itself, its component nucleotides formed from sugar, phosphate, and purine or pyrimidine, must have first arisen by a mechanism which did not involve their own pre-existence and, therefore, was not strictly reproduction.

It is very natural for a biochemist considering the question to start at this point, to take the nucleic acid-protein mechanism as given and to consider the kind of reactions and developments that may have occurred since the basic nucleic acid-protein mechanism occurred. Where I think they are wrong is in making excessive claims that this was the beginning of life. It may perhaps be taken as a *definition* of life but it cannot be that of its biochemical beginning because the compounds needed for this process, on account of their complexity, themselves require an anterior evolution.

This evolution must have been different in kind from subsequent biochemical evolution. It must have been one in which molecular reproduction did not occur in any exact sense. For instance, the extraordinary precision with which such a molecule as hemoglobin is put together today, so that the detachment of a single amino acid is a definitely detectable variation, cannot have existed for all time. It may have come about, however, in the first place without any abrupt transition. There may have been bad copying before perfect copying and before that no copying at all, just a statistical production of molecules which gradually sorted themselves out into the arrangements that we see today.

The appearance of reproduction thus makes a convenient landmark in studying biochemical evolution. We might divide it into the pre-nucleic acid stage and the post-nucleic acid stage. We can if we like use this watershed to make the first appearance of life and define *life* as anything that contains the nucleic acid-protein mechanism and non-life, or pre-life, as anything that does not. But I would prefer to call the region between "restricted organic" synthesis produced by the action of physical agents in space or on the earth, and the point at which the nucleic acid-protein mechanism first worked, a kind of intermediate stage or pre-life—what Pirie has called the "eobiotic stage." This region remains the most uncertain and inaccessible of the stages leading to life as we know it today. Although it is much too early to attempt a reconstruction of it, it may still be

valuable to give some idea of what might have happened, simply to guide the research and find out what actually did happen. It will be as important to prove false what I am suggesting as to prove it true.

From the purely biochemical point of view we can see something of the division itself. The biochemistry of organisms of today, in fact what we call biochemistry as such, deals with a series of transformations, in the first place of simple into complex chemical substances and back again into simple substances, all following regular paths, the anabolic and catabolic parts of metabolism. But it also deals with the change of these patterns, the historic evolution that occurs and often occurs quite rapidly under the influence of change of environment. Here we have biochemistry coupled with biological evolution, and very ingenious mechanisms are developed to assist that coupling, for instance, the mechanisms in immunology which enable an organism to copy or rather copy the opposite or antigen of a molecule within a very short period of time.

Looked at from the biochemical and ultra histological point of view, what is most striking is not the secondary biochemical evolution but the extraordinary stability of the parts which are concerned not only with the chemical part of life but also with its physical part. Such a structure as muscle, for instance, once evolved at some very early stage in coelenterates, persists right through to man, while cellulose, already evolved in bacterial structures, persists right through the vegetable kingdom on one side also continues to exist in the animal kingdom not only in the tunicates but even, as recently shown, in all higher chordates including man himself. This structural conservatism reaches down into the very components of cells.

The peculiar structure of the base of cilia, with its trebled rings bearing nine outer and two inner fibers, seems to be formed of every kind of cellular organism apart from the bacteria. When it does not occur in some higher plants, there is sufficient evolutionary evidence that this is a loss of function. One of the amazing discoveries of the electron microscope is that the centrioles of cells have precisely the same structure. It is a standardized component whose very complexity implies a history but, as far as we know now, its predecessors are not to be found. Yet anything as complicated as that must have simpler predecessors. It has recently been shown (Fraser *et al.*, 1962) that precisely this nine plus two structure on a smaller scale is the basis of mammalian hair structure.

The reproduction of the centrioles further shows an element of reproduction which previously we have not been able to see except possibly in the development of viruses, namely, reproduction not by splitting but by outgrowth on a molecular scale, for the new centriole is produced apparently by a growth sideways and at right angles to the old one, by some kind of pattern-making device operating on a scale intermediate between that of the molecular and the organismal level.

Later we will be able to analyze and to divide these structures and processes and show how the development of an organism can be reduced to what might be called "operating building rules" for standardized components. A very fine example is that given by T2 phage with its polygonal head and its retractile nose both constructed of apparently identical or very similar protein molecules in finite numbers, fifteen on each side of the head and thirteen around the tail. It would seem almost as if there existed somewhere a program like that of a computing machine, just a set of construction orders for putting these elementary bricks together.

But the bricks themselves obviously must be older than the structures that are built from them and it would seem that the organismal part of life is really subsequent to the evolution of its components, the protein molecules themselves.

How are we to conceive such molecular or chemical evolution? Here, because they are generalized and logical, the same principles can be applied that apply to the evolution of organisms themselves. The genius of Darwin has already given us the key in form of the idea of *natural selection*. Before there was competition between organisms there was competition between molecules and reaction networks, as first put forward by Horowitz in 1945. But we are now in a position to say a little more as to the nature of that competition because what is so surprising at first sight is the way in which biochemical molecules are, so to speak, tailored in extreme detail for the job they have to carry out. This fitness of structure for function has what may be called a purely geometrical side in the actual shape of the molecule in three dimensions, and particularly the shape of its surface, and also the physical side, namely, the adjustment of energy levels, a subject to which Szent-Györgyi (1960) himself has given a new impetus.

As to the first, we are *just* beginning to see, and have not yet in fact seen, how an enzyme molecule operates as a three-dimen-

sional unit bringing together in its tertiary structure different segments of a protein chain which have to combine to attach the molecule on which the enzyme is working. This could not be deduced from the primary or even secondary structure of the protein. Unfortunately, as yet we do not know the full tertiary structure of any protein enzyme and are still at the guessing stage but at last we are beginning to see why such a complicated arrangement as a chain of several hundred links is required to carry out apparently quite simple chemical operations such as breaking a molecule of hydrogen peroxide.

As to their electronic interactions, those of bio-molecules are quite clearly much more subtle in character. We have a glimpse of one of them in the existence of one of the most uniformly standardized molecules in the whole of biology, that of chlorophyll. As Sir Walter Raleigh said "Why blood is red and why the grass is green are mysteries that none may reach unto." The first stage of that mystery was in finding that blood contained iron and grass contained magnesium. But the subtlety of the distinction only appeared when the energy levels of excitation triplets of the two are compared and it was found that the substitution of iron by magnesium, chemically a most unusual occurrence, resulted in the possibility of having time enough to break up a molecule of water to liberate the oxygen (Rabinowitch, 1956).

By selection or by modification of a molecule so as to change its energy levels slightly, many new reactions can be carried out and it is not surprising that once such a reaction is established, it is not further modified but is built into a more complex scheme in which its peculiar functions can be fully used.

It is no accident that the substances apparently responsible for the sense of smell turn out to be very similar to those in the retina responsible for the sense of sight. Both are carotenes, very sensitive electronic receivers (Briggs and Duncan, 1961). In the former case we may be considered to taste molecules, in the latter to taste light or, if you prefer it the other way round, to see chemical molecules and see light. The specific words lose their meaning because what is detected in either case is a very small change in electronic levels. The evolution of molecules capable of responding in this way produces the first possibility of sensation in a primitive organism. The correlative of sensation, that is, motion, again seems to depend on a little molecular evolution of parts of protein or polymer molecules that are

capable of changing their shape under slight electronic changes. This translated into motion makes the senses of use to the organism. Motion without sensation or sensation without motion would be useless; both must have come about at the same time though not necessarily in the same part of the organism.

But what can be said now about the eobiontic change? I have already indicated that we cannot use the existent mechanisms as direct evidence for this; we must go back and discard the possibility of precise reduplication and rely much more on the kind of molecules that are likely to occur spontaneously and on the kind of chemical changes or physicochemical conditions that these will produce. Now we are beginning to get clues of this from experiments refining the earlier work of Müller on the synthesis of organic compounds from inorganic precursors. Apparently these syntheses are much more common and easier than we might have supposed but, also, once they are carefully analyzed they can be seen to be very regular leading to a limited group of molecules which we actually find in biochemistry.

Ten years ago we speculated whether the molecules we find in biochemical structures are what they are because they are naturally stable substances or because they happened to occur first by some historic accident and got repeated that way. That may or may not apply to the molecules of the later and more sophisticated part of biochemistry where the possibility of exact reproduction is natural, but it cannot apply in the earlier part where this cannot exist. We must assume, and we have every reason to from experimental work, that the molecules we actually do get are those which are either of specially low energy and consequently especially stable, or those that are formed with the highest probability, that is, have the highest entropy contribution to their formation.

Oro's recent work on the synthesis of purines and pyrimidines from molecules of ammonia and cyanogen (Oro, 1961a) shows very clearly the actual steps by which these substances can be formed without the intervention of any vital processes at all. We may reasonably suppose that the basic chemical molecules, which may have included things as complicated as vegetable acids, sugars, and lipids as well as the amino acids and the purines, possibly even the porphyrins, were all first produced in irregular abiotic ways. How they sorted themselves out afterwards is going to be a very intriguing problem, but we do not need to assume that it is one involving any

high improbability. Similar syntheses, segregations, and/or sequences of reactions also occur at very similar temperatures in the superficial minerals of the earth's crusts without any intervention of life in the formation of such structures as bauxites, laterites, and other clay mineral transformations involving dehydration, oxidations, and reductions but most of all segregation.

Here the processes of chromatography which we use in industry and in science may very well have occurred by themselves. The suggestion which I made years ago that the early chemical processes of life occurred among molecules absorbed on the clay particles of estuary oozes may carry further than I thought at the time. Not only may the molecules out of which life was formed have been preserved there, but they may have sorted themselves out into quite complicated extensive layers, what I called the prevital areas, which will also allow considerable differentiation and segregation. There is one obvious logical necessity for the origin of life and that is sufficient concentration of materials to allow such changes to take place. If the beginning of the concentration was on clay, the sequel is probably some kind of polymer aggregate or coacervate as Oparin postulated.

If we try to look further back, we have now got to face one of the most exciting problems that has arisen in puzzling out the story of the origin of life—where the original material came from. At the present moment, as I have discussed elsewhere (Bernal, 1961), we really have two hypotheses to choose from, both based on certain chemical and other evidence, namely, the original Oparin-Haldane hypothesis of the evolution of carbon-containing compounds from the gases in the earth's original atmosphere, or the revival of a still older idea that these compounds were brought to the earth from some extraterrestrial source. The evidence for the latter view comes from recent analyses of carbonaceous meteorites, which certainly contain adequate supplies of carbon and nitrogen compounds and the necessary free energy to lead to further transformations. Another source of supply as Oro (1961b) has suggested may have come from the comets.

A still more recent idea, though one foreshadowed by Arrhenius at the beginning of the century, is that life itself went through the earlier stages of its evolution on some other planet. Concrete evidence for this is claimed to come from the spectral and chromatographic analyses of Calvin (Calvin and Vaughn, 1960) and Briggs

(1961) who find in carbonaceous meteorites metabolites like the purines, or of those of Meinschein and Nagy (Nagy *et al.*, 1961) who, by mass spectrographic analysis, find the presence of hydrocarbons corresponding to those produced in terrestrial living organisms.

More recently still, Nagy and Claus (1961) have claimed to discern structures like algae in the Orgueil and Ivuna meteorites. If these results are established we may have to look for the origins of terrestrial life beyond the earth, but this is no longer the fantastic task it seemed a short five years ago.

However, we should beware of taking two jumps in one. Life may have originated elsewhere but it does not follow that odd remains of that life carried to our earth at some remote period, even if they survived the journey alive, could immediately propagate and become also the origin of terrestrial life. We may have to consider more than one origin for life. In view of the new evidence at hand and the far greater flood of evidence that is bound to come with the results of the first planetary voyages, it is clearly premature to do more than speculate at this stage.

With what evidence we can muster from this earth, we have enough to go on to indicate the very considerable antiquity of life here. But we are not yet in a state to make any serious estimate of the relative lengths of organic and biochemical evolution. Two pertinent considerations have to be taken into account. First that organic evolution as we know it is extremely slow on account of the excellence of the copying mechanism which is provided by the genetic apparatus. Now only the aberrations of this copying mechanism, such as mutations and crossings, can give rise to new species. Without the copying mechanism there would be what might be called continuous rather than discontinuous evolution (Oparin, 1957). Any one molecule that was formed would be part of a total complex and would comparatively rapidly eliminate older molecules that were less effective in the biochemical process. But against this there is an opposite consideration, namely, that the first part of the biochemical evolution was very much a hit-and-miss affair, with obviously far more misses than hits and the task of selecting and stabilizing the successful molecules and reactions must have been so great that it must have taken longer. Moreover, the degree of intrinsic biochemical complexity is much greater than that of extrinsic organic complexity which is only concerned with shape and

size, sensation and movement. And, we are limited in time. We have altogether some four thousand million years to play with and of that I think we may claim from the geological evidence that about one thousand five hundred or two thousand million years were certainly occupied by organic evolution. This would leave at most another one to two thousand million years for biochemical evolution.

The major point that I am trying to make in this essay is that we can now speak of an entity or process which we may call biochemical evolution and which was largely though not entirely prior to organic evolution. It continued all through palaeontological history but it was subordinate to organic evolution just as today the continuance of organic evolution is subordinate to human social evolution.

The study of biochemistry in this framework is one to be pursued not only in its own right and because of its invaluable contribution to human life and health but also as an essential but still largely unexplored past of cosmic history.

To conclude, I would like to try to draw together the ideas of the history, that is of the evolution, of biochemical structure on the one hand and of the chemical reactions which go on in organisms, on the other. I feel that, just as in actual organic evolution we always have to consider the two aspects of ontogeny and phylogeny—how the animal grows up the way it should and then how it derives from its ancestral animals which were different from itself—so also we must do for the molecules that build them up. Every molecule that we find in a living creature today has its appropriate biosynthetic pathway in that organism: that is its ontogeny. It also has its molecular ancestry and as, for instance, the protein molecules, now that their sequence of amino acids is being worked out, appear to be texts copied from older texts of a simpler kind with a certain amount of reduplication and a certain amount of altering. In this way, I imagine that all biochemical molecules could be traced back to simpler and simpler ancestors until we get back to the isolated atoms in space. It was from these in the first place that originated not only the planet on which we live but also the life on it and the humanity which is just beginning to understand the processes of its origins.

REFERENCES

- Bernal, J. D. (1961). *Nature* 190, 129.
- Briggs, M. H., and Duncan, R. B. (1961). *Nature* 191, 1310.

- Briggs, M. H. (1961). *Nature* **191**, 1137.
- Calvin, M., and Vaughn, S. K. (1960). "Space Research." North Holland Publ., Amsterdam.
- Frazer, R. D. B., Macrae, T. P., and Rogers, G. E. (1962). *Nature* **193**, 1052.
- Horowitz, N. H. (1945). *Proc. Natl. Acad. Sci. U.S.* **31**, 153.
- Nagy, B., and Claus, G. (1961). *Nature* **192**, 594.
- Nagy, B., Meinschein, W. G., and Hennessy, D. J. (1961). *Ann. N.Y. Acad. Sci.* **93**, 25.
- Oparin, A. I. (1957). "Origin of Life on the Earth," p. 117. Oliver & Boyd, London.
- Oro, J. (1961a). *Nature* **191**, 1193.
- Oro, J. (1961b). *Nature* **190**, 389.
- Rabinowitch, E. I. (1956). "Photosynthesis," Vol. 2, Part II, p. 1793ff. Interscience, New York.
- Szent-Györgyi, A. (1960). "Introduction to Submolecular Biology." Academic Press, New York.

Evolutionary Possibilities for Photosynthesis and Quantum Conversion

MELVIN CALVIN

Department of Chemistry and Lawrence Radiation Laboratory,¹ University of California, Berkeley, California

I. Introduction	23
II. Modern Photosynthetic Processes	24
Nature of the Organisms	25
III. Mechanism of the Photosynthetic Process	27
A. The Path of Carbon in Photosynthesis	27
B. Quantum Conversion in Photosynthesis	29
1. Photoinduced Redox System	31
2. Photoinduced Dehydration	34
IV. Pyrophosphate Linkage in Nonphotosynthetic Processes	36
V. Evolution of the Photosynthetic Apparatus in the Green Plant	37
Chlorophyll Structure	39
VI. Chemical Evolution	42
VII. Development of Rudimentary Catalysts	48
Pyrophosphate Formation	51
VIII. Coupling	52
References	55

I. Introduction

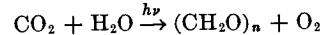
The planning of this discussion has turned out to be particularly difficult, perhaps the most difficult one that I have ever undertaken. The reason for this, I console myself, lies in the very nature of the evolutionary process itself. In physical science (and particularly in mathematical sciences) we are accustomed to a single sequence of events, in which each idea is precursor to the next, and one gradually develops a whole pattern of thought—a whole notion from beginning to end—in a single sequence. Those of you who are more familiar with the way biological material has evolved will know that this is not really the way the living organism can be described in its evolutionary history. The subject of this discussion, the problem of photosynthesis, is especially difficult to trace.

¹ The preparation of this paper was sponsored by the U.S. Atomic Energy Commission.

It turns out, as you will see as we go along, that the evolution of photosynthesis entails the fusion of a number of quite independent threads of evolution at some point in time to give rise to the modern process and the modern apparatus as we know it. In trying to describe that sequence of events, I find myself greatly increasing my respect for the novelist who writes historical novels. He has many apparently independent chains of events, giving rise to a particular incident at the end, or perhaps at the beginning of the novel, and he is very skillful at starting each of these threads and jumping from one thread to the next, bringing them along so they all come together at the right time and in the right place. I haven't yet been able to move smoothly among the various evolutionary threads that are involved here, which ultimately fuse together to give rise to the very complex process of photosynthesis. The story may appear, therefore, more confused than it really is, since I must jump back and forth between separate evolutionary threads and try to indicate their points of fusion.

II. Modern Photosynthetic Processes

With this apology over, let us begin our study of the evolutionary history of photosynthesis by first describing what we think we know of the modern process at which we must eventually arrive. The process of photosynthesis is *the* process by which living organisms are able to transform electromagnetic energy into chemical energy by inducing the reaction between carbon dioxide and water to evolve molecular oxygen and reduced carbon:



This is the over-all process of photosynthesis which has long been recognized as a process for transforming electromagnetic energy, here represented by the quantum, into chemical potential, represented by oxygen in the elementary form and the elements of carbon and hydrogen largely in the oxidation level of carbohydrate (Bassham and Calvin, 1957, 1962; Bassham, 1959, 1961).

If this were all we knew about the process of photosynthesis, we would be hard pressed to try and predict an evolutionary history which might give rise to this process. Fortunately, in the last decade or two we have learned perhaps more about the process of photosynthesis from this point on than in the previous one hundred years. This was the stage that was available to us roughly one hundred

years ago. Only slow progress was made in increasing the chemical knowledge of photosynthesis until just prior to World War II—beginning in the middle thirties and then going on after the war at an increasingly rapid rate.

What do we know today about the process of photosynthesis? Rather than try and give you a history of how the knowledge has evolved, I am going to (1) put down some of the established things that we know about photosynthesis, represented by the over-all reaction; (2) examine the types of organisms which perform this process; (3) determine what the biological apparatus is within some of the organisms (as far as we can do it); and (4) finally go further on down to the molecular level. The question of the evolution of a process of this sort also raises others: What level shall we deal with? Shall we deal with photosynthesis at the level of the whole organisms, the level of the cell, the level of subcellular particles, the level of the macromolecules, or at the level of the small substrate molecules that are involved? We should, in fact, deal with all of these, if possible, but this is another complication which makes this discussion extremely difficult.

I am going to try to pick up two aspects of the process, the mechanism itself on the substrate (and possibly submolecular) level, and the apparatus on the subcellular (or macromolecular) level.

NATURE OF THE ORGANISMS

I hardly need review for you the nature of the organisms which are capable of performing the process of photosynthesis. Quite obviously, the higher green plants, such as a wheat field or a forest, do this on a grand scale. There is, however, a whole set of other organisms besides the higher green plants which are able to do this, or parts of it, and they represent an important part of the biological scheme of things to be examined in the course of our study. These are the marine algae; both the green and the red ones are important in terms of the amount of carbon which is turned over on the surface of the earth per year, as the algae represent the largest single plant family involved in this turnover. Then, there is another group, the blue-green algae, which appear to be structurally more primitive organisms which are capable of doing the entire process of photosynthesis, that is reducing carbon and evolving oxygen. And, finally, we come to the bacteria, both the green and the red, which are capable of performing part of this conversion process. The

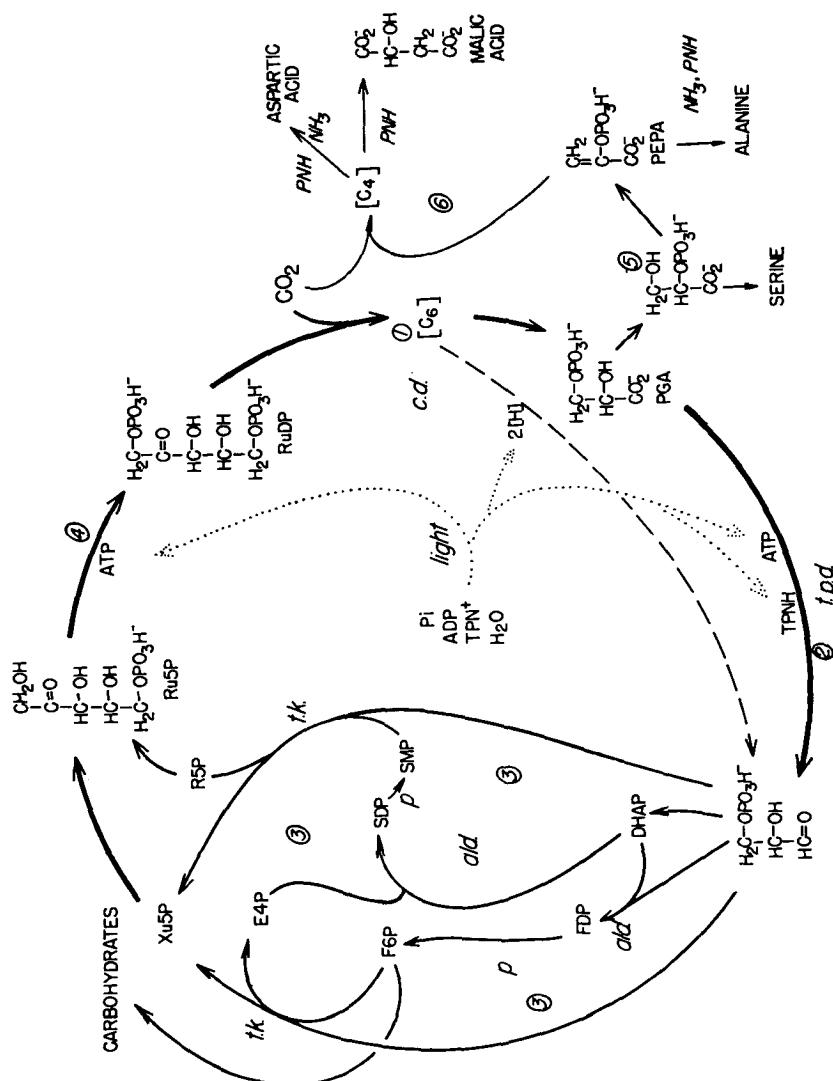


FIG. 1. For descriptive legend see opposite page.

bacteria are capable of transforming electromagnetic energy into chemical energy, but not with the evolution of oxygen. They use ultimate reducing agents other than water in order to reduce the carbon and therefore they produce other oxidants than oxygen. But the photosynthetic bacteria are able to capture electromagnetic energy from the sun and transform it into chemical potential.

These are the classifications of organisms that can do all, or part, of this conversion (energy manipulation) process. These organisms really constitute the whole gamut of biological diversity, as far as I am aware of it, which can do all, or some, of this energy conversion process and they all can do the crucial part of it—the quantum absorption and the quantum conversion.

III. Mechanism of the Photosynthetic Process

A. THE PATH OF CARBON IN PHOTOSYNTHESIS

Let us see what we know about the mechanism of the process of photosynthesis itself. Part of this knowledge is a result of the tracer work which was mentioned earlier (Bassham and Calvin, 1957, 1962; Bassham, 1959, 1961) beginning before the war. My colleague, Sam Ruben, began this work, using radioactive C^{14} , but right after the war in 1945 we took it up again using C^{14} -labeled CO_2 to examine the sequence of events and determine the sequence of compounds involved in the transformation of CO_2 into carbohydrate. The answer to these questions is now available to us, and we can draw a rather complete road map of the reduction of CO_2 . (A simplified version of the carbon reduction cycle is shown in Fig. 1.) The first step in the photosynthetic carbon cycle is the carboxylation of a sugar, ribulose diphosphate, to give phosphoglyceraldehyde, and this, in turn, can now be reduced to triose phos-

Fig. 1. Carbon reduction cycle (simplified version). (1) Ribulose diphosphate reacts with CO_2 to give an unstable 6-carbon compound which splits to give two 3-carbon compounds, one of which is 3-phosphoglyceraldehyde. The other 3-carbon compounds might be either 3-PGA, as it is known to be in the isolated enzyme system, or some other 3-carbon compound such as triose phosphate (dashed arrow). (2) PGA is reduced to triose phosphate with ATP and TPNH derived from the light reaction and water. (3) Various condensations and rearrangements convert the triose phosphate to pentose phosphates. (4) Pentose phosphate is phosphorylated with ATP to give ribulose diphosphate. Further carbon reduction occurs via conversion of PGA to phosphoenolpyruvic acid (5), and carboxylation (6), to form a 4-carbon compound (probably oxaloacetic acid). Reactions leading to the formation of some of the secondary intermediates in carbon reduction are shown.

phate using some kind of *reducing agent* as well as some *pyrophosphate-containing compound*. The triose phosphate then goes through a series of rearrangements to produce ribulose diphosphate again, and the carbon cycle can continue.

The light is required to produce these two agents: a reducing agent, here represented by [H] and a particular (pyrophosphate-containing) phosphorus compound to help the reducing agent in the reduction process. This particular phosphorus compound seems to be adenosine triphosphate (ATP) which contains a pyrophosphate linkage. This is of great importance and will be discussed in detail later on.

The major point that I want to introduce at this stage is the idea that the reduction of CO_2 through the carbon cycle and the whole sequence of enzymatic reactions that are involved in this reduction are *dark* reactions. Once we have available the products of the *light* reaction, namely, a reducing agent and some type of "high energy" phosphate, the whole carbon cycle can be operated and carbon can be taken from CO_2 into a variety of compounds, among them sugar. The sugar can be taken out of the cycle. Every time the cycle turns six times, for example, we can take out a hexose sugar molecule and still have the cycle molecules left. This, indeed, is what happens.

We recognize also that all of the eleven enzymes (catalysts) that are involved in these transformations in the carbon reduction cycle are to be found very nearly everywhere very widely distributed in the biological world—not limited solely to organisms which are converting solar energy, but also in organisms that have nothing whatever to do with the photosynthetic process. It therefore seems quite clear that at least this sequence, that is, the carbon reduction sequence, undoubtedly evolved in a separate chain of evolutionary events having little or nothing to do in the early stages with the electromagnetic energy conversion process itself (Calvin, 1957, 1959d). The electromagnetic energy conversion process itself appears to produce in a primary act, or very close to it, two materials, a reducing agent and a pyrophosphate linkage, which can then run the carbon reduction cycle.

We can already see the two quite independent evolutionary streams which were joined only very recently in evolutionary history to produce the modern green plant (Calvin, 1953, 1956a, 1959c). The carbon reduction system was one independent stream. These streams will, of course, break up into finer parts as we go along, but this is our beginning.

B. QUANTUM CONVERSION IN PHOTOSYNTHESIS

Let us now return to the photochemical process itself. Having separated out the carbon reduction system as a separate evolutionary stream, I am going to leave it since there is nothing unique about it for photosynthetic organisms except the combination of the product of the light reaction with a certain collection of enzymes, all of which can be found either separately or in various combinations in nonphotosynthetic organisms (Fuller and Gibbs, 1956; Milhaud *et al.*, 1956). Therefore, the carbon reduction cycle had a separate evolutionary history until the recent times.

Let us now see what more we can say about the quantum conversion process. We do not have anywhere near the detailed knowledge of the quantum conversion process as we do of the carbon reduction process. It is perhaps worthwhile to put down on paper before we start this discussion the structural formulas of the two molecules which we believe to be essential for running the photosynthetic carbon reduction cycle. (There are undoubtedly others, of which we are still unaware, required for oxygen evolution as well.) To run the carbon cycle we need the reducing agent, which is a pyridine nucleotide in its reduced form. An adenine and pyridine moiety are tied together by two ribose sugars and a pyrophosphate link to give the molecule known as diphosphopyridine nucleotide. Actually, in photosynthesis it seems that there is a molecule very similar to this, but involving another phosphate group on one of the ribose molecules, and so I will actually use the triphosphopyridine nucleotide (TPNH) in its reduced form as the structural formula for the reducing agent which is required to run the carbon reduction cycle.

The possibility exists that still another, and perhaps more specific, reducing agent might be used by photosynthetic organisms in the reductive splitting of the initially produced carboxylation product (Fig. 1, Step 1) (Bassham and Kirk, 1960). If so, it is almost certainly as good a reducing agent as TPNH and may or may not be structurally and kinetically related to it. If such a specific photosynthetic reducing agent functions in green plants, it will, in all probability, have been a late addition in the evolutionary development of a higher efficiency, since we already know that the cycle can operate through TPNH.

The other molecule that is essential for running the cycle and which clearly must come somewhere from the photochemical reac-

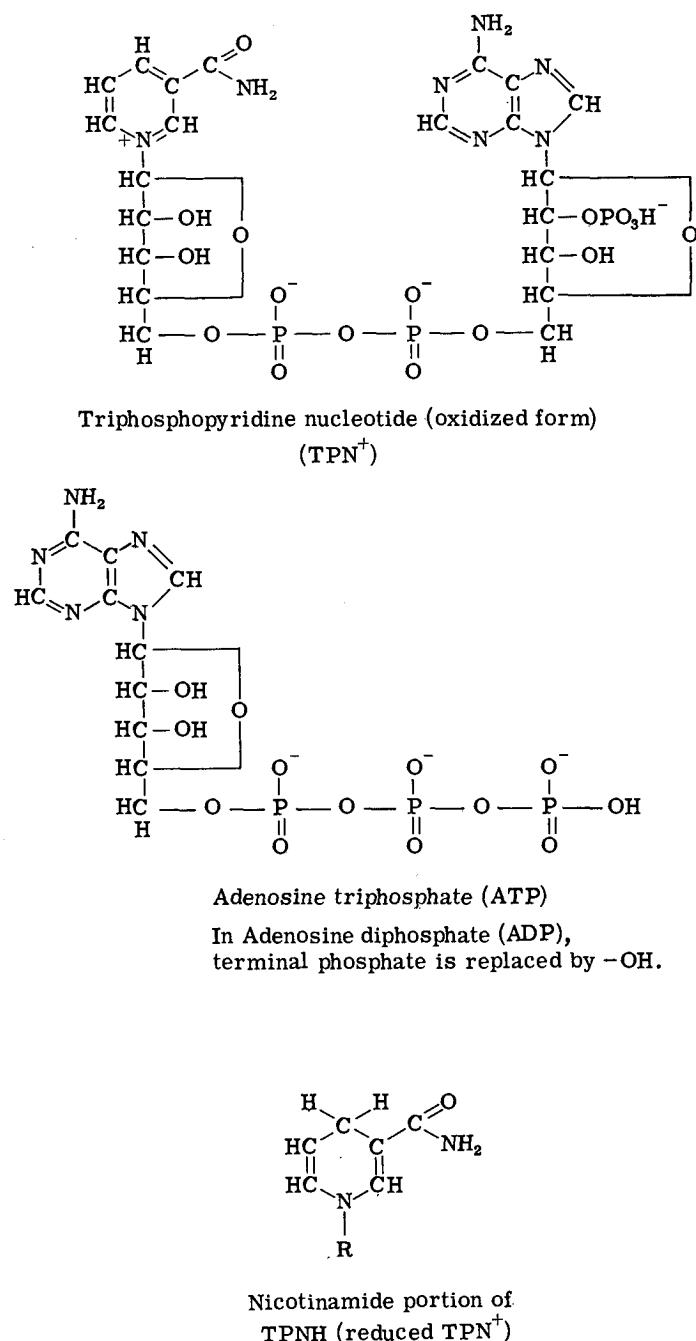


FIG. 2. Structural formulas of triphosphopyridine nucleotide and adenosine triphosphate, two of the agents required to run the photosynthetic carbon cycle.

tion is the adenosine triphosphate (ATP). Here there are two pyrophosphate linkages, and the important one for our purposes is the terminal phosphate link (Fig. 2). These are the two molecules that are required in order to move the cycle around, and clearly these must be manufactured as a result of the photochemical transformation.

How much do we know about how photochemical transformation manufactures those two substances? Here we are not so thoroughly informed, but a good deal, nevertheless, is known and some of it is of considerable importance in guiding our thinking as to what the evolutionary relationship between the photosynthetic equipment and other equipment of living organisms might be.

1. Photoinduced Redox System

The principal photochemical reaction we now know is, first, the absorption of light by chlorophyll to produce some kind of an excited chlorophyll, either a molecule or molecular aggregate. (I don't mean this to be a separate chlorophyll molecule in solution, but simply the chlorophyll as it exists in the photosynthetic equipment of the organisms.) This electronically excited molecule must then undergo some kind of transformation—for example, it may react with another molecule or molecules to produce a separation of an *oxidant* from a *reductant* (Calvin 1961a,b). I am using this language first because of a bit of confusion that has arisen in the meaning of these terms. In ordinary photosynthesis the oxidant will eventually become molecular oxygen; the *reductant* will eventually become a reduced compound, pyridine nucleotide. The pyridine nucleotide, together with the ATP for which we have not yet described a formation mechanism, will then go on to drive the carbon cycle.

These terms, *oxidant* and *reductant*, are the chemists' terms for what happens after the excited chlorophyll loses its energy to some molecule, or collection of molecules, if any redox system is directly involved. The biologist has been accustomed to writing these two things in different terms. Following van Niel (1956; Stanier, 1961; Arnon, 1959, 1961) the biologist has generally associated the term M (Fig. 3) with water and has generally called the oxidant ($[\text{O}]$ in Fig. 3) *hydroxyl* (OH), but he has been very careful, I must say, to put a bracket around it. (Those of you who know what the meaning of a bracket is will understand the significance of this; when you see a biochemist putting a bracket around something of this sort it

means that he doesn't really know what he is talking about. It is something he doesn't know and it is a general representation and not a chemical formula.) The reductant ([R] in Fig. 3), according to the biologists, has been called [H] hydrogen, and this has led many to suppose that the primary process of quantum conversion involves the splitting of the water molecule itself. What is meant

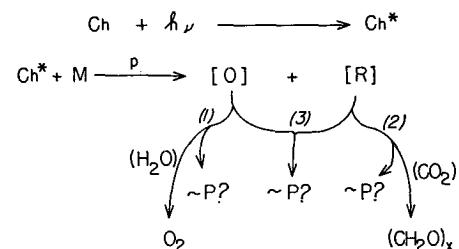


FIG. 3. Simplified photosynthesis scheme. The quantum is first absorbed by the chlorophyll molecule; then something happens (P for primary) to the excited chlorophyll to produce two chemical species ([O] and [R], for example) which can go on, one of them [O] to become molecular oxygen in some way (1), and the other one [R] leading to the reduction of CO_2 to carbohydrate (2). Along these two routes, various other energy-containing species may be created (ATP or $\sim P$). ATP would be an energy storage product. This may be created on either, or both sides. There may be back reaction (3) between the oxidants and reductants which also could create products of higher energy. The obvious one here is, of course, the pyrophosphate linkage in ATP.

by the van Niel theory, at least in chemical terms, is the creation of a reductant of some general character, whose nature we do not know as yet, and of an oxidant whose nature we do not know. These two things must ultimately come from water as given by the stoichiometry of the primary reaction of photosynthesis in the first place.

In more recent years still another terminology has entered into this discussion and it comes from quite a different source. The physicist has called the reductant the "electron" and what is left after an electron is taken away from a molecule is called a "hole" (Calvin, 1956b, 1958a, 1959a,b). These are the physicists' terms for the same phenomenon. You must not get confused about the terminology because all of these—oxidant-reductant, hydroxyl-hydrogen, electron-hole—all are different names for essentially the same thing. What we are trying to do now is to discover exactly the best way to describe these things in ultimate and intimate detail.

I introduce the terminology of the physicist because in the last

few years we have learned a number of the reactions of excited chlorophyll and one of them is an electron transfer reaction which is observable spectroscopically. An electron is transferred from iron in the divalent state to give iron in the trivalent state (Lundegardh, 1954, 1959; Kamen, 1961; Chance and Smith, 1959; Smith, 1961; Chance and Nishimura, 1960; Arnold and Clayton, 1960; Bassham, 1960); with the electron located in an as yet unknown place. It is an important recognition that this phenomenon occurs and occurs very quickly after the chlorophyll absorbs the light.² The excited chlorophyll in some way is able to extract an electron from the ferrous iron compound, at present associated with the chlorophyll in modern organisms in the form of cytochrome, to produce the ferricytochrome and an electron in some molecules as yet undesignated (Kamen, 1956, 1961). This appears to be an important connection between a molecule that is unique to photosynthetic plants, namely, chlorophyll, and certain kinds of molecules which are not unique to photosynthetic plants, namely, the iron cytochromes (iron hemes). The iron hemes have universal distribution, and this is an important fact to remember.

In addition, we now know that electrons must ultimately find their way to pyridine nucleotide. The oxidized iron, or something close to it, will eventually take electrons from water, giving rise to the ferrous iron and molecular oxygen and protons.

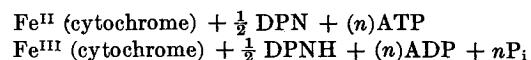
At the same time that all of these things are happening, somewhere along the line either on the way from the intermediate oxidant to oxygen [reaction (1), Fig. 3], or on the way from the intermediate reductant to the pyridine nucleotide [reaction (2), Fig. 3], or, perhaps, in a recombination reaction in which the electron falls back into the hole [reaction (3), Fig. 3] we also create ATP. The ATP is designated by $\sim P?$, which represents "high energy phosphate" linkages. The reactions in Fig. 3 indicate possibilities only, and not knowledge of three different ways (places) in which pyrophosphate could be created: (1) The fall of the intermediate oxidant toward oxygen; (2) the fall of the intermediate reductant (perhaps a sulf-

² A recent modification (Stanier, 1961) of the van Niel generalization inserts a ferrocyanochrome ahead of the water molecule as the primary electron donor to the excited chlorophyll, but does not specify the primary fate of the excited electron which must be removed from chlorophyll. The oxidized cytochrome is presumed capable of oxidizing water to oxygen, with the concomitant formation of ATP, a suggestion similar to that of Bassham (1960) and corresponding to reaction (1) in Fig. 3.

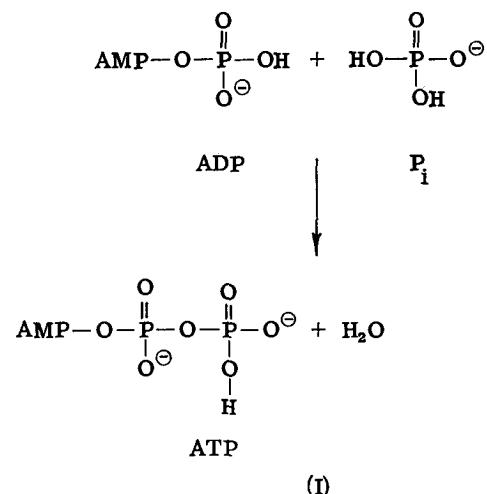
hydryl group) to the pyridine nucleotide which would, perhaps, give rise to pyrophosphate; or (3) perhaps the energy of recombination of the hydrogen-hydroxyl (electron-hole) could also give rise to a number of pyrophosphate linkages.

2. Photoinduced Dehydration

A more profound departure from the basic redox photo process is possible, particularly in the light of the recently indicated (Chance, 1961) *reversibility* of at least some of the steps of oxidative phosphorylation. Thus, there is evidence that in mitochondria it is possible to produce the reaction

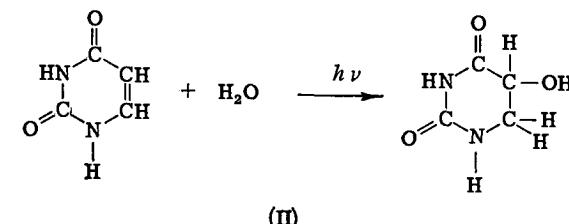


If an independent (non-redox) method of dehydration could be found for producing ATP according to the reaction shown in (I),

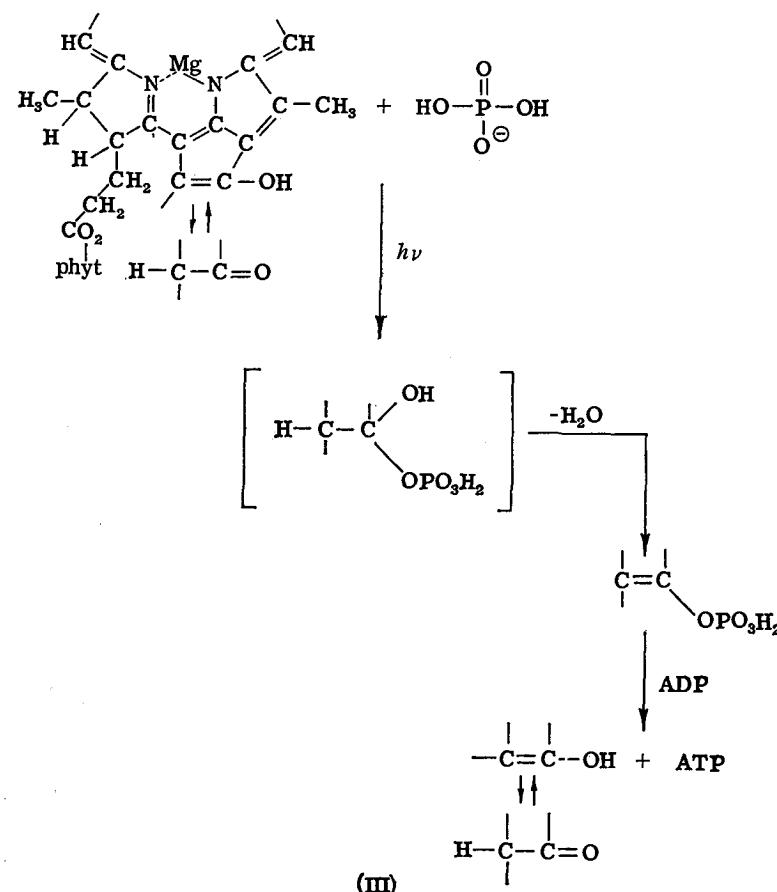


then both ATP and TPNH could be photoproduced without calling upon a photoinduced direct electron transfer reaction.

We have already a precedent for the idea that an optically excited π -electron system can have an increased affinity for water leading to its hydration by an only very slowly reversible process so that energy may be trapped in this manner [see Scheme (II)] (Shugar and Wierchowski, 1957, 1958). For example, if the 9-10 enol in chlorophyll were to add orthophosphate (when excited) an enol phosphate could be produced, which presumably would be capable of phosphorylating adenosine diphosphate (ADP) to make the



required ATP (Wasserman and Cohen, 1960). Part of this would then be used to reverse the DPNH-cytochrome reduction to produce the ultimately necessary separation of oxidant and reductant (water splitting) required for O₂ production and CO₂ reduction [see Scheme (III)].



The not inconsiderable difficulty with such a plan as this is the necessity for producing a *good many more than one* ATP for each quantum absorbed by chlorophyll. Even if a way of circumventing this difficulty were found, it remains fairly clear that such a device would be a rather recent evolutionary addition to an already highly developed biosynthetic energy-manipulating system.

IV. Pyrophosphate Linkage in Nonphotosynthetic Processes

The appearance of pyrophosphate linkage in a variety of organisms is well known. In practically all organisms, there are mechanisms for producing ATP which do not involve photosynthetic mechanisms at all. One of them is a reversal of one reaction in which ATP is used in the photosynthetic cycle (triose phosphate dehydrogenase). By running the reaction backwards (Step 2, Fig. 1) one can make ATP. A more important source is a reaction which apparently involves iron—the cytochromes, involving also the oxidation and reduction of the pyridine nucleotide. The two reactions together are involved in the creation of ATP in nonphotosynthetic organisms. This process of the oxidation of pyridine nucleotide by the passage of electrons from pyridine nucleotide back to oxygen through the iron cytochromes with the concomitant formation of ATP is known as oxidative phosphorylation. It leads to the creation of more ATP than does the substrate oxidation process. The return of a photoexcited electron of chlorophyll through all or part of a similar chain could produce the necessary ATP [see reaction (3), Fig. 3].

Thus the creation of both the reduced pyridine nucleotide and the ATP are not unique to photosynthetic processes. These processes also occur in nonphotosynthetic organisms (Calvin, 1956a). We know something about how pyridine nucleotide is created, but we know relatively little about how ATP is created in oxidative phosphorylation in which the electrons pass from reduced pyridine nucleotide through iron back to oxygen. This is one of the major problems of energy transformation in all biological organisms.

We have now split up the *photo* process of photosynthesis into two other streams of evolutionary development, the stream which gave rise to pyrophosphate (ATP) and the stream which gave rise to pyridine nucleotide. Neither of these necessarily involves the photo process directly. This leads us to the conclusion that the appearance of the photo reaction, or the coupling of the photo

reaction, with the creation of ATP and of reduced pyridine nucleotide was a very late thing in the evolutionary scheme (Calvin, 1957, 1959d). You see that we are forced, now, to consider the question of the origin of life in discussing the origin of photosynthesis. We cannot dodge that issue, and we are indeed considering it and doing so in a much more sophisticated way than has been possible up until recent times. This has been discussed more thoroughly elsewhere (Calvin, 1961c), so I shall not dwell upon it in any great detail.

I shall simply pass through some of the states that we need in order to try and focus your attention on the separate evolution of mechanisms for making ATP, mechanisms for making the molecules which are involved in the creation of ATP today, mechanisms for creating pyridine nucleotide, and, finally, at the very end, how the light-capturing molecule, chlorophyll, may have appeared and was coupled to the other energy-transforming processes. This is really the story in principle, and I now want to go through it quickly and try to give you some idea of how I think these things might have occurred.

V. Evolution of the Photosynthetic Apparatus in the Green Plant

Figure 4 shows the apparatus (the chloroplasts) in the green plant which is responsible for performing the process of photosynthesis. I have not discussed in detail the visible features of the photosynthetic apparatus, but it is perhaps necessary to say a few words here about the relationship of the tangible physical material that performs photosynthesis as it can be seen on the subcellular, but still visible, level. I will then discuss the macromolecular level (where this apparatus cannot yet be seen), and, finally, go to the substrate level where we can again deal with things in a chemical way.

Three different chloroplasts are shown in Fig. 4, illustrating the highly ordered array of layers in all of the three types of organisms: a unicellular green alga, a blue-green alga which does not have a chloroplast (the layers are still present, however, winding their way in and out through the entire cell), and a chloroplast from a higher plant (tobacco) showing the layering of the green material very cleanly. The layers (lamellae) themselves are constructed of arrays of macromolecular subunits which we now think we can see



FIG. 4. Chloroplasts from a unicellular green alga, from tobacco, from tobacco, and mitochondria from guinea pig pancreas. (a) Mitochondria from guinea pig pancreas. (b) Chloroplast with several mitochondria, *Chlamydomonas* (Sager). (c) Entire cell of *Anabena* showing lamellae (Vatter). (d) Tobacco chloroplasts. Maintained 24-36 hours in the dark before fixing with permanganate (Weier).

(Park and Pon, 1961). Figure 5 gives a model for chloroplast lamellar structure, and Fig. 6 is an electron micrograph of frozen dried spinach chloroplast supernatant purporting to show the substructure of the lamellae.

Figure 4 shows the high degree of order in the chloroplasts, and, furthermore, that this high degree of order exists in other elements in the cell, such as the mitochondria, which perform other functions

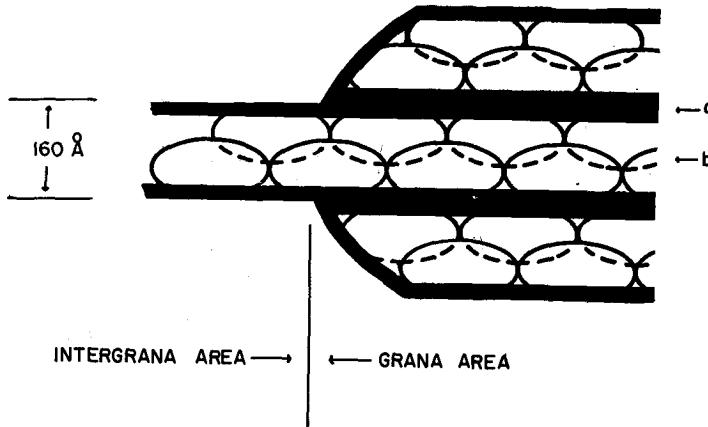


FIG. 5. Model for lamellar structure within a spinach chloroplast (Park and Pon, 1961). (a) Osmium-staining layer of the lamellar structure. Thickness 30 Å in the intergrana regions and 60 Å in the grana regions. (b) Particles forming the granular inner surface of the two layers making up the lamellar structures. The packing of oblate spheres would not be as simple as illustrated in the figure since the central axis of both layers would not be in the same vertical plane shown here.

(formation of ATP by oxidative phosphorylation of pyridine nucleotide) (Calvin, 1958b). The purpose of Fig. 4 is to show the similarity of structure between the photosynthetic apparatus and material which is not photosynthetic, and to show also that it is a highly ordered array in all cases. This highly ordered array must be achieved in some systematic way from molecules which themselves are ordered by virtue of the atoms of which they are made.

CHLOROPHYLL STRUCTURE

The actual detailed structure of the one molecule unequivocally associated with the capture of light and its transformation, i.e.,

chlorophyll, is shown in Fig. 7. This shows the structure of some of the different kinds of chlorophyll that are known: The first is protochlorophyll which appears in etiolated plants grown in the dark. When such plants are placed in the light, the protochlorophyll is converted to chlorophyll (Smith, 1960). The principal difference between protochlorophyll and chlorophyll is the addition of two

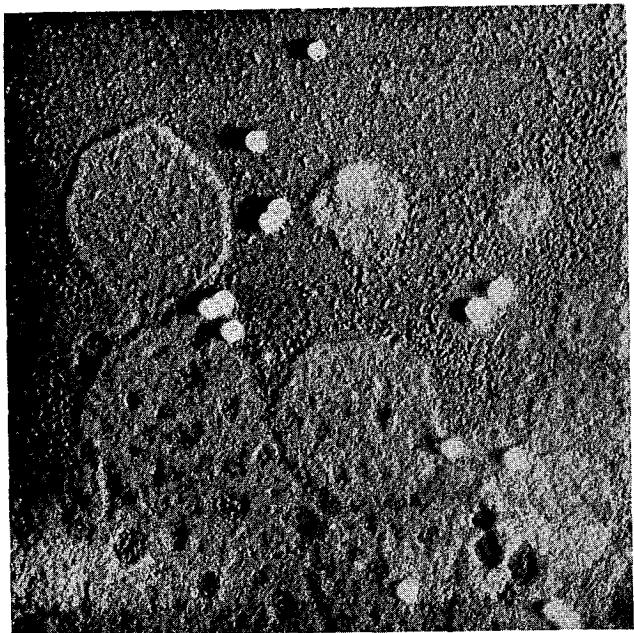


FIG. 6. Frozen dried spinach chloroplast sonicate; 880 Å diameter PSL (polystyrene latex) markers.

extra hydrogen atoms at the double bond in ring D. Bacteriochlorophyll is the molecule which is responsible for the capture and conversion of light in the purple and green bacteria, and differs from green plant chlorophyll in having a second, dihydropyrrole ring in it.

We must devise some way of making those ordered chloroplast structures, which were seen in Figs. 4, 5, and 6, and we must envisage some way of evolving this particular molecule, chlorophyll, belonging to the general class of tetrapyrrolic substances known as porphyrins. These two things—ordered array within the cells and the

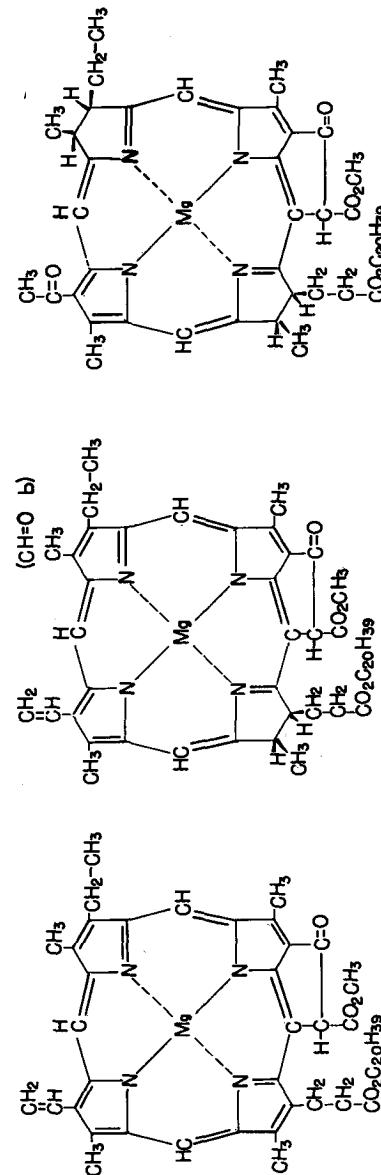


Fig. 7. Structure of protochlorophyll, chlorophyll, and bacteriochlorophyll.
BACTERIOCHLOROPHYLL
CHLOROPHYLL α (b)
PROTOCHLOROPHYLL

development of chlorophyll itself—are two essential features of our evolutionary scheme for the process of photosynthesis.

The structural feature, the appearance of order and structure, is something common to the evolution of all living organisms, and belongs to the general discussion of how ordered structures may be evolved from nonliving material. This is really part of the problem of chemical evolution and the origin of life.

VI. Chemical Evolution

I wish to discuss briefly the beginnings of chemical evolution, starting with the molecules of a primitive atmosphere which you

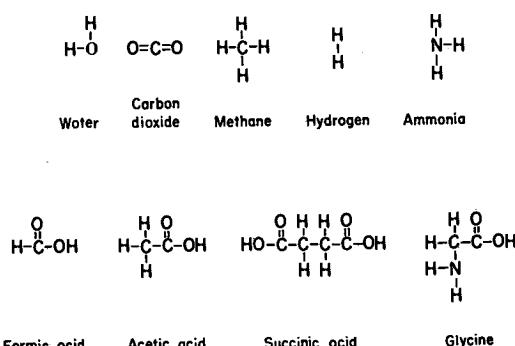


FIG. 8. Primeval and primitive organic molecules.

heard about earlier, being subject to a primitive photosynthesis using the far ultraviolet or radiation from the radioactivity of the earth's crust to transform them. The earliest molecules on the surface of the earth are presumed to be those shown in Fig. 8 (top row), particularly methane, ammonia, and water. If these molecules are subjected to radiation of energy great enough to break the bonds of carbon-carbon, carbon-hydrogen, hydrogen-hydrogen, nitrogen-hydrogen, hydrogen-oxygen, which can be done by ionizing radiation (Garrison *et al.*, 1951) such as the β -rays of K^{40} which are plentiful in the earth's crust (Swallow, 1960), or with ultraviolet light of wavelengths shorter than 2200 Å (Groth and von Weyssenhoff, 1960), then the atoms which are so formed may reorganize to form more complex molecules, a few of which are shown in the bottom row of Fig. 8. You already recognize these molecules as being the present-day substrate materials (formic acid, acetic acid, succinic

acid, and glycine) upon which all living organisms operate. Glycine shown here is the only nitrogen-containing compound in the bottom row of Fig. 8, and it is the simplest of the amino acids, of which the proteins are constructed. By exchanging one of the carbon-bound hydrogen atoms of the glycine for any of a group of other atoms, some twenty different amino acids can be built up.

In the first experiment of this type in 1950 in which we used the cyclotron as a source of ionizing radiation (Garrison *et al.*, 1951), we started with carbon dioxide, hydrogen, and water, and we were able to get by random transformation processes reduced carbon compounds such as formic acid, acetic acid, and succinic acid. In later experiments, in which ammonia was added to the initial mixture following Miller (1955, 1960; Miller and Urey, 1959), glycine was obtained. Still more recently, in the last 6 months, we have performed this same type of experiment again, but instead of depending upon ordinary analytical methods to find these randomly occurring compounds, we have used C^{14} -labeled methane in the primitive gas mixture, thus providing radioactive carbon atoms which could be followed around. The discharge from a 5-mev electron linear accelerator was passed through the mixture of methane, ammonia, and water, and we took the water solution containing the product from this bombardment and spread it out on a piece of filter paper in a systematic way (Palm and Calvin, *in press, a*).

Figure 9 shows the results of one of these bombardment experiments. It is a photograph of the darkened X-ray film which results when a paper chromatogram containing radioactive products is placed on top of an X-ray film. Wherever there is a black spot on the film a particular compound has been located. We can tell what the nature of the compound is by where it is located on the film with respect to its origin. All of the different nonvolatile radioactive compounds which result from one particular bombardment are shown in Fig. 9, and about a dozen compounds have separated out.

We have been able to identify in this way some half-dozen compounds,³ including glycine, alanine, and various other amino acids and sugars, some fatty acids and some hydroxy acids—the very things of which today's living matter is composed. One of the compounds, representing about 60% of the total, is urea. We find in neutral and acidic fractions a large number of compounds, includ-

³ HCN was identified in the aqueous solution by a separate procedure (Palm and Calvin, *in press, a*).

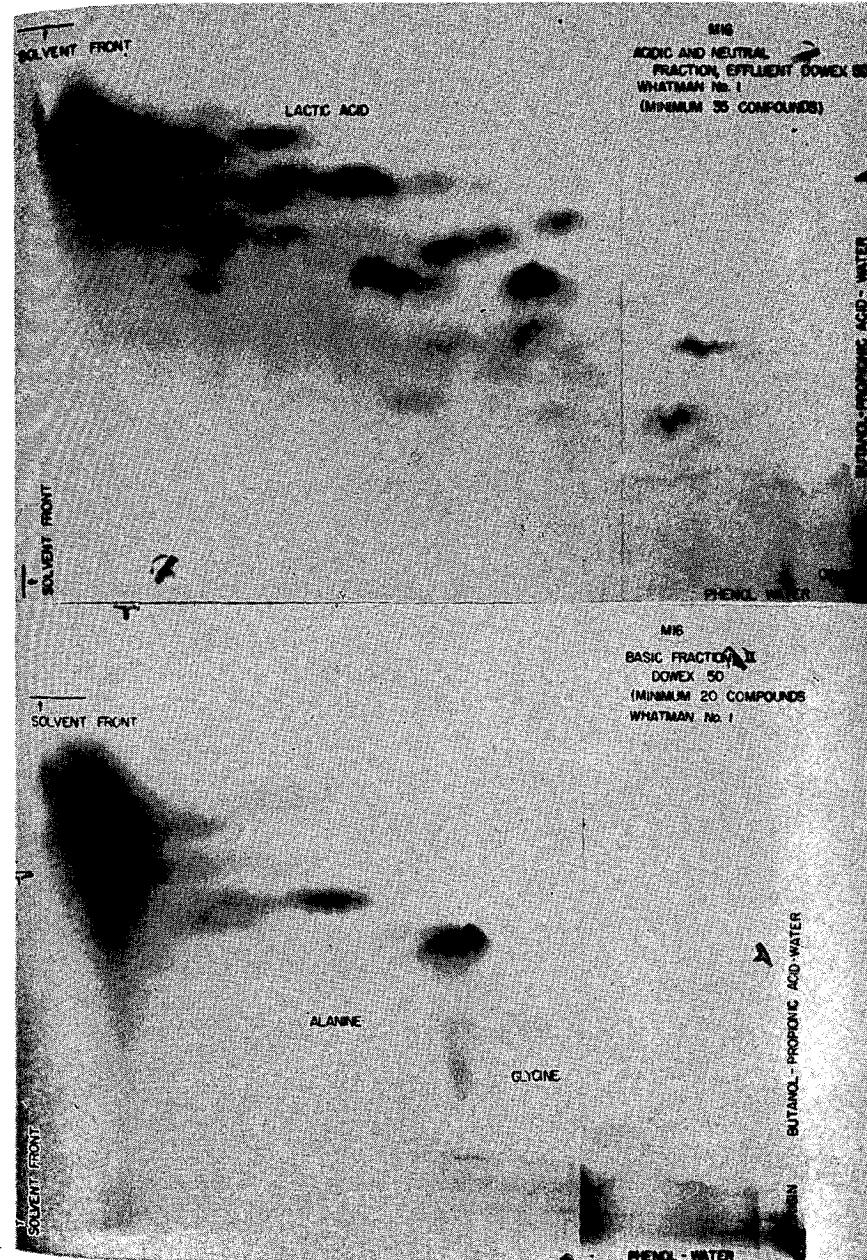
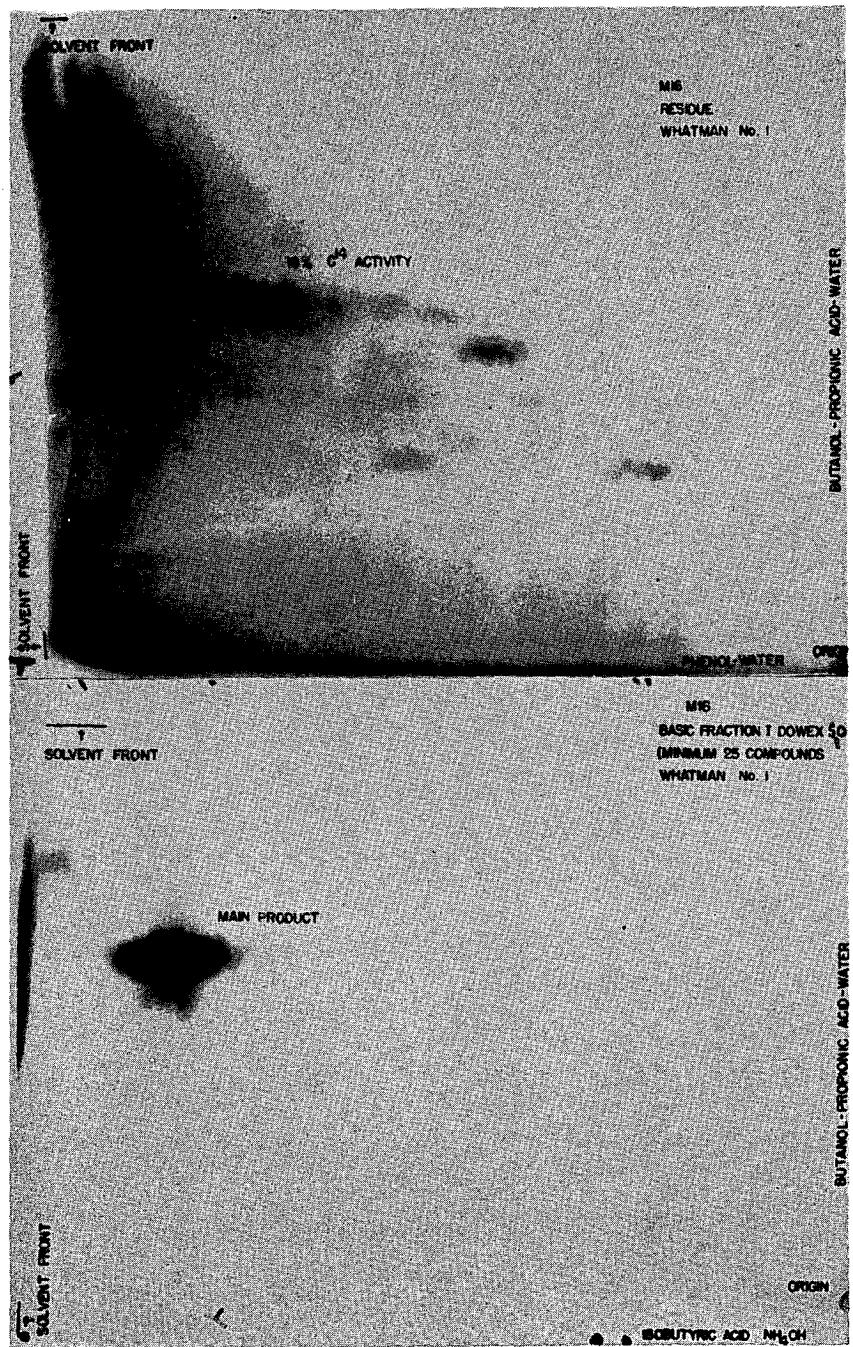


FIG. 9. Radioautograph of paper chromatogram showing compounds which result after irradiation of a mixture of C¹⁴ labeled methane (C¹⁴H₄), ammonia, and water with 5-mev electrons.

ing lactic acid and sugars (Palm and Calvin, in press, *a*). You can also see that alanine and glycine (amino acids) represent a very small amount of the total. There appears to be present in this irradiated mixture a number of undetermined bases, including heterocyclics. Thus, such random processes as these may give rise to all of the simple compounds that are needed by present-day living organisms (Calvin, 1961c,d,e).

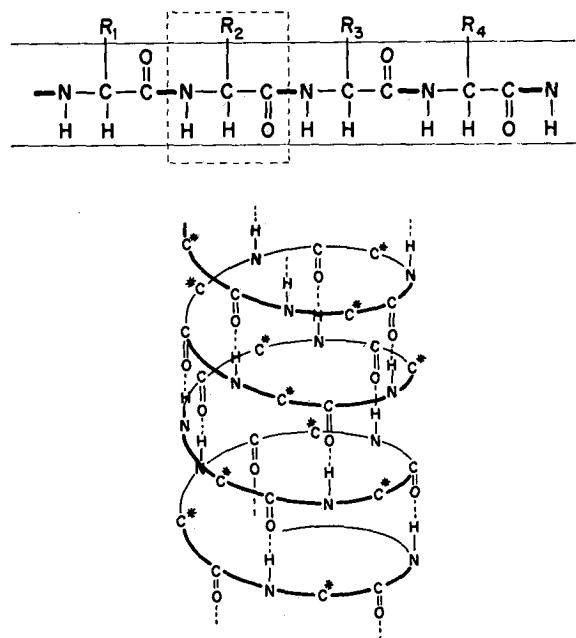


FIG. 10. Protein structure.

Having made these simple compounds (particularly the amino acids) by the random methods, we can build them up into proteins in various ways. Aside from the more or less laborious and specific methods involving special protective or activating groups, at least two simpler methods, possibly applicable to primitive conditions, have successfully been demonstrated in the laboratory recently. The first involves heating amino acid mixtures in molten glutamic acid together with some polyphosphoric acid to produce a mixed polypeptide resembling protein (Fox *et al.*, 1959; Fox, 1960; Fox and Harada, 1961). The second involves heating the amino acid in an

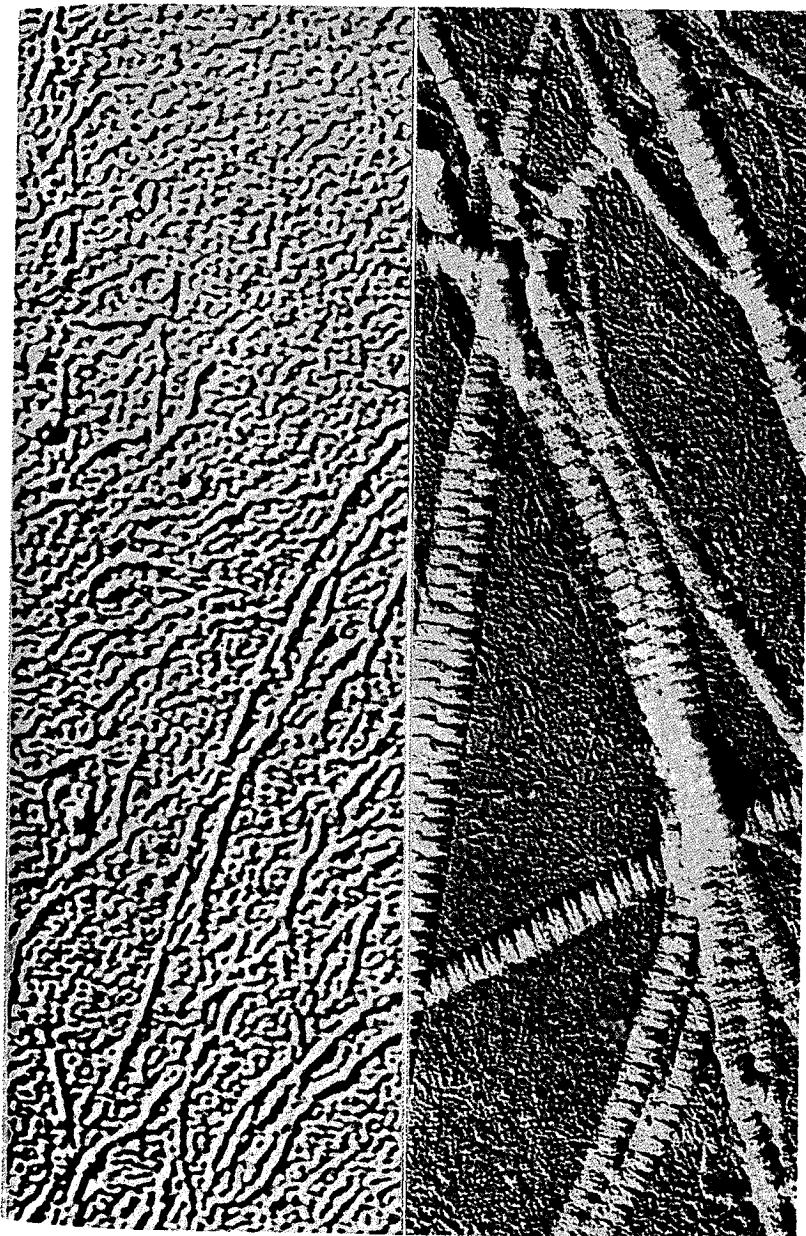


FIG. 11. Electron micrographs of collagen filaments.

aqueous ammonia solution to produce a polypeptide of intermediate size (Oró and Kimball, 1961).

The proteins themselves can take on a specific structure which is shown in Fig. 10. The helical structure is built-in into the linear array of the amino acids because of the particular arrangement of carbon, hydrogen, nitrogen, and oxygen atoms in such a chain. Figure 11 shows how the helical structure can take on visible order. The upper photograph is an electron micrograph of a protein which is a component of collagen. When the protein filaments are aggregated, as shown in the lower photograph, they do so in a specific ordered array because of the particular arrangement of amino acids in the proteins. Here you can begin to see the appearance of the *visible* order that must be generated to create mitochondria, chloroplasts, and other subcellular particles. This generation of order is, of course, common to all living things and is not unique to photosynthesis. One can generate order, beginning from the primitive molecules (Fig. 8) of an early earth's atmosphere, through proteins (Figs. 10 and 11) into the subcellular material itself (Fig. 5).

VII. Development of Rudimentary Catalysts

Let us now turn to the question of the generation of the porphyrins which seem to be central not only in the capture of light, as represented by chlorophyll, but to the appearance of ATP in present-day organisms and perhaps to the appearance of ATP in primitive organisms as well.

Figure 12 shows that starting with the primitive function of iron for the decomposition of hydrogen peroxide, which will be formed in the seas either by ultraviolet radiation or by K^{40} radiation, the iron catalysis can be improved by a factor of a thousand if it is built into a porphyrin. If we now transform this iron further by encasing the heme into a folded protein and make the molecule of catalase, the catalytic function is improved by another factor of ten million for this particular peroxide decomposition reaction (Calvin, 1961d).

This fact is of great importance because I believe that peroxide appeared in the primitive seas of the earth at the very earliest stages as a result of both the ultraviolet radiation at the top of the atmosphere and of the K^{40} radioactivity in the earth's crust. This peroxide can now serve as an evolutionary selection pressure (Gerschman, 1959) to improve the catalytic function of iron from the bare iron to the iron heme to the iron heme-protein combination.

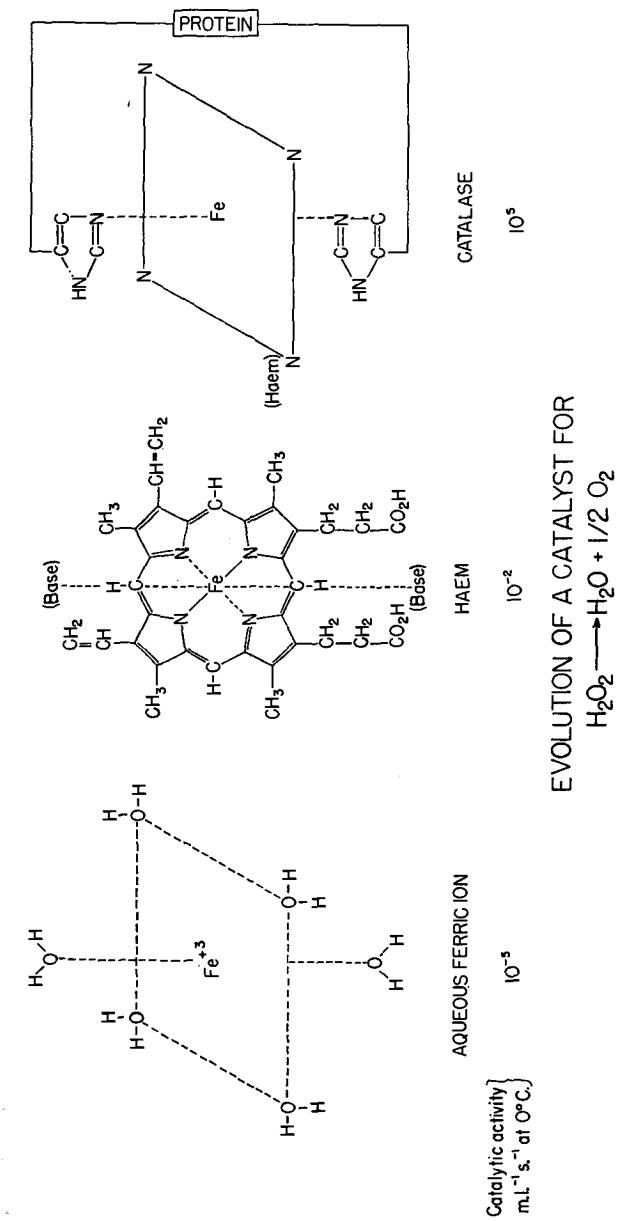


FIG. 12. Evolution of a catalyst.

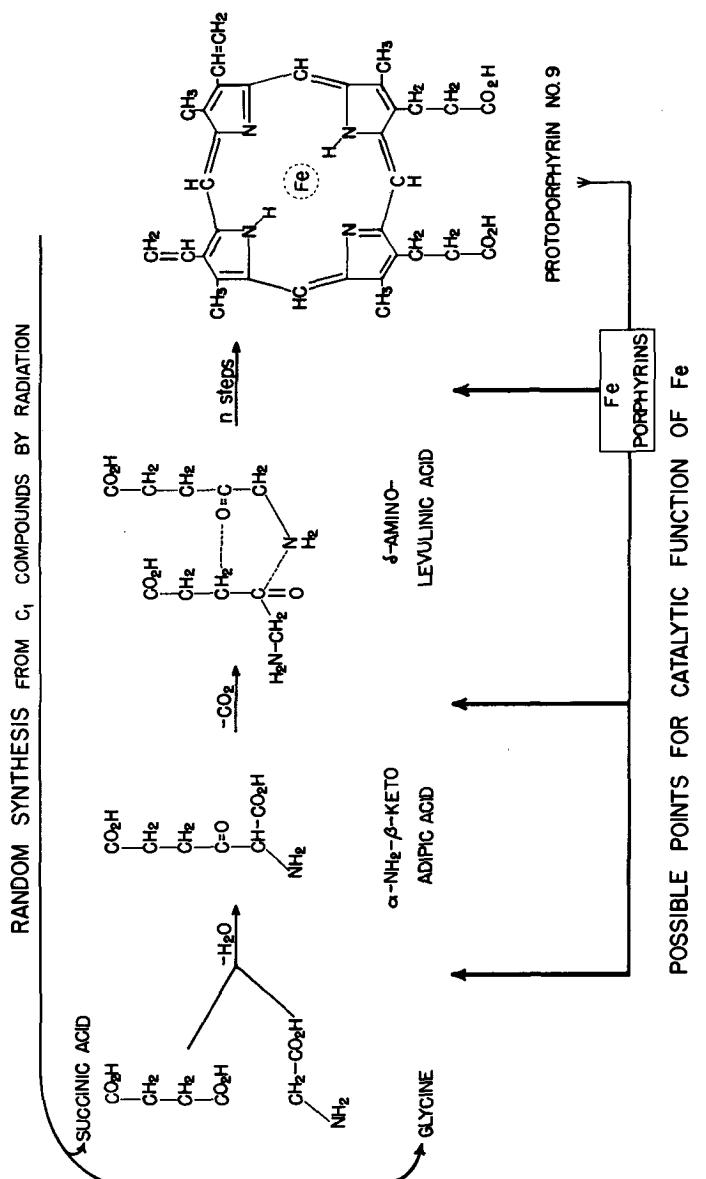


FIG. 13. Biosynthesis of porphyrin and the evolution of the catalytic functions of iron.

The way in which this can occur is shown by having a look at the way in which hemes are synthesized by modern living organisms (Fig. 13). We start with succinic acid and glycine, which were made by random synthesis from the primitive earth's atmosphere, and by combining these two substances we make the α -amino- β -keto adipic acid which then decarboxylates to give the δ -aminolevulinic acid, two of which can combine to form the heterocyclic pyrrole ring. Then there follows a series of oxidation and condensation steps to give rise to the tetrapyrrole ring (Shemin, 1954-55). This reaction is a spontaneous one which involves a number of oxidation steps, several of which are almost certainly catalyzed by iron. The oxidation is achieved either by oxygen or peroxide under the influence of iron and presumably better achieved by iron in a porphyrin than by bare iron. Therefore, once the porphyrin is formed, more of it will be formed because of this autocatalytic self-selection mechanism (Calvin, 1957, 1959c,d).

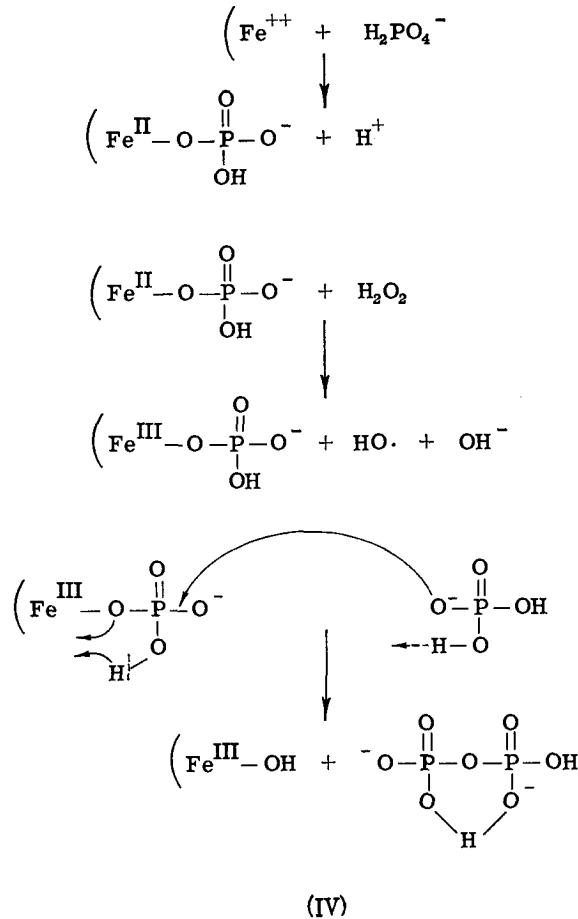
PYROPHOSPHATE FORMATION

This idea is important because the mechanism of the formation of pyrophosphate seems to involve the oxidation of iron. In the last few months we have been able to demonstrate that one can generate pyrophosphate in aqueous media by simply allowing hydrogen peroxide to oxidize ferrous iron in the presence of orthophosphate (Barltrop, 1961). In this reaction, a certain amount of orthophosphate is converted into pyrophosphate. The reaction may be written as shown in (IV).⁴ I believe this to be evidence of the primitive way in which the highly evolved oxidative phosphorylation which takes place today began. The complexing of phosphate by ferrous iron, followed by the withdrawal of an electron from the ferrous iron to make ferric iron, the elimination of a water molecule to make pyrophosphate, and reduction of the ferric iron to ferrous, completes a cycle for the formation and the liberation of the pyrophosphate linkage. This is now demonstrated in a simple system, and I think it will not be long before we will be able to demonstrate it in the highly evolved iron systems that are used in oxidative phosphorylation, both in plants and animals, and which are also used in photosynthetic phosphorylation, probably in a similar manner.

You can see here a driving force which will give rise to the por-

⁴ The half-circle around the iron symbol is introduced to represent any other coordinated atoms or groups.

phyrin molecule. The driving force is the peroxide present in the ocean and the usefulness of transforming orthophosphate to pyrophosphate in aqueous solutions so the pyrophosphate can then be used to assist the combination of amino acids to make proteins. This



was the evolutionary sequence which gave rise first to the porphyrin and second to a mechanism for manufacturing pyrophosphate.

VIII. Coupling

As yet we have suggested no mechanism for using light to perform these processes. All that would be required in the later stages was to find a way of removing the electron from the iron, not with

hydrogen peroxide but with light, in order to couple the photochemical reaction to what we now know to be nonphotochemical processes.

I think this event happened very late in the evolutionary scheme, and the evidence for it lies in the fact that the chlorophyll molecule is today manufactured by a sequence of reactions almost identical with the sequence of reactions used to manufacture the heme (Granick, 1948, 1950, 1955; Bogorad, 1960), but just before the iron is put into the heme (protoporphyrin IX), a branching occurs

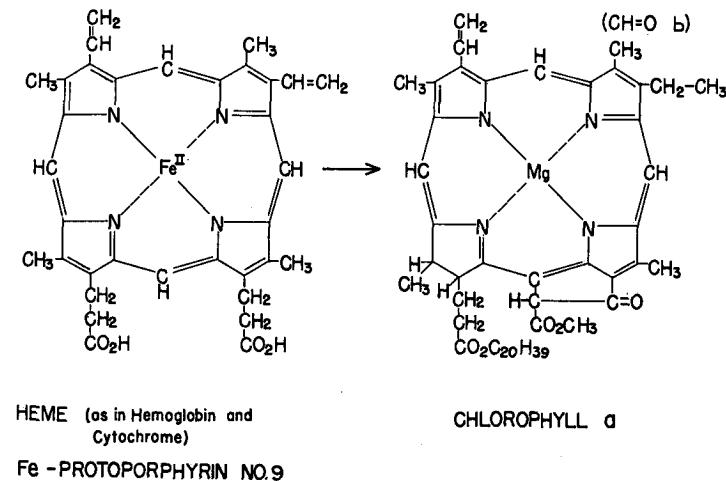


FIG. 14. Structural relations between heme and chlorophyll

leading to the chlorophyll molecule in which magnesium is situated (Fig. 14). I think the reason for that reaction is, first, that the light-absorbing ability of the heme itself is very poor. Although heme is red, it does not have anywhere near the light-absorbing capacity of chlorophyll, and one of the reasons for the evolutionary selection of magnesium chlorophyll (magnesium chlorin) is the fact that the absorption of light by a magnesium chlorin is several thousand times greater than that of the iron porphyrin. Secondly, something very special about the electronic structure of the magnesium and of the packing together of the chlorophyll molecules in a crystal lattice, leading to the separation of electrons from the chlorophyll (Calvin, 1961a), is better achieved by the chlorin than it is by the porphyrin.

If the dehydration-phosphate activation idea (by the 9-10 enol of chlorophyll) turns out to play a role, we would then have a third powerful selective factor favoring the chlorophyll structure.

The emerging likelihood that the products of *two different* quantum conversion acts can collaborate to produce the products of photosynthesis more efficiently than either one alone must be considered (Emerson *et al.*, 1957; Govindjee *et al.*, 1960; Govindjee and Rabinowitch, 1960, 1961; Duysens *et al.*, 1961; French, 1961). One of these processes seems to be electron transfer *from* reduced cytochrome (Calvin, 1961a,b). It has been suggested that the other is electron transfer *to* oxidized cytochrome (Duysens *et al.*, 1961). An alternative pair of transfers would be *to* chlorophyll (from cytochrome) and *from* chlorophyll (to quinone or disulfide) (Calvin, in press). The experimental question as to whether *either one* of these two different quantum acts *alone* could accomplish the whole of photosynthesis, albeit at reduced efficiency, has yet to be unequivocally answered. In any case, the collaboration is surely a late addition.

The mechanism and the detailed chemical and physical reasons for the advantage of the chlorophyll over the porphyrin remains for the future to discover. It is of interest to examine the paleontological record to see if it might be possible to (1) confirm the notion that heme (and its oxidative function) preceded the appearance of larger amounts of oxygen in the earth's atmosphere for whose presence oxygen-producing photosynthesis seems to be the only competent geochemical process; and (2) if confirmed to date, the appearance of chlorophylloous pigments. The presence of *both* heme and chlorophyll fossil molecules in petroleum and other organic minerals has long been known (Treibs, 1936). The principal hope of distinguishing between these two origins lies in the possible presence of a carbon substituent on the δ -carbon atoms of these substances derived from chlorophyll with its isocyclic ring. The relative stability of other possible distinguishing features, and even the structure of some of the bacterial chlorophylls (*Chlorobium*), are not yet known to us. Presumably bacterial photosynthesis, producing as it does only ATP (no oxygen), is a more primitive process and, therefore, the pigments there involved might be expected to have appeared earlier. As yet, no porphyrin at all has been unequivocally found in Pre-Cambrian formations although the presence of fossil forms strikingly resembling in morphology the blue-green algae have been described

by Barghoorn (Tyler, Barghoorn and Barrett, 1957; Barghoorn, 1961).

As early as 1937 Hans Fischer, in discussing chlorophyll, said: "In historical development we regard hemin as the older dye-stuff . . .," but he did not give explicit reasons. These were undoubtedly based on structural chemical relationships. In view of our modern knowledge of the present-day biosynthetic relationship, Fischer will probably turn out to be right.

REFERENCES

- Arnold, W., and Clayton, R. K. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 769.
- Arnon, D. I. (1959). *Nature* **184**, 10.
- Arnon, D. I. (1961). In "Light and Life" (W. D. McElroy and B. Glass, eds.), pp. 489-565. Johns Hopkins Press, Baltimore, Maryland.
- Barghoorn, E. S. (1961). Private communication at Woodring Conference on Major Biologic Innovations and the Geologic Record, Shenandoah National Park, Virginia, June 1961.
- Barltrop, J. A. (1961). Private communication.
- Bassham, J. A. (1959). *J. Chem. Educ.* **36**, 548.
- Bassham, J. A. (1960). *Radiation Research, Suppl.* **2**, 497.
- Bassham, J. A. (1961). *J. Chem. Educ.* **38**, 151.
- Bassham, J. A., and Calvin, M. (1957). "The Path of Carbon in Photosynthesis," Prentice-Hall, Englewood Cliffs, New Jersey.
- Bassham, J. A., and Calvin, M. (1962). "The Photosynthesis of Carbon Compounds," W. A. Benjamin, New York.
- Bassham, J. A., and Kirk, M. R. (1960). *Biochim. et Biophys. Acta* **43**, 447.
- Bogorad, L. (1960). In "Comparative Biochemistry of Photoreactive Systems" (M. B. Allen, ed.), pp. 227-256. Academic Press, New York.
- Calvin, M. (1953). *Idea and Expt.* **2**, No. 4, June.
- Calvin, M. (1956a). *Am. Scientist* **44**, 248.
- Calvin, M. (1956b). *J. Chem. Soc.* p. 1895.
- Calvin, M. (1957). *U.S. Atomic Energy Comm.* UCRL-3915, August 1957.
- Calvin, M. (1958a). In "Radiation Biology and Medicine" (W. D. Claus, ed.), pp. 826-848. Addison-Wesley, Reading, Massachusetts.
- Calvin, M. (1958b). *U.S. Atomic Energy Comm.* BNL 512 (C-28), 160.
- Calvin, M. (1959a). *Rev. Modern Phys.* **31**, 147.
- Calvin, M. (1959b). *Rev. Modern Phys.* **31**, 157.
- Calvin, M. (1959c). *Evolution* **13**, 362.
- Calvin, M. (1959d). *Science* **130**, 1170.
- Calvin, M. (1961a). In "Light and Life" (W. D. McElroy and B. Glass, eds.), pp. 317-355. Johns Hopkins Press, Baltimore, Maryland.
- Calvin, M. (1961b). *J. Theoret. Biol.* **1**, 258.
- Calvin, M. (1961c). *Ann. Internal Med.* **54**, 954.
- Calvin, M. (1961d). *Chem. Eng. News* **39**, 96 (May 22).
- Calvin, M. (1961e). Condon Lectures. "Chemical Evolution." Oregon State Board of Higher Education, Eugene, Oregon.

- Calvin, M. (in press). *Advances in Catalysis* **14**.
- Chance, B. (1961). *Nature* **189**, 719.
- Chance, B., and Nishimura, M. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 19.
- Chance, B., and Smith, L. (1959). *Nature* **175**, 803.
- Duysens, L. N. M., Ames, J., and Kamp, B. M. (1961). *Nature* **190**, 510.
- Emerson, R., Chalmers, R., and Cederstand, C. (1957). *Proc. Natl. Acad. Sci. U.S.* **43**, 135.
- Fischer, H. (1937). *Chem. Revs.* **20**, 41. Quotation on p. 66.
- Fox, S. W. (1960). *Science* **132**, 200.
- Fox, S. W., and Harada, K. (1961). *Science* **133**, 1923.
- Fox, S. W., Harada, K., and Vegotsky, A. (1959). *Experientia* **15**, 81.
- French, C. S. (1961). *Proc. 5th Intern. Biochem. Congr., Moscow*, in press.
- Fuller, R. C., and Gibbs, M. (1956). *Plant Physiol.* **31**, Suppl., xxi.
- Garrison, W. M., Morrison, D. C., Hamilton, J. G., Benson, A. A., and Calvin, M. (1951). *Science* **114**, 416.
- Gerschman, R. (1959). *Proc. 21st Intern. Physiol. Pharmacol. Congr., Buenos Aires, 1957*, p. 222.
- Govindjee, and Rabinowitch, E. I. (1960). *Biophys. J.* **1**, 73.
- Govindjee, Rabinowitch, E., and Thomas, J. B. (1960). *Biophys. J.* **1**, 91.
- Granick, S. (1948). *J. Biol. Chem.* **172**, 717.
- Granick, S. (1950). *Harvey Lectures, Ser.* **44**, 220.
- Granick, S. (1955). In "Porphyrin Biosynthesis and Metabolism," Ciba Foundation Symposium. Churchill, London.
- Groth, W. E., and von Weyssenhoff, H. (1960). *Planet. Space Sci.* **2**, 79.
- Kamen, M. D. (1956). In "Enzymes: Units of Biological Structure and Function" (O. H. Gaebler, ed.), p. 483. Academic Press, New York.
- Kamen, M. D. (1961). In "Light and Life" (W. D. McElroy and B. Glass, eds.), pp. 483-488. Johns Hopkins Press, Baltimore, Maryland.
- Lundegardh, H. (1954). *Physiol. Plantarum* **7**, 375.
- Lundegardh, H. (1959). *Biochim. et Biophys. Acta* **35**, 340.
- Milhaud, G., Aubert, J. P., and Miller, J. (1956). *Compt. rend. acad. sci.* **243**, 102.
- Miller, S. L. (1955). *J. Am. Chem. Soc.* **77**, 2351.
- Miller, S. L. (1960). In "Aspects of the Origin of Life" (M. M. Florkin, ed.), pp. 85-97. Pergamon Press, New York.
- Miller, S. L., and Urey, H. C. (1959). *Science* **130**, 245.
- Oró, J., and Kimball, A. P. (1961). *Arch. Biochem. Biophys.* **93**, 166.
- Palm, C., and Calvin, M. (in press, a). "Primordial Organic Chemistry. I."
- Palm, C., and Calvin, M. (1961). University of California Radiation Laboratory Report UCRL-9519, 30-39, January 1961.
- Park, R. B., and Pon, N. G. (1961). *J. Mol. Biol.* **3**, 1.
- Shemin, D. (1954-55). *Harvey Lectures, Ser.* **50**, 248.
- Shugar, D., and Wierzchowski, K. L. (1957). *Biochim. et Biophys. Acta* **23**, 657.
- Shugar, D., and Wierzchowski, K. L. (1958). *Postepy Biochem.* **4**, 243.
- Smith, J. H. C. (1960). In "Comparative Biochemistry of Photoreactive Systems" (M. B. Allen, ed.), pp. 257-278. Academic Press, New York.
- Smith, L. (1961). In "Light and Life" (W. D. McElroy and B. Glass, eds.), pp. 436-443. Johns Hopkins Press, Baltimore, Maryland.
- Stanier, R. Y. (1961). *Bacteriol. Revs.* **25**, 1.

- Swallow, A. J. (1960). "Radiation Chemistry of Organic Compounds," p. 244. Pergamon Press, New York.
- Treibs, A. (1936). *Angew. Chem.* **49**, 682.
- Tyler, A. S., Barghoorn, E. S., and Barrett, L. P. (1957). *Bull. Geol. Soc. Am.* **68**, 1293.
- van Niel, C. B. (1956). In "The Microbe's Contribution to Biology," pp. 155-182. Harvard Univ. Press, Cambridge, Massachusetts.
- Wasserman, H. H., and Cohen, D. (1960). *J. Am. Chem. Soc.* **82**, 4435.

On Dating Stages in Photochemical Evolution

HANS GAFFRON

*Fels Fund, Department of Biological Sciences
and the Institute of Molecular Biophysics,
The Florida State University, Tallahassee, Florida*

I. Current Beliefs about Biopoiesis	59
II. Organic Catalysis and Reproduction	65
III. On the Efficiency of Photochemical Reactions During Evolution	67
IV. An Artificial Flavo-Mangano-Heme Photooxidase	71
V. The Appearance of Oxygen as the Result of a Mutation in a Photo-reducing Organism	74
References	87

I. Current Beliefs about Biopoiesis

Having found and collected innumerable facts of life, the biologists of the 19th century began to search for an idea which would encompass and explain them all. Finally, with Darwin, this search ended successfully. Later biologists found that his conception of evolution might explain even the very origin of life, and they are now searching for some facts to support so convenient an explanation.

Yet the recent discovery of the uniformity and sameness of the fundamental metabolic reactions which sustain life in all organisms, though amazing and beautiful in itself, seemed to preclude a rational scientific approach to the problem of the origin of life, for it pointed to the appearance on earth of one single primeval organism—the one individual ancestor cell of all cellular organisms that have come afterwards. Several arguments in favor of such a unique historical event appeared at first rather convincing. Now we are equally convinced that these arguments are beside the point. Our task is not to describe the first appearance of the one particular and only successful organism, but to understand and study the conditions which have led, more or less inevitably, to the formation of an entire class of lifelike phenomena. Formulated in this manner the program, though formidable, does not appear insuperable (for

TABLE I

SHORT SUMMARY OF CURRENTLY HELD BELIEFS ON THE SUCCESSION OF EVOLUTIONARY ERAS

Era	Determining environmental conditions	Main sources of energy	Outcome
I	Anaerobic, strongly reducing	UV; heat; electrical discharges	Simple radicals. Accumulation of organic molecules in the oceans. Acetate, glycine, uracil, adenine, etc. (Oparin-Haldane's "Organic soup")
Loss of hydrogen			
II	Anaerobic; traces of oxygen	UV; heat	Complex organic substances, carotenes, nucleotides, peptides, polyphosphates, pigments, <i>porphyrins</i> . Organo-metalloc and photo-catalysis. Primitive surface catalysis. Intermolecular oxidoreductions
Visible light			
Ozone layer; loss of UV			
III	Mainly anaerobic; some carbon dioxide; traces of oxygen	Visible light	Evolution of "synthetic cycles." Multiple replication. Specific catalysis and <i>photochemistry</i> on surfaces of large organic molecules. Primitive enzymes. Genes
First organisms			
IV	As in III	Visible light via photoreduction	Metabolic units replicating at the expense of "food." <i>Photophosphorylation</i> . <i>Photoconversion of acetate</i> . Diminishing supply of original organic food. <i>Photoreduction</i> of carbon dioxide at the expense of organic and inorganic hydrogen donors
Fermentation			
Appearance of oxygen in large amounts			

ON DATING STAGES IN PHOTOCHEMICAL EVOLUTION 61

TABLE I (Continued)

Era	Determining environmental conditions	Main sources of energy	Outcome
V	Mainly aerobic with anaerobic pockets (first more, then less, carbon dioxide)	Visible light via photosynthesis Respiration	Autotrophic plants. New food supply. Evolution of respiring organisms. Autoxidation, photooxidation. Differentiation of multicellular organisms. Plants and animals completely dependent upon the <i>photosynthesis</i> with free carbon dioxide. Equilibration to present conditions. Continuous turnover of a nearly constant volume of organic matter

more about this see Gaffron, 1960a). This readjustment of our working hypothesis is mainly due to new findings in astronomy. According to Hoyle (1960), Shapley (1960), Dauvillier (1958), it is extremely improbable that life has evolved solely on this earth. Astronomical observations force us to think in terms of untold millions of earthlike planets in the cosmos, and therefore of natural organic evolution, ending in the phenomena of life, self-awareness and thought, as a part of evolutionary history everywhere. Seen this way organic evolution may be said to resemble a chess game played first with two dozen elementary particles and a hundred elements, and later with twenty amino acids and four nucleotides. The chances are extremely small, to continue with the simile, that any fully developed chess game of this kind can long remain a recognizable copy of another play (Reid, 1959). What we are interested in, however, are the fundamental rules of the game of evolution and though the middle or end plays, such as our Darwinian evolution during the last billion years, are bound to be awfully complex and always different, the beginnings ought to be practically routine, nearly the same everywhere, and perhaps sufficiently simple as to let us succeed when we earnestly try to have a hand at this game ourselves.

As everybody knows, this discussion began with the imaginative essays by Haldane and Oparin in the thirties, and since then has expanded enormously. Table I contains an outline of the hypotheses on the course of organic evolution which seem to be fashionable

and generally accepted at the present time. I have based this summary mainly on some articles in the volume on "The Origin of Life on the Earth" (Oparin, 1959) and on others which I may have read since then. It contains nothing new. Its purpose is merely to facilitate references to certain phases in the hypothetical history of the evolution of life.

Obviously there are two practical ways, pointing in opposite directions to approach our problem experimentally. We may start from the very beginning, era I. There we take the more or less established knowledge of the geochemists and astronomers as a firm basis, and start experimenting with increasingly sophisticated model reactions of the kind we imagine might have pushed evolution forward during a specified geological period. Or we begin at our time, the aerobic era V, and dig into the past of such metabolic reactions which may have a long intelligible evolutionary pedigree. If we choose to explore the history of synthetic photochemistry on earth, we can immediately go back to the borderline which divides this latest era from the last anaerobic one since era IV ended with the appearance of oxygen produced by photosynthesis. Era V saw the greater part of the Darwinian evolution.

For reasons we shall come back to later, Darwinian evolution proceeding by the familiar way of heredity and mutation in cellular organisms must have begun before the date just set, that is, deep within the anaerobic era IV. The origin of life must lie therefore much further back still. By the time photosynthesis with the release of oxygen made its appearance, the great work was long done. The date signifying the end of the history of the origin of life, meaning the moment the first "modern" cells appeared, precedes, therefore, that of the beginning of complete photosynthesis by a long time. This statement immediately gives the opportunity to ask the question, which of the main metabolic reactions we know so well were already fully developed at that time? Respiration? Of course not. Fermentation? Yes, but probably not in the way we usually think of it in connection with the glycolysis of carbohydrates. Photosynthesis? Yes, but without the trick to evolve oxygen from water that is in its anaerobic form which we call photoreduction. What makes it possible to attempt such guesses? Studies of some contemporary unicellular organisms have revealed metabolic pathways which appear to be just a little simpler, less complex, less efficient than those found in the majority of aerobic

organisms, and we have no reason to believe that these reactions have become simpler merely by reason of later evolutionary losses. In some algae the simpler pathway can be found existing side by side with the latest development.

Algae and bacteria are commonly spoken of as primitive organisms because they occupy the lowest rank on our own evolutionary scale, but this scale applies only to multicellular differentiation. From a biochemist's point of view, the unicellular organisms are hardly less complex than any others. With them the fundamental reactions like fermentation, respiration, and photosynthesis are fully developed. After two billion years of era V there are no true primitive organisms left on earth. However, there is one important difference between the lower and higher organisms. The latter have lost most of such traits which living things had when our aerobic era V began and which they brought as a specific inheritance from the preceding anaerobic era IV. Despite the passing of a couple of billion years, the anaerobic photosynthetic bacteria and unicellular algae possess metabolic characteristics which strike us as belonging to the pre-aerobic era IV. Of these organisms, which are sitting on the fence as it were, there are two broad classes: those that thrive only in the absence of oxygen, while not being killed outright by mild aerobic conditions, and those which are already aerobic organisms but don't mind at all being put under anaerobic conditions where they continue to metabolize actively. It is among these organisms where we like to go looking for memories of earlier evolutionary days.

Recent studies of species of green and blue-green algae and dinoflagellates have uncovered an intriguing variety of transitions and recombinations between aerobic and anaerobic metabolic patterns, and below I shall describe a new example of how it has become possible to ascertain experimentally that the most important of fundamental reactions, photosynthesis, evolved by a stepwise acquisition of already specialized enzyme systems.

The photochemical core, on the other hand, the pigment complex with chlorophyll as the main component, in all probability was functioning efficiently, that is, transforming light energy into chemical action, long before the first cell or even the first true enzyme was completed. In other words—as I try to indicate in Table I—we assume that porphyrin photochemistry has not only accompanied but decisively influenced the results of evolutionary

processes from very early times and became important during era II. Can we prove this? Here I have to register a disappointment. The model reaction which establishes an easy spontaneous synthesis of porphyrins under conditions of era I or II has, to my knowledge, not yet been announced. Meanwhile the famous Miller-Urey experiment is nine years old. Impressed, as all of us were, by the easy way aliphatic and amino acids are formed by electrical discharges in a methane-ammonia atmosphere (Miller, 1955; Pavlovskaya and Pasynskii, 1959) and on the other hand by the surprisingly simple way the living cell constructs porphyrins from the very products of Miller's reactions (Shemin, 1956), I suggested in 1955 at the Tucson Conference on the Use of Solar Energy that porphyrins which are excellent photosensitizers in laboratory experiments, might have played a decisive part early in evolution—practically the same they play at the present time (Gaffron, 1957, 1958).

The transition between era I and era II was brought about by the loss of all excess free hydrogen. Thereupon the direct decomposition of water vapor in the ultraviolet produced sufficient oxygen and ozone to shield the earth's surface from further effects of the first important source of energy. It seemed that only much later, after modern photosynthesis appeared, light could resume the evolutionary role originally played by ultraviolet, but now at lower frequencies and consequently with a lower energy content per quantum absorbed. If, however, porphyrins can be formed spontaneously from the component of Miller's brew it is obvious that evolution must have been speeded up considerably by these fluorescent pigments, particularly because daylight is available in much greater amounts than ultraviolet. At the same time the action of visible radiation is much milder, less destructive, less "sterilizing" and yet sufficient to promote any number of electron (hydrogen) transfers, oxidoreductions and outright oxidations between small or large organic molecules. I still believe that this is not only a likely but a necessary hypothesis to account for the speed of evolution.

In the meantime an impressive number of biologically important chemicals have been synthesized by means clearly compatible with the environmental conditions of era I. Let us assume then, for argument's sake, that all *small* molecules serving as building stones and catalysts for the basic metabolism of living cells could have arisen spontaneously, given the time. What is the difference in

meaning of the new experiments as compared with, say, Wöhler's synthesis of oxalic acid or urea a hundred and thirty years ago? Only that they now are modelled intentionally to suit the conditions we believe to have existed during eras I and II. As to the way life originated on earth, they provide no more insight and direction than the finding of a heap of stones and bricks provides for the understanding of the designs, construction rules and mechanical laws which make it possible to build a bridge or a cathedral.

The fact that these simple chemicals may have been available in quantity is, of course, very comforting, meaning an invitation to approach the true problem with vastly greater confidence. A negative outcome of all attempts to produce the right kind of organic chemicals under geochemical conditions would have been a hard blow. But it should be remembered that the substances which did accumulate during era I were as dead then as they are now when we can produce them (and many more complex ones) in any desired quantity.

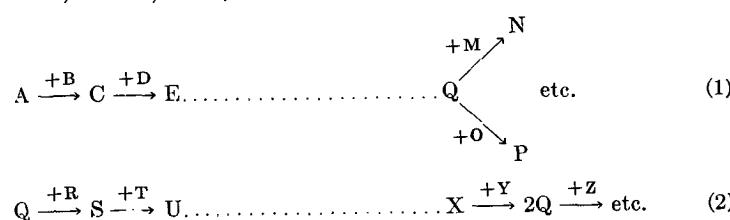
II. Organic Catalysis and Reproduction

When we say that what happened during the first and last era of our timetable is now understood in principle it does not mean, of course, that concerning these periods there are no more important scientific problems left to solve. It is only a way of saying that such problems can be formulated in an intelligible way, while concerning the puzzle we know is hidden among the events of era II and particularly era III we are struggling even with the task of formulating it properly.

The most famous riddle is the appearance of proteins and nucleic acids in systems that reproduce themselves. Proteinlike and nucleic acid-like high polymers can be obtained by very unrefined means. But a catalyst of the family of known enzymes which does something quite specifically on account of its detailed internal structure and configuration, can never have arisen spontaneously by the forces at work during era I. This seems to be a tenet now agreed upon (Anker, 1961). As long as a selection of specific other enzymes are around, however, the *de novo* synthesis of a particular enzyme from its smallest components may not require the participation of the entire cell (see, in particular, Schramm *et al.*, 1962).

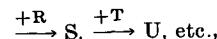
Only the reproduction process in its entirety still does, because the synthesis of one catalyst depends on that of other catalysts,

which in turn . . . and so on. There seems to be no other way out than to assume, as Anker does, that catalysts as well as reproductive cycles were originally much simpler, necessarily less specific and less efficient. How enormous the biologically induced acceleration of a certain metal catalysis can be, provided there are suitable proteins around, has been pointed out by Calvin (1956a,b). This still leaves us with the riddle of protein formation. Therefore artificial homemade metal-complexing agents such as quinolines, combined with equally artificial molecular surfaces as produced by detergents, a combination which, according to Lowe and Phillips (1961) catalyzes metalloporphyrin formation, are very much worth our attention. It is the complete disappearance of earlier stages which makes the riddle that confronts us especially difficult. When we try to reinvent them we should not expect or hope that what we first achieve is "natural," i.e., historically correct. Any substance Q which can be produced by a succession of steps permissible under the limitations imposed by era I or era II and which subsequently can initiate and become part of a sequence of reactions where it reappears again is quite obviously self-reproducing (Allen, 1957; Ycas, 1955; Anker, 1961).



In (2) all substances Q, S, U X are "reproducing" at the expense of food and with the aid of intermediate metabolites symbolized by R, T Y. Polymerization of preformed similar or identical constituents is not self-reproduction, nor is autocatalysis meaning acceleration of type (1) reaction by catalytic action of the product Q. Also the question of the particular source of energy to keep either class of synthetic reactions going is secondary to the general principle.

Biochemistry is traditionally concerned with the nature of the compounds in the sequence



and their chemical interactions. Evolutionary biochemistry has the task of reinventing sequences of type (1) leading to all sorts of Q capable of initiating reactions of type (2). Further evolutionary steps have deepened the problem, for in the cell hundreds of self-reproducing cycles have become interwoven like the words of a crossword puzzle. At the present time Q_1 is unable to reappear unless Q_2 is ready to do likewise, and so on. This oversimplification of a central problem about which so many details are known and for which we have no demonstrably valid general rule—a condition which parallels the state of the question on the origin of species around 1800—should serve us here only to point out that simple reproducing cycles may have been functioning at an early stage as Anker has emphasized.

The two types of reaction—the straight synthesis and the cyclic replication—could have appeared simultaneously if conditions were ready for (2) and waiting, as it were, for (1). Then (2) being more efficient (and speculating about evolution we assume that it was) would have taken over, causing reaction (1) to disappear from view. Or the two reaction types appeared at different times separated by many millions of years and across the change from one era into another. For example, the straightforward formation of compounds to serve later as Q's may have depended on short wavelength ultraviolet light available during era II, while the appearance of cycles had to wait for the conditions present during era III. This automatically would have closed the books on the initial evolution of Q, leaving a contemporary observer with the simpler edition of the puzzle we face today. Once the "cycles of the Q's" had come into existence they multiplied at the expense of the free food, while simultaneously evolving into more and more complex systems. When the original organic matter was nearly exhausted, only a new kind of photochemistry could supply both energy and suitable carbon compounds. Thus the course of evolution first selected, and then depended upon, a special type of organic photochemistry.

III. On the Efficiency of Photochemical Reactions During Evolution

A major part of the problem concerning the evolution of metabolic reactions is how they became more and more efficient, faster and more specific as compared with similar reactions *in vitro* while proceeding at lower temperatures. The attempt to understand the

evolution of photosynthesis adds to the problem of catalytic enzymes another of a peculiarly contrasting type, namely how not to lose but to preserve the potential high rate and efficiency of a photochemical reaction. The quantum yield of a light absorption act which produces an excited high energy state is one—by definition.

The probability with which organic dyestuffs can subsequently initiate a chemical reaction after their absorption of a light quantum varies from zero to (practically) unity. Light-excited chlorophyll is one of the molecules which may, *in vitro*, given the proper reactants, produce a chemical reaction 95 times out of a hundred. (This we shall call a quantum yield of one, leaving aside possible chain reactions.) The same is true for some of the porphyrins of which we postulate that they have been formed during era II (Gaffron, 1927). Present-day photosynthesis in living cells is so complex that eight quanta, i.e., eight primary processes in eight separate chlorophyll molecules, contribute to the effective release of one molecule of oxygen from water. [For the latest reviews of the notorious quantum yield controversy see Kok (1960) or Gaffron (1960b).] We can make it sound simpler by saying that two light quanta are needed to produce the equivalent of a hydrogen radical, plus one hydroxyl in the state roughly equivalent to half a hydrogen peroxide. Two is the smallest number of quanta of visible light that contains sufficient energy to do the job, since one is not enough. Conditions can be found where each and every chlorophyll molecule in a living chloroplast is doing its duty. Despite the presence of a number of other associated pigments, essentially only one type of molecule, chlorophyll (a magnesium porphyrin), is photoactive in the world and by its photochemistry supplies all the food there is. This can mean only one thing. The advantages chlorophyll has held or gained over likely competition during the evolution of life must have been considerable, and at least as great as that of iron porphyrins in the realm of oxidoreduction catalysis. Its predominant position was definitely established before the last era V, for chlorophyll photosynthesis caused the change from era IV to V. Leaving aside numerous experiments on inorganic photocatalysts, and the many speculations based upon them, one might think of at least three reasons why chlorophyll has won in competition with other organic pigments. First, only the structure of chlorophyll permits a nearly 100% efficient internal conversion of absorbed light energy into excited states leading to useful chemistry (quantum yield of

~ one). Second, only the structure of chlorophyll combines a minimum of internal losses by heat or radiation with the function of an oxidoreduction catalyst, thus making the coupling to other enzyme systems extremely efficient. Third, the porphyrins came first and in quantity, and being as good but more stable than other pigments, lorded over the evolution of living photochemistry ever since. The first reason is not very compelling. Pigments with an entirely different chemical structure may also produce a quantum yield of one in laboratory experiments. The second reason is also not strong because flavin, so different from chlorophyll, is the paradigm of a natural oxidoreduction dye with pronounced photochemical activity, excellently suited for action on protein surfaces, as we shall see in the model to be described below. Reason three, then, is our choice explanation. It is so very simple yet supported by the circumstance that science has found no class of living beings devoid of porphyrins. Furthermore, in a struggle for survival among light-excited pigment molecules, chlorophyll is likely to survive longest since it is apt to photooxidize, if only traces of oxygen are present, other pigments first before it succumbs itself to photooxidation. It is not difficult to invent still more reasons why chlorophyll dominates the organic world—such as the need of overlapping absorption bands among several nearly related porphyrin pigments in one complex for an efficient charge transfer, etc., but these would be late refinements in the course of evolution. And the plainest hypothesis suffices for the moment.

The contention is, then, that the intrinsic potentiality for a photochemical reaction approaching a quantum yield of one has been put to good use all through the organic chemical evolution beginning with era II through era IV when photosynthesis became what it is now. There must have been two major triumphs in the evolution of living photochemistry. One is the avoidance of useless back reactions among the immediate products of the primary light step. One-third of the absorbed energy is tucked away by the plant as sugar and oxygen. To achieve this result each and every precursor substance on the way either to carbohydrate or to free oxygen is prevented from disappearing into side reactions or from recombining with its counterpart without delivering useful work. The other is to combine the effects of two light quanta in order to achieve what the energy of one quantum alone could not possibly deliver, the photolysis or photooxidation of water.

The laboratory experiments with a quantum yield of one are always exergonic, the primary reaction products part for good. It is like pushing someone over a precipice—the probability of his coming back immediately is small; whenever the over-all photochemical reaction is endergonic the quantum yields are very disappointing—the primary reaction products roll back from their higher energy levels like the stone of Sisyphus, and recombine for lack of an intervening molecular trapping device. In the living cell these problems have been solved by arranging the pigment molecules on the surface of a macromolecular structure, and by making the evolution of oxygen, the most difficult task of all, a photochemical two-step arrangement (see Rabinowitch, 1956; Kok, 1960; Franck, 1960; Gaffron, 1960b). When in the course of evolutionary history were these two fundamental inventions made? The evolution of photosynthesis must have taken a decisive turn the moment protein-like substances appeared. Here our problem merges with that of the evolution of proteins in general. And it is as a contribution to the problem of the beginning interplay of photochemical and enzymatic reactions that I shall describe below a peculiar photochemical model reaction.

The evolution of the photosynthetic apparatus came to an end, that is into the shape it has now, when it acquired the capacity to produce free oxygen. Many types of reactions had to join hands one by one to produce the final result. Which among them was the last addition that spelled the end of the anaerobic era? Everybody knows that the green plants release the oxygen from water. Hence the obvious answer that may occur even to a layman will almost certainly be that the last evolutionary step must have been the introduction of water, H_2O , to serve as hydrogen donor where up to then this task had been accomplished by substances like H_2 , H_2S , or butyric or lactic acids.

This answer would have been perfectly acceptable thirty years ago when van Niel (1931) put research on photosynthesis on the right track for the first time. However, since many years we possess additional and sufficient evidence to show that this answer is too superficial and does not explain what we know to be true. Nevertheless it is being right now presented again in the same appealing and unsophisticated form as a new thought (Losada, Whatley, and Arnon, 1961).

The last section of this article is therefore devoted to more experi-

mental evidence which, like other facts we have, points to a long evolutionary history of the water-decomposing mechanism, reaching perhaps back to the pre-cellular era.

IV. An Artificial Flavo-Mangano-Heme Photooxidase

The following is based on the outcome of studies of interaction among some of those factors which are commonly used to fortify chloroplast reactions. (1) *Ascorbic acid*. The recipes for various experiments with chloroplasts call for ascorbic acid to keep a certain balance between the oxidation stages of other electron transport agents (Losada, Trebst, and Arnon, 1960). (2) *Flavin (or FMN)* is a factor playing a part in one type of cyclic photophosphorylation (Arnon, Whatley, and Allen, 1959). (3) *Catalase*. In experiments done in this laboratory Mehler (1951) used Keilin-Hartree's (1945) catalase-ethanol mixture to remove peroxide formed during the reduction of free oxygen by illuminated chloroplasts. (4) *Manganous ion*. Gerretsen (1948) showed that manganous ion influenced potentials which arose during the illumination of green plant extracts. Kessler (1955) found that the strongest effect of a lack of manganous ion on photosynthesis could be traced to reactions concerning the release of oxygen (see also Lowenstein, 1958). (5) *Chloroplast preparations from Helianthus annuus*.

It is known that flavins are light sensitive and decompose when illuminated aerobically or anaerobically. It is also well known that fluorescent dyestuffs photooxidize ascorbic acid. One example is photooxidation by flavin, which, however, is a more complex reaction because of the high photosensitivity of the sensitizer itself. Catalase-ethanol prevents accumulation of hydrogen peroxide should it be formed in the course of such a photooxidation. Thus if we illuminate a solution of the substances mentioned under (1), (2), (3) above in the presence of air, the ascorbic acid is oxidized rapidly to dehydroascorbic acid. Results of this kind are so well known that we may call them trivial.

The moment manganous ion is added to the mixture as the fourth component, the aspects of the reaction change remarkably and are now by no means trivial. The photooxidation of ascorbic instead of stopping as usual at the dehydroascorbic acid level continues. A second equivalent of oxygen is taken up and the dehydroascorbic is decomposed also. The two oxidation steps do not occur simultaneously or follow each other in immediate succession. They are

neatly separated by a third intermediate reaction which takes a definite time to run to completion without an uptake of oxygen. This three-step phenomenon appears only when the components (1)-(4) mentioned above are present (Habermann and Gaffron, 1961, 1962). A more detailed analysis which takes too long to summarize here, supports the following description (see Fig. 1). Ascorbic acid, the original substrate, is photooxidized at the surface of a

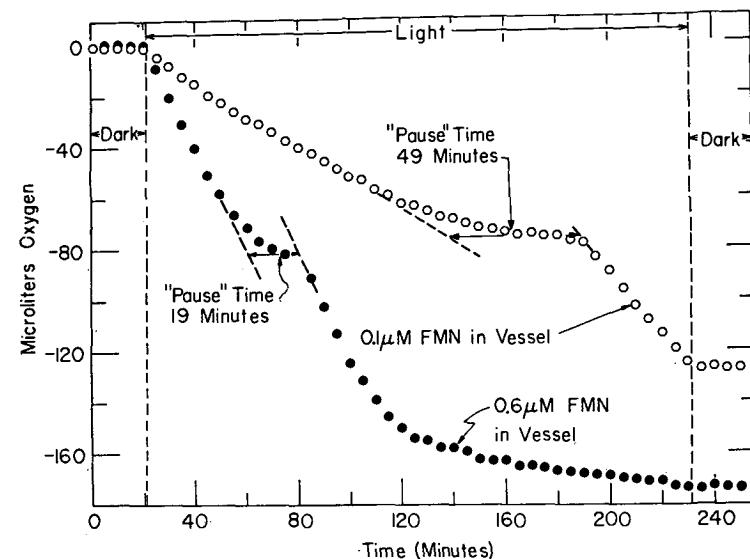


FIG. 1. Photooxidation of ascorbic acid in three separate consecutive steps by a combination of flavin (FMN), catalase, and Mn^{++} (after Habermann and Gaffron, 1961, 1962).

flavin-mangano-catalase complex until the supply is nearly exhausted. Meanwhile the oxidation product which, as yet, is *not* stable dehydroascorbic acid, accumulates. Only after the concentration of the original ascorbic acid has become very low begins the second step, the photoreaction without oxygen which transforms the first oxidized intermediate. The latter is converted into dehydroascorbic acid proper. Again it happens that the product of this conversion, the dehydroascorbic acid, is left to accumulate. It must wait its turn until the preceding form has mainly disappeared. Now the third reaction begins—the second photooxidation. As long as the light intensity is low and determines the rate this second photo-

oxidation, that of dehydroascorbic acid, is likely to proceed even faster than the initial photooxidation. It is this second photooxidation which demands specifically the presence of catalase and manganous ion, in addition to flavin. It would not go otherwise. Ascorbic acid can be photooxidized, though with lower quantum yield, in pure solution as mentioned above, and the existence of a precursor form to dehydroascorbic acid seems to have been observed before (Nason, Wosilait, and Terrell, 1954; Kern and Racker, 1954).

The sudden increase in complexity and specificity of an ordinary photooxidation by the addition of a metal protein can be extended further by coupling it to a much larger inhomogeneous molecular cluster—namely to chloroplasts or their broken parts. The preparations were made from *Helianthus annuus* (sunflower) leaves and, for all we know, they were completely inactive in those tests for the various photometabolic studies which are presently en vogue.

The addition of chloroplasts did not change the character of the described triple reaction sequence. It merely shifted the range of the effective light absorption into the red corresponding with the spectrum of chlorophyll. "Native" chlorophyll, while relieving the flavin completely of its role of sensitizer, did not replace it in its chemical function. FMN could not be left out of the combination. The nature and kinetics of the energy transfer between the two pigments remain to be explored—as so many other facets of this experiment.

What interests us here in the context of the problem of photochemical evolution is the considerable change which the introduction of only one protein species brings about in the interplay of five simpler substances, iron porphyrin, manganous ion, flavin, oxygen, and ascorbic acid. It imposes rank, order, and specificity. A photoactive pigment which absorbs visible light must be present, while ultraviolet light likely to destroy some of these molecules by direct action must be absent. In our timetable the period when many kinds of organic substances were already present and were now gradually subjected to the change-over from ultraviolet to visible light as the main energy source is given as era II.

From the point of view of evolution, the unsatisfactory aspect of our model is the lack of "synthetic" qualities. It is overall merely a photooxidation. We would like to replace its strong exergonic character by that of a light driven formation of peroxide or an oxidation-reduction, with some demonstrable gain in ΔF (Krasnovsky, 1948,

1960). Another task would be to replace the natural hemoprotein by an artificial high molecular product (Fox, 1960; Fox and Harada, 1961; Lowe and Phillips, 1961). This would much increase the importance of our pseudo-metabolic light reaction as a means to gain more insight into what might have happened during era II, the early middle ages of the history of life.

V. The Appearance of Oxygen as the Result of a Mutation in a Photoreducing Organism

Literally hundreds of publications on the biochemical aspects of photosynthesis have appeared during the last two decades (Rabinowitch, 1956; Pirson, 1960; Gaffron, 1960b). One major point discernible even before 1940 has now been definitely solved. It is that the assimilation and reduction of carbon dioxide in plants is the result of an association of several enzyme systems. They resemble closely those metabolic processes which in non-photosynthetic organisms constitute parts of respiration or fermentation or other more general chemo-synthetic processes. These enzymatic systems cluster around the one part of the photosynthetic apparatus which is unique to it: the light-absorbing chlorophyll complex. By comparison our knowledge of the latter has advanced little beyond what has been surmised, discussed, and theoretically postulated many years ago.

The famous work of Calvin and Benson, which established in minute detail the path of the transformation of carbon dioxide into sugar via more than a dozen different steps, gives only incidental and rather remote information about the products of the primary light absorption process (Bassham and Calvin, 1960). The reason is the looseness of the connection between such enzymatic reactions and the photochemistry proper. It is surprising in how many ways the complete process of photosynthesis can be divided into separate partial reactions without interfering with the course of the one reaction line which sets the photosynthetic organism apart from the rest of the living world, namely the absorption of light quanta and the extremely efficient fixation of this electromagnetic energy in the form of potential chemical energy. Easiest to cut off is the Calvin-Benson cycle, the synthesis of carbohydrates. In such a case the chlorophyll complex is perfectly able to release oxygen, that is, to decompose (or oxidize) water in the light, provided a number of physiological and unphysiological substances are simultaneously

reduced. Or it is possible to block the evolution of oxygen quite specifically by various means and to show that the photochemical reduction of carbon dioxide continues unharmed provided another outlet exists to take care of the dehydrogenation products of water. If both these activities, reduction on the one hand and oxidation and oxygen evolution on the other, are eliminated, the harnessing of light energy may still proceed. It is a particular trick of the photochemical system that any back reactions arising from an incipient accumulation of intermediates when they are not used for external oxidoreductions are so guided as to yield energy-rich phosphate (for example ATP) whenever the necessary enzymes and phosphate acceptors are present (Arnon, 1961). When this last way for biochemical action is cut off, spectroscopic evidence still proves the existence of initial oxidoreductions within the pigment complex (Chance and Nishimura, 1960; Witt *et al.*, 1960, 1961). We cannot ask for better evidence in favor of the hypothesis that the complete mechanism as it exists in the green plant is the result of a prior acquisition of one synthetic property after another (Krebs and Kornberg, 1957).

Here we have therefore a fundamental metabolic process which could give us clues as to the events that occurred long before Darwinian evolution started. Ever since van Niel in his studies on the comparative biochemistry of photosynthetic organisms drew our attention to the similarities which exist between the oxygen-evolving plants and the anaerobic but photosynthetic (photoreducing) bacteria, it has been assumed that the latter constituted an earlier, slightly more primitive form of life. And it is generally agreed that the last evolutionary step in the development of photosynthesis was the reaction which permitted the release of free oxygen. For what follows we ought to keep in mind that bacteria and plants share the main types of photochemical catalysts, such as chlorophylls, cytochromes, and carotenes. Carbon dioxide is fixed via phosphoglyceric acid. Cyclic photophosphorylation exists in both. Among the dissimilarities we should note the predilection of the bacteria for an acetate-fat metabolism as compared with the primarily carbohydrate-oriented intermediate metabolism of the plants. The contrast between anaerobic and aerobic environmental conditions in which these two classes of organisms prefer to live is less sharp when we find that some purple bacteria, though a minority, can tolerate or even grow with the aid of a respiratory mechanism,

while some of the green algae tolerate oxygen, yet are unable to grow in the dark heterotrophically, since they have no useful respiration. Others depend on acetate and light for growth (Pringsheim and Wiessner, 1961). In short, all known unicellular organisms that use light energy have so many traits in common and the differences from species to species or strain to strain overlap so much that it is difficult to assign to such biochemical deviations a truly fundamental importance. Thus we have only a small choice of traits to choose from if we want to find the reason why some of these organisms cannot produce oxygen in the light while others do it easily.

If we want to undertake in earnest, that is by experimental means, this trip back into the history of the photochemistry in living cells, we have to start with what is clearly accessible and formulate our question precisely. Which particular partial reaction was the finishing touch in the mechanism for the release of oxygen? On the correct answer hinges the interpretation of the mechanism for what is loosely called the photolysis of water, a word signifying that light is essential for the release of oxygen which arises from water. There is no metabolic process known by which oxygen is continually produced from water in total darkness. Thus here we have a truly fundamental point. Whoever prefers to allude to the likely mechanism of this reaction may call it the dehydrogenation or the photooxidation of water. It is unfortunate that such semantic predilections are often presented as if they truly contributed something new to the discussion (cf. Baur, 1930).

In contrast to the other partial reactions of photosynthesis the evolution of oxygen has so far not been separated as a recognizable partial reaction system from the photochemistry. We do not know where the reaction starts nor the intermediate which is decomposed by the manganese enzyme. So closely is it bound to the chlorophyll complex that the capacity to evolve oxygen remains intact after chloroplast preparations have been washed and have lost the ability either to fix carbon dioxide, or reduce triphosphopyridine nucleotide (TPN) or store energy-rich phosphate bonds. This chloroplast activity also survives dry-freezing and extraction and reactivation with lipid solvents. But oxygen evolution disappears together with all synthetic power when the chloroplasts are mechanically broken down into particles containing less than 400 chlorophyll molecules.

Thus we find the reaction with water more closely bound to the activity of the pigment complex than any other of the partial reac-

tions. It is for this reason that we say the dehydrogenation of water is the first achievement of the primary process (Stoll, 1932). Nothing certain is known about the specific mechanism. Kinetic and chemical as well as spectroscopic data have implicated several cytochromes, carotenes, plastoquinone, and manganous ion besides the light-absorbing pigments. Except for manganese the same kind of substances are present in the active particles extracted from purple bacteria.

Why then are purple and green bacteria incapable of evolving oxygen from water? The question is as old as the thirty-year-old discovery that they just don't do it, and so is the obvious answer: not water but other substances serve as hydrogen donors. Why should they, under these circumstances, release oxygen? But this is a tautology and not an explanation. The point is that something new has been added to an anaerobic mechanism to produce a result which deviates in one crucial point from what the organisms already were able to achieve before.

In countless experiments, it has not been possible to induce any photoreducing bacterium (Thiorhodaceae and Athiorhodaceae), as well as green bacteria (*Chlorobium*) to produce oxygen like the plants, not even in traces. Their reduction of carbon dioxide stops the moment the supply of an appropriate hydrogen donor becomes exhausted.

On the other hand many algae can be easily persuaded to give up the production of oxygen in the light during photosynthesis, provided there is some hydrogen available which they then use to assimilate carbon dioxide exactly as before. Thus in these algae at least—and there are now many forms known to contain the necessary hydrogenase—the evolution of oxygen is something which can be taken up or put aside without any noticeable modification in the fundamental photochemical process. Two interpretations are possible: we may say that each time the alga reverts to the metabolic form of its anaerobic ancestor it stops to oxidize water and changes over to a reduction of TPN (or of phosphoglyceric acid) by a more direct photocatalytic hydrogen (or electron) transport. Or the switch we observe is merely a premature reduction of intermediate precursors of oxygen ("photoperoxides" alias "oxidized cytochromes," etc., etc.). The latter has been my view since the phenomenon was first seen and it entails unavoidably the hypothesis that water is oxidized not by a remote and removable dark reaction but as an immediate

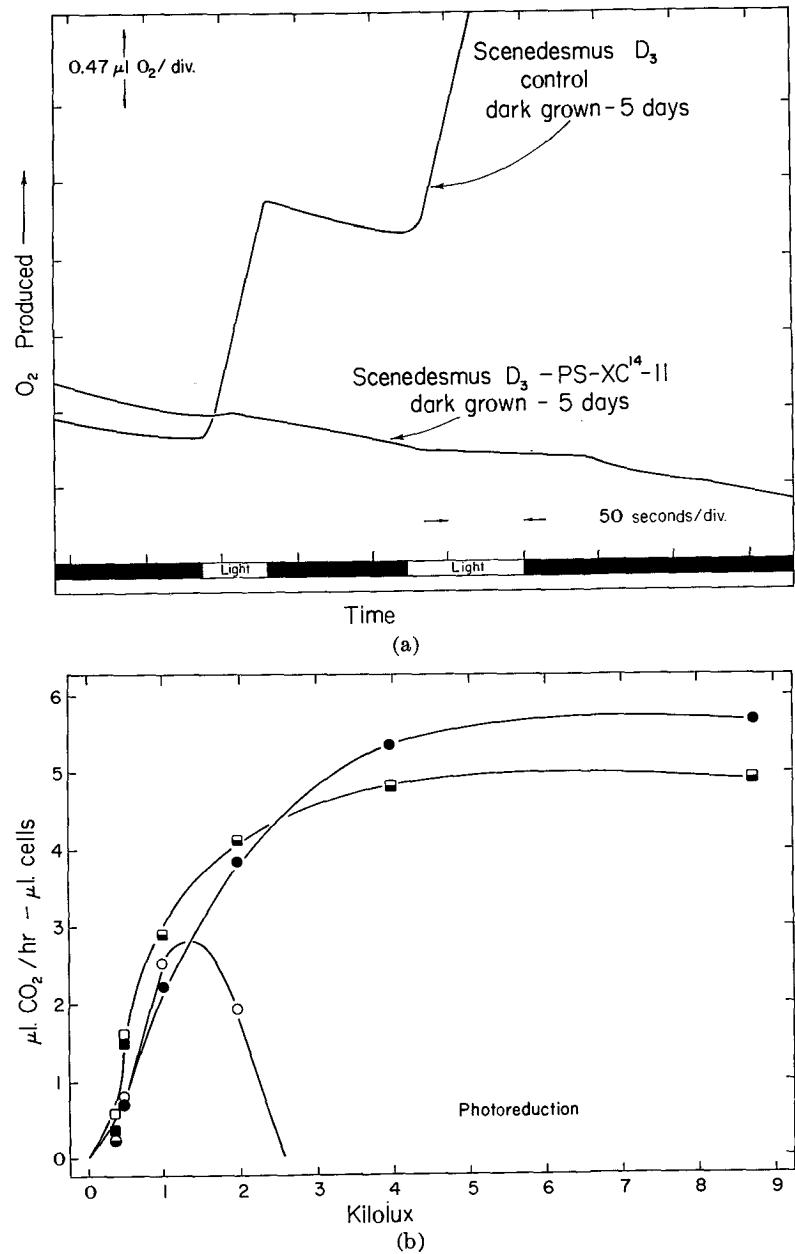


FIG. 2. For descriptive legends see opposite page.

ON DATING STAGES IN PHOTOCHEMICAL EVOLUTION 79

consequence of the photochemistry. It has been possible to inspect more closely the circumstances which cause a separation of the oxygen-evolving set of reactions from the rest of the photochemical oxidoreduction.

The first, least severe interference consists simply in the activation of a hydrogenase under anaerobic conditions. The mechanism for the release of oxygen remains fully active. The specific activity of the hydrogenase, the hydrogen partial pressure, and the light intensity (= concentrations of oxidizing photoproducts) determine together what percentage of the oxygen of photosynthesis will be prevented from appearing. At low light intensities only very little oxygen escapes, at high light intensity nearly the expected amount. Provided all the oxygen is blown away in a stream of hydrogen this situation remains perfectly reversible and stable. The obvious explanation is that we witness a competition for a common intermediate (Horwitz and Allen, 1957; Gaffron; 1960b, p. 188).

The next degree is reached by blocking one of the competing pathways. Normally the path to the hydrogenase is closed because the latter is inactive in air. After anaerobic adaptation to hydrogen it is easy to block completely the other side, the way to oxygen, by specific poisons. Under these circumstances oxygen does not appear at all and the rate of photoreduction can thus be increased until a saturation rate is reached characteristic for the condition of the hydrogenase system in the particular kind of algae under investigation. A very similar result is obtained by a reverse procedure—not by adding an inhibitor but by removing a factor necessary for oxygen evolution—namely manganese ion. Manganese-deficient

FIG. 2. Genetic separation of the oxygen-evolving system from the photochemistry and the reduction of carbon dioxide in a mutant of the green alga *Scenedesmus D₃* (after Bishop, 1962b).

FIG. 2a. Polarographic traces of oxygen consumption and production by dark-grown normal *Scenedesmus* and mutant 11. Gas phase = 4% CO₂ — 96% N₂; Temperature = 20°C; Cell volume = 15 μl; pH = 6.5; Phosphate buffer = 0.05 M; Final volume = 2 ml; Light intensity = 4 kilolux.

FIG. 2b. Rates of photoreduction of dark-grown normal *Scenedesmus* and mutant 11 at various light intensities with and without 3-(3,4-dichlorophenyl)-1,1-dimethyl-urea (DCMU) (2×10^{-6} M). Temperature = 25°C; Phosphate buffer = 0.05 M; pH = 6.5; Cell volume = 30 μl/vessel.

○—○ = control; ●—● = Control + DCMU;
□—□ = Mutant 11; ■—■ = Mutant 11 + DCMU.

algae photosynthesize poorly but give an excellent rate of photoreduction (Kessler, 1957). Both the poison treatment and the deficiency effect are perfectly reversible. The plants act like bacteria only as long as they are prevented from making use of their faculty to release oxygen.

The third degree of separation is final, irreversible. Here the separation is accomplished on the basis of genetics. Bishop (1962a,b) has searched for and found an X-ray-induced mutation in *Scenedesmus* with a property specified in advance. The desired mutant was to be normal in all respects except one: the ability to release oxygen in the light. The mutant strain has indeed the normal complement of pigments. It respires, grows, and reduces carbon dioxide exactly like any other green alga. But growth has to be maintained by heterotrophic nourishment in air (light or dark) and the photochemical reduction of carbon dioxide takes place only under adapted conditions in hydrogen. Figure 2 compares the activities of normal and mutant *Scenedesmus*. The rate of photoreduction is the highest so far observed for *Scenedesmus* while the production of oxygen in the light is nil.

As those who have done mutation experiments know, it will take a longer time to find the corresponding back mutation, a spontaneous reappearance of the ability to release free oxygen. But it will be interesting to try.

What has the story of this mutation to do with our evolutionary problem? I believe it means that when everything was ready and set one single mutation step ended the anaerobic era, and that it was *mutatis mutandis* the reverse of the X-ray effect which in Bishop's strain obliterated the ability to evolve oxygen.

One look at an alga and a purple bacterium suffices to see that both types of organism have after one or two billion years of evolution drifted apart considerably. But we cannot help thinking that the deviation is least in the photochemical mechanism. Even here we find more than one conspicuous difference and each one of them may be important or decisive for the ability to oxidize water and release oxygen. Algae absorb the visible red light not the infrared as the bacteria do. This may give them a crucial advantage of 10 kcal in the effective excitation level of their pigments. But this difference alone is insufficient as an explanation since green, red-absorbing, sulfur bacteria do not evolve oxygen. The set of cytochromes in both organisms are not identical (Kamen, 1956; Hill and Scaris-

ON DATING STAGES IN PHOTOCHEMICAL EVOLUTION 81

brick, 1951); yet this difference may again not be decisive though probably essential on an intermediate level. Two characteristics, however, seem to be the exclusive property of the algae, the so-called Emerson effect and the need for a relatively high level of manganous ion. The requirement for manganous ion is at least a hundred times less in purple bacteria than in aerobic algae, and in

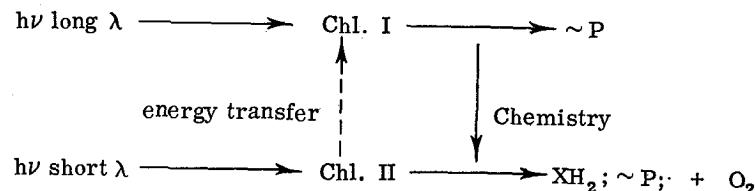


FIG. 3a. Emerson Effect. Two photochemical reactions which originate in two differently bound chlorophyll *a* molecules cooperate in the release of oxygen. The effective excitation level of Chl. I corresponds to absorption of light with wavelengths *longer* than λ 700 m μ . The effective excitation level of Chl. II corresponds to absorption of light of wavelengths *shorter* than λ 700 m μ . Excitation of Chl. II leads by energy transfer also to excitation of Chl. I. Therefore light absorbed only by Chl. II suffices for photosynthesis. Light absorbed exclusively by Chl. I does not.

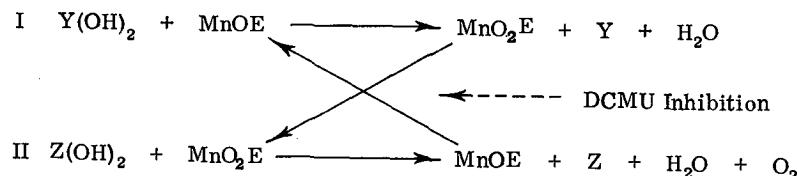


FIG. 3b. Without a manganese enzyme (Kessler, 1955) there is no evolution of oxygen during photosynthesis. The reduction of carbon dioxide may continue only in the anaerobic form of photoreduction. The role of the enzyme MnOE is here depicted as catalyzing a dismutation between the "photoperoxides" originating in the photochemistry of chlorophylls I and II.

the latter it disappears when they shift into photoreduction. Furthermore the astonishing interplay of two excitation levels in differently bound chlorophyll *a* discovered by Emerson in algae is reportedly absent in purple bacteria (Emerson, Chalmers, and Cederstrand, 1957; Emerson and Rabinowitch, 1960; Duysens, Ames, and Kamp, 1961).

Figures 3a and 3b may serve to illustrate what the Emerson effect is about and how I would like to interpret the function of the

manganese enzyme. Emerson (Emerson *et al.*, 1957; Emerson and Rabinowitch, 1960) discovered that in green plants light of a wavelength longer than $700 \text{ m}\mu$ is inactive, though absorbed. For itself alone its energy is lost for photosynthesis. If however light of shorter wavelength $< \lambda 680 \text{ m}\mu$ is given simultaneously, the light at $\lambda 700 \text{ m}\mu$ becomes fully effective. The tentative interpretation is that two forms of the same chlorophyll *a* have to be brought into two different but cooperating excitation levels and yield different primary oxidation products of which at least one has a lifetime of seconds. Both are needed to accomplish the evolution of oxygen from water. The manganese enzyme catalyzes the dismutation which in its absence does not take place. The "photoperoxides" then disappear via side or back reactions (French and Fork, 1961).

Franck prefers another interpretation (private communication). According to him the Emerson effect does not signify a different and necessary role of a chlorophyll absorbing at $\lambda 690\text{-}710 \text{ m}\mu$ as given in Fig. 3b but merely the more or less accidental oxidation of chlorophyll in its *n*- π excited state under aerobic conditions—in other words a side reaction. Our newest experiments with deep red light under anaerobic conditions and with Bishop's photoreducing mutant reveal a better yield than in air. Whether this should be taken as a support of Franck's way or would be compatible with a more detailed analysis of our scheme we are not yet prepared to say. Manganese analyses have also shown that the manganese content of normal and photoreducing *Scenedesmus* cells is not significantly different. But we do know that without a manganese enzyme the plant will not evolve oxygen, and this enzyme the purple bacteria seem to lack.

It stands to reason that every one of the discernible differences between the photoreducing bacteria and the photoreducing algae has been the result of at least one mutation step which prepared the way for the production of those oxidized intermediates which on contact with the manganese enzyme release oxygen. But until this enzyme was added as a last step neither the increase in the energy levels of the excited chlorophyll by the shift in color nor the addition or substitution of one or more of the cytochromes, nor the incorporation of water somewhere into the oxidoreduction mechanism were sufficient to bring about the final shift from photoreduction to photosynthesis with the release of oxygen. The addition of the manganeseous enzyme, or a closely related as yet unknown catalyst, to the

already existing "green" chlorophyll system was the mutation that ended the anaerobic era.

What does this all mean in regard to the participation of water in photosynthesis? Arnon *et al.* (1961) would like us to believe that the precursors to free oxygen arise all the way from water not earlier than at the parting of the ways marked by the adaptation phenomenon, simply by the substitution of H_2O for H_2A (the hydrogen donors capable of serving in various photoreductions).

To quote Arnon *et al.* (1961): ". . . The differences between bacterial and plant photosynthesis seem to centre on the electron donors that are consumed in the reduction of pyridine nucleotide. With some electron donors, as, for example, hydrogen gas, the reduction of pyridine nucleotide by photosynthetic bacteria is a dark reaction. . . . Oxygen evolution is peculiar to green plants because it is the product of an auxiliary photochemical reaction, the oxidation of water (OH^-), that is found only in green plants. Water as an electron donor is therefore not involved in the two main events of non-cyclic photophosphorylation in plants and bacteria, the photoreduction of pyridine nucleotide and the coupled formation of adenosine triphosphate."

It seems inconceivable that the oxidation of water as an enzymatic dark reaction at this spot could compete so successfully with the oxidation of hydrogenase hydrogen as it would have to in order to bring about the (reversible) oxygen evolution in adapted algae.

The improbability of this explanation need not be argued at length, for we have an experimental fact of fundamental importance which it would be rather unscholarly to ignore. This is the constancy of the number of light quanta required to assimilate one molecule of carbon dioxide, whether this happens in the anaerobic sulfur purple bacteria or in aerobic or adapted plants, while the permanent gain in ΔF in these reactions varies from 4 to 120 kcal. Evidently the sequence of events, the stoichiometry of the internal mechanism in all photosynthetic organisms known, requires a minimum of eight quanta per CO_2 —or of two quanta per reducing hydrogen. And the only explanation for this observation which so far has made sense is the one van Niel (1935), Franck (1957, 1958), and myself (1940) have often enough pointed out, namely the involvement of water in the photochemical mechanism, as a reaction apart and independent of the eventual release of oxygen. I recommend in particular a rereading of van Niel's article of 1949.

Evidently the basic mechanism exists already in the anaerobic bacteria. It is one of the evolutionary stages which had to occur first before the other changes leading to green plant photosynthesis already mentioned above could become effective. It follows that the invention of bringing water into the photochemical mechanism must have preceded the final development of the bacterial type of photometabolism and that this had a definite selective advantage over any simpler one quantum hydrogen transfer process of the kind we can artificially produce in the laboratory. Let us survey the hypothetical history of chlorophyll reactions as follows:

Sensitized isomerizations; hydrogen transfers in homogeneous solution. Quantum yield approaching unity only in a small number of cases.

Coupled hydro-dehydrogenations on surfaces. Incipient separation of efficient from inefficient reactions.

Extension of effective range by the introduction of new reactants under the influence of the first specific proteins (hemoproteins).

Complexing with cytochromes. Water as intermediary oxidoreductant. Two-quanta mechanism. Coupling with independently evolved enzymatic oxidoreduction systems.

Evolution of back reactions into phosphorylation systems. Coupling with ATP requiring metabolic reactions. Acetate assimilation. Fixation and reduction of carbon dioxide. Complete photoreduction in first phototrophic organism.

Gain of 10 kcal per quantum absorbed by shift of color and diversification among the chlorophylls. Emerson effect. Accessory pigments. Other cytochromes (*f*). Other quinone. *Manganese enzyme. Liberation of oxygen.* Survival of carotene-protected organisms.

Evolution of respiration. The oxidation chain from oxygen over cytochrome, flavin, pyridine nucleotide to organic substrate may have come from the phosphorylation mechanism in the photosynthetic apparatus, and the dicarboxylic acid cycle from the mechanism by which such acids are decomposed in the light into carbon dioxide and free hydrogen.

In this sequence the beginning of the two-quanta photolytic process, which uses water to create, with the aid of cytochromes (Kamen, 1956), not only strong reductants but also strong oxidants, is tucked away deep in the anaerobic era, perhaps many millions of evolutionary years away from the appearance of the first green alga. The argument that Arnon (1960) has repeatedly given.

ON DATING STAGES IN PHOTOCHEMICAL EVOLUTION 85

to dislodge van Niel's unified concept of photosynthesis from its pre-eminent position and to substitute instead phosphorylation as the unifying principle is, it seems to me, a scientific retrogression. It cannot explain half of the facts we already have before us. It takes many steps to oxidize four molecules of water in such a way that at the end one molecule of oxygen is set free. And the evidence is plainly against the assumption that all the necessary evolutionary steps occurred during the transformation of the bacterial type into the plant type of photosynthesis. On the other hand, if the two-quanta process with water appeared in the middle of an early anaerobic era, it brought with it the unique advantage for the performance of typical oxidation reactions of a kind which without light, chlorophyll, and water are known to occur only under aerobic conditions (Schenck, 1960). To quote the latest example studied by Scher and Proctor (1960): "The experiments to date with benzoate and other aromatic compounds indicate that bacterial photosynthesis involves the production of an oxidant equivalent to molecular oxygen produced as a consequence of a photochemical reaction." It does not help us much to say "light is the equivalent of ATP" since that can be said of numerous enzymatic steps in intermediate metabolism. But when we say: "Under anaerobic conditions visible light and water are the equivalent of oxygen," we refer to something rather special (cf. Goldfine and Bloch, 1962).

We all believe that respiration developed after photosynthesis became what it is. But there might be more than simply a succession in time. One could make a case for respiratory mechanisms to have evolved from parts of the already existing rather complex photosynthetic apparatus. The oxidative chain may have broken away from the pigment complex, where it served originally as one of the phosphorylation mechanisms. The scheme of Fig. 4 makes use of Krogman and Vennesland's (1959) discovery of an oxygen-requiring photophosphorylation as a suggestive intermediate step. Gest's discovery of a total decomposition of some dicarboxylic acids into carbon dioxide and hydrogen by anaerobic purple bacteria in the light suggests that also the respiratory dicarboxylic acid cycle had its origin in an anaerobic mechanism (Gest, 1962).

After the appearance of oxygen a protective adaptation against the danger of photooxidation became necessary. There is no doubt that presence of carotene in the pigment complex, though not essential for the photochemistry proper, provides such a protection

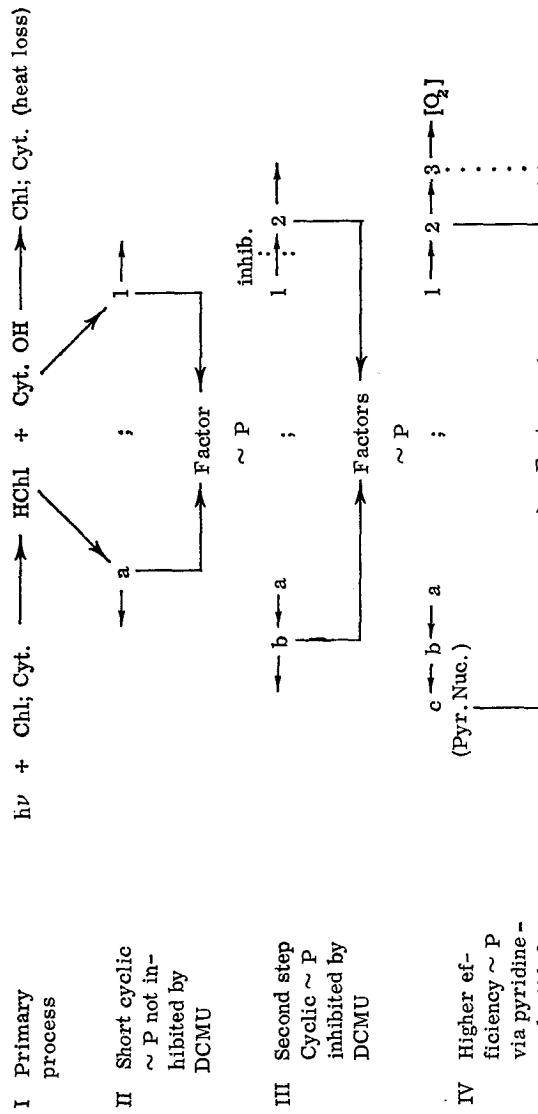


Fig. 4. Partial evolution of respiration from back reactions in a photochemical system reacting with water.

ON DATING STAGES IN PHOTOCHEMICAL EVOLUTION 87

(Stanier, 1961). Carotenoids are plentiful in all the anaerobic photoreducing bacteria. Originally present for other reasons, they insured the survival of those photosynthetic organisms that happened to contain them. We know that *in vitro* carotenes prevent photo-oxidation by quenching the triplet state of chlorophyll. If we use this as a model to explain the protective action *in vivo* we create a new difficulty. If carotenes render the triplet state of chlorophyll harmless before oxygen has had a chance to react it must be closely associated with the chlorophyll complex. The protection depends upon the immediate quenching of the triplet state—how then is photosynthesis at all possible in the presence of carotene? A way out is to remember the properties of the photosynthetic unit. Only one in several hundred chlorophyll molecules is engaged in chemical action, the others have merely the task to absorb light and let the light energy of the first singlet excited state migrate to the "working" chlorophyll molecules. Yet each one of these hundreds of light-absorbing molecules is a potential source of trouble in the presence of oxygen. A "derailment" into the triplet state on the spot is not dangerous, however, as long as such triplet states are selectively quenched by carotene. Why were carotenes already present at the dawn of the aerobic era? This seems again to be a case where a substance which already had a function assumed another, perhaps more important one when the survival of the cell was at stake. Our knowledge about what else carotenes may do in the chlorophyll complex is sparse. Bishop (1962c) has isolated from spinach leaves a birefringent β -carotene protein compound having no discernible function. It remains to be seen whether this compound is a new acquisition of the green plants or whether the same or a similar compound can be found in the purple bacteria. This would move its birthday back into the last of the pre-Darwinian anaerobic eras.

REFERENCES

- Allen, G. (1957). *Am. Naturalist* **91**, 65.
- Anker, H. S. (1961). *Perspectives in Biol. and Med.* **5**, 86.
- Arnon, D. I. (1960). *Sci. American* **203**, 2.
- Arnon, D. I. (1961). *Bull. Torrey Botan. Club* **88**, 215.
- Arnon, D. I., Whatley, F. R., and Allen, M. B. (1959). *Biochim. et Biophys. Acta* **32**, 47.
- Arnon, D. I., Losada, M., Nozaki, M., and Takawa, K. (1961). *Nature* **190**, 601.
- Bassham, J. A., and Calvin, M. (1960). In "Handbuch der Pflanzenphysiologie" (A. Pirson, ed.) Vol. 5, p. 884. Springer, Berlin.

- Baur, E. (1930). *Helv. Chim. Acta* **12**, 788.
- Bishop, N. I. (1962a). *J. Gen. Phys.* **45**, 592A.
- Bishop, N. I. (1962b). *Nature* (in press).
- Bishop, N. I. (1962c). Abstracts, Biophysical Society.
- Calvin, M. (1956a). *Am. Scientist* **44**, 248.
- Calvin, M. (1956b). *Naturwissenschaften* **43**, 387.
- Chance, B., and Nishimura, M. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 19.
- Dauvillier, A. (1958). "L'Origine Photochimique de la Vie," Masson, Paris.
- Duysens, L. N. M., Amesz, J., and Kamp, B. M. (1961). *Nature* **190**, 510.
- Emerson, R., Chalmers, R. V., and Cederstrand, C. (1957). *Proc. Natl. Acad. Sci. U.S.* **43**, 133.
- Emerson, R., and Rabinowitch, E. (1960). *Plant Physiol.* **35**, 477.
- Fox, S. W. (1960). *Science* **132**, 200.
- Fox, S. W., and Harada, K. (1961). *Science* **133**, 1923.
- Franck, J. (1957). In "Research in Photosynthesis" (H. Gaffron, ed.), p. 142. Interscience, New York.
- Franck, J. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 941.
- Franck, J. (1960). In "Handbuch der Pflanzenphysiologie" (A. Pirson, ed.), Vol. 5, p. 689. Springer, Berlin.
- French, C. S., and Fork, D. C. (1961). Proceedings of Vth International Congress of Biochemistry, Moscow. Preprint No. 73.
- Gaffron, H. (1927). *Ber. deut. chem. Ges.* **60**, 755.
- Gaffron, H. (1940). *Am. J. Botany* **27**, 282.
- Gaffron, H. (1957). In "Rhythmic and Synthetic Processes in Growth" (D. Rudnick, ed.), p. 127. Princeton Univ. Press, Princeton, New Jersey.
- Gaffron, H. (1958). In "Transactions of the Conference on the Use of Solar Energy, Tucson, Arizona, 1955," Vol. 4, p. 145. Univ. of Arizona Press.
- Gaffron, H. (1960a). *Perspectives in Biol. and Med.* **3**, 163; in "Evolution after Darwin" (Sol Tax, ed.), p. 39. University of Chicago Press, Chicago, Illinois.
- Gaffron, H. (1960b). In "Plant Physiology" (F. C. Steward, ed.), Vol. IB, p. 3. Academic Press, New York.
- Gerretsen, F. C. (1948). *J. Biol. Chem.* **176**, 299.
- Gest, H., Ormerod, J. G., and Ormerod, K. S. (1962). *Arch. Biochem. Biophys.* In press.
- Goldfine, H., and Bloch, K. (1962). "Oxygen and Biosynthetic Reactions." In press.
- Habermann, H., and Gaffron, H. (1961). In "Progress in Photobiology" (B. Chr. Christensen and B. Buchmann, eds.), Proceedings of the 3rd International Congress on Photobiology, 1960, p. 586. Elsevier, Amsterdam.
- Habermann, H., and Gaffron, H. (1962). *Photochem. and Photobiol.* **1**. In press.
- Hill, R., and Scarsbrick, R. (1951). *New Phytologist*, **50**, 98.
- Horwitz, L., and Allen, F. L. (1957). *Arch. Biochem. Biophys.* **66**, 23.
- Hoyle, F. (1960). "Nature of the Universe." Harper, New York.
- Kamen, M. D. (1956). In "Enzymes: Units of Biological Structure and Function" (O. H. Gaebler, ed.), p. 483. Academic Press, New York.
- Keilin, D., and Hartree, E. F. (1945). *Biochem. J.* **39**, 293.
- Kern, M., and Racker, E. (1954). *Arch. Biochem. Biophys.* **48**, 235.

ON DATING STAGES IN PHOTOCHEMICAL EVOLUTION 89

- Kessler, E. (1955). *Arch. Biochem. Biophys.* **59**, 527.
- Kessler, E. (1957). *Planta* **49**, 435.
- Kok, B. (1960). In "Handbuch der Pflanzenphysiologie" (A. Pirson, ed.), Vol. 5, p. 566. Springer, Berlin.
- Krasnovsky, A. A. (1948). *Doklady Akad. Nauk U.S.S.R.* **60**, 421.
- Krasnovsky, A. A. (1960). *Ann. Rev. Plant Physiol.* **11**, 363.
- Krebs, H. A., and Kornberg, H. L. (1957). "Energy Transformations in Living Matter (A Survey)," p. 275. Springer, Berlin.
- Krogman, D. W., and Vennesland, B. (1959). *J. Biol. Chem.* **234**, 2205.
- Losada, M., Trebst, A. V., and Arnon, D. I. (1960). *J. Biol. Chem.* **235**, 832.
- Losada, M., Whatley, F. R., and Arnon, D. I. (1961). *Nature* **190**, 601.
- Lowe, M. B., and Phillips, J. N. (1961). *Nature* **190**, 262.
- Lowenstein, J. M. (1958). *Biochim. et Biophys. Acta* **28**, 206.
- Mehler, A. H. (1951). *Arch. Biochem. Biophys.* **33**, 65.
- Miller, S. L. (1955). *J. Am. Chem. Soc.* **77**, 2351.
- Nason, A., Wosilait, W. D., and Terrell, A. J. (1954). *Arch. Biochem. Biophys.* **48**, 233.
- Oparin, A. I. (1959). "The Origin of Life on the Earth," Proceedings of the First International Symposium, Moscow. Pergamon Press, New York.
- Pavlovskaya, T. E., and Pasynskii, A. G. (1959). In "The Origin of Life on Earth" (A. I. Oparin et al., eds.), p. 151. Pergamon Press, New York.
- Pirson, A., ed. (1960). "Handbuch der Pflanzenphysiologie," Vol. 5, Springer, Berlin.
- Pringsheim, E. G., and Wiessner, W. (1961). *Arch. Mikrobiol.* **40**, 231.
- Rabinowitch, E. I. (1956). "Photosynthesis and Related Processes," Vol. 2, Part 2. Interscience, New York.
- Reid, C. (1959). In "The Origin of Life on Earth" (A. I. Oparin et al., eds.), p. 619. Pergamon Press, New York.
- Schenck, G. O. (1960). *Z. Elektrochem.* **64**, 997.
- Scher, S., and Proctor, M. H. (1960). In "Comparative Biochemistry of Photo-reactive Systems" (M. B. Allen, ed.), p. 387. Academic Press, New York.
- Schramm, G., Grötsch, H., and Pollmann, W. (1962). *Angew. Chem.* **74**, 53.
- Shapley, H. (1960). *Perspectives in Biol. and Med.* **3**, 222; in "Evolution after Darwin" (Sol Tax, ed.), p. 23. University of Chicago Press, Chicago, Illinois.
- Shemin, D. (1956). In "Currents in Biochemical Research" (D. E. Green, ed.), p. 518. Interscience, New York.
- Stanier, R. (1961). *Bacteriol. Revs.* **25**, 1.
- Stoll, A. (1932). *Naturwissenschaften* **20**, 955.
- van Niel, C. B. (1931). *Arch. Mikrobiol.* **3**, 1.
- van Niel, C. B. (1935). *Cold Spring Harbor Symposia on Quant. Biol.* **3**, 138.
- van Niel, C. B. (1949). In "Photosynthesis in Plants" (J. Franck and W. E. Loomis, eds.), p. 437. Iowa State College Press, Ames, Iowa.
- Witt, H. T., Moraw, R., Müller, A., Rumberg, B., and Gieger, G. (1960). *Z. physik. Chem. (Frankfurt) [N. F.]* **23**, 133.
- Witt, H. T., Müller, A., and Rumberg, B. (1961). *Nature* **192**, 967.
- Ycas, M. (1955). *Proc. Natl. Acad. Sci. U.S.* **41**, 714.

Origin and Evolution of Bioluminescence

W. D. McELROY AND H. H. SELIGER

*Department of Biology and McCollum-Pratt Institute,
Johns Hopkins University, Baltimore, Maryland*

I. Introduction	91
II. Evolution of the Electron Transport Process	92
III. The Struggle for Anaerobic Conditions—Oxygen Utilization	96
IV. Summary	100
References	101

I. Introduction

The emission of light by organisms, bioluminescence, is not restricted to any one group of animals or plants (Harvey, 1952). The random distribution of a light-emitting system among the bacteria, fungi, numerous invertebrate groups and even fish suggest that it is an offshoot of a chemical reaction fundamental to all organisms. The widespread occurrence of luminous organisms may be the important key in understanding its origin and evolution.

Bioluminescence is an example of a chemiluminescence process in which a reduced substrate (luciferin) is oxidized by molecular oxygen thereby creating an excited state which emits light on return to a ground state (McElroy and Seliger, 1961). The high efficiency of this oxidative chemiluminescent reaction is due to the presence of a catalyst, luciferase. Luciferins from various luminous organisms are quite different although they have one feature in common: they are highly fluorescent. Some examples of luminescent organisms are listed in Table I.

LH_2 refers to the oxidizable substrate, luciferin, and E refers to the enzyme, luciferase. Both luciferin and luciferase are different for the various forms. Recently Johnson and Haneda have demonstrated that the luminous fish (*Apogon*) appears to have the same light system as *Cypridina*. The structure and synthesis of firefly luciferin have been accomplished recently (McElroy and Seliger, 1961) and have been shown to be different from the *Cypridina* and the bacterial luciferins.

TABLE I
SELECTED EXAMPLES OF LUMINOUS ORGANISMS

Organism	Nature of reactants	Peak light emission ($m\mu$)
Luminous bacteria	$\text{FMNH}_2 + \text{RC}\begin{array}{c} \text{O} \\ \diagdown \\ \text{H} \end{array} + \text{O}_2 + \text{E}$	495
Fireflies	$\text{LH}_2 + \text{ATP} + \text{Mg} + \text{O}_2 + \text{E}$	562
<i>Cypridina</i> (crustacean)	$\text{LH}_2 + \text{O}_2 + \text{E}$	460
<i>Odontosyllis</i> (polychaete worm)	$\text{LH}_2 + \text{O}_2 + \text{E}$	510
<i>Pholas dactylus</i> (luminous clam)	$\text{DPNH} + \text{FMN} + \text{O}_2 + \text{E}$	480
<i>Omphali flavida</i> (fungus)	$\text{DPNH} + \text{X} + \text{O}_2 + \text{E}$	530
<i>Renilla reniformis</i> (sea pansy)	$\text{LH}_2 + \text{AMP} + \text{O}_2 + \text{E}$	Blue
<i>Gonyaulax polyhedra</i> (protozoan)	$\text{LH}_2 + \text{E} + \text{O}_2$	470
<i>Apogon</i> (fish)	$\text{LH}_2 + \text{E} + \text{O}_2$	460

In the present paper we will develop the hypothesis that these light-emitting reactions in organisms were, in fact, detoxifying processes for the removal of oxygen which was necessary for the survival of early anaerobic forms of life. We propose that the use of organic-reducing substances to remove oxygen by direct reduction led to the formation of an excited state which could emit light. We believe that these reactions were the basis for the origin and evolution of luminescence in organisms.

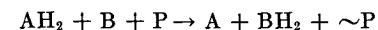
II. Evolution of the Electron Transport Process

Most investigators assume that at the time of origin of the primitive forms of life on earth the biosphere was devoid of oxygen.¹ Although traces of oxygen may have been produced by short wave-

¹ For a discussion of the nature of the earth's atmosphere and various authors' ideas on the origin of life and associated processes the reader is referred to the Proceedings of the First International Symposium on "The Origin of Life on the Earth," Pergamon Press, New York, 1960.

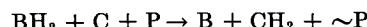
length ultraviolet radiation it is doubtful that it remained as free oxygen for any length of time. The extensive reducing environment and the large excess of such compounds as Fe^{++} would insure the immediate reduction of any free oxygen.

Although the stepwise oxidative reactions are the primary ones concerned with the efficient supply of energy in modern aerobes, it seems most unlikely that oxygen functioned as the terminal electron acceptor in the primitive forms of life. Free oxygen must have appeared in significant quantities at a much later date as the environment gradually changed from a reducing one to an oxidizing one. Accompanying these changes in evolution was the development of the extensive enzyme systems coupling the primary photochemical reactions with water which acted as the hydrogen donor. Green plant photosynthesis is a consequence of these evolutionary events. Thus it seems certain that the primitive organisms were true anaerobes. The early energy-yielding and energy-coupling processes, however, must have involved dehydrogenation and condensation reactions which were catalyzed first by primitive catalysts, and much later by enzymes. The important point is that life arose and evolved only when such entities were capable of coupling the energy of dehydrogenation to the necessary synthesis of all those compounds which were required for duplication. As we visualize this energy-coupling process today it seems most likely, as suggested by Lipmann, that inorganic phosphate was important in the trapping of energy in the form of a pyrophosphate bond. Thus the first organism must have made use of the rich organic-reducing environment to couple dehydrogenation to energy liberation and utilization. The first energy-trapping reaction must have been of the type indicated in the reaction below:



An organism could make use of the energy from the oxidation of AH_2 for growth and reproduction as long as AH_2 was present in sufficient quantity in the environment. There are a number of ways in which the oxidized product A could be reduced but we will defer a discussion of this until later. Needless to say, once AH_2 became exhausted the organism would then be forced to rely upon other dehydrogenation reactions. The successful organisms must have been capable of using less reduced compounds for coupled oxidation-reduction reactions. We visualize the subsequent dehydrogenation

and energy coupling as follows:



During such an evolutionary process it is evident that by using the highly reduced compounds first the anaerobes were gradually creating a more oxidizing environment. At each successive step an electron acceptor was used which was approaching the potential of oxygen itself. Thus, we visualize a stepwise succession of dehydrogenation reactions in which energy liberation was closely linked to synthesis by means of the pyrophosphate bond. If a modest amount of AH_2 could be reformed within the cellular environment by whatever means, it is clear that such a selective evolutionary process could lead to the establishment of a series of electron transport systems which are capable of liberating energy. At some stage in this evolutionary process the reformation of reduced primary substances must have occurred in which sunlight was the source of energy. Life may have originated and evolved to its high level because of the presence of complex organic molecules which were capable of reasonably long-lived excited states. Creation of the excited states by light quanta established for the first time the primitive photochemical event which sometime later in evolution was capable of restoring the highly reducing environment. It seems reasonably certain that the present highly efficient photosynthetic apparatus is much too complicated to have been present in these primitive organisms. A less complicated inorganic metal complex and later an organic metal complex was the most likely system for creating a reducing system driven by light energy. It was the coupling of excitation energy to the electron transport process which created an efficient and inexhaustible supply of reducing energy. A number of organic compounds could have functioned in this manner under anaerobic conditions. For example, in the absence of oxygen, FMN is reduced by light to FMNH_2 .

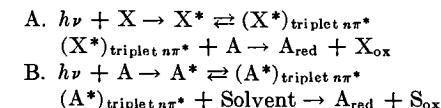
In the absence of oxygen in the primitive atmosphere the spectrum of the radiant energy reaching the oceans extended well below the present 2900 Å cut-off so that even the pyridine nucleotides could have been raised to excited states in great abundance. It is a rather significant point that the photochemical production of the leuco form of dyes is favored when oxygen is excluded from the environment. Further the quenching effect by oxygen of the fluorescence of many organic compounds is pictured as an interaction

between the ground triplet state of the oxygen molecule and the triplet state of the excited organic molecule. This would leave the latter in a state of high vibrational energy which by interaction with the environment returns to the ground state. The freezing of a solution or adsorption of a substrate onto a lattice such as in a mitochondrion has two effects: there is a reduction in the vibrational degrees of freedom due to the bonding, which sharpens the band structure and there is also a reduction in the vibrational interaction of the excited state with dissolved oxygen. For this reason even organic molecules which show no fluorescence in solution at room temperature will exhibit phosphorescence at low temperatures or when dissolved in a "glass" at room temperature.

The nitrogen heterocyclics of which the adenine molecule is an example are characterized by the presence of nonbonding electrons which give rise to energy levels in the excited molecule lower than the $\pi \rightarrow \pi^*$ levels. It is presumably the triplet $n \rightarrow \pi^*$ level to which the singlet $\pi \rightarrow \pi^*$ level eventually decays, and this triplet level is quenched by oxygen or by interaction with the environment.

In order for a molecule to make efficient use of the available light energy it must have a system of relatively loosely bound π -electrons which can be raised to excited states. These excited states must be reasonably well separated from the vibrational ground states in order that the energy is not dissipated immediately to the environment and in order that the now highly reactive molecule can interact with a suitable donor or acceptor. The molecules involved in the electron transport process are not the cyclic carbon compounds but the nitrogen heterocyclics with their $n \rightarrow \pi^*$ triplet levels.

Porter has shown that in the strict absence of oxygen, triplet-triplet energy transfer is very efficient, and it is therefore conceivable that initially the light excitation was coupled in this way to a suitable reducing substance giving the general reactions:



The advantages of this mechanism over the corresponding singlet-singlet energy transfer in a weakly structured system are the reduction of energy-wasting fluorescence and the longer lifetimes of the triplet states.

Although at the present time, in the well-ordered mitochondrial electron transport apparatus, the reduction of FAD by DPNH and the subsequent reduction of the cytochrome pigments by reduced FAD probably do not involve excited states of these molecules, it is suggestive that the energy-level differences between the blue-fluorescent pyridine nucleotides, the yellow-green-fluorescent flavines and the red-fluorescent porphyrins are of the order of 10 kc per mole which is sufficient for the generation of energy-rich phosphate bonds. The nature of the singlet and triplet excited states of these cofactors may have been an important factor in their early selection for electron transport processes. The advent of the lipoprotein mitochondrial matrix for the stepwise reduction involving the terminal oxygen molecule was almost certainly a much more recent development in biological time, relegating the light-energy utilization role to the more efficient aerobic chlorophyll systems.

III. The Struggle for Anaerobic Conditions—Oxygen Utilization

As a result of the utilization by the anaerobes of the highly reducing substances, and the sensitized photochemical dissociation of water by the action of light on various pigments, it seems quite likely that the amount of oxygen in the environment was gradually increasing, concurrent with the exhaustion of the organic material suitable for anaerobic metabolism. Some authors believe that during this period there developed those chemoautotrophs which were able to oxidize directly ammonia, hydrogen sulfide, or ferrous ion to satisfy their energy requirements. However, if the earlier organisms were strict anaerobes it is evident that growth would be inhibited by the presence of oxygen. Unfortunately, we do not understand why anaerobes fail to grow in the presence of oxygen. There are some organisms that can grow in the presence of a small amount of oxygen but there are others which cannot grow at all. Obviously, oxygen in some way prevents growth as long as it is present. However, this is an inhibiting rather than a lethal action since removal of oxygen will allow the initiation of growth. The present evidence suggests that oxygen may prevent the growth of strict anaerobes by a number of mechanisms. The formation of hydroperoxides is considered to be one of the primary factors. There are some organisms among the strict anaerobes that can take up oxygen and form hydrogen peroxide. In spite of this oxygen uptake and the accumulation of hydrogen peroxide the anaerobe will still grow if placed in an envir-

onment devoid of oxygen. In other words the organism cannot grow in the presence of oxygen but the exposure to it does not prevent it from growing later when the oxygen is removed. In addition there are some anaerobes in which oxygen uptake has not been demonstrated and in which hydrogen peroxide has not been detected. These organisms show the same inability to grow in oxygen. Thus we are forced to the conclusion that oxygen is inhibiting but not lethal at least over a limited time exposure.

At the time of the appearance of low concentrations of oxygen there was a struggle to survive in a changing environment and it is our contention that the organisms that were successful were those that were able to reduce molecular oxygen directly and quickly. At first this process was an ordinary chemical reduction which in some cases could have been quite slow. However, the most successful organisms were those that acquired a catalyst that accelerated the reduction process. This new enzyme would be an oxidase which catalyzed the oxidation of a reduced organic substrate by molecular oxygen. The oxidase would have the properties of what we presently call luciferase. The direct reaction of the reduced substrate with molecular oxygen is generally a strongly exothermic reaction. We therefore propose that the oxidative reaction left the product or intermediate molecule in a highly excited state. If the specific molecule was capable of fluorescence the oxidation would have resulted in what we term bioluminescence. Thus the struggle to maintain anaerobic conditions led to the selection of organisms having specific oxidases (luciferase) which catalyzed the rapid removal of oxygen. We propose, therefore, that those organisms that successfully survived the exposure to low oxygen tensions were all potentially luminescent because of the nature of the detoxifying reaction. It seems significant that all luciferase systems that have been carefully studied can catalyze the utilization of reducing substrate (luciferin) at very low oxygen tensions. The firefly, bacteria, and *Cypridina* systems are at least one hundred to one thousand times more sensitive to oxygen than is hemoglobin.

The reduction of oxygen in most cases led to the production of hydrogen peroxide or some organic peroxide radical. Although most organisms can tolerate a small amount of hydrogen peroxide it is clear that the really successful ones were those that were capable of removing the peroxide by a reducing agent to form water. This must have been the selective pressure which led to the adaptation of a

catalyst to use hydrogen peroxide and additional reducing material to form water and an oxidized product. We propose that again the successful organisms were those capable of catalyzing a peroxidase type of reaction. Coupled with the luciferase reaction this led to the elimination of oxygen with the production of water, oxidized product, and light. Undoubtedly during this part of the evolutionary process the development of the iron porphyrin type of catalyst for peroxides occurred. We note for example that iron in the Fe^{++} state is a very effective catalyst for the decomposition of hydrogen peroxide. If iron, however, is incorporated into a special structure such as a tetrapyrrole ring then the catalytic activity is increased by several hundred times. Calvin has presented excellent arguments as to why hydrogen peroxide might favor the formation of such iron pyrrole type compounds.

Simultaneous with the gradual loss of the reducing environment and the appearance of an oxidizing one, a catalase type enzyme would be favored over the peroxidase system, the former decomposing hydrogen peroxide into water and oxygen. This would lead to the selection of those systems which were capable of reducing oxygen stepwise to form water without free hydrogen peroxide being an intermediate. We feel that it was during this period that the iron-porphyrin enzyme systems, possessing the ability to activate oxygen for accepting electrons directly, came into existence. The oxygen could therefore act as the electron acceptor without the intermediate formation of hydrogen peroxide. It was this adaptation which led to the completion of the oxidative phosphorylation system as we see it today. The utilization of oxygen by this primitive aerobic form was then the last major step in the establishment of the electron transport process. By using oxygen as the terminal electron acceptor the entire organic reducing environment was susceptible to oxidation. This must have led to the rapid and final development of the heterotrophic organisms which appeared during this phase of evolution.

Photochemical excitation energy processes must have been highly successful in this environment, using first the electron donors of an inorganic (Fe^{++}) or organic type and finally, with the exhaustion of these, water. Thus we think of the present green plant photosynthesis as being one in which water became the ultimate electron donor leading to the formation of some OH type compound which was capable of giving rise to molecular oxygen. The emergence of an

organism capable of using simple carbon compounds such as CO_2 , water, and light as the energy source was the primary evolutionary event which led to the creation of a rich organic environment on earth for a second time (see Fig. 1).

With such systems established and with the appearance of the true aerobes, it is evident that the direct reduction of oxygen and the accompanying luminescence were no longer of selective advantage. Therefore, with the appearance of the aerobes, luminescence would begin to disappear. However, since light emission was originally intimately connected with the essential energy-liberating electron transport processes it is likely that a number of the light-emitting

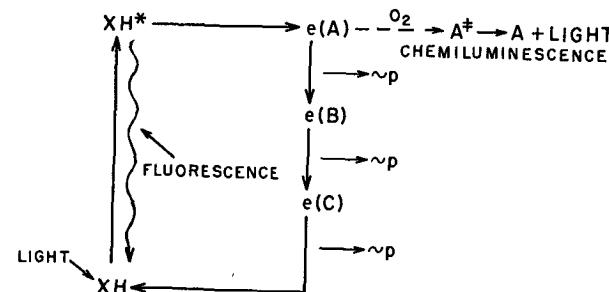


FIG. 1. Coupling of excitation energy to the electron transport process.

systems would persist. It is therefore our argument that bioluminescence is a vestigial system in the evolutionary process and that there is at present no selective advantage insofar as the primary excitation process is concerned. It is true, however, that during the evolution of various species the luminescent system has been adapted for secondary purposes which have selective advantage. The identification of the female firefly by the male is an excellent example. No such argument involving sexual reproduction advantages can be made for the bacteria. Where luminescence has been put to a use, this use can be important in explaining selective advantage at a secondary level. The primary process which involves an interaction of oxygen with a reducing substance to create an excited state cannot readily be explained except by a more fundamental metabolic system common to all organisms.

The fact that the luminescent systems which have been carefully analyzed have been shown to have entirely different emission spectra

is an argument in favor of the idea that originally organisms used a variety of complex organic molecules for the reducing system for removal of oxygen and for the creation of an excited state. It seems highly likely that the reducing systems of the potential level of the pyridine nucleotides were the ones used early for energy liberation during dehydrogenation. The establishment of this system probably served as the basis for the evolution of other electron transport steps. This may explain why in photosynthesis we first observe the net reduction of pyridine nucleotides by chlorophyll even though thermodynamically others might be favored. The flow of electrons propelled by photochemical processes seems always to involve a highly reducing state. If the initial removal of oxygen by the anaerobes made use of reducing systems of this potential level one might expect to observe the high-energy blue bioluminescence emission most frequently. As the reducing power of the environment decreased other less reducing chromophoric groups could have become involved. It may be significant for our argument therefore that most luminescent forms are unicellular (a large number are also photosynthetic) and their emission is in the blue region of the spectrum. It may be significant that a high proportion of luminous organisms appear in the oceans. The second most frequent forms have their emission in the yellow green region and as far as we know are all multicellular or complex multinuclear structures. Red light emission is quite rare and is known to occur in only one form—the South American railroad worm.

IV. Summary

Bioluminescence is randomly distributed throughout the animal, plant, and microbial world. The chemical reaction which creates the excited state in all forms is oxidative in nature making use as it does of molecular oxygen as the ultimate electron acceptor. Except for two species, none of the oxidizable substrates (luciferins) from these various organisms will react with any but its own enzyme (luciferase), leading to the suggestion that luminescence is not uniquely associated with any single organic substance. It has been suggested, however, that despite the apparent diversity of the various luciferins they are part of a chemical reaction that is fundamental to all organisms.

We propose that in the evolution of organisms from an anaerobic to an aerobic form of life there was, initially, a struggle for anaerobic

existence. Because of the toxic nature of oxygen to anaerobes, we suggest that the organisms that survived initially were those that were able to reduce molecular oxygen directly and quickly. This struggle for anaerobic conditions led to the selection of organisms having specific oxidases (luciferase) which catalyzed the rapid removal of oxygen; all primitive forms were therefore potentially luminescent. The gradual selection and evolution of electron transport processes in which oxygen was reduced stepwise to form water gave rise finally to the aerobic forms. With the appearance of the latter the luminescent, oxidative reaction was no longer of selective advantage. Thus we argue that bioluminescence is a vestigial system of organic evolution, but that through secondary evolutionary processes the luminescent system has been preserved in various and unrelated organisms by virtue of the fact that it has been adapted for other useful purposes.

REFERENCES

- Harvey, E. N. (1952). "Bioluminescence." Academic Press, New York.
McElroy, W. D., and Seliger, H. H. (1961). In "Symposium on Light and Life" (W. D. McElroy and H. B. Glass, eds.), p. 219. Johns Hopkins Press, Baltimore, Maryland.

On the Problems of Evolution and Biochemical Information Transfer

ALEXANDER RICH

*Department of Biology,
Massachusetts Institute of Technology,
Cambridge, Massachusetts*

I. Introduction	103
II. An Outline of Nucleic Acid Function	104
III. Chemical Evolution and the Origin of Life	111
IV. Did Life Originate with Protein Molecules?	112
V. Polynucleotides as the Origin of Living Systems.	113
VI. The Trend Toward Increased Complexity	117
VII. Changes in the Composition of the Nucleic Acids	119
VIII. Has the Number of Amino Acids Increased During Evolution?	120
IX. An Evolutionary Increase in Nucleic Acid Content	121
X. Why Are There Two Nucleic Acids?	123
XI. Extraterrestrial Life	124
XII. Conclusions	125
References	125

I. Introduction

Perhaps the most striking characteristic in the development of biochemical understanding during the past decade has been the discovery that nucleic acids play the central role in the transmission of molecular information. Although a great deal of data had been compiled on pathways of intermediary metabolism by the early 1950's, relatively little was known about those reactions which govern the over-all flow of metabolism. In contrast to that situation, we now have a reasonably coherent explanation of the replication and transmission of genetic information and the process whereby this information leads to the ultimate regulation of metabolism. At the very heart of the system are the nucleic acids, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). These molecules now occupy such a central role in biochemistry that it is difficult to realize the extent to which their properties and functions were unknown as briefly as a decade ago.

Inevitably the development of a more comprehensive under-

standing concerning the role of the nucleic acids is accompanied by a large number of questions concerning the manner in which these molecules came to occupy such a commanding position in biochemical systems. In short, what has been their evolutionary history? Unfortunately, we have little experimental data at the present time with which to answer these questions. Nonetheless, a study of the over-all pattern of genetic replication and genetic expression leads one to ask certain direct questions and to suggest possible evolutionary routes which might lead to the present-day system. Several speculations and hypotheses about the evolution of the nucleic acids are developed in this article. Some of them may prove useful in stimulating further discussion and they may also suggest directions for an experimental approach to the problem.

II. An Outline of Nucleic Acid Function

A brief and simplified outline of the functions of the nucleic acids is schematically illustrated in Fig. 1. The DNA molecule is the major carrier of genetic information. When we use the term "information" in conjunction with the nucleic acids, it is often synonymous with the word "sequence." The nucleic acids are built in a linear chain containing four major nucleotide components and we believe that the "information" resides in the order of these four units. This is analogous to the way in which "information" in a sentence is determined by the ordering of the letters. Thus we can say that the molecular language of the nucleic acids is written with a four-letter alphabet.

Broadly speaking the DNA molecule has two functions: first, it must reproduce itself during cell division so that its information can be transmitted to daughter cells; and secondly, it must express this information by influencing the metabolism of the cell. According to our present understanding, the major mechanism for metabolic control is exercised through regulating the supply of protein molecules which act as catalysts. In the presence of these enzymes, chemical reactions proceed readily; in their absence these reactions proceed very slowly if they occur at all. Only in the past few years has the mechanism underlying these two functions of DNA become clear. The work of Kornberg and his collaborators (see Bessman *et al.*, 1958) has demonstrated that the replication of DNA is governed by the activity of an enzyme, DNA polymerase. Given a primer of DNA, the enzyme has the property of separating the twisted

chains in the double-stranded molecule and organizing a series of complementary nucleoside triphosphate bases along each strand which are then polymerized to produce two identical molecules.

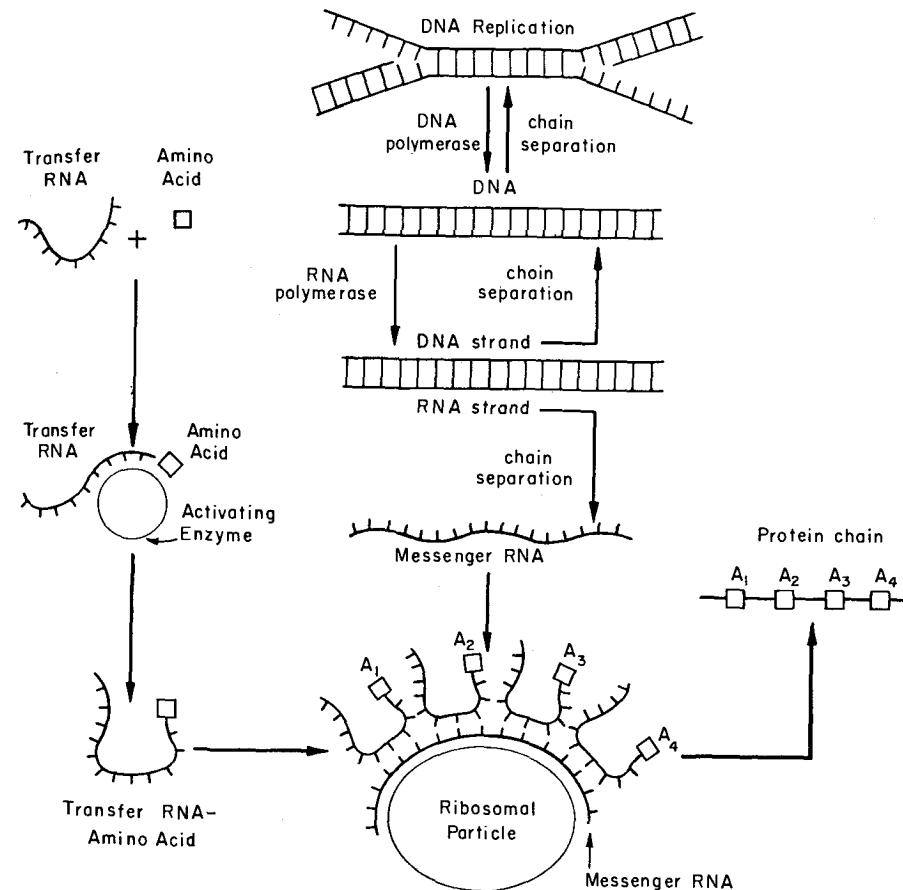


FIG. 1. A schematic outline of nucleic acid function. The ladder-like figures represent two-stranded nucleic acids and the bases are represented by the short cross lines.

The experiments of Meselson and Stahl (1958) demonstrate *in vivo* the separation of the two strands of the parent DNA molecules. The basic requirement in this replication mechanism is the "reading" or recognition by a deoxynucleoside triphosphate of a complementary site on one of the separated DNA chains. Thus the adenine

residue is hydrogen bonded specifically to thymine, and the guanine residue specifically to cytosine.

Recently Weiss (1960) and Hurwitz *et al.* (1960), as well as other investigators, have shown that another enzyme, RNA polymerase, is able to produce a complementary RNA strand using the DNA molecule as a template. If two-stranded DNA is used, both chains are active in producing RNA complements. However, it has been shown that a single deoxynucleic acid strand is capable of acting in this system to produce its complementary RNA strand. In this reaction the significant interaction also involves the "reading" of a DNA polynucleotide strand by individual RNA nucleoside triphosphate units. The basis of this action is the specificity of the hydrogen bonding interactions between adenine and uracil, guanine and cytosine. Following the polymerization of the RNA strand *in vivo*, it is believed that the RNA is liberated from the DNA and is then released into the cytoplasm of the cell. This has been called Messenger RNA.

There are a great number of unknown factors in both of the polymerization reactions described above. For example, we know very little about the regulating control mechanism. In addition, although it seems clear that both strands of the DNA molecule are active in producing their complementary DNA strands under the action of DNA polymerase, it is by no means clear at the present time that both strands of the DNA act to form their complementary RNA strands. In an isolated enzyme system it is possible to show that both of the DNA strands are capable of making their complementary RNA strand, but it has yet to be demonstrated how the reaction proceeds *in vivo*. The significance of producing one or two strands of RNA becomes apparent when one considers the function of the material which is released.

Recently we have come to understand a great deal concerning the mechanism of protein synthesis. The most remarkable feature of the protein synthetic process is the extent to which the nucleic acids have a central role in organizing the assembly of amino acids into completed protein chains. This is also shown diagrammatically in Fig. 1. Short lengths of RNA, called Transfer RNA, containing some 90 nucleotides have been shown to react with the amino acids to make an RNA-amino acid complex. This reaction is carried out on the surface of an activating enzyme. At the present time the evidence suggests that there is a specific Transfer RNA molecule

for each amino acid, as a specific activating enzyme. The activating enzyme is specific in that it only accepts a particular amino acid (actually, an amino acid-adenylate), and a particular Transfer RNA molecule, presumably because it can detect or "read" a segment along the polynucleotide chain.

The newly polymerized Messenger RNA which is released from the DNA molecule contains a sequence of bases complementary to the sequence found in the DNA molecule. This material is metabolized very rapidly and has been detected only recently through the use of rapid pulse labeling techniques (Brenner *et al.*, 1961; Gros *et al.*, 1961). The Messenger RNA can be followed from the DNA to the ribosomal particles which are the sites of protein synthesis. These particles, roughly spherical in size with a molecular weight of approximately 4 million, have the ability to take up Messenger RNA as well as the Transfer RNA-amino acid complex. Polymerized polypeptide chains come out of the ribosomal particle and proceed to coil up into specific protein molecules.

Our understanding of the protein synthetic process is quite limited. Nonetheless, recent work by Nirenberg (see Nirenberg and Matthaei, 1961) has provided significant insight into the mechanism. He has shown that one can use synthetic polyribonucleotides which will act as synthetic Messengers, and the product is a synthetic polypeptide. Thus, for example, polyuridylic acid produces polyphenylalanine, while other ribonucleotide polymers or copolymers stimulate the addition of different amino acids. It is quite likely that this reaction occurs in the ribosomal particle through an interaction in which some of the purine and pyrimidine bases attached to the Transfer RNA find an appropriate site on the Messenger RNA so that there is a "reading" process which results in the correct alignment of the activated amino acids. In the case of polyuridylic acid, the Transfer RNA for phenylalanine probably has a series of adjacent adenine residues which combine with the uracil residues. After alignment, the amino acids are polymerized together starting from the amino end to form a protein chain.

A great deal of effort has been directed toward understanding the nature of this reading process, and an important quantity in this discussion is called the "coding ratio"; that is the number of nucleotides on the Messenger RNA molecule which interact with the nucleotides of a given Transfer RNA to determine the specificity of that site. Assuming that all four nucleotides are active, a pair of

nucleotides could only define 16 amino acids. Since there are about 20 major amino acids, it is believed that at least three nucleotides are needed in this process. Recent genetic experiments by Crick and his collaborators (1961) strongly suggest that the coding ratio is three. However, it is quite likely that the ratio will be determined in the very near future using systems such as those described above with synthetic Messenger RNA. In this regard it should be pointed out that although it is commonly assumed that the coding ratio is the same for all amino acids, this is not necessary, since Nature may have developed a system in which amino acids fall into different classes utilizing a different number of nucleotides for coding the different classes. However, this probably is unlikely in view of the genetic experiments cited above (Crick *et al.*, 1961).

We have mentioned the possibility that Messenger RNA may be made *in vivo* as complementary copies of one or both strands of DNA. If both strands are active, then the DNA would produce two RNA strands which are complementary to each other. Only one of these might be active in protein synthesis, and the other strand might be a component of the control or regulatory system. However, if both strands are coded for protein synthesis, we are faced with the possibility that proteins are manufactured in pairs related to each other by the two complementary Messenger RNA strands which are believed to determine amino acid sequence in the ribosomal particle. An alternative possibility is that the two complementary strands *both* synthesize the same protein; this could exist if there were a special type of degeneracy in the code. A given amino acid might have more than one triplet of nucleotides as its code letters, but they would have to be complementary. Thus, leucine might be AUG as well as CAU where the letters stand for the nucleotides in Transfer RNA which determine the code for leucine. Under such special conditions, complementary strands of Messenger RNA would synthesize identical proteins. It is quite likely that we will soon be able to answer such questions concerning the code through the use of synthetic Messenger molecules.

Using this brief outline, we can draw attention to a few generalities about nucleic acid function. First of all, the most important operation is the "reading" of one nucleic acid unit by another, which arises out of the specificity of the purine-pyrimidine interactions. Although many different types of hydrogen bonding interactions are possible among purines and pyrimidines (Rich, 1959), the most

important of these appears to be the pairs which are seen in DNA. Thus, the AT (or AU) pairing (Fig. 2a) and GC pairing (Fig. 2b) provide the specificity in DNA replication, RNA production, and probably are also used in the alignment of the activated amino acids on Messenger RNA during protein production. This is the most characteristic feature of the nucleic acids, and the stereochemistry of the molecules is especially well adapted for this type of interaction. The flat, unsaturated purine and pyrimidine rings interact with considerable van der Waals stabilization when they are stacked on top of each other with their flat surfaces opposed. The sugar-phosphate backbone has a preferred configuration such that the molecules can form helical aggregates around the centrally situated pile of purines and pyrimidines. Finally the detailed system of two or three hydrogen bonds between the purines and pyrimidines have sufficient specificity to determine the choice of a complementary base.

Information-containing molecules may be defined as polymers in which the sequence of residues is not random or regular, but rather has a specific and often complex pattern. All but a few of the proteins would fall in this class, and many of the nucleic acids but probably not all of them. Thus, ribonucleotide polymers which may be formed in bacterial cells through the action of polynucleotide phosphorylase probably have sequences which are close to random, since they are not made from a template. Possibly one can make a hypothetical generalization, namely that *all* information-containing polymers in biology are made by transferring sequence information from the nucleic acids. It is probable that this is true for the proteins, and perhaps this hypothesis may prove useful in understanding some aspects of polysaccharide synthesis. Several different types of sugar molecules are found in polysaccharides, and some of these polymers may contain information in the sense that they are not random or regular copolymers. This may be true for some of the cell wall polysaccharides or the blood group-specific polysaccharides. It may be that further investigation will disclose an RNA-dependent polymerization mechanism not too unlike that seen in protein synthesis.

We can use the term "information transfer reactions" to include those molecular interactions which are important in determining the sequence of information-containing polymers. According to the outline above, all of the information transfer reactions discovered up to the present involve the interaction of nucleic acids with other

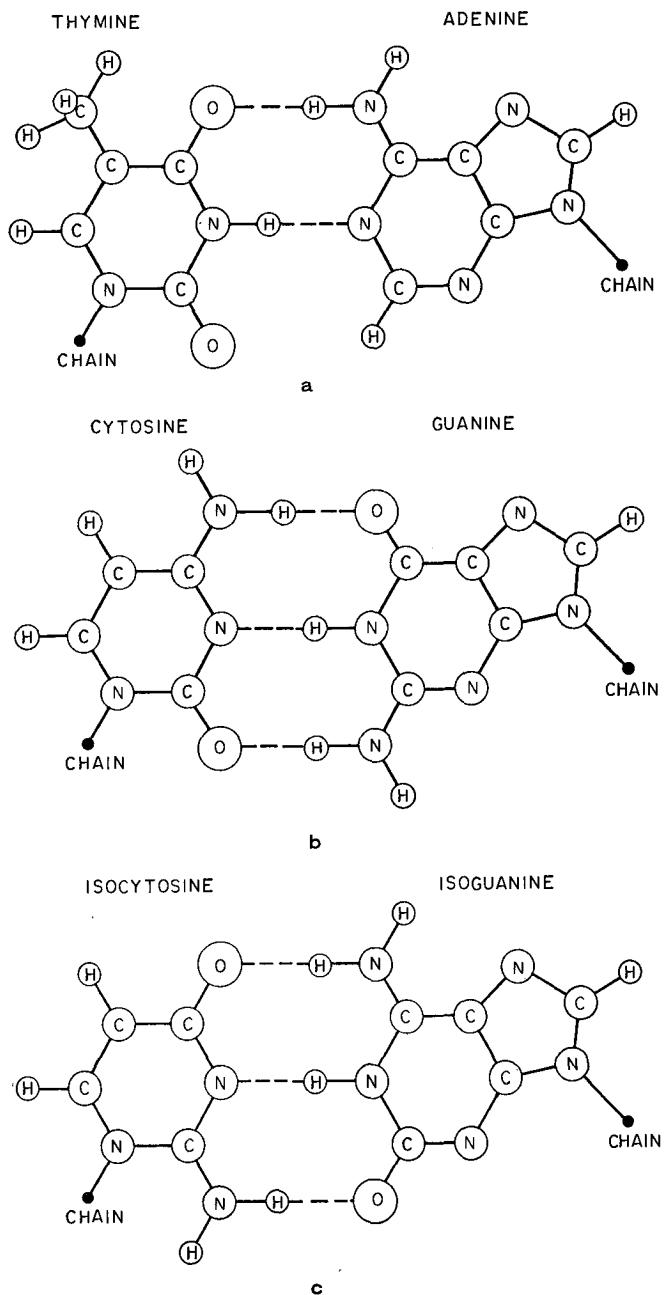


FIG. 2. Hydrogen bonding between (a) adenine and thymine, (b) guanine and cytosine, and (c) isoguanine and isocytosine.

nucleic acids or their components. This statement has one outstanding exception, namely the interaction of the Transfer RNA with its specific amino acid on the activating enzyme. This reaction does not necessarily involve a nucleic acid-nucleic acid interaction, and it is, of course, crucial in determining the correct sequence of amino acids within protein molecules.

III. Chemical Evolution and the Origin of Life

There are two levels on which one can begin to discuss the problem of the origin of life. The primary level consists of trying to understand the natural phenomena which produced a variety of molecular species during the early years in the evolution of our planet. Recently, several experimental approaches to this problem have been made by Miller and Urey (1959), and others. In general these experiments are designed to study the formation of biologically important substances such as amino acids in a primitive reducing atmosphere which is believed to have existed in the early stages of the evolution of the earth. Under the influence of an electric discharge through such an atmosphere, a variety of amino acids have been isolated. These experiments thus point to a particular route or reaction mechanism which would have made possible the accumulation of a large variety of organic molecules which gradually increased in concentration in a primitive sea.

During this period of what has been termed "chemical evolution," there would be a steadily increasing number of complex molecules, because the accumulation of simpler substances, such as some of the more stable amino acids, means that they could then become reactants leading to the formation of other organic substances. As yet there has been only limited experimental work in studying possible routes for the synthesis of nucleotide units by means of such reactions. However, recent experiments have demonstrated that primitive atmospheres are also capable of producing adenine (Oro, 1961) or uracil (Fox and Harada, 1961). Thus it is not unreasonable to believe that molecules as complex as nucleotides could have been made by nonenzymatic processes which utilize either radiation energy from the sun or electric discharges in a primitive atmosphere and it is likely that they will be observed with further experimentation. For the purpose of this discussion, it will be assumed that primitive seas some three billion years ago contained molecules of this complexity at a stage which antedates a

molecular organization which we would call "living" by present-day standards.

The second level to be considered in the origin of life is the development of replicating, information-containing polymer molecules. What system one prefers to call "living" is, of course, arbitrary, but perhaps a reasonable description of a primitive "living" system might be one in which there is an autocatalytic replication of an information-containing polymer which utilizes monomeric components from the environment and incorporates them into replicas of itself. This system, once started, would then be subject to all the modifying influences of the environment and we could call this the beginning of life as we know it.

There are two characteristic types of large molecules which are found in living systems, the proteins and the nucleic acids, and they can each in a sense serve as the basis for a theory concerning the origin of living systems. In each case, one imagines that energy derived from the sun, either in terms of radiation or electric discharges serves directly or indirectly to create conditions which bring about the polymerization of amino acids or nucleotides to form polymer molecules. The events which might happen subsequent to the formation of either primitive proteins or nucleic acids are quite distinct and we can examine them separately.

IV. Did Life Originate with Protein Molecules?

According to this theory emphasis is placed on the importance of the catalytic role played by primitive protein molecules in creating the necessary environment for self-replication. It is assumed that polypeptides are made by a random polymerization of amino acids and among these are found a few molecules which are catalytic and can act as enzymes in assembling the nucleotide units which we assume to be present in the primitive sea. At the same time another enzyme catalyst is formed, also by random assembly, which facilitates the replication of the primitive polynucleotide chain. Thus it acts as a primitive nucleic acid polymerase, organizing complementary mononucleotides along the chain leading to the formation of a two-stranded molecule. This can be followed by a separation of the two individual strands so that the process can continue. However, such a system has not yet coupled the synthesis of protein enzymes with the informational content of the nucleic acid, and this cannot be accomplished by a random amino acid polymerization.

Here is the greatest weakness in a theory of this type, since it does not really explain this most difficult step, the evolution of nucleic acid-controlled protein synthesis. However, another type of theory can be described as the basis for the origin of life which places the nucleic acids in a more central role.

V. Polynucleotides as the Origin of Living Systems

Theories of the biochemical origin of life on this planet were seriously developed about 20 to 30 years ago. At that time, it was quite clear that proteins were the most characteristic molecules in living systems and that their specific catalytic properties were essential to the functioning of biochemical systems. Accordingly much attention was devoted to ways of developing primitive protein molecules through nonliving agents. However, the almost explosive development of our understanding about the role of the nucleic acids during the past decade has made it imperative for us to reformulate theories about the origin of life in order to place the nucleic acids in proper perspective. As discussed above, the sequence of amino acids in proteins is, in a sense, a derivative of the sequence information encoded in the order of nucleotides in some part of the nucleic acids. Accordingly, it may be more reasonable to consider a theory of the origin of life in which the nucleic acids were developed as the primary agents.

Here we imagine a large number of nucleotide monomers among other molecules which are floating freely in a primitive sea, having been created by chemical reactions promoted by the products of ultraviolet radiation or electric discharges. We postulate that these nucleotide units can be assembled in random chains to make primitive, single-stranded polynucleotides in the absence of any protein catalysis. This is not a completely absurd hypothesis since we are able to do this chemically at the present time, although under somewhat special conditions. Methods for producing polynucleotide chains have been developed recently by Khorana and his associates (see Tener *et al.*, 1958) as well as by Schramm *et al.* (1961). Both of these polymerizations depend upon the use of strong dehydrating agents in a nonaqueous medium. In the presence of the dehydrating agents, water is split out of the nucleotides and they are joined together to form polynucleotides. In the case of the synthesis by Khorana, molecules of molecular weight up to five thousand or so have been formed while Schramm's recent synthesis has yielded

macromolecules of molecular weight greater than fifty thousand. Further experiments of this type should tell us more about the conditions under which such a polymerization occurs. However, we now know that it is possible to make polynucleotide molecules in the absence of a catalytically active protein.

The next step is the development of replicating nucleic acids. We postulate that the primitive polynucleotide chains are able to act as a template or as a somewhat inefficient catalyst for promoting the polymerization of the complementary nucleotide residues to build up an initial two-stranded molecule. This is in a sense the same system as that whereby a single DNA strand is able to make new RNA or DNA strands by the stepwise addition of complementary nucleotides which are subsequently polymerized together. However, we postulate that this may happen in a primitive environment in the absence of protein catalysts. We might ask if we have any evidence that a reaction of this type could occur. It would be of great interest if we could demonstrate experimentally that single polynucleotide chains have a tendency to bind or act as condensing sites for their complementary bases or nucleotides. Experiments of this type could be carried out at the present time. However, a very interesting effect was noted recently by Schramm *et al.* (1961) in the chemical polymerization of polyuridylic acid. He found that the addition of polyadenylic acid increased the rate of polymerization of polyuridylic acid over tenfold. This suggests that the polyadenylic acid is acting as a condensing site for the complementary uridylic acid residues which are then polymerized more rapidly. Again, further experiments could be carried out on this type of nonenzymatic polynucleotide catalysis to determine the limits of specificity.

What is being pointed out here is an important difference between proteins and nucleic acids. There are significant stereochemical reasons why the polynucleotides can act as their own catalysts for self-replication. However, there are no analogous reasons for believing the polyamino acids have this ability to reproduce themselves.

At this stage we imagine a system which contains nucleotide monomers together with polymers which, under some conditions, are able to make two-stranded versions of themselves using a system of complementary hydrogen bonds. This nonenzymatic replicating system would undoubtedly be very inefficient and slow. It is possible that the two-stranded form of the nucleic acid might then be thermally denatured by heat from the sun and separated into

single strands again. The process can then continue again with the manufacture of an increasingly larger number of nucleic acid chains. The concentration of these nucleic acid polymers might become appreciable since the rate of chain destruction might be as slow as the rate of chain production.

Now we must consider a possible modification of this primitive chain replicating action. We imagine that the mononucleotide units which are about to be incorporated into a polynucleotide chain occasionally have other small molecules attached to them. Thus, for example, if the oncoming nucleotide unit is a ribose nucleotide, the 3'- and 5'-hydroxyl groups might be used in polymerization, but there is an additional hydroxyl group found at position 2'. It might have attached to it another type of molecule which would then remain attached to the growing polynucleotide chain. In particular this residue might be an organic acid linked to the hydroxyl group by an ester linkage. It might also be an amino acid attached through the same linkage. At this stage we theorize this would be the beginning of a prototype system in which the polymerization of a nucleic acid molecule is coupled with the assembly of a series of amino acids. The amino acids might then be subsequently polymerized once they are organized in a linear assembly by their attachment to the adjoining nucleotides. In a sense it is this juncture in any theory of the origin of life which will present the greatest difficulty, since in order to develop the type of information-transferring system seen in biochemical systems, we must have a coupling of the nucleotide sequence with an amino acid sequence in proteins. However, to do this we must evolve specificity in amino acid sequence in order to develop those catalytic functions which make the system operate. In short we must try to develop a method for evolving the activating enzyme function.

Let us explore this point in more detail. We could imagine that a random polynucleotide chain, acting as a condensing agent and randomly polymerizing amino acids, could produce a molecule which specifically attaches glutamic acid to a particular nucleotide such as the adenine unit (or to a triplet of nucleotides, or even larger). The chance development of this catalytic molecule might give rise to large numbers of adenylic-glutamic units, but these would not aid in the recreation of the original activating enzyme, since the latter was produced by a random association of amino acids with nucleotides. Hence, such a mechanism could not "go critical" and

lead to the synthesis of specific protein molecules governed by the polynucleotide sequence. Thus we must look to alternative methods for developing an activating enzyme.

Let us imagine one way in which this could have developed. For example, the initial form of this contaminated nucleic acid synthesis might have individual amino acids attached to individual nucleotides. Thus the coding ratio for such a primitive system is one. If four nucleotides were used in this primitive system, then such a mechanism could produce only random polypeptide chains if any amino acid could be attached to a particular nucleotide. However, we might imagine conditions in this primitive environment in which there was some selectivity. For example, the adenine nucleotide might have a preference for attachment to glutamic acid, and so on for the other three bases. This might lead to the development of primitive polypeptide chains which have predominantly four different amino acids, namely those four which have the greatest preference for attachment to the four individual nucleotides. In such a hypothesis, the basis of selectivity must lie with some intrinsic stereochemical property or reactivity of the given amino acid with a given nucleotide in the absence of protein catalysts. At the present time there is no information about specificity of this type. It would be of interest to see whether there are any affinities among the different nucleotides toward certain of the amino acids. For example, are certain amino-acyl nucleotides formed preferentially in a chemical synthesis? This problem could be subjected to investigation in the laboratory at the present time. Alternatively, we cannot overlook the possibility that other components of the primitive environment, such as the surface of minerals or clays, could serve in the role of a selective catalyst.

The first important step in the development of this system must be the creation of prototype polypeptide catalysts which act as activating enzymes to reinforce the crude selectivity which prevailed initially, as in combining adenylic acid with glutamic acid in the example cited above. Once such primitive activating enzymes appear, the system starts to exhibit specificity, and there begins to be a relation between nucleotide sequence and amino acid sequence.

It is very difficult to overestimate the pivotal role of the activating enzymes in any evolutionary theory. These are the molecules which relate polynucleotides to polyamino acids and the entire mechanism of specific synthesis hinges on their evolutionary devel-

opment. These enzymes had to be among the first made in an evolving biochemical system; before they appeared there would only have been random polyamino acid polymers.

To get the nucleic acid-protein system fully coupled, we imagine that a small number of randomly arranged polynucleotide chains begin to form polypeptide chains which act as activating enzymes, so that the nucleic acids are "read" in a unique manner. If the initial coding ratio were 1, then four such enzymes would be necessary if four nucleotides were present in this primitive nucleic acid. If the primitive nucleic acid were composed of only 2 nucleotides, as discussed below, then only two activating enzymes would be necessary. This is a minimal number, since more than one enzyme is needed to obtain specificity.

Once such a system is developing specific amino acid sequences which are coupled to nucleic acid sequences, then we may consider that "life" has started. Some of these primitive nucleic acid chains may have a nucleotide sequence which leads to a polymerase enzyme, and the production of that type of nucleic acid would be stimulated preferentially over others; in short, natural selection on the molecular level can commence. If the primitive coding ratio were 1, we can see that considerable advantages would develop with a mutant in which nucleotides were taken as pairs (or triplets) in the activating step, since they could then incorporate other amino acids and develop polypeptide chains of greater chemical subtlety.

The major thesis in this development is the idea that what we may call "life" started with the coupling of nucleic acid polymerization and amino acid polymerization. This process made an important step forward when specificity of protein synthesis was developed in the system through the discovery of activating enzymes. However, the prototype of this reaction may have been in a form quite different from that which we observe today.

VI. The Trend Toward Increased Complexity

The study of evolution on the molecular level is still very new. Only in recent years have we developed the tools which are necessary to determine the sequence of amino acid residues in the proteins. Thus we can now begin to study their molecular evolution, and it is quite likely that this type of effort will be continued for many years. However, at the present time, we do not have guide lines for our thinking, that is, general concepts about molecular evolution

which may prove of utility in understanding and interpreting the results.

We have a considerable body of information about macroscopic evolution in terms of the changes, with time, of organisms and the development of new species. From this we see that the mechanism of evolution is such that slow refinements are usually made within a given organism in relation to its environment which lead toward increased efficiency of reproduction. Occasionally there are discontinuities which are opportunistically utilized in the exploitation of a new environment or of an altered physiological function. In general, the trend in macroscopic organic evolution is toward increased complexity. We can trace the development of increasingly subtle physiological functions and greater control leading to a progressive extension of the organism's ability to utilize the environment and maximize reproductive ability.

It is reasonable to ask whether a similar principle applies on the molecular level. Thus, has there been an increase in the complexity of molecular organization and function during the course of evolution? It is perhaps not unreasonable to believe that this may have occurred and accordingly we can inquire about the various ways in which this may have developed, especially in relation to the biological information transfer system.

In the previous section we described a primitive, self-replicating nucleic acid system which has started to develop a feed-back mechanism in which a modification of the replicating ability of the nucleic acids is associated with the polymerization of a polypeptide chain, which can then be liberated to form a catalytically active primitive enzyme. Once this system is able to form the activating enzymes it would then tend to become more directional in its development. Some classes of nucleic acids which, through chance, have developed a nucleotide sequence which coded for a set of functional proteins, would tend to become more numerous. Perhaps an initial reason for this is simply one of concentration. If a given nucleic acid started producing a nucleic acid polymerase enzyme in addition to its activating enzymes, larger amounts of this type of nucleic acid would be produced because of the higher concentration of the enzyme in its vicinity.

There are several parts of this molecular system in which increasing complexity could manifest itself. We can list these and discuss possible historical alterations which may have occurred in the

system. Such a listing may be of some value in that it may help to interpret data on biochemical evolution as it is uncovered. In addition, we are on the threshold of space exploration and we may have an opportunity to experimentally study more primitive life forms.

VII. Changes in the Composition of the Nucleic Acids

We may ask why the nucleic acids have four units in them at the present time (for this discussion we will equate thymine and uracil, and ignore the difference between sugars). In order to contain information, it is obvious that two bases would be enough; for example, simply adenine and thymine. A primitive organism whose nucleic acid contained only two complementary bases could still develop a similar type of biochemical system for information transfer. The sequence of the adenine and thymine residues along one chain carries information in many ways analogous to the dot-dash system used in the Morse code. In such a hypothetical organism, if one had to code for twenty amino acids, it would require five bases in order to obtain a minimum of twenty combinations. Nonetheless, in principle, a system of this type would be functional.

Is it possible that there may have existed an early form of life which used nucleic acids with only two bases? For example, it has been shown by Kornberg and his collaborators that the DNA polymerase enzyme, when left without a primer in the presence of the nucleotide triphosphates, will make an AT copolymer. This is a DNA-like molecule with a regular repeating sequence of adenine and thymine. One is tempted to ask whether this could be an atavistic form of a nucleic acid. It is also worth noting that the crab contains a DNA molecule which is almost entirely an AT copolymer (Sueoka, 1961). In this regard it will be of great interest to study the base composition of nucleic acids as a function of phylogenetic evolution to see whether there is any significant trend. Preliminary steps have already been taken in this direction by Sueoka (1961).

One can also ask why has Nature used nucleic acids built with four bases instead of six, for example. In a more general way, one can ask whether it is possible for nucleic acids to be built out of additional complementary purine-pyrimidine pairs which would still maintain the same geometry and specificity in the hydrogen bonds such as is now seen in DNA. The answer is not particularly clear since a great deal depends upon the tautomeric form of the bases. However, if we restrict ourselves to amino rather than imino

groups and keto rather than enol groups, another possible type of complementary pair is that between isoguanine and isocytosine, as illustrated in Fig. 2c. If this is the stable tautomeric form of these two molecules, then they could also be incorporated into a nucleic acid molecule and would have the requisite specificity. That is, they would hydrogen bond only with each other but not with any of the other four bases. It is important to note that this requires that the hydrogen atom in isoguanine be located on N-1 of the ring rather than on N-3. It is not clear whether this condition is fulfilled in isoguanine. However, it is sufficient to note that there are *very few* additional complementary pairs which one can fit into the DNA structure which would have the necessary specificity. It is quite clear, however, that a nucleic acid built out of six bases using the same principle of complementarity could function suitably in a biological system. To code for twenty amino acids in such a system, it would only be necessary to have two bases to contain the necessary information, rather than three.

VIII. Has the Number of Amino Acids Increased During Evolution?

Due to the recent discovery by Nirenberg (Nirenberg and Matthaei, 1961) that we can use synthetic polyribonucleotides as a synthetic form of Messenger RNA to direct the synthesis of polypeptides, it seems quite likely that we will have the entire amino acid code deciphered very soon. The recent work by Crick and co-workers (1961) on a series of acridine orange mutants makes it quite likely that the code is a triplet, and it also points out the possibility that the code is degenerate. Thus there may be more than one triplet of nucleotides which specifies a given amino acid.

It is usually stated that there are twenty amino acids; however, this ignores the fact that several additional amino acids are found in nature but their distribution is limited. Examples are the hydroxyproline or hydroxylysine amino acids which are found in collagen and related proteins. It should be relatively easy to determine whether or not these minor amino acids have separate code letters, i.e., groups of nucleotides assigned to them. Thus, for example, if one found that hydroxyproline had a particular soluble RNA molecule to which it was attached by means of a special activating enzyme, then it would be quite likely that there is a special code letter which is unique for hydroxyproline. Similar investigations

can be extended to include the other minor amino acids. The results of these investigations may show that there are only twenty amino acids which can be incorporated into proteins through the activating enzymes. But perhaps a more likely finding is that the number twenty is not unique, and that there are code letters attached to some of the other so-called trace amino acids. This could represent one of the ways in which increasing complexity has been built into biochemical systems. Thus the number of amino acids used by biological systems to build up proteins may be a time dependent function. In a very primitive biochemical system, only a small number of amino acids may have been used and gradually, as the system evolved, finer degrees of functional control were developed through utilizing other amino acids with slightly different chemical side chains attached to them. Such a hypothesis can be examined experimentally.

It should be noted that a degenerate coding system for amino acids is well adapted for this type of evolutionary change. For example, suppose leucine were represented by four sets of nucleotide triplets. A mutational error might assign one of these triplets to a new amino acid, say hydroxyproline, but leucine would continue to be used by the organism. Accordingly the system would tend to move toward a state of less degeneracy and a wider distribution of amino acids. It is quite clear that considerable selective advantage would accrue with the development of a coding system which allows for a gradual increase in the number of amino acids which can be utilized.

It is important to recognize that the imminent decoding of the relation between the Messenger RNA and the polypeptide chain which it helps to synthesize will throw considerable light on the possibilities outlined above.

IX. An Evolutionary Increase in Nucleic Acid Content

If we survey the DNA content of a variety of organisms, we note that the mass of DNA increases with the complexity of the organisms. In a similar fashion it is not unreasonable to believe that the amount of genetic information contained by individual primitive cells may have increased through the very early stages of biochemical evolution. For example, there may have existed at one time a primitive organism containing enough nucleic acid to code for five or ten proteins. This could be possible in a specialized en-

vironment in which the concentration of many of foodstuff molecules were built up by natural processes as discussed earlier. In this system, the nucleic acid replicates in a manner similar to the replication of DNA as we now understand it. Thus there is a building up of complementary chains and the separation of the two original nucleic acid strands of the parent molecule.

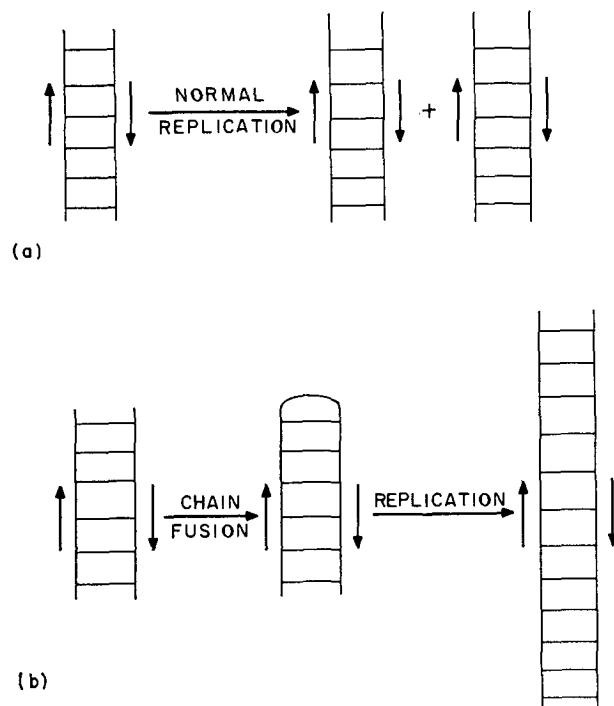


FIG. 3. (a) The replication of one nucleic acid molecule to produce two identical daughter molecules. (b) The effect on replication of joining the ends of the parent molecule.

We might ask what would happen in such a system if, under the action of a cosmic ray or other radiation, the ends of the two-stranded DNA molecule were joined together with a normal covalent bond. There would result, as shown in Fig. 3, a doubling in length of the original nucleic acid molecule. During the next cycle of replication the DNA would have twice the original length. It has the same informational content as the original nucleic acid but there

are two copies present. An accident of this type may seem to offer no selective advantage. However, during the process of mutation, alterations in these bases may occur in one of these copies but not in the second one which codes for the same protein. This means that a mutation which might be lethal for the organism would not be lethal in this duplex state. Accordingly an anomaly of this type allows for the development of a variety of new protein molecules which can, in a sense, explore the environment more readily than the original organism with only a single copy of its genetic information. Here the luxury of surplus, redundant information appears to provide a selective advantage in evolution. If this process goes on several times, eventually the organism ends up with many genetic copies of an original prototype protein molecule. These copies might then evolve along somewhat separate evolutionary lines and give rise to classes of molecules which, though similar, are in fact different in many ways. It is possible that this may be the origin of certain classes of related proteins such as myoglobin and the α - and β -chains of hemoglobin, in which there are many similarities in both amino acid sequence and secondary structure.

It might be possible to carry out experiments to determine whether or not the process outlined in Fig. 3b is possible. Thus, a homogeneous solution of DNA could be subjected to radiation, and an attempt could be made to detect a small number of longer molecules after treating with the DNA-polymerase system.

X. Why Are There Two Nucleic Acids?

It is quite remarkable that contemporary biochemical systems have two nucleic acids, DNA and RNA, which differ only by a systematic hydroxyl group and an occasional methyl group. Despite the great chemical similarities, the molecules nevertheless have quite different functions in the cell. DNA acts as the major carrier of genetic information, while the RNA molecule is used to convert this genetic information into actual protein molecules. Because of the close chemical similarities, we are tempted to ask whether they could have originated historically from a common stem nucleic acid molecule which then specialized in the course of evolution to produce the two different classes of nucleic acids which we see today. To discuss this further we should note that the RNA molecule is also able to carry genetic information, as, for example, in the RNA-containing viruses. Thus, it may be reasonable to speculate that the

hypothetical stem or parent polynucleotide molecule was initially an RNA-like polymer which was able to convey genetic information as well as organize the amino acids into a specific sequence to make proteins. This implies that the RNA polynucleotide strand had the ability to replicate itself and produce a complement in a manner somewhat similar to that which is found in DNA. It is possible that an enzyme of this type may have been observed already. By this view, DNA may be regarded as a derivative molecule which has evolved in a form such that it only carries out part of the primitive nucleic acid function. It specializes in the molecular replicating cycle that is part of the mechanism for transmitting genetic information. DNA is metabolically less reactive than RNA, perhaps because of the absence of the hydroxyl group on carbon 2. The loss of this hydroxyl group may have made it impossible for the DNA molecule to have attached to it the amino acids which are used in protein synthesis. However, considerable selective advantage may be derived from the development of two different classes of nucleic acids, one of which is less active metabolically and specializes in self-replication. In a sense, this tends to preserve the primary copy of the genetic information. It will be of considerable interest to study the available simple life forms to see whether some of them may exist with only one type of nucleic acid rather than two types. It is possible that the RNA containing viruses may be regarded as present-day examples which may have degenerated evolutionarily from such a primitive life form.

It should be noted that the changes which are discussed in many of the sections above may only have occurred during the very early stages of evolution. What we see today is the net result of a complex competition among many different types of systems.

XI. Extraterrestrial Life

We can make a few generalizations from our description of the molecular basis of life as we know it on this planet. The most characteristic feature of a living system is the fact of its enormous complexity on the molecular level. There are a large number of chemical reactions going on which involve the manufacture of many specific protein catalysts which are used as regulators for facilitating these reactions. However, the information necessary to make these large molecule catalysts is stored in polymer chains in a much more compact form than the mass of catalyst molecules themselves. Indeed

one might say that a characteristic of life is the storage of information in polymer molecules and its subsequent expression in terms of building other polymer molecules which have specific functions, such as making catalysts, cell walls, and other specialized structures.

We are now on the verge of being able to carry out an exploration of space to look into the possibility of life existing on other planets or, eventually, other stellar systems. If life exists in these other sites, it is possible that it may exist in quite a different form from our own. Thus, for example, we could imagine that a living system evolving in a somewhat different environment, may have selected different classes of molecules to be built into polymers both for storing genetic information as well as for carrying out the large number of necessary chemical reactions. Indeed the molecular basis of other life forms may be entirely distinct from ours. However, it is quite likely that there would be one major feature in common with terrestrial life, namely, the use of polymer molecules to store the large amount of information which is needed to adequately define the complexity of events which go on in a system which we call "living." However, the polymers might be quite different and the molecular composition of other living systems might bear no further resemblance to terrestrial life. In particular, the nucleic acids which play such a central role in transmitting information in our form of life, may not be found at all in other forms of life but may be replaced by a different type of polymer.

XII. Conclusions

In this essay the essentials of biochemical information transfer have been outlined briefly. An attempt has been made to discuss various problems which arise in trying to understand how this information transferring system may have evolved. Since we have an almost complete lack of factual information concerning such events, an enterprise such as this must, of necessity, be speculative. Nonetheless some of these reflections may serve as a stimulus for additional experimental approaches which will enable us to place biochemical evolution on a firmer foundation.

REFERENCES

- Bessman, M. J., Lehman, I. R., Simms, E. S., and Kornberg, A. (1958). *J. Biol. Chem.* **233**, 171.
Brenner, S., Jacob, F., and Meselson, M. (1961). *Nature* **190**, 576.

- Crick, F. H. C., Barnett, L., Brenner, S., and Watts-Tobin, R. J. (1961). *Nature* **192**, 1227.
- Fox, S. W., and Harada, K. (1961). *Science* **133**, 1923.
- Gros, F., Hiatt, H., Gilbert, W., Kurland, C. G., Risseeuw, R. W., and Watson, J. D. (1961). *Nature* **190**, 581.
- Hurwitz, J., Bresler, A., and Drenger, R. (1960). *Biochem. Biophys. Research Commun.* **3**, 15.
- Meselson, M., and Stahl, F. W. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 671.
- Miller, S. L., and Urey, H. C. (1959). *Science* **130**, 245.
- Nirenberg, M. W., and Matthaei, J. H. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1588.
- Rich, A. (1959). *Rev. Modern Phys.* **31**, 191.
- Schram, G., Grötsch, H., and Pollmann, W. (1961). *Angew. Chem.* **73**, 619.
- Sueoka, N. (1961). *J. Mol. Biol.* **3**, 31.
- Tener, G. M., Khorana, H. G., Markham, R., and Pol, E. H. (1958). *J. Am. Chem. Soc.* **80**, 6223.
- Weiss, S. B. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1020.

Life in the Second and Third Periods; or Why Phosphorus and Sulfur for High-Energy Bonds¹

GEORGE WALD

*The Biological Laboratories, Harvard University,
Cambridge, Massachusetts*

It is one thing to ask questions and another to answer them; but of the two, the former is by far the rarer and more difficult—that is, if the questions are of the right sort, capable of being answered, and leading to the next questions. Szent-Györgyi has been having such a dialogue with Nature all his life; from inside, of course, not from outside, since he is part of Nature, like all of us. It is hard to know which is Socrates and which the Athenians, which is gadfly and which stung; but if one stays within earshot, one is almost sure to be stung too.

Some time ago, talking with Szent-Györgyi, I was a little startled to hear him ask, why phosphorus plays the peculiar role it does in organisms. It was a question I had asked myself, as part of a larger question: why any of the atoms that compose organisms occupy the position they do.

This is clearly a proper question, for there is nothing random about the choices. There are the atoms that form the bulk of the organic molecules: C, H, N, and O, and in special instances S and P; there are the major monatomic ions Na⁺, K⁺, Mg⁺⁺, Ca⁺⁺, and Cl⁻; there are the trace elements, mostly transition elements, and hence adapted to fill the roles in which we mainly find them, as nuclei and ligands in metallo-organic complexes and oxidoreduction enzymes: Fe, Mn, Co, Cu, Zn. All these elements are relatively light: the 16 elements mentioned fall within the first 30 in the Periodic System. Of the remaining 62 natural elements, only two—I and Mo—have restricted roles as trace elements in certain organisms. By and large, the lightest elements tend to be the most available; but

¹ The general lines of this discussion were included in a paper called "Toward a Universal Biochemistry," presented to the American Philosophical Society in Philadelphia at its Fall Meeting in November, 1960.

except for the dominant position in the sea of the monatomic ions mentioned above, I think it is clear that something other than simple availability has governed the choice of the elements of which organisms are principally composed.

That "something other" is *fitness*, the peculiar combination of properties that makes these elements most suitable to play the parts that organisms demand of them. As one pursues this argument, it is to realize, I think, that organisms have had little choice among the elements. For the most part, they had to end up as they did. For that reason I think that not only the Periodic System, but the constitution of organisms, is probably about the same everywhere in the universe.

The question we ask is therefore meaningful; and it is within this larger context that I came to ask the question Szent-Györgyi asked: why phosphorus?—except that, as we shall see, I would prefer to ask at once, why phosphorus and sulfur?—for sulfur is part of the same question. And why *not* silicon?—that's part of the question too. It's all a matter finally of being in the Second and not in the Third Period of the Periodic System.²

About 99% of the living parts of living organisms are made of the four elements, H, O, N, and C. Some time ago I asked myself, why these?—and after a time thought I knew the answer. *These are the smallest elements in the Periodic System that achieve stable electronic configurations by adding respectively 1, 2, 3, and 4 electrons.* C, N, and O follow one another directly in the Second Period. Next comes fluorine, F; but fluorine, though it too gains a stable electronic configuration by adding one electron, has no part in making organisms, because in any of its properties that might matter for this reason it is outdone by the smaller element, hydrogen. Adding electrons is part of the story, but smallness is the rest. The only thing better than being in the Second Period is to be in the First.³

² There is some equivocation in numbering the periods of the Periodic System. I shall call the First Period H and He; the Second Period running from Li to Ne, and the Third Period from Na to Ar. I shall speak therefore of C, N, and O as in the Second Period, and Si, P, and S as in the Third.

³ It should be added that the tendency to add electrons, carried to an extreme, ends in forming anions rather than molecules. Their strong tendency to form anions is one reason why fluorine—the most electronegative of the elements—plays no role in biochemistry, and why chlorine plays its principal role as an ion. This is also why I have paid so little attention to chlorine in the discussion of multiple bonding.

What it means to add electrons is clear enough. Adding electrons by sharing them with other atoms is the mechanism for forming chemical bonds, and hence molecules. But what has smallness to do with it? Two things: (1) the smallest atoms ordinarily form the tightest, most stable bonds; and (2) they alone regularly form stable multiple bonds. Many years ago G. N. Lewis (1923, p. 94) commented that "the ability to form multiple bonds is almost entirely, if not entirely, confined to elements of the first period of eight, and especially to carbon, nitrogen and oxygen." Indeed he suggested that were it not for these elements, the concept of the multiple bond might never have been invented. Incidentally the rare occasions when one does encounter multiple bonds outside this trio involve most frequently sulfur and phosphorus.

All of this comes to a head in the comparison of carbon with silicon. In those parts of the earth even remotely available to living organisms, silicon is about 135 times as plentiful as carbon. In the surface layers of the earth, including the atmosphere and hydrosphere, silicon constitutes 16.08% of all atoms, carbon only 0.119%. Coming directly under carbon in the Periodic System, it shares with carbon the property of tending to gain four electrons, and so to form four covalent bonds by sharing electrons with other atoms. Why then is life on earth based upon the relatively rare element carbon rather than on the prevalent silicon?

First, the strength and hence stability of bonding. These are shown in the following table, from which it is clear that the interatomic distance (bond length) is much smaller in a C-C than in an Si-Si bond, and the bond energy of the former is almost twice that of the latter:

	Interatomic distance (Å)	Bond energy (kcal per mole)
C-C	1.54	80-83
Si-Si	2.34	42.2

It is clear that to the degree that the business of making organisms includes the capacity to form tight, stable bonds, and eventually long, stable chains of atoms, carbon has a large intrinsic advantage over silicon.

The importance for organisms of the capacity to form multiple

bonds can best be illustrated by comparing CO_2 with SiO_2 . In CO_2 , carbon is bonded to each of the oxygen atoms by double bonds, each involving the sharing of two pairs of electrons. By this means, each of the atoms of CO_2 achieves the complete octet of outer shell electrons found in the neighboring inert gas, neon. All the combining tendencies are satisfied, and the molecule, free and independent, goes off in the atmosphere as a gas, and readily dissolves in and combines with water, the forms in which living organisms obtain and use it.

In SiO_2 , on the contrary, Si is joined to oxygen by single bonds, leaving two unpaired electrons on the silicon and one on each of the oxygen atoms. Unable to pair by forming multiple bonds, these pair instead with the unpaired electrons on neighboring molecules of silicon dioxide. This process, repeated endlessly, ends in a huge polymer, in a sense a huge supermolecule of silicon dioxide. This is the essential structure of quartz—an extraordinarily dense, hard, inert material, which can be broken only by breaking covalent bonds.

Silicon has a third fundamental disability relative to carbon. Like carbon, silicon has a strong tendency to combine with itself to form chains. Potentially it should be possible for silicon compounds to exist in a variety and complexity rivaling the carbon compounds. From the point of view of life, however, the former have a fatal disability: Si-Si bonds are unstable in the presence of water, ammonia, or oxygen.

The reason for this is fundamental to our discussion. In such a Second-Period element as carbon, the outer shell contains one $2s$ and three $2p$ orbitals, each potentially capable of holding a pair of electrons of opposed spin, so completing an octet. (In forming four tetrahedral covalent bonds, the four second-shell orbitals hybridize to form four identical hybrid sp orbitals.) The formation of four covalent bonds—in whatever mixture of single and multiple bonds—in a sense completes the octet, filling all available second-shell orbitals, and that is the end of it.

In silicon, however, the outermost electron shell is the third. When the one $3s$ and three $3p$ orbitals have been filled, as the result of chemical combination, the atom has achieved a measure of stability, its electron configuration—though distorted in that the atomic orbitals have been replaced by molecular orbitals—approaching that of the inert gas argon. The completion of an octet of electrons, however, does not saturate the third shell. It possesses in addition

five $3d$ orbitals, holding potentially 5 further pairs of electrons. That is, though when the third shell is outermost its stable number of electrons is 8, as in argon, the formation of a fourth shell, as in the Fourth Period, ends in bringing the third shell number from 8 to 18, as in the elements from zinc to krypton. After silicon has in effect completed an octet by forming four covalent bonds, its outermost shell can still potentially accept further electrons in the empty $3d$ orbitals.

The molecules to which silicon chains are most susceptible have the particular characteristics of small size and the possession of lone pairs of electrons. The relatively large interatomic spacing of silicon chains (see above) seems to allow such molecules to come in close enough for their lone electron pairs to occupy empty $3d$ orbitals, and to disrupt the chain. Indeed much the same attack is made on linkages between silicon and other atoms than itself. For example, whereas methane (CH_4) is stable to water and sodium hydroxide, silane (SiH_4) is attacked by these substances to form sodium silicate and gaseous hydrogen ($\text{SiH}_4 + 2 \text{NaOH} + \text{H}_2\text{O} \rightarrow \text{Na}_2\text{SiO}_3 + 4 \text{H}_2$).⁴

We can state therefore three powerful reasons, one relative and two absolute, why silicon is unsuitable as a basis for forming organisms: (1) It forms much weaker bonds than carbon, both with itself and with other atoms. (2) Its reluctance to form multiple bonds results in the formation of huge, inert, covalently bonded polymers, so removing all but traces of silicon from circulation. (3) The instability of silicon chains and compounds in the presence of oxygen, ammonia, and water should suffice to disqualify them. I cannot imagine life existing at all apart from water, or going very far without oxygen; and both conditions a priori rule out basing the constitution of living organisms upon silicon.

It is precisely these disabilities of silicon that help to create a special opportunity for phosphorus and sulfur. I am thinking here of the opportunity for these elements to play their most distinctive role in organisms, as agents of group and energy transfer. They have of course other functions, but no others that make such special demands, or that single these atoms out so particularly.

Two properties should most facilitate this type of function: on

⁴ This is of course an example of what is called a nucleophilic attack, and the hydroxyl ion is here a typical nucleophilic reagent. I avoid this special terminology in the present paper in order to concentrate upon the phenomena themselves.

the one hand the capacity to form a wide variety of linkages of small and large energy potential, i.e., an adequate coinage; and added to this, an intrinsic instability of linkage, so facilitating exchange. The first property lies in the province of thermodynamics; as for the second, thermodynamics sets the stage, but what actually happens lies in another province, that of reaction mechanisms and kinetics. We will be mainly concerned with this latter type of consideration in discussing the special virtues of phosphorus and sulfur.

The thermodynamic aspects of energy and group transfer—the business of "low-energy bonds" and "high-energy bonds," as biochemists commonly employ these terms—have been carefully explored in the twenty years since Lipmann (1941) first formulated this type of concept (cf. Kalckar, 1946, 1947; Oesper, 1950; Hill and Morales, 1951; George and Rutman, 1960; Huennekens and Whiteley, 1960; Rutman and George, 1961). It is by now well recognized that these are poor terms; that one is dealing here, not with the energy localized in a bond—a bond energy in the strict sense—but with the change in free energy that accompanies a transfer reaction, i.e., with the differences in free energy between the reactants and their products. This free energy change in any given instance is compounded from changes in resonance stabilization, ionization, electrostatic forces—indeed all the consequences of reaction under the particular circumstances in which it occurs. It is in this sense that high- and low-energy compounds are at present defined on the basis of the standard free energy change that accompanies the hydrolysis of a particular bond; a ΔF° of -1 to -3 kcal per mole characterizes a low- and -5 to -10 kcal per mole a high-energy bond.

By far the widest variety of energy and group transfer reactions in biological systems is carried out by organic phosphates. Sulfur forms three types of "high-energy" complex, one consisting of acyl esters of thiols (e.g., acetyl CoA); another of mixed anhydrides of phosphoric and sulfuric acids [e.g., adenosine-phosphoryl-sulfate for sulfonations (cf. Lipmann, 1958)]; and a third of sulfonium compounds [e.g., S-adenosyl methionine for methylations (cf. Cantoni, 1953, 1960)]. This short list almost exhausts the known categories of "high-energy" compounds. One must add an—as yet—very limited class of acyl imidazoles, in which the acyl group is attached to a nitrogen atom (Stadtman and White, 1953; Stadtman, 1954); and

the highly important class of activated amino acids, in which the amino acid carboxyl group is joined in ester linkage with the 2'- or 3'-OH of ribose in the terminal adenylic acid of transfer-RNA (Zachau *et al.*, 1958; Hecht *et al.*, 1959; Preiss *et al.*, 1959). These latter instances are as yet the only ones I know in which group transfers are negotiated by something other than sulfur or phosphorus compounds.

What lends sulfur and phosphorus this special position? Or to put this question a little more in the context of our general argument, what properties do sulfur and phosphorus have that oxygen and nitrogen—their congeners in the Periodic System—lack, that fit them better to perform this type of function?

I think the answer to this question is to be sought in three directions, all multiply interconnected: (1) S and P form more open and in general weaker bonds than O and N. (2) S and P can expand their covalent linkages beyond 4 on the basis of their 3d orbitals. (3) To a unique degree among Third-Period elements, S and P retain the capacity to form multiple bonds.

(1) The interatomic distances (bond lengths) are larger in covalent linkages of S and P than in those of O and N. In making such comparisons one must consider comparable molecules, since the length of any particular bond varies somewhat with the remaining structure of the molecule. A first approach to this situation is gained in such tabulations of covalent bond radii (Table I) as given by Pauling (1960, pp. 224, 228):

TABLE I
COVALENT BOND RADII (\AA)

	C	N	O
Single bonds	0.772	0.74	0.74
Double bonds	0.667	0.62	0.62
	Si	P	S
Single bonds	1.17	1.10	1.04
Double bonds	1.07	1.00	0.94

The interatomic distances of bonded atoms, obtained by adding together the appropriate bond radii, are reasonably reliable for

molecules in which the atoms in question possess the numbers of covalent bonds ordinarily associated with their position in the Periodic System (i.e., C, 4; O, 2; P, 3; etc.), and in which the bonds do not possess too much ionic character or in which the bond order does not depart too greatly from a whole number.

In general the larger bond radii encountered in the Third Period go also with the formation of weaker bonds—bonds of lower energy—than those formed by their congeners in the Second Period. This point was made earlier in comparing Si-Si with C-C. So, for example, the bonds formed by these atoms with hydrogen have the energies shown in Table II (Pauling, 1960, p. 85):

TABLE II
BOND ENERGIES (kcal per mole)

C-H	98.8	Si-H	70.4
N-H	93.4	P-H	76.4
O-H	110.6	S-H	81.1

The first point in our argument therefore is that in going from the Second to the Third Period, from N and O to P and S, one goes in general to more widely open and weaker bonds, hence bonds intrinsically more susceptible to attack, and more ready to undergo cleavage and exchange reactions.

(2) We have already pointed out with reference to silicon that elements of the Third Period possess *d* in addition to *s* and *p* orbitals, and so have place to hold electrons beyond the normal outer-shell octet. It is this property that most distinguishes S and P from their congeners in the Second Period. The five 3*d* orbitals of S and P could hold potentially 5 pairs of electrons. It seems however that the possibilities of forming stable linkages are more restricted. Ordinarily phosphorus does not go beyond 5 covalent bonds, as in PCl_5 , and sulfur does not go beyond 6, as in SF_6 . Some formalisms prefer to represent such molecules otherwise than possessing 5 or 6 equivalent covalent bonds; but it is significant that, however represented, they call upon properties not shared by N and O, which do not form molecules of comparable valence.

We can expect therefore that as with compounds of silicon, the

presence of unoccupied 3*d* orbitals in S and P invites attack by molecules possessing lone pairs of electrons, the lone pairs occupying in an intermediate stage the 3*d* orbitals, with an exchange reaction as the eventual result.

(3) Sulfur and phosphorus, to a unique degree among Third Period elements, retain the tendency to form multiple bonds, otherwise so characteristic of carbon, nitrogen, and oxygen. Having stressed earlier in this paper that silicon does not markedly display this property, it is interesting to ask why sulfur and phosphorus possess it.

The tendency of elements to form multiple bonds seems to me to be associated somehow with small size as such. I have not been able to find a clear formulation of this view in the literature, nor can I formulate it myself. Coulson (1953, p. 178) ascribes "the experimental fact that multiple bonds are practically confined to the first two rows of the Periodic Table" to increasingly strong repulsive forces between nonbonding electrons as atoms grow larger. If one thinks of the bond as a pair of electrons shared between two atoms, this demands that the bonded atoms approach each other closely. Under such circumstances the electrons not engaged in bonding repel one another strongly, and the larger the number of such electrons the stronger the mutual repulsion. This is thought to be the reason for the relative weakness of bonds between the heavier elements.⁵ A multiple bond, since it involves still closer approach between atoms than a single bond, evokes still stronger repulsions. For this reason only atoms of small kernel, containing relatively few electrons, might permit the formation of stable multiple bonds.

This cannot be the whole story, however, for it would seem to imply that in the series Si, P, S, both bond strengths and the tendency to form multiple bonds should decline continuously, since the number of nonbonding electrons and hence the mutual repulsions

⁵ The large repulsion between nonbonding electrons has also been suggested to account for the abnormally low bond energies of N-N (38.4 kcal per mole), O-O (33.2), and F-F (36.6). N, O, and F form strong bonds with other elements in part because they are so highly electronegative; their bonds tend to have considerable ionic character. The bond energies of P-P (51.3 kcal per mole), S-S (50.9), and Cl-Cl (58.0) are relatively strong, because though the presence of more nonbonding electrons could lead to even stronger interatomic repulsions, these repulsions are relieved by the availability in these elements of 3*d* orbitals (cf. Pauling, 1960, p. 144).

are rising; whereas just the opposite is the case. The energy values of single bonds are: Si-Si, 42.2; P-P, 51.3; S-S, 50.9; and Cl-Cl, 58.0 kcal per mole; and S and P have more, not less tendency than Si to form multiple bonds.

It may be therefore that a small atomic radius as such promotes the tendency to form multiple bonds. If so, that raises an interesting consideration. The atomic radii of the elements do not rise continuously with atomic number; on the contrary, they decline within each period of the Periodic System, jumping to a higher level at the opening of each new period. The reason for this is that the atomic radius is set principally by the number of electron shells, which of course does not change within each period. As one ascends each period, however, the increasing positive charge on the nucleus draws the electrons in closer toward it. For this reason within each period the atomic radii grow smaller as the atomic number rises.

It is not as strange as it first seems therefore that sulfur and phosphorus may make double bonds more readily than silicon, for they are *smaller*. The order of covalent bond radii in Second and Third Period elements is as given in Table III (Pauling, 1960, pp. 224, 228):

TABLE III
COVALENT BOND RADII (\AA) VS. ATOMIC NUMBER

	F	O	N	C	Cl	S	P	Si
Single bond	0.64	0.74	0.74	0.772	0.99	1.04	1.10	1.17
Double bond	0.60	0.62	0.62	0.667	0.89	0.94	1.00	1.07
Atomic number	9	8	7	6	17	16	15	14

Selenium has the same single and double bond radii as silicon, bromine has somewhat smaller radii. All the other elements have larger bond radii. One might suppose from such a series that phosphorus is the largest (not heaviest!) element that readily forms multiple bonds; and that from silicon on only vestiges of this tendency remain.

Occasionally one encounters the notion that it is more rigorous to write the P-O linkages in H_3PO_4 as three single covalent bonds and one dative bond, rather than three single bonds and one double

bond.⁶ This is however an almost meaningless distinction, and has little basis in fact or theory. I think it arises in part through a supposed analogy with nitrogen, which after forming three covalent bonds, as in trimethylamine, $(\text{CH}_3)_3\text{N}$, forms a fourth dative bond as in trimethylamine oxide, $(\text{CH}_3)_3\text{N} \rightarrow \text{O}$. The analogy is not apt, however, since unlike nitrogen, phosphorus can form a fifth covalent bond by accepting a pair of electrons in its $3d$ orbitals. A double bond in addition to three single bonds is clearly possible in this instance. Indeed the measurements of bond length make this the best *single* formula that can be written for phosphoric acid. Similarly sulfuric acid is best written with two single and two double S-O bonds, making six covalences in all (cf. Pauling, 1959, p. 240).

TABLE IV
BOND LENGTHS IN TETRAHEDRAL OXYANIONS (\AA)^a

Ions:	SiO_4^{4-}	PO_4^{3-}	SO_4^{2-}	ClO_4^-
Single bonds				
Calculated	1.76	1.71	1.69	1.68
Observed	1.63	1.54	1.49	1.46
Contraction	0.13	0.17	0.20	0.22

^a Cruickshank (1961).

Actually of course no *single* formula properly represents these molecules. They are resonance hybrids in which all the P-O and S-O bonds are considerably contracted compared with single covalent bonds. Cruickshank (1961) has estimated the amount of such contraction in the tetrahedral oxyanions of Third-Period elements (Table IV).⁷

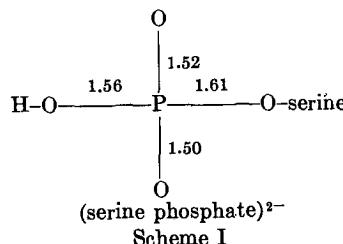
⁶ See for example Huennekens and Whiteley (1960; footnote on p. 114): "The P=O double bond (i.e., of phosphoric acid) is more correctly written as a 'semi-polar' double bond: P \rightarrow O." Compare this with Pauling's discussion (1960, p. 320), concluding "that the available evidence indicates that the older valence-bond formulas . . . with the double bonds resonating among the oxygen atoms, making them equivalent, and with the bonds considered to have partial ionic character, represent the ions (of the oxygen acids) somewhat more satisfactorily. . . ."

⁷ I am much indebted to Professor William Lipscomb of Harvard University for directing my attention to this paper.

The single bond lengths were calculated here on the basis of the Schomaker-Stevenson formulation (1941), in which a correction term is introduced for the partial ionic character of the bond, owing to differences in electronegativity of the participant atoms. The further contraction of the bonds observed here is taken to be evidence of distributed double-bond character, and involves in each instance the $3d$ orbitals of the central atom.

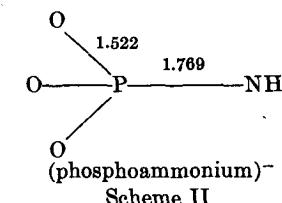
Cruickshank (1961) has also estimated the rearrangement of bond lengths that accompanies the attachment of another atom or group to one or more of the oxygen atoms in such ions. He states it to be "a simple empirical rule" that the average X-O distances in XO_4^- tetrahedral ions remain equal to the "observed" distances in Table IV. The attachment of another atom or group to O may lengthen that X-O bond by amounts up to 0.15 Å, but the other X-O bonds simultaneously contract so as to preserve the average.

So for example in the tetrahedral ion PO_4^{3-} , each P-O bond is 1.54 Å long. In H_3PO_4 however, the P-OH bonds are about 1.57 Å long, the P-O bond 1.52 Å. If instead of H a large organic radical is attached to O, this produces a much larger asymmetry. In serine phosphate, for example, the P-O bond lengths are as in Scheme I.



Sulfate exhibits similar relationships. In the ion, SO_4^{2-} , the S-O bonds are 1.47 Å long. An S-OH bond is longer, 1.56 Å for example in $KHSO_4$. In the ethyl sulfate ion, however, $(C_2H_5)_2O-SO_3^-$, the S-O(C_2H_5) bond is 1.603 Å long, the three S-O bonds 1.464 Å long.

Molecules in which phosphate P is bonded directly to N, as in phosphocreatine and phosphoarginine, exhibit similar relationships. So for example in the phosphoammonium ion the bond lengths are distributed as shown in Scheme II.

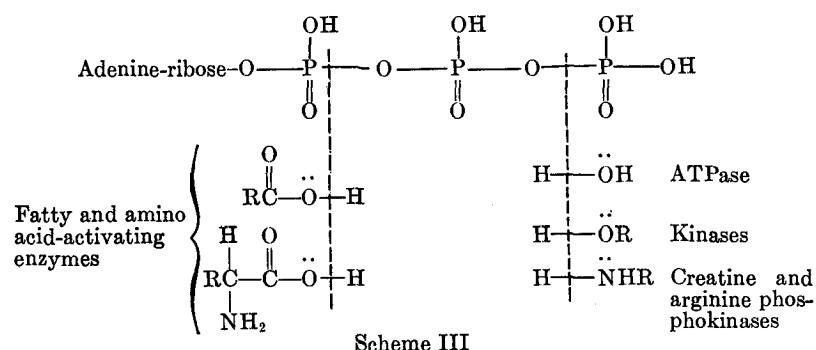


In the substituted phosphates and sulfates, therefore, the lengthened bond where an organic or other radical is attached to O makes a preferred opening; the unoccupied $3d$ orbitals on P or S provide a berth. Any molecule that can donate a lone pair of electrons should find easy access, and a ready means of attachment to the S or P nucleus. This is a condition that invites attack by water, ROH, RNH₂, and similar molecules; and to some degree this vulnerability must be shared by all organic compounds of sulfur and phosphorus.

This discussion implies that in general the exchange reactions in which organic S and P compounds participate begin—whatever rearrangements may follow—with the addition of a lone electron pair to the S or P nucleus, occupying an empty $3d$ orbital. The hydrolysis of an S or P bond for example would begin with the attachment of a lone electron pair of oxygen to S or P. Barring further rearrangements one would expect that in the products of such a hydrolysis, a hydroxyl group derived from water would remain attached to S or P.

Apparently all known enzymatic hydrolyses of organic phosphates do end in this way, as has been shown by carrying out such reactions with H_2O^{18} . All the known kinase reactions, in which molecules with alcohol-OH groups replace water as reagents, also yield this result. In all such reactions the existing P-O or P-N bond is broken, and the -OH or -OR of the attacking molecule remains attached to P (cf. Cohn, 1959).

The reactions of ATP seem to involve a similar mechanism. Here enzymatic hydrolysis (ATPase reaction) and exchange (kinase) reactions with ROH or RNH₂ cleave the terminal phosphate bond, the HO-, RO-, or RN- remaining with the amputated phosphoric acid. In another important class of reactions involving the activation of fatty and amino acids, ATP is cleaved so as to separate inorganic pyrophosphate, yielding the fatty acid or amino acid adenylate as the other product of the exchange (Scheme III).



In all these cases the reaction is probably initiated by a lone pair of O or N attacking to P. Why this attack involves the first phosphoric acid in some cases and the terminal phosphoric acid in others is not understood. (No case of pyrophosphoryl transfer has yet been observed.) It should not be forgotten that these are enzymatic reactions, and the enzyme presumably has much to do with where the attack occurs.

To leave the subject at this point, as I am about to do, is hardly to begin it. However such exchange reactions may in general be initiated, they do not always end as simply as just described. So, for a prominent example, acyl-S-CoA transfers the acyl group so as to leave, not a donator of lone pairs, but a hydrogen atom on S: the product of course is HS-CoA. None of this is especially mysterious, though it does present further problems. The mechanisms of numbers of enzymatic reactions involving the cleavage of organic phosphates have been carefully studied by Cohn, Koshland, and others (cf. Cohn, 1959); and nonenzymatic transfer reactions involving acyl phosphates have recently been explored by Di Sabato and Jencks (1961). Clearly also one should have to invoke quite different mechanisms from those I have discussed to deal with the transfer reactions of acyl imidazoles and the amino acid-ribosyl linkages on transfer-RNA.

To summarize the whole argument: I have tried to find a basis for the biological selection of S and P for group and energy transfer reactions in (1) the fact that they form more open and usually weaker bonds than their congeners in the Second Period, O and N; (2) their possession of 3d orbitals, permitting the expansion of their valences beyond four; and (3) their retention of the capacity to form multiple bonds, a property otherwise characteristic of C, N, and O.

The capacity to form multiple bonds contributes principally to the thermodynamics of energy transfer. Particularly when combined, as in P and S, with the possibility of forming 5 and 6 covalent bonds, this introduces a wide range of resonance possibilities among the precursors and products of exchange reactions that greatly increases the variety and extent of the energy changes that can occur. To use an earlier phrase, these properties ensure an adequate coinage.

The relatively wide spacing and weakness of S and P bonds, together with their tendency to add lone electron pairs in their unoccupied 3d orbitals, induces an intrinsic instability and vulnerability to attack by other molecules that promote exchange reactions.

I think that as with all the other elements of which organisms are principally composed, sulfur and phosphorus were selected on the basis of fitness: among all the elements of the Periodic System they apparently possess to a unique degree properties that lend themselves to group and energy transfer. It was to find those properties that organisms had to go from the Second into the Third Period.

REFERENCES

- Cantoni, G. L. (1953). *J. Biol. Chem.* **204**, 403.
- Cantoni, G. L. (1960). In "Comparative Biochemistry" (M. Florkin and H. S. Mason, eds.), Vol. I, p. 181. Academic Press, New York.
- Cohn, M. (1959). In Symposium on Enzyme Reaction Mechanisms. *J. Cellular Comp. Physiol.* **54**, Suppl. 1, 17.
- Coulson, C. A. (1953). "Valence." Oxford Univ. Press, London and New York.
- Cruickshank, D. W. J. (1961). *J. Chem. Soc.* p. 5486.
- Di Sabato, G., and Jencks, W. P. (1961). *J. Am. Chem. Soc.* **83**, 4393.
- George, P., and Rutman, R. J. (1960). *Progr. in Biophys. and Biophys. Chem.* **10**, 2.
- Hecht, L. I., Stephenson, M. L., and Zamecnik, P. C. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 505.
- Hill, T. L., and Morales, M. F. (1951). *J. Am. Chem. Soc.* **73**, 1656.
- Huennekens, F. M., and Whiteley, H. R. (1960). In "Comparative Biochemistry" (M. Florkin and H. S. Mason, eds.), Vol. I, p. 107. Academic Press, New York.
- Kalckar, H. M. (1946). In "Currents in Biochemical Research" (D. E. Green, ed.), p. 229. Interscience, New York.
- Kalckar, H. M. (1947). *Nature* **160**, 143.
- Lewis, G. N. (1923). "Valence" Chemical Catalog Co., New York.
- Lipmann, F. (1941). *Advances in Enzymol.* **1**, 99.
- Lipmann, F. (1958). *Science* **128**, 575.
- Oesper, P. (1950). *Arch. Biochem.* **27**, 255 (1950).

- Oesper, P. (1951). In "Phosphorus Metabolism" (W. D. McElroy and B. Glass, eds.), Vol. I, p. 523. Johns Hopkins Press, Baltimore, Maryland.
- Pauling, L. (1959). "General Chemistry," 2nd ed. W. H. Freeman, San Francisco.
- Pauling, L. (1960). "The Nature of the Chemical Bond," 3rd ed. Cornell Univ. Press, Ithaca, New York.
- Preiss, J., Berg, P., Ofengand, E. J., Bergmann, F. H., and Diekmann, M. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 319.
- Rutman, R. J., and George, P. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1094.
- Schomaker, V., and Stevenson, D. P. (1941). *J. Am. Chem. Soc.* **63**, 37.
- Stadtman, E. R. (1954). In "Mechanisms of Enzyme Action" (W. D. McElroy and B. Glass, eds.), p. 581. Johns Hopkins Press, Baltimore, Maryland.
- Stadtman, E. R., and White, F. H. (1953). *J. Am. Chem. Soc.* **75**, 2022.
- Zachau, H. G., Acs, G., and Lipmann, F. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 885.

Molecular Genetics

A Biochemical Approach to Genetics

VERNON M. INGRAM

*Division of Biochemistry,
Department of Biology, Massachusetts Institute of
Technology, Cambridge, Massachusetts*

With the discovery by Pauling, Itano, Singer, and Wells (1949) of the abnormal electrophoretic behavior of sickle cell hemoglobin and of the mode of inheritance of this abnormality (Neel, 1949; Beet, 1949), it became possible to provide chemical evidence for the effect of gene mutations on biochemical structure. The idea that there might be such a causal relationship was, of course, not new, but here was a concrete example. When it was eventually shown (Ingram, 1957, 1961) that the abnormal physical behavior of sickle cell and other mutant hemoglobins might be ascribed to specific amino acid substitutions, the expression of these mutations was even more precisely defined (Table I). The effects of gene mutations on primary protein structure are now being investigated in a number of systems, both in mammals (bovine β -lactoglobulin) and especially in bacterial systems: alkaline phosphatase (Levinthal, 1960) and tryptophan synthetase in *Escherichia coli* (Yanofsky and Crawford, 1959 and Helinski and Yanofsky, 1961), to mention only two. In this article we will try to develop some conclusions which may be drawn from the (rather meager) experimental facts.

To date, sickle cell anemia remains the only inherited human disease where all or most of the clinical manifestations may fairly be said to be caused by the effect of a gene mutation on the structure of a protein; in fact, it is caused by a specific amino acid substitution, the replacement of a glutamic acid by a valine residue, in the protein hemoglobin. This is truly a "molecular disease" (Pauling's term) which is a phenomenon explored in another chapter in this volume. Undoubtedly we may expect some other clear-cut instances, but it remains to be seen how many of the inherited diseases can be explained in such a simple manner. It is our current belief that at the basis of most inherited diseases there lies a change in the primary structure of a protein; however, this change may not be directly expressed, but only become apparent through the disturb-

TABLE I
SUMMARY OF THE KNOWN AMINO ACID SUBSTITUTIONS IN THE ABNORMAL HUMAN HEMOGLOBINS

Hemoglobin	α -Chain												Reference
	1	2	16	30	57	58	68	141	141	141	141	141	
Hb A	+	—	—	—	+	—	NH ₂	+	—	—	—	—	+
Hb I	Asp.	NH ₂	Glu.	Asp.	Tyr.	—	—	—	—	—	—	—	Murayama (1960)
Hb GHonolulu	Asp.	—	—	—	—	—	—	—	—	—	—	—	Hill (unpublished)
Hb Norfolk	Asp.	—	—	—	—	—	—	—	—	—	—	—	Baglioni (1961)
Hb MBoston	Asp.	—	—	—	—	—	—	—	—	—	—	—	Gerald and Efron (1961)
Hb GPhiladelphia	Asp.	—	—	—	—	—	—	—	—	—	—	—	Baglioni and Ingram (1961)
Hemoglobin	α -Chain												Reference
	1	2	3	6	7	26	63	67	121	146	146	146	
Hb A	+	—	—	—	—	—	—	—	—	—	—	—	Ingram (1957)
Hb S	Val.	His.	Leu.	...Glu.	Glu.	...Glu.	...His.	...Val.	...Glu.	...Glu.	...Glu.	...Glu.	Hunt and Ingram (1958, 1959a) Hill and Schwartz (1959)
Hb C	—	—	—	—	—	—	—	—	—	—	—	—	Hunt and Ingram (1959b) Gerald and Efron (1961)
Hb GSanJose	—	—	—	—	—	—	—	—	—	—	—	—	Gerald and Efron (1961)
Hb E	—	—	—	—	—	—	—	—	—	—	—	—	Hill and Schwartz (1959)
Hb GMilwaukee	—	—	—	—	—	—	—	—	—	—	—	—	Hunt and Ingram (1958, 1959a) Hill and Schwartz (1959)
Hb D β Punjab (= D γ)	—	—	—	—	—	—	—	—	—	—	—	—	Baglioni (unpublished)

ances in a sequence of biochemical events. It is also likely that many inherited diseases will turn out to be due to the complete absence or to the inadequate production of an essential protein, which is normal insofar as its structure is concerned.

In this connection, it is interesting to note that Smithies (1961) believes that he has evidence for a deletion overlapping two cistrons which results in the loss of several amino acids and the joining of the remainders of the peptide chains. These studies refer to variants of the human serum protein haptoglobin.

In a quite different sense the expression of gene mutations as chemically altered proteins is of the greatest importance in evolution. These effects are instances of "micro evolution," providing the organism with altered proteins or enzymes which might be unfavorable or favorable under the pressures of natural selection. The alterations found so far (as in Table I) are however the smallest conceivable alterations in the primary structure of a protein; other types are possible and indeed likely.

Inherent in all these arguments is the assumption that the amino acid sequence of the protein—its primary structure—determines the coiling of the peptide chains and thereby the chemical nature of the surface of the protein to which its most important biochemical properties are due. This assumption is likely to be true for a small protein up to a molecular weight of, say, 20,000, whether or not the final structure is stabilized by disulfide bridges. Haber and Anfinsen (1962) have shown the effect of tyrosine-tyrosine interaction in promoting the refolding and reactivation of reduced pancreatic ribonuclease. Such a simple mechanism becomes increasingly unlikely with increasing molecular weight and complexity, since surely a large molecule has a number of energetically equivalent tertiary structures; these, however, may be by no means equivalent in their biological specificity. And how is it determined which of these configuration will be favored? It may very well depend on the interaction of the folding chain with some other protein, itself genetically determined, or with some simple substrate, present due to the action of (genetically determined) enzymes.

Furthermore, the primary alteration in a mutated protein may affect not only that protein, but also others with which it is interacting. Genetically, the simultaneous expression of a single mutation as a bouquet of effects is well known as "pleiotropism." We can see a possible molecular mechanism for pleiotropism in the interactions

of a single mutated protein with a number of different biochemical systems.

Thus rather subtle and widespread gene effects are possible. The tendency in this field of "molecular genetics" has been to concentrate on exploring very simple experimental systems in order to illustrate clearly some general principle. We will rapidly have to become more sophisticated, however, if we want to understand the molecular basis of evolution or molecular disease, fundamentally similar phenomena. We need think only of the recent work on complementation (Fincham, 1960) or the current discussion of the different forms of thalassemia (Ingram and Stretton, 1959; Itano and Pauling, 1961; Rucknagel and Neel, 1961).

So far we have mentioned only the biochemical manifestations of "point mutations," those mutations small enough to cause only the substitution of a single amino acid by another. Other, and larger effects are possible, but have not yet been observed, except by Smithies (1961). A comparison of the amino acid sequences of the α - and β -chains of human hemoglobin have caused Braunitzer *et al.* (1961) to observe a remarkable degree of similarity in the two sequences. One is led to postulate (Braunitzer *et al.*, 1961; Ingram, 1961; Gratzer and Allison, 1960) that the two chains have evolved from a common ancestor and that the differences between them arose from the kind of mutation already discussed. But this is not enough. In order to obtain the maximum degree of correspondence in sequence between the two chains, Braunitzer has to postulate that a few amino acids in each chain have no partners in the opposite chain; there are "gaps." Of course, these are not real gaps, since the polypeptide chains are continuous covalent structures. Perhaps the "gaps" indicate the places where amino acids present in the ancestral chain have been lost during evolution. Such a loss might have occurred through the deletion of some genetic material, similar to the deletion of functions already well known to the geneticist, but smaller in extent; in the case of hemoglobin, the function is not lost, but shortened.

If we believe that the deoxyribonucleic acid (DNA) of the genetic material carries a code for the specific hemoglobin sequence, then we must distinguish between two kinds of deletion: *Clean deletions* of DNA (Fig. 1) which extend exactly from the beginning of one coding unit to the end of the same or some following coding unit; thus, no "nonsense" is produced. *Dirty deletions* of DNA

(Fig. 1) which start and/or finish *within* coding units. These latter may either leave "nonsense" at either end which might lead to a break in the peptide chain or to an interruption of protein synthesis; alternatively, by a mechanism similar to that proposed by Crick *et al.* (1961), for certain forms of mutagenesis, the code subsequent to the dirty deletion might read entirely differently, having now a new starting point. In this latter case, the affected portion of the peptide chain would be very strange and not recognizably hemoglobin. On both counts, dirty deletions would be harmful and therefore eliminated in the course of evolution. The "gaps" we see in the hemoglobin chains of today are likely to derive from clean deletions, since there are many homologous amino acid sequences both before and after the "gaps."

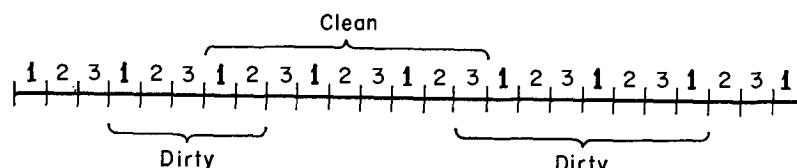


FIG. 1. Diagrammatic representation of the "triplet code" of DNA. The extent of "clean" and "dirty" deletions of genetic material is indicated by the braces.

What of additions to the genetic material leading to an increased length of peptide chain? We cannot point to any biochemical illustrations of this phenomenon amongst the abnormal human hemoglobins, or anywhere else, but we might equally well interpret the "gaps" in Braunitzer's hemoglobin chains as "additions" of amino acids in the homologous chain. The α - and β -chains of human hemoglobin are unequal in length (141 and 146 residues), yet similar in amino acid sequence at each end. This state of affairs could have arisen either by a succession of deletions or by the addition of genetic material within the gene. If the former explanation were the correct one, we would deduce that the mode of action of evolution is to change and shorten chains and that the ancestral chains were all longer than their modern counterparts; this is an unlikely situation. It is more probable that additions, *clean* additions, also occurred.

We see several ways in which gene mutations can alter directly or indirectly the structure of proteins and thus provide the raw

material for evolutionary change. Biochemical work on systems illustrating these processes has only just commenced and much remains to be done.

Perhaps one of the most fascinating problems in molecular biology is the mechanism by which genes are either activated or remain dormant. Although this is of course a general phenomenon of fundamental importance in cell differentiation, it can be very simply defined in the hemoglobin system. Presumably all vertebrate cells contain the *genetic equipment* to make hemoglobin, yet only the immature red cells do so. Conversely, these cells make almost exclusively hemoglobin and not much of anything else. The mechanism becomes even more clearly spotlighted when one considers that the human fetus makes fetal hemoglobin, different from the adult hemoglobin, and involving different genes.

$$\begin{aligned} \text{Fetal hemoglobin (F)} &= \alpha_2^A \gamma_2^F, \\ \text{Adult hemoglobin (A)} &= \alpha_2^A \beta_2^A, \end{aligned}$$

where α^A , β^A , γ^F are the chemical symbols for the normal α -, β -, and γ -peptide chains of these hemoglobins (see Schroeder, 1959). It should be noted that α -chains are common to both proteins; they are controlled by the same gene. Each of these types of chain is thought to be controlled in its primary structure by the corresponding structural gene. The genotype

$$\frac{\alpha^A}{\alpha^A} \frac{\beta^A}{\beta^A} \frac{\gamma^F}{\gamma^F}$$

is potentially capable of supplying the information to make both hemoglobin A and hemoglobin F. However, the young fetus makes only $\alpha_2^A \gamma_2^F$ and the normal adult makes almost exclusively $\alpha_2^A \beta_2^A$. The switch from fetal to adult protein—which is by no means confined to the hemoglobins—affects apparently only the γ and β structural genes. The mechanism by which this switch takes place is under active experimental investigation, since a solution of this puzzling problem will enable us to gain insight into the mechanism of cell differentiation.

REFERENCES

- Baglioni, C. (1961). *Federation Proc.* **20**, Pt. I, p. 254.
 Baglioni, C., and Ingram, V. M. (1961). *Nature* **189**, 465.
 Beet, E. A. (1949). *Ann. Eugenics* **14**, 274.
 Braunitzer, G., Hilschmann, N., Rudloff, V., Hilse, K., Liebold, B., and Müller, R. (1961). *Nature* **190**, 480.

- Crick, F. H. C., Barnett, L., Brenner, S., and Watts-Tobin, R. J. (1961). *Nature* **192**, 1227.
 Fincham, J. R. S. (1960). *Advances in Enzymol.* **22**, 1.
 Gerald, P. S., and Efron, M. L. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1758.
 Gratzier, W. B., and Allison, A. C. (1960). *Biol. Revs.* **35**, 459.
 Haber, E., and Anfinsen, C. B. (1962). *J. Biol. Chem.* In press.
 Helinski, D., and Yanofsky, C. (1961). *Federation Proc.* **20**, Pt. I, 255.
 Hill, R. L., and Schwartz, H. C. (1959). *Nature* **184**, 641.
 Hunt, J. A., and Ingram, V. M. (1958). *Nature* **182**, 1062.
 Hunt, J. A., and Ingram, V. M. (1959a). *Nature* **184**, 640.
 Hunt, J. A., and Ingram, V. M. (1959b). *Nature* **184**, 870.
 Ingram, V. M. (1957). *Nature* **180**, 326.
 Ingram, V. M. (1960). In "Genetics" (H. E. Sutton, ed.), pp. 65-176. Josiah Macy Jr. Foundation, New York.
 Ingram, V. M. (1961). *Nature* **189**, 704.
 Ingram, V. M., and Stretton, A. O. W. (1959). *Nature* **184**, 1903.
 Itano, H. A., and Pauling, L. (1961). *Nature* **191**, 398.
 Levinthal, C. (1960). In "Genetics" (H. E. Sutton, ed.), p. 100. Josiah Macy Jr. Foundation, New York.
 Murayama, M. (1960). *Federation Proc.* **17**, Pt. I, p. 78.
 Neel, J. V. (1949). *Science* **110**, 64.
 Pauling, L., Itano, H. A., Singer, S. J., and Wells, I. C. (1949). *Science* **110**, 543.
 Rucknagel, D. L., and Neel, J. V. (1961). "Progress in Medical Genetics," p. 178 (A. G. Steinberg, ed.), Vol. 1, Grune & Stratton, New York.
 Schroeder, W. A. (1959). *Fortschr. Chem. org. Naturstoffe* **17**, 322.
 Smithies, O. (1961). *Second Intern. Conf. Human Genetics, Rome*.
 Yanofsky, C., and Crawford, J. P. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 1016.

Enzymatic Mechanisms in the Transmission of Genetic Information

SEVERO OCHOA

*Department of Biochemistry, New York University School of
Medicine, New York*

I. Introduction	153
II. Replication of Genetic Material	154
III. Protein Synthesis	155
IV. Synthesis of Messenger RNA	158
V. "Stencil" RNA	160
VI. RNA Synthetase	161
References	165

I. Introduction

It is known since the classic work of Beadle and Tatum that the genes determine the synthesis of specific enzymes in a 1:1 ratio, and the work of Caspersson and Brachet strongly suggested that ribonucleic acid (RNA) is a mediator in the transmission of the genetic information from deoxyribonucleic acid (DNA) to protein. The studies of Avery, McLeod, and McCarthy, and those of Hotchkiss, provided conclusive proof that DNA is the genetic material through the demonstration that bacterial transformation is brought about by DNA. This view received support from the work of Hershey and others showing that infection of bacteria by bacteriophages, leading to multiplication of these viruses inside the cell, is brought about by injection of the viral DNA into the bacterial body.

In the case of the RNA-containing viruses, e.g., tobacco mosaic virus (TMV) and other plant viruses, poliomyelitis, influenza, RNA is the genetic material. We owe this demonstration to the work of Schramm and his collaborators and Fraenkel-Conrat and co-workers who showed that TMV-RNA is infectious and causes the production of virus by the plant. This work was followed by elegant studies (Gierer and Schramm, Fraenkel-Conrat *et al.*) on the production of TMV mutants by treating TMV-RNA with nitrous acid. This treatment results in deamination of the purine and purimidine bases bearing NH₂ groups, e.g., deamination of cytosine to uracil. With

the determination of the amino acid sequence of the protein coat of TMV, in the Tübingen and Berkeley laboratories, it became possible to isolate nitrous acid mutants of TMV in which the protein differs from that of the wild type strain in the replacement of one single amino acid by another, e.g., proline by leucine, threonine by serine. Thus, deamination of one RNA base is responsible for the replacement of one amino acid by another. These experiments provide strong support for the idea that the genetic code resides in specific nucleotide sequences and that a sequence of three or more nucleotides codes for a specific amino acid. There is considerable interest at present in the manner in which the DNA code determines the production of a characteristic protein. Closer examination of this problem requires a brief survey of present knowledge on the mechanisms of replication of genetic material and protein synthesis.

II. Replication of Genetic Material

The work of Watson and Crick (1953) on the molecular structure of DNA, as deduced from X-ray diffraction studies, was of paramount importance for an understanding of the mode of replication of genetic material. They concluded that DNA consists of two helical strands wound around each other and held together by hydrogen bonds between complementary pairs of bases, namely adenine and thymine and guanine and cytosine. This was in line with Chargaff's results showing that in any DNA the amount of adenine equals that of thymine and the amount of guanine equals that of cytosine, the amount of purines being equal to that of pyrimidines (Chargaff, 1955). The proposal of Watson and Crick received experimental support a few years later through the finding (Warner, 1957) that synthetic homopolynucleotides, prepared with polynucleotide phosphorylase (Ochoa, 1960), e.g., polyadenylic and polyuridylic acid, interact in solution to form double-stranded helical structures held by hydrogen bonds between the complementary pairs adenine and uracil.

Watson and Crick postulated that duplication of the genetic material, i.e., DNA, might occur by separation of the strands followed by formation of two new strands around each parental chain. This view received striking confirmation from the work of Meselson and Stahl (1958). These workers grew *Escherichia coli* in a medium enriched with N¹⁵ so that "heavy" N¹⁵-labeled DNA was produced. Because of its higher density, this DNA could be separated from

"light," N¹⁴-containing DNA, and from DNA of intermediate density by equilibrium centrifugation in cesium chloride gradients. When fully labeled cells were placed in N¹⁴-medium and lysates examined at various times thereafter it was found that, until one generation time elapsed, half-labeled DNA molecules accumulated while fully labeled DNA was depleted. After one generation time all the DNA consisted of half-labeled hybrid molecules. After two generation times half-labeled and unlabeled molecules were present in equal amounts. This indicates that after one generation each DNA molecule consisted of one heavy and one normal strand while after two generations two out of four molecules consisted of one heavy and one normal and the other two contained exclusively normal DNA strands.

The discovery of DNA polymerase substantiated the above results at the enzyme level. Kornberg (1960) and collaborators showed that this enzyme synthesizes DNA from the deoxyribonucleoside triphosphates with liberation of inorganic pyrophosphate. The enzyme is inactive unless all of the four triphosphates are present and requires DNA as a primer. The primer is replicated to yield DNA of the same base composition and sequence through a base-pairing mechanism. The DNA acts therefore as a template for its own replication, each strand giving rise to a complementary strand with formation of two double strands after one synthetic cycle is completed.

III. Protein Synthesis

Work in Zamecnik's laboratory (Zamecnik, 1960) showed that the ribonucleoprotein particles, or ribosomes, associated with the endoplasmic reticulum are the site of protein synthesis. Subsequent experiments with cell-free systems demonstrated a requirement for ribosomes, cytoplasmic supernatant, adenosine triphosphate (ATP), or an ATP-generating system, and Mg⁺⁺. The need for ATP indicated that the amino acid must be activated prior to the formation of peptide bonds, and shortly thereafter Hoagland demonstrated the presence of amino acid-activating enzymes in the cytoplasmic supernatant fluid. In the presence of ATP, Mg⁺⁺, amino acids, and hydroxylamine, these enzymes lead to the formation of amino acid hydroxamates concomitantly with the splitting of ATP to adenosine 5'-phosphate (AMP) and inorganic pyrophosphate. In the absence of hydroxylamine the reaction leads to the formation of amino acyl adenylate-enzyme complexes with release of pyrophosphate.

Due to its reversibility this reaction brings about an incorporation of exchange of P^{32} -labeled pyrophosphate into ATP; this exchange provided a sensitive assay for amino acid activation. The real nature of the reaction was disclosed by Holley's observation that in the presence of an RNA fraction from the cytoplasmic supernatant there was an exchange not only of pyrophosphate but also of AMP into ATP. Further investigation by Berg (1961), Lipmann, and others led to the discovery that the supernatant fraction contains an RNA of small size (molecular weight about 30,000) which accepts amino acid residues in the activation reaction. This RNA, previously referred to as soluble RNA (sRNA), is now commonly known as amino acid acceptor or transfer RNA. It is now known that there is a specific activating enzyme and a specific acceptor RNA for each amino acid and several activating enzymes have been isolated in highly purified form. Methods have also been developed for the enrichment of specific transfer RNA's. All the transfer RNA's have a common feature; they are terminated by an adenylic acid residue, bearing two unesterified hydroxyls at carbon 2' and 3' of the ribose moiety, preceded by two cytidylic acid residues. The chains are initiated in all cases by guanylic acid. It is also known that the nucleotide sequences immediately preceding the three terminal nucleotides vary for different transfer RNA's. In the activating reaction the amino acyl residue is esterified to one of the hydroxyls of the terminal adenosine of the transfer RNA. This represents the active form of the amino acid which, in a subsequent step involving one or more soluble enzymes and requiring guanosine 5'-triphosphate (GTP), is transferred to the ribosomes where it becomes incorporated into a polypeptide chain. The nature of the transfer reaction is as yet poorly understood. Much work has been done on the composition and structure of the ribosomes from animal, bacterial, and plant cells. They consist roughly of equal parts of protein and RNA. The RNA is of rather high molecular weight (of the order of 1 to 2 million). It differs from transfer RNA not only in its much larger molecular weight but also in the fact that it contains only the four common nucleotides adenylic, guanylic, uridylic, and cytidylic acid whereas the transfer RNA contains a relatively high proportion of pseudouridylic acid residues as well as smaller amounts of ribothymidyllic acid and other nucleotides with methylated bases. The ribosomes exist in several states of aggregation depending on the concentration of Mg^{++} in the suspending medium.

Thus, ribosomes with sedimentation coefficients of 50 S, 70 S, and 100 S are found. It appears that the 100 S ribosomes are the ones concerned with protein synthesis.

It was believed until recently that the ribosomal RNA was the actual template in protein synthesis. The first indications that this might not be so came from experiments of Volkin (Volkin *et al.*, 1958) who shortly after infection of *E. coli* with T-even phages in the presence of P^{32} -labeled orthophosphate observed the formation of a very short-lived RNA whose base composition was similar to that of the phage DNA. These and other observations of Spiegelman and co-workers (Nomura *et al.*, 1960; Hall and Spiegelman, 1961) led to the belief that the actual template is not the ribosomal RNA but an RNA that is made under genetic control, i.e., under DNA direction, just prior to protein synthesis. It had in fact been difficult to correlate a template function for ribosomal RNA with the observations of Chargaff, Belozersky (Belozersky and Spirin, 1958), and others that the composition of ribosomal RNA, which makes up about 80% of the total RNA of the cell, was remarkably constant in a great number of bacterial species and did not reflect the base composition of the corresponding DNA which varied widely in different species. The occurrence of a short-lived template RNA, made under genic control as required for the synthesis of specific proteins, would explain the observations on induction and repression of enzyme synthesis. Jacob and Monod (1961) postulated the existence of such an RNA which they referred to as messenger RNA. This concept is supported by the recent elegant experiments of Brenner *et al.* (1961) and Gros *et al.* (1961). The former authors have shown that no new ribosomes are synthesized in *E. coli* following T2 phage infection and that an unstable RNA of the same base composition as that of the viral DNA is formed shortly after infection and becomes associated with the old ribosomes; these are the site of new (phage) protein synthesis.

A dramatic confirmation of the messenger RNA concept was recently afforded by the demonstration (Nirenberg and Matthaei, 1961; Lengyel *et al.*, 1961; Wood and Berg, 1962) that some natural RNA's (e.g., TMV-RNA) and synthetic polyribonucleotides prepared with polynucleotide phosphorylase markedly stimulate the incorporation of amino acids into acid-insoluble products by a system of *E. coli* cytoplasmic supernatant and ribosomes; this effect depends on the presence of transfer RNA. With polyuridylic acid

as messenger the system synthesizes polyphenylalanine, phenylalanine being the only amino acid whose incorporation into acid-insoluble products is stimulated by this polymer. There are indications that with TMV-RNA as messenger, a protein with the immunological properties of TMV protein is synthesized. Various copolymers prepared with polynucleotide phosphorylase, containing uridylic acid together with other nucleotide residues, stimulate the incorporation of a number of amino acids besides phenylalanine. The nucleotide composition of triplets coding for nineteen amino acids has been determined in this way (Speyer *et al.*, 1962; Lengyel *et al.*, 1962). The observation that rat liver ribosomes can be substituted for their *E. coli* counterparts for the synthesis of polyphenylalanine in the presence of polyuridylic acid strikingly emphasizes that the nature of the resulting protein is determined by the messenger RNA and not by the ribosomes (Lengyel *et al.*, 1961).

In summary the protein-making machinery of the cell consists of twenty transfer RNA's, each specific for one amino acid and as many activating enzymes, each bearing specificity toward an amino acid and its corresponding transfer RNA. These are present in the cytoplasmic supernatant along with one or more soluble enzymes concerned with the transfer of the activated amino acid residues to the ribosomes and with the formation of polypeptide chains on the ribosomes. The directing agent which specifies the sequence of amino acids along the polypeptide chain is the messenger RNA. This is made under genetic control and is a copy of a gene, i.e., of a DNA chain coding for a given protein. The DNA determines the nucleotide sequence of the RNA messenger. There is increasing evidence in favor of a triplet code, i.e., each three nucleotides in the messenger coding for one amino acid, and the amino acids are assumed to be placed in position on the template, one after another as peptide bonds are being made, by a base-pairing mechanism involving a given triplet of the messenger and a complementary nucleotide triplet in the transfer RNA.

IV. Synthesis of Messenger RNA

Nothing is known of the mechanism of synthesis of transfer RNA. Its "adapter" triplet and the specific features for interaction with a given activating enzyme are apparently predetermined and fixed. A multiplicity of proteins can be synthesized with one set of activating enzymes and transfer RNA's depending on the sequence

of triplets along the messenger RNA chain. The translation of the DNA code into the corresponding code of the RNA messenger becomes therefore the focal point in transmission of genetic information. This requires a mechanism for making RNA replicas of DNA,

TABLE I
ILLUSTRATING EQUAL BASE COMPOSITION OF RNA SYNTHESIZED
BY RNA POLYMERASE AND ITS DNA TEMPLATE

DNA		Base composition of DNA (%)			
Strand 1	Strand 2	Base	Strand 1	Strand 2	Total
A	T				
G	C				
C	G	A	33.3	11.1	22.2
T	A				
A	T	G	33.3	22.2	22.7
A	T				
C	G	T	11.1	33.3	22.2
G	C				
G	C	C	22.2	33.3	27.7

RNA		Base composition of RNA (%)			
Strand 1a	Strand 2a	Base	Strand 1a	Strand 2a	Total
U	A				
C	G				
G	C	A	11.1	33.3	22.2
A	U				
U	A	G	22.2	33.3	27.7
U	A				
G	C	U	33.3	11.1	22.2
C	G				
C	G	C	33.3	22.2	27.7

and the discovery of RNA polymerase provided such a mechanism (Furth *et al.*, 1961; Weiss and Nakamoto, 1961; Stevens, 1961).

Like DNA polymerase RNA polymerase has been found in bacterial and animal cells. It catalyzes the DNA-dependent synthesis of RNA from ribonucleoside triphosphates and, as in the case of the former enzyme, the presence of all four triphosphates is required for activity which is accompanied by release of inorganic pyrophos-

phate. The enzyme is completely inactive in the absence of DNA; it can be primed by DNA from a number of sources whether single- or double-stranded. However, depending apparently on the source of the enzyme, there may be differences in the degree of effectiveness of native (double-stranded) and denatured (single-stranded) DNA. The enzyme from *Azotobacter vinelandii* shows a preference for native DNA (Burma *et al.*, 1961). Naturally occurring, single-stranded DNA, e.g., ϕ X-174 phage DNA is a good primer for RNA polymerase. With this DNA as primer the enzyme makes complementary RNA chains (Chamberlin and Berg, 1962). This is a clear indication that, like DNA polymerase, RNA polymerase operates through a base-pairing mechanism. In agreement with Volkin's observations on the phage-induced RNA, the base composition of the RNA synthesized with double-stranded DNA as primer is the same as that of the DNA with uracil replacing thymine. This indicates that both DNA strands are replicated since, as illustrated in Table I, replication of only one strand would not result in RNA of equal base composition to that of the DNA primer. Nearest neighbor frequency studies indicate that the sequence of ribonucleotides in the RNA synthesized by RNA polymerase is the same as that of the DNA primer. This conclusion is strengthened by the finding that when the isolated synthetic RNA is heated in the presence of the DNA primer and subsequently slow-cooled, double-stranded DNA-RNA hybrids are formed (Geiduschek *et al.*, 1961). Thus, each DNA strand must give rise to a complementary RNA strand.

V. "Stencil" RNA

The mechanism of the RNA polymerase reaction is not clear. It would appear that on replication through a base-pairing mechanism hybrid double-stranded molecules (one DNA and one RNA strand) might be formed as a first step but there is no evidence so far that this is the case. The possibility of hybrid formation between RNA and DNA chains was first pointed out by Rich (1960) who obtained hybrids of polyriboadenylic acid and short-chain polythymidylic acid prepared synthetically by Khorana. Later Schildkraut *et al.* (1961) obtained hybrids of polyribocytidylic and polydeoxyguanylic acid. More recently the natural occurrence of hybrid DNA-RNA molecules in T2 phage-infected *E. coli* has been reported by Hayashi and Spiegelman (1961), and Warner (unpublished) has indications for the occurrence of similar hybrids in *Azotobacter*.

Although, as already mentioned, no hybrid formation has yet been detected following RNA synthesis by RNA polymerase, it is conceivable that only a small fraction of the DNA primer was utilized and any hybrid formed may have escaped detection. Hybrid formation appears, however, as a most likely step in this reaction. If so, the question arises as to how the hybrid double strands are pulled apart with release of RNA. A likely possibility, which is susceptible of experimental verification, is that the hybrids may be used as templates by DNA polymerase with formation of one DNA double strand and release of single-stranded RNA. If this were the case, the synthesis of RNA replicas of DNA by RNA polymerase would be tied up with DNA duplication and would occur simultaneously with cell division. A mechanism would then be required for formation of messenger RNA at times other than cell division, particularly in animal and plant cells. Such a mechanism could be provided by an enzyme capable of replicating the RNA made by RNA polymerase. In this view RNA polymerase would make an RNA stencil which could be utilized by the cell in-between divisions for the production of copies, i.e., for the production of messenger RNA. The recent discovery of an enzyme in spinach leaves which appears to catalyze an RNA-dependent synthesis of RNA might provide the missing link.

VI. RNA Synthetase

In a search for enzymes that might be concerned with the synthesis of TMV-RNA Reddi (1961) in our laboratory isolated an enzyme from spinach which catalyzes the incorporation of nucleotides from the ribonucleoside triphosphates into acid-insoluble products. Although the enzyme has been but partially purified, it exhibits optimal activity, like DNA polymerase and RNA polymerase, in the presence of all four nucleoside triphosphates. Preparations of this enzyme of low RNA content have been obtained which have little or no activity, under otherwise optimal conditions, in the absence of added RNA. As shown in Table II activity is restored by RNA. DNA is completely inactive. This is true whether native or denatured DNA is used. Preliminary results with use of ATP labeled with P^{32} in the α -phosphate and nonlabeled guanosine triphosphate (GTP), uridine triphosphate (UTP), and cytidine triphosphate (CTP), indicate that the distribution of P^{32} label among the 2'- and 3'-nucleoside monophosphates (obtained by alkaline hydrolysis

of the RNA isolated at the end of incubation) differs depending on whether yeast RNA or TMV-RNA is used as primer. This result suggests that the enzyme may replicate the RNA primer.

It is conceivable that the function of such an enzyme might be to make copies of the "stencil" RNA made by RNA polymerase. Although RNA synthetase is present in the cytoplasmic supernatant of spinach leaf, the possibility that it is released from the nucleus when the cells are broken up cannot be excluded. A number of observations on RNA turnover following administration of radioactive precursors (P^{32} -labeled purine and pyrimidine bases) *in vivo*,

TABLE II
EFFECT OF RNA ON ACTIVITY OF RNA SYNTHETASE^a

Nucleic acid addition (μ g/mg protein)	AMP incorporation (m μ moles/mg protein)
None	0.2
Yeast RNA (100)	3.6
<i>Azotobacter</i> DNA (100)	0.3
TMV-RNA (40)	2.3
TMV-RNA (80)	2.9
TMV-RNA (120)	3.8
TMV-RNA (160)	7.1

^a From data of K. K. Reddi (unpublished).

and experiments with intact and enucleated cells, indicate that most if not all RNA synthesis takes place in the nucleus with later appearance of RNA in the cytoplasm.

Since most of the nuclear RNA is present in the nucleolus and nuclear RNA has the highest turnover rate of all the cell's RNA, it may be assumed that copying of the "stencil" RNA takes place in the nucleolus. This view finds support in the experiments of Amano and Leblond (1960) who injected tritiated cytidine into mice and observed the labeling pattern at various times by autoradiography of liver and pancreas cells. Early labeling was greater in the chromatin than in the nucleolus. From the intersection of the specific radioactivity curves (only the nucleolar curve intersecting the cytoplasmic one) they concluded that the reactions leading to RNA synthesis in the chromatin and the nucleolus are independent of each

other and that only RNA from the nucleolus migrates to the cytoplasm. Upon selective destruction of the nucleolus of HeLa cells by means of an ultraviolet microbeam, Perry (1960) found that the uptake of tritiated cytidine was decreased by 30% in the nucleus and by 65% in the cytoplasm by comparison with controls in which neighboring nuclear areas were irradiated. In experiments on isolated pea seedling nuclei incubated with ATP, GTP, UTP, and tritium-labeled cytidine, followed by breaking up of the nuclei and separation of nuclear fractions by fractional centrifugation, Rho and Bonner (1961) found incorporation of label first into the chromatin and later in the nucleolus. These observations would support the view that the synthesis of RNA, first occurring on the chromatin, is catalyzed by RNA polymerase with DNA as template and that the later appearance of RNA in the nucleolus may be the result of synthesis catalyzed by an enzyme of the RNA synthetase type with the previously synthesized RNA, i.e., "stencil" RNA, as template. When the cell divides all RNA synthesis must be brought under genetic control and the "stencil" RNA must disappear. This is in line with the disappearance of the nucleolus during cell division (Mazia, 1961).

A schematic illustration of the hypothesis presented in this essay is shown in Fig. 1. All the events are depicted as nuclear events. Stages 1 and 2 represent the formation of two hybrid DNA-RNA strands by RNA polymerase. This is followed in stages 3 and 4 by DNA duplication, assumed to be catalyzed by DNA polymerase using the hybrid double strands as templates, with liberation of single-stranded "stencil" RNA and cell division. Since the RNA strands of each daughter cell are copies of only one of the DNA strands of the mother cell, each RNA strand undergoes self-replication (stage 5) under the influence of RNA synthetase; the "stencil" RNA of each daughter cell now contains all the information of the original DNA. Stage 6 depicts the here postulated localization of this RNA in the nucleolus of the daughter cells.

The question whether both DNA strands are meaningful or whether one of the strands bears a correct protein code while the other carries a "nonsense" one cannot be answered with the information available at present. Marmur and Lane (1960) annealed transforming DNA carrying a certain marker with nontransforming DNA of the same bacterial species and found no decrease in the transforming activity. This result shows that genetic information

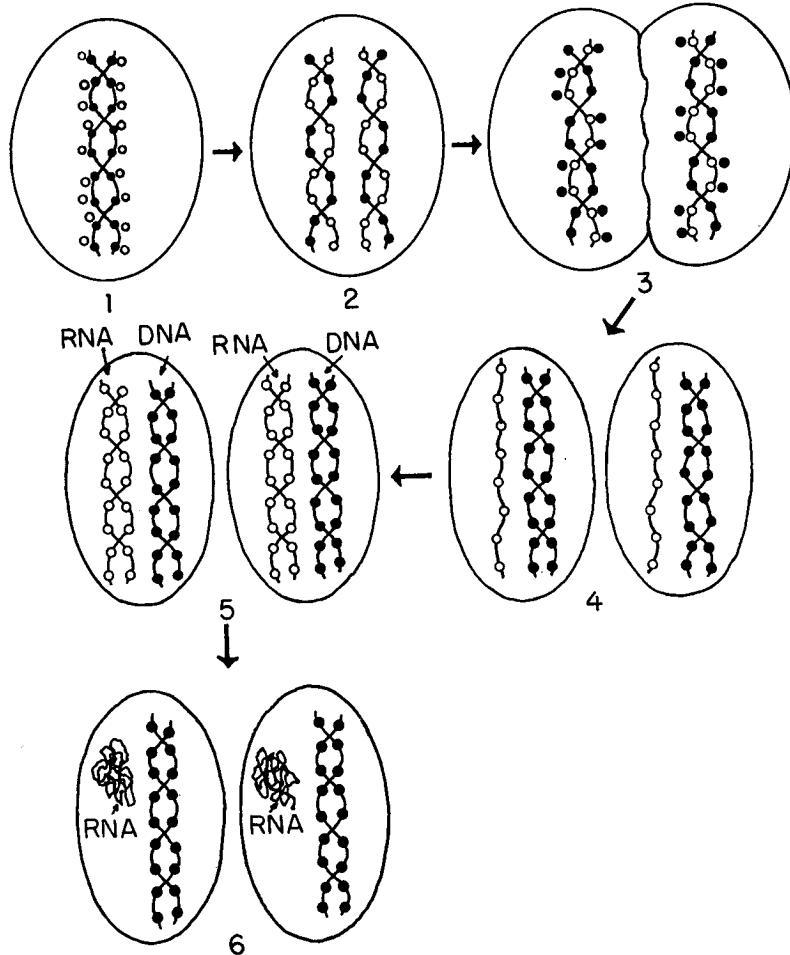


FIG. 1. Schematic illustration of the hypothetical synthesis of "stencil" RNA and accompanying nuclear events. For explanation see text.

is carried independently by each DNA strand. Poor penetration of the single-stranded material into the cells is responsible for the large loss of activity following denaturation of transforming DNA, for on heating P^{32} -labeled transforming DNA, the P^{32} uptake by the cells and the transforming activity fall off at the same rate (Lerman and Tolmach, 1959). About 5% of the residual transforming activity of heated DNA resides in single strands (Guild, 1961). However,

there is no direct evidence that each of the DNA strands can carry different genetic information except for the experiments of Herriot (1961) on the production of heterozygous DNA by annealing of DNA's bearing different transforming markers.

REFERENCES

- Amano, M., and Leblond, C. P. (1960). *Exptl. Cell Research* **20**, 250.
 Belozersky, A. N., and Spirin, A. S. (1958). *Nature* **182**, 111.
 Berg, P. (1961). *Ann. Rev. Biochem.* **30**, 293.
 Brenner, S., Jacob, F., and Meselson, M. (1961). *Nature* **190**, 576.
 Burma, D. P., Kröger, H., Ochoa, S., Warner, R. C., and Weill, J. D. (1961). *Proc. Natl. Acad. Sci. U.S.* **57**, 749.
 Chamberlin, M., and Berg, P. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 81.
 Chargaff, E. (1955). In "The Nucleic Acids" (E. Chargaff and J. N. Davidson, eds.), Vol. 1, p. 307. Academic Press, New York.
 Furth, J. J., Hurwitz, J., and Goldman, M. (1961). *Biochem. Biophys. Research Commun.* **4**, 362, 431.
 Geiduschek, E. P., Nakamoto, T., and Weiss, S. B. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1405.
 Gros, F., Hiatt, H., Gilbert, W., Kurland, C. G., Risebrough, R. W., and Watson, J. D. (1961). *Nature* **190**, 581.
 Guild, W. R. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1560.
 Hall, B. D., and Spiegelman, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 137.
 Hayashi, M., and Spiegelman, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1564.
 Herriot, R. M. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 146.
 Jacob, F., and Monod, J. (1961). *J. Mol. Biol.* **3**, 318.
 Kornberg, A. (1960). *Science* **131**, 1503.
 Lengyel, P., Speyer, J. F., and Ochoa, S. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1936.
 Lengyel, P., Speyer, J. F., Basilio, C., and Ochoa, S. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 282.
 Lerman, L. S., and Tolmach, L. J. (1959). *Biochim. et Biophys. Acta* **33**, 371.
 Marmur, J., and Lane, D. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 453.
 Mazia, D. (1961). In "The Cell" (J. Brachet and A. E. Mirsky, eds.), Vol. 3, p. 77. Academic Press, New York.
 Meselson, M., and Stahl, F. W. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 671.
 Nirenberg, M. W., and Matthaei, J. H. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1588.
 Nomura, M., Hall, B. D., and Spiegelman, S. (1960). *J. Mol. Biol.* **2**, 306.
 Ochoa, S. (1960). *Angew. Chem.* **72**, 225.
 Perry, R. P. (1960). *Exptl. Cell Research* **20**, 216.
 Reddi, K. K. (1961). *Science* **133**, 1367.
 Rho, J. H., and Bonner, J. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1611.
 Rich, A. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1044.
 Schildkraut, C. L., Marmur, J., Fresco, J. R., and Doty, P. (1961). *J. Biol. Chem.* **236**, PC 2.

- Speyer, J. F., Lengyel, P., Basilio, C., and Ochoa, S. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 63, 441.
 Stevens, A. (1961). *J. Biol. Chem.* **236**, PC 43.
 Volkin, E., Astrachan, L., and Countryman, J. L. (1958). *Virology* **6**, 545.
 Warner, R. C. (1957). *J. Biol. Chem.* **229**, 711.
 Watson, J. D., and Crick, F. H. C. (1953). *Nature* **171**, 964.
 Weiss, S. B., and Nakamoto, T. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1400.
 Wood, W., and Berg, P. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 94.
 Zamecnik, P. C. (1960). *Harvey Lectures, Ser.* **54**, 256.

A "Book Model" of Genetic Information-Transfer in Cells and Tissues¹

JOHN R. PLATT²

*Department of Physics, University of Chicago,
Chicago, Illinois*

Abstract	167
I. Introduction	168
II. One-Dimensional Sequence of Information	169
III. Sequential Read-Out	170
IV. One Line to Read (Not Two)	171
V. Copying without Reading, and Vice Versa	173
VI. Pages Open or Closed	173
VII. Only a Few Pages Are Open	174
VIII. Page Headings for Reference	176
IX. Sequential Reference	182
X. Differentiation: Cross-References from Other Books	183
XI. Stored Configurations	185
XII. Summary	186
References	187

Abstract

The expression of genetic information in cells and whole organisms is like the reading out of a complex instruction manual, but the analogy extends to more details than is generally realized. The information is linearly arranged in "words" that are "read out" sequentially in time. There is one copying mechanism (DNA polymerase) for reprinting the whole book, and another (RNA polymerase) for selective read-out into cell chemistry. The read-out is by "paragraphs" (genes) and by "pages" (operons) that can either be "closed" (repressed) or "opened" (induced), according to contingent "instructions" (repressor-corepressor complexes) from "references" (regulator genes) on earlier pages or in the "books" of adjacent tissues.

¹ Work supported in part by a U.S. Public Health Service Grant at the Marine Biological Laboratories, Woods Hole, Massachusetts, and by an Atomic Energy Commission Contract with the University of Chicago.

² Visiting Professor in Biology, Massachusetts Institute of Technology, 1961-1962.

One of the most important implications is that the "pages" are "referred to" by "page numbers" or "page headings" (inducers for operator-genes) that must be widespread low-information codes very different in kind, and different chemically, from the rest of the species-specific high-information "text" (structural-genes) on the "page." Between "readings," the "books" can be "locked away" in compact "storage" forms (phage heads, chromosomes). This analysis points up several questions, such as the "reading" of the two strands in a double-helix, the many roles of protein-nucleic-acid "translator" molecules, and the diverse mechanical configurations of the chains as they are "opened" and "closed" and "stored."

I. Introduction

The presently known facts of genetic information-transfer have been reviewed by Jacob and Monod (1961) and summarized by Rich (1962a). Let us recapitulate the main results, as described in the current genetic-biochemical language.

"DNA makes RNA makes protein." That is to say, the long hereditary deoxyribonucleic acid (DNA) chains [or sometimes ribonucleic acid (RNA) chains], containing specific sequences of bases, can be copied in short sections by Messenger-RNA chains; which are then copied (or, better, "translated") into the amino-acid sequences of enzymatic proteins.

What determines whether a particular enzyme is manufactured or not?

Jacob and Monod show that a "regulator-gene," perhaps distant from the enzyme locus, makes a "repressor-substance." This substance can complex with certain cellular constituents called "co-repressors" or "inducers." It then interacts, or fails to interact, with a particular "operator-gene" on the DNA chain so as to "repress" or "derepress" (that is, "induce") the manufacture of Messenger-RNA from neighboring "structural-genes," among which is the gene determining the structure of the enzyme in question.

Several enzymes, frequently those involved in successive steps of the synthesis of the same cellular product (histidine, for example, or arginine, or orthophosphate), may be under the control of the same repressor-gene. In bacteria, they are often clustered in sequence on the genetic map and are under the control of a single operator-gene as well. An operator-gene and the enzyme structural-genes

that it controls, together make up a single information unit called an "operon."

In this picture, the co-repressor that acts to suppress the formation of an enzyme is frequently the excess product (histidine, for example) of the reaction-sequence that the enzyme takes part in. On the other hand, an inducer that acts to get an enzyme formed is frequently the substrate (glucose, for example) of a degradative process that the enzyme mediates. In short, the system has "feedback control" to keep enzymes from being produced unnecessarily, except when mutations damage this regulatory system. Since most enzymes are needed only on specific occasions or for specific contingencies, their operons stay most of the time in the "repressed" state.

The present remarks are a series of reflections on these curious and interesting relationships. They were inspired by the realization that there is a simple and familiar analogy to any chain of information that is expressed, section by section, in this all-or-none fashion. It is the everyday example of the information-sequence in a book—or more precisely, perhaps, a very complex "instruction manual"—that can only be open to one page, or a few pages, at a time.

For many years, of course, writers and speakers on genetics have used the linguistic and bookish metaphors of "codes" and "commas" and "words" that can be "read out." But the point here is that this casual analogy can apparently be extended now to many additional details of genetic storage and read-out. It therefore becomes interesting to discuss more carefully the question of just how exact the analogy is, and at what points it begins to break down. If it is accurate in a sufficient number of details, a serious comparison between the genetic chain and an instruction book could become a useful device for teaching and for explanation in genetics and embryology, it could give us a more general and coherent view of the genetic process, and it could point up research questions concerning structural and biochemical analogies that we might otherwise neglect.

We can see these possibilities better by examining the various parallels one by one. We will begin with the most familiar ones.

II. One-Dimensional Sequence of Information

The information in a book (except for the pictures) is ordered in a one-dimensional array; so is the genetic information. From the

beginning, Morgan's hypothetical "genetic map" was assumed to be linear in order to represent the experimentally transitive relations among the recombination probabilities of the genes. Benzer has now proved conclusively the linearity of the microgenetic map within a gene (the r_{11} section of T2 phage) all the way down to the level of the point-mutations (single base changes) (Benzer, 1959). Electron-microscope pictures of DNA chains deposited on a backing also show that they may extend unbranched for thousands or tens of thousands of angstroms (Beer, 1961).

It is true that the genetic map in T2 and T4 phage appears to be circular, but this may be a purely formal effect of some kind of continuous replication rather than the result of physical looping of the DNA (Stahl, in press). A pulling-out of a helix or double-helix chain into side branches during some part of the life cycle is also conceivable (Platt, 1955), but it would not change the linearity of the genetic information-sequence. Consequently, although our present picture of the genetic chain is more like the continuous line of writing on a scroll than the divided lines of a Western book, the one-dimensional information-relation in either the book or the scroll is evidently an accurate representation.

III. Sequential Read-Out

It has now been shown that the amino-acids are not added to the Messenger-RNA template at random positions before uniting to form a protein, but are added sequentially from one end of the sequence to the other (Bishop *et al.*, 1960; Dintzis, 1961). It has also been shown that the addition or deletion of a single base in the DNA chain shifts the amino-acid "read-out" forward or backward by one base for dozens or hundreds of bases "to the right" on the genetic map, which indicates sequential read-out in long sections within a gene at the DNA level also (Crick *et al.*, 1961) (though of course, some sections might be read "to the right" and others "to the left").

As many authors have said, this is like reading out a special kind of writing, in which the "letters" (the 4 different bases in a DNA or RNA chain) are grouped, probably 3 at a time, into "words" (the 20 or so different amino-acids); but where the words are run together with no spaces or commas between them, so that a displacement by one letter backward or forward makes all the words "afterward" be read wrong.

To force the book analogy physically close to such a scheme, we would then have to do without separate lines or spaces or punctuation on a page. The letters would have to be printed in one continuous string from beginning to end, perhaps winding back and forth across the page (if this is necessary for compactness) in the old "boustrophedon" style that is still found sometimes in childrens' reading-puzzles. But a normal book gives us this same directional and sequential continuity so far as the sense is concerned.

IV. One Line to Read (Not Two)

In the Watson-Crick double-helix picture of DNA, each of the two complementary strands carries an information sequence that could be read out into one of two complementary Messenger-RNA sequences. Are both of these Messenger-RNA's actually formed? If so, does each make a different protein, of the same length but with different and useful functions? (Rich, 1961, 1962b). The latter, at least, has always seemed highly unlikely to those who have considered the question, and there is now direct chemical evidence against it.

For example, Speyer *et al.* (1962) and Matthaei and associates (1962) have reported the synthetic RNA base-triplet "words" that "make" the different amino acids and have found that the "meaningful" triplets are all non-complementary to each other with, at most, two or three exceptions. The meaningful triplets all contain a U, or are U-rich; the complementary ones would necessarily be U-poor.

These results point to a series of conclusions, as suggested by Rich (1960, 1961). One is that most or all of the complementary RNA base-triplets (the U-poor ones) may not make amino acids. If so, in any section of a DNA double-helix, only one strand would need to be read, say the meaningful strand (U-poor in DNA) that gives a U-rich Messenger-RNA chain. (Note that on the one-strand-read-out assumption, it was never necessary to consider more than half the base-triplets—that is, 32 out of the possible 64—for coding the 20 amino acids.)

The DNA strand that is not read (or is read into non-protein-forming Messenger) should still not be regarded as "nonsense," but as something that could be called "compsense." For it still contains the information sequence, although in a complementary form that would require an additional replication (in a given generation) be-

fore it could be turned into protein. On the other hand, its presence eliminates one replication step between generations before daughter protein can be made; and it undoubtedly helps protect the other strand from breakage or chemical change. (On the book model, like the thin glazed sheets sometimes put in to protect expensive "illuminated" pages.) The single-strand phage, ϕ X-174, and the RNA viruses have shown us for some time, of course, that the second strand is not informationally necessary to genetics.

There is a difficulty in assuming that one DNA strand is a "sense" strand and the other "compsense," if "sense" means only the U-poor triplets. For this would imply (Rich, 1961) that the strands should then show easily detectable differences under electrophoresis or centrifugation, contrary to what has been reported. But these results could be reconciled if each of the strands contain alternate sections of "sense" and "compsense" which tend to balance the purine-pyrimidine ratios. A similar suggestion was made earlier, on other grounds (Platt, 1955). In evolution, a "sense" section might become joined to a "compsense" section through failure of the second strand to separate from the end of the first strand at some replication (see also Rich, 1962a).

Messenger-RNA chains might terminate naturally at the end of a "sense" section, which would thus provide a natural end to a gene or a protein. If each strand "points the same way" (in each base-ribose-phosphate unit) for its entire length, but opposite to the other strand, as the Watson-Crick model of DNA suggests, then the "sense" sections on one strand would have to be read out "to the right," while the intervening sections where the "sense" is on the other strand would be read out "to the left." But both could be read out by the same read-out enzyme or RNA polymerase.

In any case, whether the read-out in every section is always in the same direction or not, we see that the genetic chains are close to the book model in that they form only a single linear sequence of sense, and are to be read out from one chain only, and not in duplicate, at each point. [Is the rate of read-out in different organisms also roughly constant, as it is for different books? Possibly. At least the lengths of the lifetimes required for a phage, a bacterium, and a man to develop all of their genetic information bear a very rough order-of-magnitude relation to the total lengths of their genetic DNA chains (U. Liddel, personal communication, 1961).]

V. Copying without Reading, and Vice Versa

We can copy a book—photographically, for example—without reading it. But monks and stenographers have also copied many a manuscript, letter for letter, without "realizing the sense" of it. In genetic information, too, we see that the "copying" process does not involve "reading." In copying, the enzyme DNA polymerase can go along replicating the two strands of a DNA double-helix from one end to the other, like a "zipper" (Rich, 1962a). No Messenger-RNA or protein is made. The whole chain is accessible to this enzyme—although it may be in this condition only during part of the life cycle, perhaps after its accessibility to "reading" has been blocked or finished. (Copying a book, too, interferes with reading it.)

Conversely, we can read a book, or sections of it, without making a permanent copy. "Reading" DNA chains "to get the sense out" requires a different enzyme, RNA polymerase, which goes along the chain making a complementary Messenger-RNA chain. This enzyme, like us, does not read everything, but only those accessible or "derepressed" sections that have been "opened" for it. And it does not make a permanent copy, since the Messenger-RNA is short-lived and seems to be destroyed after making its protein. (The protein, of course, is *not* a "copy" of the instructions in the present sense. It cannot make further copies of itself, for example. It is instead the "working equipment" that the instruction manual shows how to assemble and set running.)

VI. Pages Open or Closed

The genetic book is not a "scroll," which can be opened for reading to sections of any length, with any beginning or end points. The evidence for "induction" and "repression" of clearly marked-off sections requires us to think of it at this point as a Western "book" with separate "pages" or, better, "double-pages" which are definitely opened or definitely closed at any given time (Jacob and Monod, 1961).

A "double-page" thus corresponds to "a *unit* in the transfer of information"—an "operon"—which is "opened" or "shut" by the action of an "inducer" or a "repressor" on a given "operator-gene." It may contain a single "operator-and-structural-gene," or it may contain an operator-gene together with two or more distinguishable

structural genes, governing two or more separate proteins (Jacob and Monod, 1961).

We might think of each of these structural genes as a "paragraph." (The paragraph even has "sentence" or "clause" subsections, the genetic "cistrons," which make protein fragments whose sections are complete enough to produce successful catalysis by mutual "complementation.") The page labeled "Histidine synthesis" in the *Salmonella* book, for example, seems to contain some 8 paragraphs of this kind, with the structural instructions for producing 7 different proteins that catalyze consecutive steps in the histidine synthesis (Jacob and Monod, 1961). (The eighth paragraph is the "heading" or "operator-gene.") The page can be closed, and all of these structural read-outs stopped, by excess histidine. It does not become partially closed; the process usually seems to have an "all-or-none" character. A page is "open" or "shut"; which is why, as we know, a vertebrate cell can do kidney chemistry or cartilage chemistry, but not something graded in between.

The lengths of pages and books. The known enzyme protein units seem to contain 100 to 500 amino-acids, or the same number of base-triplet "words" on their DNA chains. A "page" with 1 to 10 paragraphs may then contain of the order of 300 to 3000 words, say 2000 as a round number. This is comparable to the page lengths in large books or encyclopedias.

A bacteriophage or virus with a DNA chain, say, 200,000 bases long, then has about 60,000 words or roughly 30 pages in its genetic instruction book. The book for a bacterium is 10 to 100 times larger. And for a man, the genetic information in the 46 chromosomes of each somatic cell is not so much a book as a very large encyclopedia with 46 volumes, about 6×10^9 base-pairs, 2×10^9 words, and a million pages; or an average of about 20,000 pages per volume. The reason for dividing large books into many separate volumes, in the biological as in the literary case, might be that it makes all this information mechanically easier and faster to reprint and handle. Also it may be mechanically easier, if many pages need to be open at the same time, to have them open in separate volumes.

VII. Only a Few Pages Are Open

A cell is particularly like a book not only because the pages can be opened or closed as needed, but because only a few pages need to be open at a given time. For one thing, different enzymes are

wanted for different environmental conditions, for example in a cell that can subsist on many different sugars. As genetic information, they may represent evolutionary adaptations that have been useful in the past, but only a few of them may be needed to survive on a particular substrate. Many enzymes also may be needed to make certain products only at one stage of the life cycle and not at another.

Jacob and Monod correspondingly emphasize the need for repressive regulation of most enzyme-making genes. They point out that "constitutive mutants," in which the control of a particular enzyme has broken down, may turn out 6% or more of their protein in the form of this enzyme (not the enzyme-product, but the enzyme!). Obviously "cells could not survive the breakdown of more than two or three of the control systems." That is, the cell cannot "hold open and read out continuously" more than two or three of its hundreds of pages.

The limitation on the fraction of its information that a cell can express at a given time is even more dramatic in a multicellular animal. Clearly the ovum, before its first division, cannot express the full informational content of the animal, and it is generally recognized that most of its genes are not functioning. Likewise in the adult animal, no single cell is expressing the information represented by all the other types of tissues. The information is "there" (at least in the somatic cells of higher animals) but it is repressed. This is proved by the classical embryological experiments on the "induction" of cells to form some of the other tissues that they had not been "destined" for (skin into eye, muscle into cartilage).

Recently, in fact, it has been recognized that this selection of only a part of the genetic information to be active chemically at a given time can be seen under the microscope. The "Balbiani puffs" and swellings at particular spots on the stretched-out chromosome chains (in insects) are now known to be sites of synthetic activity (Beerman, 1959). And different sites along the chromosome are found to be swollen and active in different tissues, or in the same tissue at different stages of development.

A human somatic cell contains a few hundred times the DNA of a large bacterial cell. If it can do at one time about the same number of different chemical syntheses as the bacterial cell can do, then it must be using a hundredfold or thousandfold smaller fraction of its total genetic information. (It may be no accident that

the total number of differentiated cells that have so far been distinguished in man is of the order of a hundred.) We begin to realize that it would probably be energetically and chemically impossible for a tissue cell to use all the different kinds of information that the whole animal uses, unless the tissue cell itself became as large and as structurally differentiated as the whole animal.

The books in each cell may be open to only a relatively few pages at any one time.

VIII. Page Headings for Reference

We now come to a feature of the book analogy that has not received much genetic or biochemical emphasis. It is that every page of a book, except a very short book, always contains *two* different kinds of information:

(a) A long sequence of words, highly specific to the book, and so containing a large amount of information—that is, the text.

(b) A short sequence, common to many books, containing only a little information—that is, the page number or descriptive reference heading.

The specific text and its paragraphs we have seen in the structural genes. "Page number" or "reference heading" is probably a good metaphor for the operator gene. The way in which a regulator-gene at one place "refers" with its repressor-substance (if the right inducer is present) to an operator-gene to open an operon page, seems closely analogous to the steps and elements involved in a "conditional-transfer" instruction in referring from one page to another of a complex instruction book.

An example may make this clearer. As we are reading along on a certain page of our instruction book, we might come to the "conditional-transfer instruction" (that is, the regulator gene) that gives, in effect, this sequence of directions:

"See if histidine is present." (That is, let the RNA-polymerase or some other repressor-read-out enzyme make the repressor-substance-for-histidine from this regulator-gene information chain.)

"Then if histidine is not present" (that is, if the repressor substance does not now form a complex with any free histidine),

"turn to 'page 137,' the 'histidine-synthesis' page" (that is, let the uncomplexed repressor find the operator-gene that it fits—in its uncomplexed form—in some complementary way),

"and start making histidine from the instructions there" (that

is, let the RNA-polymerase get past this "derepressed" or "open" operator-gene and start peeling off the 7 kinds of Messenger-RNA for the histidine-synthesis enzymes from the 7 paragraphs of instructions).³

In this analogy, perhaps it makes little difference whether we think of the operator-gene as a page number or as a reference heading. (In computer-programming terminology, it is an "address.") The "number" idea might be more appropriate whenever the reading is "sequential" from page to page down the chain; for example, perhaps, in the first sections read out after mitosis or fertilization. (Though a chain that is read exactly sequentially, without any forward or backward referencing at all, would not even need page numbers.)

On the other hand, in most cell chemistry, and certainly in multicellular organisms, the pages will not be opened sequentially but according to function, and the operator-gene must then play a role close to that of a "short reference heading." Short, of course, because the operator-gene-code, like the reference heading, must be written in a language common to many organisms—the language of the common inducers and co-repressors—or else it must be "translated" by the repressor-substance into such a language. This means that the "recognition-section" of the repressor-substance—and the section of the operator-gene, also, that recognizes this repressor—only needs to be a short, "low-information," "heading" that responds to, and corresponds to, "histidine" or "glucose" or some other one of a relatively small number of widespread substances.

Thus a book—or a bacterium—with 900 pages can specify which page is open by using a number (an "address") with 3 digits or 10 bits, even though each page contains 1000 to 5000 bits of information. And each of these 3-digit numbers, or each of the 2-word page headings written on the same line with them in the anatomy book, is common to many books.

³ The reader may be puzzled or amused by my choice of verb forms in writing down these "directions." They are chosen to represent the fact—the deviation from the "human-reading-a-book-model"—that there is no extra "little man" in the cell to "do" the reading or interpreting of the DNA instructions. The cell "reads itself," where "read" has become an intransitive verb, and means that the reading is essentially synonymous with the action of "turning pages" and "carrying out the instructions read." We have no customary verb forms in English for this "automatic" reading-which-is-action, in which the book does not require an external reader, and is not read, but reads.

This low-information feature must continue to characterize the more complex biological "inducers" by which one tissue affects another, and the repressors that recognize them, even though we do not know so much in detail about how each one acts. The hormones, for example, are one class of inducers. They are small and diffusible molecules common to many species, producing such similar tissue-development reactions that they must have low-information and low specificity compared to the structural proteins. The same is true of the more obscure tissue-inducer-chemicals. Structures such as eye-elements or feathers, of types specific to a given host-species, can be induced from the skin of the host in the wrong places, by embryo-transplants of inducer-tissues from an entirely different species or even class of animal. The structural genes induced are specific to the host; the inducer-chemical is specific only to the type of inducer-tissue. The inducer-chemicals, whatever they are, must therefore be common to many species or classes.

A similar organ-specific, low-information interpretation of a more subtle interaction seems to be demanded by the experiments of Moscona (1961a,b) and Weiss on reaggregation of separated ("trypsinized") tissue cells. They showed that chick kidney cells and mouse kidney cells "recognize" each other "as kidney" (at least at first, before any immune-reactions set in). The cells, meeting at random, form bonds with each other and then form organized tissue, reshaping and rearranging themselves into a hybrid pseudo-kidney, with kidney-like tubules and tubule-walls and cilia and other details. By contrast, when separated kidney cells are mixed with cartilage cells, either from the same species or from different species or classes, they "reject" each other and form separate kidney and cartilage tissues. This suggests assuming that some diffusible and widespread, but organ-specific, "kidney-recognition" molecules can diffuse into cells that have once been opened to the kidney "pages" and can open again the synthetic page whose "heading" reads: "Kidney, tissue formation with neighbors." But the other, "cartilage-recognition" molecules, diffusing in, are unable to open any synthesis at that heading. (Possibly they are blocked even from diffusing in; but this would simply involve the sub-case of the formation of "recognition-permeases," which is probably a complication that we do not need to go into here, since it seems likely that its genetic aspects do not involve any further new principle.)

Chemistry of "page heading" and repressor molecules. Evidently,

the "repressor-substance" must "recognize," and be capable of specific links with, two kinds of molecular groups: the operator-gene, and the co-repressor or inducer chemical. This suggests that it must itself have two functional groups, one of them probably protein, since the repressor is known to be stereo-specific for the inducer or co-repressor. It is true that the repressor-substance, at least during formation, does not test for protein, but it is hard to reconcile the stereo-evidence with any other alternative; and Jacob and Monod conclude that perhaps "the repressor itself synthesizes the 'induction protein' and remains thereafter associated with it."

By implication, the other, first, functional group in the repressor-substance could then be a nucleic-acid chain, like a Messenger-RNA that is not destroyed. It might be complementary in base-sequence to the regulator-gene DNA from which it is made. It might also be a template for the protein it becomes attached to. Or it might be complementary in base-sequence to the operator-gene DNA it can become attached to; or conceivably it might be all of these, although that is perhaps too much to expect. (On the other hand, an identity of both the regulator-gene and the operator-gene base-sequences would be like the identity of the two page-numbers on the line where the page is referred to and on the heading of the page itself.)

But whatever the details, this indication that the repressor-substance or "page-heading translator" is a compound of protein-plus-nucleic-acid, suggests immediately some important biochemical generalizations into which a number of other pieces of evidence can also be fitted.

For it is clear, on thinking about it, that if we are to have two different kinds of codes related to each other in cellular biochemistry—the base-sequence codes of the nucleic-acid chains, and the amino-acid-sequence codes of the protein chains—we must have, at every point of interaction between them, "translator molecules" containing both codes and able to "speak" both languages. The repressor-substances are molecules of this sort that translate from the inducer-code (whether all inducers are proteins or not) to the operator-gene DNA language. (If we suppose that their interaction is not with the operator-gene DNA but with some unknown "operator-gene protein," it only pushes the translation problem back onto this protein. In fact we see that when the repressor-substance is attached to the operator-gene, it is operator-gene protein.)

The general translation requirement suggests that we look for

other sets of protein-nucleic-acid translation-molecules. Two genetic-translation sets come to mind immediately. One is the set of Transfer-RNA molecules ("soluble-RNA"), each of which consists of an RNA-chain attached to its specific amino-acid. These are the molecules that transfer the base-sequence information from a Messenger-RNA template into a protein chain (Rich, 1962a). Another set is the set of postulated transfer-enzymes that "read" some base-sequence on these Transfer-RNA molecules in order to attach the proper amino-acids to them.

[It is instructive to spend a moment inquiring into the possible problem of "infinite regress" here. Will another second-order set of transfer-enzymes (or "language-teachers"?) need to be postulated to "read" the bases on the first set and attach the right amino-acid groups to them? And a third-order set to read these, and so on? To state the problem is to see it is absurd—and that it leads us to a chemical conclusion! For clearly at some stage, there must be either a chain of bases that does not need a further transfer-enzyme but that complexes spontaneously with a particular amino-acid; or a chain of amino-acids that does the same with a particular base. Where does this stage occur? The translation from RNA base-triplets to amino-acids already appears to be general for several organisms (Nirenberg and Matthaei, 1961; Speyer *et al.*, 1962). Therefore the Transfer RNA's themselves—or at most, their postulated transfer-enzymes—may already be universal. This can only mean that their base-and-amino-acid relationships are founded, like base-complementarity in the nucleic-acids, on steric and chemical reasons for "fitting," rather than on the accidents of biological survival. This in turn suggests a novel conjecture that seems strange but that might be worth exploring. A Transfer-RNA molecule contains some 80 bases but needs only 3 of them to "fit" the Messenger base-triplet; is it possible that this extra length is needed to make it a kind of "nucleic-acid-antibody" having a specific wrap-around relationship with its own amino-acid and with the other groups it must be fitted to? This would eliminate the need for a further transfer-enzyme.]

There may be another important set of protein-nucleic-acid "translator" molecules, namely those involved in embryological induction. There seems to be general agreement now that the "inducers" that diffuse from one cell to another are very likely RNA-protein compounds. Lash (1962) and Hommes *et al.* (1962) analyzed

chick-embryo spinal-cord inducer for cartilage-information in the somites. They found that it contains monophosphates of the bases guanine and cytosine, plus a hexose, a hexosamine, and some 15 amino-acids.

An embryological "inducer" acts in many ways like an "order" to a cell to "open" to a particular "heading" even if it is doing something else. This feature, together with the presence of the nucleic acid, suggests that it may not be like the small substrate-inducers complexing with a "repressor," but that instead it may be itself a repressor-competitor. If this were so, the need for endogenous "translation" of the exogenous inducer-compound would vanish. Certainly such a "translation-service" could often be eliminated between the cells of a species. And clearly, in that case, the interspecific generality found for tissue RNA-inducers must indicate an interspecific generality of the "operator-gene codes."

Whatever the nature and distribution of these various "translator" molecules, the translation requirement seems to demand that there be appreciable chemical and physical differences between regulator-genes, operator-genes, and structural-genes. To generate nucleic-acid-protein repressor-substances, for instance, demands some special kind of read-out of the regulator genes. Can this be done just by a special kind of base-sequence, or is a different kind of read-out-polymerase needed? This read-out might have to be done early in the life of a cell so that these "repressor-headings" could be attached to the operator-genes before any structural read-out is ready to begin. Does this mean that regulator-genes and operator-genes are specially accessible? If many regulator-genes have to be read out in this way "before the cell needs them," these repressor-syntheses might represent a major fraction of the early synthetic activity of the chains. The operator genes of course must have a different kind of chemical specialty; for they must evidently say "open" or "closed" to the RNA-polymerase that makes Messenger-RNA, but they need not themselves be copied into Messenger-RNA.

In the light of these considerations, it would not be surprising if the regulator and operator genes are marked off from the other information in the genetic book by some kind of chemical commas, or by a different helicity or configuration. Possibly the "bending" of DNA chains when they fold into the compact "storage" forms might occur most easily at the "page-number" genes, increasing

further the resemblance to a book. Some physical "closure" of this kind, during storage or during reading-periods, might be the simplest way to obtain chemical inaccessibility of the closed pages; but obviously during read-out, at least, the page-numbers or reference headings themselves would still have to be left accessible. What could satisfy these requirements more easily than physical folding, with the page-headings "out"?

And chemically, the page-heading sections of the chain, in any stages when they really do have attached repressor-protein, should be easy to distinguish from the structural sections of the DNA chain. This might be a simple reason for the alternating dark and light nucleic-acid and protein-containing bands of successive "genes" visible in the stretched-out salivary chromosomes of *Drosophila*. [See Stedman and Stedman (1950) and Bloch (1962) on the possible page-heading role of the histones on DNA.] We see that, with this idea of "reading page-headings," the book analogy is fruitful in suggesting links among an especially large number of phenomena and in throwing a different light on several current directions of research.

IX. Sequential Reference

On the book model, our picture of individual cell development comes to be like the reading of a complex instruction manual with forward-references and back-references and appendixes.

In a protozoan after mitosis or in an ovum after fertilization, we may then think of the genetic book as opening automatically, or being opened, to "Page 1." (In diploid cells, there are two books to be opened together and read out together, not always agreeing with each other; but let us omit this added complication here.) All the genetic information from this first operon, page 1, begins to be read out into structural Messenger-RNA, perhaps by RNA-polymerase already present. Soon the regulator-information from this page or from later pages begins to be read out (by this or another enzyme) into repressor-page-headings (reference-slips?) which can diffuse to their operator-sites on later pages. Some of the enzymes or some of the repressors manufactured from this page of instructions may be made only once; others might be made over and over again as long as the page is open, perhaps depending on whether the Messenger-RNA chains carry "cancel" or "repeat" instructions on their ends.

When "Page 1" has finished being read, or perhaps when enough of its enzyme products have accumulated as "inducers" demanding further chemical treatment, "Page 2" may be opened. "Page 1" might then be closed by the same "inducers" acting as "co-repressors" on it; or it might be closed later by others. At some point, the conditional repressors may begin to be specific for food or poisons in the environment, and "Page 3," "Page 12," and "Page 64" may be opened by some inducers, not sequentially, but according to functional "heading." And so on, with further page-reading and forward and backward references, or references to special appendixes, up to some point where "enough" growth products have accumulated, the pages are closed or the RNA-polymerase is "turned off," and the DNA-polymerase "copying" begins, in preparation for mitosis.

For the metazoan, the whole book is not finished by the first cell division, but only "Chapter 1." When the book reopens after mitosis, it may reread certain parts of that chapter again, skipping the bits about "reaction to fertilization," but repeating many of the instructions for growth and cell division. But the cytoplasm is no longer the same as it was before, and the accumulated products and inducers in these daughter cells may also begin to open up new pages, the pages of "Chapter 2." And so on, again, to the end of that chapter; and on, through successive chapters in successive divisions.

On this book-reading model, the rule that "ontogeny recapitulates phylogeny" would have to be translated into the statement that many of the early chapters in the genetic books remain almost unchanged for millions of years.

X. Differentiation: Cross-References from Other Books

As each cell begins to have neighbors, so that the chemical gradients produce different concentrations of inducers and co-repressors in the different cells, the various daughter cells in a given generation no longer open to the same chapters at the same time, but to different chapters. In mosaic embryos and similar cases, where the daughter cells seem to lose some potentialities irreversibly from the very first division, each one might have received only a partial copy of the whole genetic book. But it is not necessary to assume this, if only the inducers in the neighboring cells are sufficiently different to start them reading out from different pages, each leading to a

different sequence of developmental steps or page references that never converge; and if there is little exchange of inducers.

Differentiation of pluripotent cells, where each cell's response does depend on neighboring-cell inducers—with one cell giving one set of inducers and co-repressors to another and getting a different set back from it—has a different book-analogy.

The situation now is more like several copies of the anatomy book lying side by side on the table, being read from by several medical students dividing up an assignment. One says, "I'll look up all the information on bone if you look up all the information on blood and you look up all the information on cartilage." Each gives cross-reference slips to the others, so that the students—or the cells—exchange information and look up pages for each other; but the different books are open to different pages and chapters.

Where there is easy diffusion of inducers between the cells, one can see physically at what point specialization must begin. For there is no reason then why any cell should start reading chapters different from the others, or different from what it would have read if it were isolated, until the 8- to 16-cell stage is reached, where some cells begin to be surrounded, and the gradients can begin to produce a chemical difference between the "center" and the "outside" cells. But from that point on, different cells follow different destinies, each referring on from page to page following its particular sequence determined by its history and its neighbors. (Such "center-outside" considerations may be what limit the number of identical twins that can be born fully formed, since the early-division cells must be separated before they have lost their "unsurrounded" totipotency.)

There is obviously no biological reason why any of the somatic pages should ever contain repressor-orders reading: "Close up the open pages and go back to the one-cell stage." The sense of the information-flow from page to page is unidirectional. It may contain instructions for programming some cyclic sequences, but none for programming a reversal. In fact, the pages in differentiated tissue cells are probably always "self-inducing" for much of their chemistry, with inducer-messages from their open operons to themselves that read: "Stay open to these pages."

Such self-stabilizing enzyme mechanisms that stay "on," once "on," have a "flip-flop" character, a "steady-state" kind of "biological memory" that has been discussed by Szilard (1960). "Dedifferentiation" is a common phenomenon in tissue cultures; but before

it can be carried all the way back to the early fertilized-egg stage, we will have to learn how to turn off the "flip-flop." This may require treatment with co-repressors and inducers for earlier chapters; conceivably this might be done fairly simply some day by isolating a tissue cell from its neighbors and then surrounding it with growing cells from progressively earlier developmental stages.

XI. Stored Configurations

A final analogy between genetic information and a book is that both may be transformed part of the time into compact "wrapped-up" or "stored" configurations "in the bookcase." Such configurations, not accessible either to read-out or copying, are found in the packing of DNA into a phage head, and in the condensation of the cellular chromosomes into the short thick forms visible during mitosis. The shrinkage in length in each case is by factors of hundreds, of thousands, and the purpose of these forms must be to make a package hard to damage while it is in transit from one "reading-room" to another.

We know surprisingly little about these compact forms, except that the chains do not get mechanically tangled or broken either while they are condensing or while they are being pulled out to their full length again. Perhaps this simply requires every section to slide in the same direction "along its own length," like slippery nylon or like a snake getting out of a knot; because no tangling or knotting is possible with such a motion.

There are several possible configurations for the compact forms. One is a "coil of coils" or a "coil of coils of coils," like the forms used for the tungsten filaments in lamp bulbs. Another is a "transfer-twist" form, with the strands of a one-strand or two-strand helix being pulled out into numerous "lamp-brush" side arms (Platt, 1955). Both these forms are derivable by small perturbations of a double-helix, but they can lead to excessive twisting between the ends unless special conditions are met, or compensatory reverse twists also occur.

Long chains could also be packed into "watch-spring" coils lying on top of each other. But a stiff helix will not bend into a uniform curve as easily as into a group of straight segments separated by sharp bends. (This can be seen by manipulating an extended household steel tape, of the kind that has a quarter-cylindrical spring "set"; because it also bends most readily where a bend has already

started.) Perhaps therefore the DNA will simply tend to fold back and forth on itself in this way. There are also numerous more complicated alternatives, including the possibility that a double-helix might unwind in packing, with its single strands stabilized by quite different chemical attractions in the compact form—attractions which the form itself proves the existence of. [Rich (1962b) has emphasized the possibility of back-and-forth folding, and its analogy to the 100-angstrom-segment folding found experimentally in many long-chain synthetic polymers.]

The related questions of the mechanochemical forces that draw up the chains will be interesting in their own right. Does a wave of drawing-up propagate along the length, or is the drawing-up simultaneous at all points? Is the process as simple as "salting-out" with ions (Rich, 1962b) which could "go" spontaneously with purified DNA *in vitro*? Or does it involve the cooperation or pulling together of attached protein chains, comparable to the observed shrinking of the attached spindle-protein fibers during mitosis? X-ray and chemical studies and model studies will be necessary. We have begun to understand something about helix-coil transitions, and something about the behavior of chain polyelectrolytes, but the problem of these 1000-fold changes in length could carry us into wholly new aspects of mechanochemistry (Platt, 1961).

Nevertheless we see that the genetic books can be and are stored away in their "bookcase" between readings, however it may be done.

XII. Summary

In summary, the book model gives us a new way of looking at several current research questions. Some are the coding questions of "sense" and "compsense" strands and the direction of read-out in a double-helix chain. Others are the biochemical-genetic questions of "addresses," "page-headings," and operator-genes; of chemical differences between these and the other parts of the genetic chain; of protein-nucleic-acid translator molecules; and of the coding and site of action of embryological inducer molecules. And others are the problems of the diverse physical configurations of the DNA or RNA genetic chains as the various pages are "opened" or "closed" or are shrunk to the compact "storage" forms—configurations that may differ widely, in the large and perhaps in the small, from the long straight, nonreacting, double-helix forms of the conventional picture.

In general, the analogy between genetic information-transfer and a complex instruction-manual would seem to give us a coherent and fairly accurate schema for relating various kinds of phenomena. Both at the teaching level and at the research level, it seems to put into better perspective many of the biophysical and biochemical details of genetic expression in cell development and tissue differentiation.

ACKNOWLEDGMENTS

I am indebted to Professors A. Rich and E. Bell of the Department of Biology at M. I. T. for numerous instructive conversations on these questions.

REFERENCES

- Beer, M. (1961). *J. Mol. Biol.* **3**, 263.
- Beerman, W. (1959). ("16th Growth Symposium," p. 83. Ronald Press, New York.
- Benzer, S. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 1607.
- Bishop, J., Leahy, J., and Schweet, R. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1030.
- Bloch, D. P. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 324.
- Crick, F. H. C., Barnett, L., Brenner, S., and Watts-Tobin, R. J. (1961). *Nature* **192**, 1227.
- Dintzis, H. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 247.
- Hommes, F. A., van Leeuwen, G., and Zilliken, F. (1962). *Biochim. Biophys. Acta* **56**, 320.
- Jacob, F., and Monod, J. (1961). *J. Mol. Biol.* **3**, 318.
- Lash, J., Hommes, F. A., and Zilliken, F. (1962). *Biochim. Biophys. Acta* **56**, 313.
- Matthaei, J. H., Jones, O. W., Martin, R. G., and Nirenberg, M. W. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 666.
- Monod, J., and Jacob, F. (1961). *Cold Spring Harbor Symposia Quant. Biol.* **26**, 389.
- Moscona, A. (1961a). *Exptl. Cell Research* **22**, 455.
- Moscona, A. (1961b). *Sci. American* **205**, 142.
- Platt, J. R. (1955). *Proc. Natl. Acad. Sci. U.S.* **41**, 181.
- Platt, J. R. (1961). *J. Theoret. Biol.* **1**, 342.
- Rich, A. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1044.
- Rich, A. (1961). ("19th Growth Symposium," p. 3. Ronald Press, New York.
- Rich, A. (1962a). This volume, p. 103.
- Rich, A. (1962b). *Proc. Pontifical Acad. Vatican*, p. 137.
- Speyer, J. F., Lengyel, P., Basilio, C., and Ochoa, S. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 441.
- Stahl, F. W., in press.
- Stedman, E., and Stedman, E. (1950). *Nature* **166**, 780.
- Szilard, L. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 277.

Molecular Disease, Evolution, and Genic Heterogeneity¹

EMILE ZUCKERKANDL² AND LINUS PAULING

*Division of Chemistry and Chemical Engineering,
California Institute of Technology, Pasadena, California*

I. Generalities	189
A. The Notion of Molecular Disease and Its Field of Application	189
B. Molecular Disease, Evolution, and Environment	191
II. Hemoglobin, Its Multiplicity and Evolution	194
A. Introduction to the Hemoglobin Molecule	194
B. Hemoglobin Heterogeneity	195
C. The Evolution of the Hemoglobin Chains	198
D. The Destiny of Duplicate Genes and the Function of Genic Multiplicity	206
III. Three Types of Molecular Diseases	213
A. Interference with Function	214
B. Interference with Normal Intermolecular Relations	217
C. Interference with Synthesis	218
IV. Fighting Molecular Disease	220
References	222

I. Generalities

A. THE NOTION OF MOLECULAR DISEASE AND ITS FIELD OF APPLICATION

Life is a relationship between molecules, not a property of any one molecule. So is therefore disease, which endangers life. While there are molecular diseases, there are no diseased molecules. At the level of the molecules we find only variations in structure and physicochemical properties. Likewise, at that level we rarely detect any criterion by virtue of which to place a given molecule "higher" or "lower" on the evolutionary scale. Human hemoglobin, although different to some extent from that of the horse (Braunitzer and Matsuda, 1961), appears in no way more highly organized. Molecular disease and evolution are realities belonging to superior levels

¹ Contribution No. 2774 from the Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California.

² On leave from the Centre National de la Recherche Scientifique, Paris.

of biological integration. There they are found to be closely linked, with no sharp borderline between them. The mechanism of molecular disease represents one element of the mechanism of evolution. Even subjectively the two phenomena of disease and evolution may at times lead to identical experiences. The appearance of the concept of good and evil, interpreted by man as his painful expulsion from Paradise, was probably a molecular disease that turned out to be evolution. Subjectively, to evolve must most often have amounted to suffering from a disease. And these diseases were of course molecular.

Relationships between molecules, which define states of health and disease, may be altered by environmental factors, or by factors of aging, or by inherited internal factors. The two last types of factors are partly the same, inasmuch as aging is itself determined by genetic factors. The two first types of factors are also partly the same, inasmuch as aging is due to the cumulative effect of external agents. The term molecular disease in its more restricted sense, the only useful one, relates to the third type of factors, to altered relationships between molecules traceable to altered genes.

To the extent to which we have grounds to believe today that inheritance is linked to nucleic acids and that the primary products of nucleic acids, beside perhaps other nucleic acids, are proteins and no other types of molecules, the notion of molecular disease relates exclusively to the inheritance of altered protein and nucleic acid molecules. Abnormal glycogens, for instance (Cori, 1954), are to be traced to abnormal enzymes, proteins responsible for their production.

The abnormal nucleic acid and the abnormal protein produced under its control represent two aspects of the same reality. A great many more protein molecules are present at a given time than nucleic acid molecules responsible for their production, and this is one of the reasons why it is much easier at the moment to study the phenomenon of molecular disease from the protein end.

An abnormal protein causing molecular disease has abnormal enzymatic or other physicochemical properties. Changes in such properties are necessarily linked to changes in structure. It seems unwarranted at the present time to draw a basic distinction between two types of structural changes as causes of molecular disease, change in folding of the polypeptide chains and changes in the sequence of amino acids, the building stones of the chains. It becomes increasingly probable that the changes in spatial configuration (conforma-

tion) are direct expressions of changes in sequence (Crick, 1958; Mizushima and Shimanouchi, 1961). Therefore a molecular disease can probably be defined at the molecular level in a way that is potentially complete by determining the alteration of the amino-acid sequence of a protein (or the nucleotide sequence in the corresponding nucleic acid). This statement applies of course to a given intracellular and extracellular environment. If there are changes in the environment, the spatial configuration of a protein may be altered without any change in amino-acid sequence and a pathological condition may ensue. Such a change in environment results either from external influences, which do not concern us here, or ultimately from the change in amino-acid sequence in other proteins, to which the molecular disease must then be traced.

B. MOLECULAR DISEASE, EVOLUTION, AND ENVIRONMENT

The study of molecular diseases leads back to the study of mutations, most of which are known to be detrimental. All loss mutations in a broad sense of the word—*involving either the total loss of a protein or the loss of protein function through a structural alteration of the protein*—are molecular diseases. Loss mutations, on the other hand, are among the conditions of adaptation of the organism to changes in its environment and adaptation, the conditions of evolution. A loss of function, when compatible with survival thanks to the nature of the environment, may make cellular energy and genic raw materials available for the acquisition of new functions. More highly evolved organisms have lost powers of synthesis that more primitive organisms possess (Lwoff, 1943). It thus appears possible that there would be no evolution without molecular disease. A maintenance of molecular health, although in the interest of the individual, is opposed to evolution. However, only a small fraction of the molecular diseases that occur are used by and turned into evolution.

A bacterium that loses by mutation the ability to synthesize a given enzyme has a molecular disease. The first heterotrophic organisms suffered from molecular diseases, of which they cured themselves by feeding on their fellow creatures. At the limit, life itself is a molecular disease, which it overcomes temporarily by depending on its environment. Every vitamin we need today bears testimony to a molecular disease our ancestors contracted sometimes hundreds of millions of years ago. These molecular diseases are not experienced as such under normal circumstances, because our environment con-

stantly supplies palliative drugs. Conversely, if phenylalanine happened to be present only in low amounts in our usual diet, the mutation leading to phenylketonuria, characterized by the inability to convert toxic amounts of phenylalanine into tyrosine, would also not be experienced as a molecular disease, whereas it actually is one under the prevailing circumstances. We might say that evolution is based in part on the appearance of molecular diseases whereof the environment can cure the symptoms. Since our remote ancestors must have been autotrophic, we may consider ourselves as degenerate autotrophic organisms. Whereas, in order to achieve superiority, it is not sufficient to be degenerate, it is however necessary.

In many cases the notion of molecular disease is thus closely linked to the nature of the environment. It is not so in other cases, such as a structural change of the hemoglobin molecule leading to the loss of its ability to combine reversibly with oxygen. Both types of cases are similar in that chemicals available in the environment either cannot be used (oxygen or vitamin precursors from which vitamins are built) or cannot be disposed of (phenylalanine). The difference between the two types of molecular diseases resides in the fact that in one case, that of vitamin need and of phenylketonuria, the environment can make up for the lost biochemical reaction either by furnishing its product (the vitamin) or by ceasing to furnish its starting material (phenylalanine), and in the other case it is not that products are needed or that toxic substances must be excluded, but the process itself of making the product is essential to the organism. Thus when oxygen cannot be carried to the tissues efficiently it would be of no avail to furnish the tissues with oxidation products. The oxidation must be carried out by the organism itself, mainly because living matter requires chemical energy to be set free at the right time in the right place. Thus molecular diseases are defined in relation to the environment when the requirement involved is that of a substance or of less specific forms of energy such as heat, and they are not so defined when the requirement involved is that of a process fundamental to the existence of living matter, that is, of a high degree of specificity in the release of energy in relation to time and space. Life can get everything out of the environment except a degree of specificity approaching its own.

Considering molecular disease and the environment in relation to evolution, we are faced with a two-way relationship. Evolution has probably been influenced not only by heritable changes in the

organism that the environment could prevent from being deleterious, but also by changes in the environment that molecular changes in the organism could prevent from being deleterious. Molecular disease can be selected for as a defense against diseases caused by external agents. For example, the past incidence of malaria has been shown to be positively correlated with the abnormal hemoglobin, HbS, present in sickle-cell anemia; with thalassemia, another type of genetically controlled hemoglobin disease; with glucose-6-phosphate dehydrogenase deficiency; and with color blindness (Allison, 1957; Siniscalco *et al.*, 1961). Apparently the presence in the environment of the agent of a highly dangerous disease, *Plasmodium falciparum*, favors the conservation and spreading in the human species of molecular diseases that afford protection against the infectious agent by unknown mechanisms. The molecular diseases, at least in the heterozygous ("trait") condition, are less lethal than the infectious disease. Observations of this kind extend the interaction pattern between molecular disease and environment. A "molecular disease" may be maintained in the species because certain agents in the environment render it innocuous; or it may, on the contrary, be maintained because it renders relatively innocuous certain agents present in the environment. On account of this latter effect it seems possible that external disease-causing agents, notably infectious agents, have played a role in evolutionary sequences of noncompensated degenerative nature, such as those leading to parasitism.

The sickle-cell gene increases the life expectancy of the individual in the heterozygous state, while in the homozygous state it decreases it probably at least as radically as does malaria. When two carriers of sickle-cell trait marry, on the average half of their offspring will again be heterozygotes. The other half will have a decreased life expectancy, because of either sickle-cell disease (sickle-cell gene homozygotes) or malaria ("wild-type" homozygotes). In malaria-infested countries the sickle-cell gene will thus have a tendency to spread in the population. This would hardly be the case if the mutant gene were advantageous in the homozygous instead of the heterozygous condition. A newly appearing mutation that would be retained only in the homozygote would usually have no chance of establishing itself in the population. We may point out that the replacement in a population of a given gene by a mutant gene may often require two successive mutations, except when the population is very small (close inbreeding). After the first mutation the mutant is selected for in

the heterozygous state. This allows the mutant gene to establish itself, but at the same time the corresponding wild-type gene is preserved. To eliminate the wild-type gene a second mutation must now occur, such that the doubly mutated gene is most advantageous in the homozygous state. This double mutant would not cause what would appear as a molecular disease, while the single mutant that precedes it might. In this sense molecular disease may be a frequent intermediary step in those evolutionary sequences that lead to the total replacement of a gene by a mutant allele, and that require that the heterozygous condition be at first advantageous. During this phase the homozygous condition may indeed often be deleterious.

II. Hemoglobin, Its Multiplicity and Evolution

A. INTRODUCTION TO THE HEMOGLOBIN MOLECULE

If one examines molecular disease in relation to evolution it is unavoidable at the present time to center the discussion on the hemoglobin molecule. So far this molecule is the only one that has been studied in many pertinent respects: amino-acid sequence, structure of the site directly involved in function, structural changes leading to molecular disease, normal structural multiplicity, different rates of synthesis of structurally distinct "editions" of the molecule, and their change in the course of time.

All vertebrates save the most primitive ones seem to have hemoglobins composed of four polypeptide chains, linked to each other by bonds much weaker than the peptide bonds that are instrumental in lining the amino acids up unidimensionally within the chains. Each of these chains is composed of slightly less than 150 amino acids and carries one heme group that contains the iron atom capable of binding oxygen reversibly. The string of amino acids winds about in space in a highly specific fashion that may be common to all vertebrate oxygen-carrying pigments, since even sperm whale myoglobin (muscle hemoglobin) shows a similar conformation in spite of the fact that its amino-acid sequence differs very considerably from the sequences so far found in blood hemoglobins (Watson and Kendrew, 1961). Nature has produced a great many such hemoglobin and myoglobin chains that differ in their amino-acid sequence and therefore in various physicochemical properties and yet apparently remain similar in their over-all conformation and in the fundamental charac-

MOLECULAR DISEASE, EVOLUTION, AND THE GENE 195

teristics of their relation with the heme group. Not only does the amino-acid sequence of these chains always vary in different animal species, except perhaps in some extremely closely related ones, but any one individual of any given species produces a number of different hemoglobin chains, in part successively and in part simultaneously.

So far the tetrahemic hemoglobin molecule of higher vertebrates has always been found to be normally made up of two kinds of chains that combine two by two. Thus human adult hemoglobin, HbA, contains two so-called α - and two so-called β -chains (Rhinesmith *et al.*, 1957, 1958). This type of hemoglobin is predominant only from the time of birth on, while during intrauterine life by far the greatest proportion of the hemoglobins produced is represented by fetal hemoglobin, HbF, composed of two α - and two γ -chains (Schroeder and Matsuda, 1958; Hunt, 1959, Schroeder *et al.*, 1959b, Shelton and Schroeder, 1960). There exists another human hemoglobin chain during postnatal life, but normally never in more than small amounts, the δ -chain. Two α -chains combine with two δ -chains to form the minor component known as HbA₂ (Kunkel and Wallenius, 1955; Kunkel *et al.*, 1957; Ingram and Stretton, 1961). This is as far as the list of structurally distinct *normal* human hemoglobin chains goes at present. Occasionally, in the diseases called the thalassemias, one or two of the chains are present in subnormal amounts. Sometimes a relative excess of the partner chain is produced, which then associates with its own kind to form tetramers. This leads to the formation of abnormal hemoglobin such as HbH, composed of four β -chains (Jones *et al.*, 1959), or Hb "Bart's," composed of four γ -chains (Lehmann, 1959). Thus the association of two different types of chains is not an absolute requirement for the formation of chain tetramers but probably a matter of preferential affinity. All the known normal human hemoglobins have one type of chain in common, the α -chain, but in other species this is not necessarily so, as has been shown for chicken (C. J. Muller, 1961).

B. HEMOGLOBIN HETEROGENEITY

It is remarkable that hemoglobin-chain heterogeneity has been found in all species. So far only vertebrates have been examined, but of widely different classes, ranging from mammals to fish and Cyclostomes (reviewed by Gratzer and Allison, 1960; also Huisman *et al.*, 1960, and unpublished results from this laboratory). The

Cyclostomes belong to the most primitive group of vertebrates, whose living representatives, hagfish and lampreys, seem to have hemoglobins composed of single, unassociated polypeptide chains (Svedberg, 1933; Roche and Fontaine, 1940; Lenhert *et al.*, 1956), presumably of the same general type as those which in higher forms associate into tetramers. In this respect Cyclostome hemoglobins resemble the myoglobins. Even though lamprey hemoglobin chains do not normally associate into higher molecular units, these animals also possess several distinct types of chains (Andinolfi *et al.*, 1959). Apparently the multiplicity of hemoglobin chains is not an evolutionary consequence of their association, but their association is an evolutionary consequence of their multiplicity. This conclusion is confirmed by the observation that there exist also several distinct types of myoglobins in all individuals (Rossi-Fanelli and Antonini, 1956; Rumen, 1959; Rossi-Fanelli *et al.*, 1960; Edmundson and Hirs, 1961), while myoglobin polypeptide chains usually do not associate to yield molecular units of a higher order. [Evidence of the presence of dimers in solutions of some invertebrate myoglobins has recently been reported by Manwell (1958b, 1960)]. It appears justified at the present time to extrapolate from these and other observations to polypeptides and proteins in general and to state that proteins and polypeptides of all kinds may usually be expected to coexist within every individual in structurally distinct "editions." Several of these may be synthesized in the same cells, as fetal and adult hemoglobins often are (Kleihauer *et al.*, 1957; Itano, 1956); or they may be produced in different tissues. This generalization rests now on a number of investigations, among which are those of Markert and Møller (1959) and of Kaplan *et al.* (1960). The latter authors showed that in vertebrates as well as invertebrates lactic dehydrogenases extracted from different tissues of one animal are different from one another. By analogy with hemoglobin we may suppose that for most kinds of proteins there will be found in every organism, in addition to major components that succeed each other in time or coexist in different tissues, structurally distinct minor components. The importance of these minor components is probably negligible from the point of view of function, but not of evolution. Molecular diseases will of course relate to the quantitatively important "major" components only. From what precedes it will be recognized that the greater the role of minor components in evolution, the smaller that of molecular disease.

If this picture is correct and it probably is, the number of distinct proteins in the human organism that has been estimated by one of us (L.P.) to be of the order of 100,000 will have to be multiplied by a factor of presently unknown magnitude.

It must be pointed out that a number of minor hemoglobin components apparently do not differ in amino-acid sequence from one of the major ones (Jones, 1961). Some arise apparently through a secondary combination with some other molecule, such as glutathione (C. J. Muller, 1961). Others may be oxidation or denaturation products or chromatographic artifacts, or unusual combinations of hemoglobin chain dimers or monomers, or they may be hemoglobin polymers. Finally, it remains probable, though the contention still awaits experimental confirmation, that some minor components might have an altered amino-acid sequence without being produced under the control of an altered gene. Such components would express "errors" in the synthesis of normal hemoglobin chains (Pauling, 1957a).

If we consider structurally different hemoglobin chains found in one given species, man, we may divide them into two groups. The chains in one group differ from each other by more than one amino-acid substitution. Thus the number of changes in sequence, when the α -, β -, γ -, and δ -chains are compared with each other, varies from about 6 to a little less than 80 (computed from Braunitzer *et al.*, 1960a,b; G. Braunitzer, 1961, personal communication; Konigsberg *et al.*, 1961; Schroeder *et al.*, 1961; W. A. Schroeder and R. Shelton, personal communication; Ingram and Stretton, 1961). These chains which show marked differences have all been found to be controlled by distinct genetic loci (Itano, 1957; Smith and Thorbert, 1958; Cepellini, 1959a,b, Ingram and Stretton, 1961), and they are present in all normal human individuals. In the second group we find chains that differ from one of the others by only one amino-acid substitution. In all cases that have been examined, the gene that controls such a chain has been found to be an allele of the gene in control of the nearly identical chain. Each of these different alleles occurs only in small proportions of the population in different areas of the world. They are the abnormal hemoglobins. While some go unnoticed by their carriers, others lead to characterized molecular diseases in the homozygous condition (see for instance Itano and Pauling, 1957; Itano, 1957; Neel, 1959; Ingram, 1961a). Of course chains differing by more than one amino-acid substitution and controlled by allelic genes may be discovered, but they will presumably remain a small

minority, and, for reasons that will become clear presently, we may expect that chains differing by more than a very small number of changes in amino-acid sequence or, more accurately, chains that have been affected by more than a very small number of mutational events will generally be traceable to distinct genetic loci.

On the other hand, when we compare hemoglobin chains from different species, chains controlled by corresponding genetic loci may differ considerably in amino-acid sequence. This is of course only a presumption, since we have no means of matching genetic loci of different species. For instance, there seem to be two differences between the human and gorilla α -chains (Zuckerkandl and Schroeder, 1961), yet there is no reason to suppose that the genic loci controlling their production are not homologous.

We may venture the following generalization. While in different species markedly different hemoglobin chains may conceivably be and probably quite often are controlled by homologous loci (by genes that would be shown to be allelic if fertile crosses between the species were possible), within one species a greater difference between chains is associated with greater independence in their genetic control. In this respect it is suggestive that the α - and β -chain genes, among the most different within the species, have been shown to be on separate chromosomes, or at least not to be closely linked (Smith and Thorbert, 1958), while the β - and the δ -chain genes, which resemble each other most, appear to be linked (Cepellini, 1959b).

C. THE EVOLUTION OF THE HEMOGLOBIN CHAINS

The foregoing observations can be understood at once if it is assumed that in the course of time the hemoglobin-chain genes duplicate, that the descendants of the duplicate genes "mutate away" from each other, and that the duplicates eventually become distributed through translocations over different parts of the genome. Different non-allelic genes are thus thought to have arisen from an original mother gene. Since it seems justified to consider effective (i.e., viable) translocations as phenomena that occur more rarely than effective amino-acid substitutions, one would expect that genes related to a common ancestor but not closely linked differ from each other by a number of mutational changes, in accordance with observation. Ideas of this kind have been evolved by Bridges (1935), Metz (1947), and notably Lewis (1951), and have been applied to hemoglobin evolution by Itano (1957), ably developed by Ingram (1961b), and

elaborated quite independently by ourselves in 1960. It seems likely that the *intraspecific* multiplicity of proteins of a given type is to be explained in these terms. Considerable *interspecific* differences between proteins of a certain type may, on the other hand, as stated before, be compatible with homology of genic loci and not require the intervention of gene duplication. As species gradually get to be more different from each other, so presumably do the genes at the homologous loci.

All we can check at present are homologies of chain structure as expressed by correspondences between amino-acid sequences in hemoglobin chains, and such homologies, whether inter- or intra-specific, suggest a common evolutionary origin. An alternate hypothesis would be convergence by selection for functionally adaptive hemoglobin-chain structures. While convergence may play a significant role, this role is most likely confined to a relatively small number of features of amino-acid sequence. The over-all similarity must be an expression of evolutionary history. This is indicated by the gradually increased amount of differences found when human hemoglobin is compared with hemoglobins from progressively more distant species (Zuckerkandl *et al.*, 1960; C. J. Muller, 1961). The difference between human and fish hemoglobins is such that no common features, except the presence of free lysine, could be detected by the comparison of peptide patterns obtained by spreading the products of a tryptic digestion two-dimensionally over filter paper by successive electrophoresis and chromatography. The absence of common features in these patterns in no way implies the absence of significant stretches of similar amino-acid sequences, but nevertheless expresses qualitatively a degree of difference. A comparable result has been obtained by comparing mammalian and fish insulin (Wilson and Dixon, 1961). Insulin, on the whole, seems less variable than hemoglobin, even taking into account its smaller molecular weight.

At the other extreme we may compare human and gorilla adult hemoglobins. From the amino-acid analysis of separate gorilla α - and β -chains it appears that there are only two differences in the α -chain and one in the β -chain. The amino-acid analysis of isolated tryptic peptides from gorilla hemoglobin, two thirds of which has been completed (unpublished), has so far furnished no evidence of further changes. It is therefore possible that the gorilla β -chain and the human normal and abnormal β -chains form one single population. (As mentioned before, the abnormal human chains differ from

the normal ones by only one amino-acid substitution.) Since gorillas get along well with their hemoglobin, as they prove by existing, it is not likely that the gorilla β -chain, if it were present in humans, would cause molecular disease. The required oxygenation properties of hemoglobin must be rather similar in the two species that are otherwise so much alike. Thus, if the gorilla β -chain occurred in a human family the physician's attention would probably not be attracted to it. Moreover, it would probably go unnoticed in general surveys, because the nature of its difference with the human β -chain—probably a substitution of a lysyl for an arginyl residue—seems to be of the kind that current scanning techniques do not detect. Conversely, it is also possible that the human β -chain occurs in some gorillas.

Some of the hemoglobin chains coexisting within one individual differ from each other as much as or more than corresponding chains may be expected to differ in the most distantly related vertebrates. While human β - and γ -chains are only moderately different, they are much more different than gorilla and human β -chains. Therefore with respect to hemoglobin an adult man resembles an adult gorilla much more than his own human embryo. Morphological observation also suggests this relationship, which is now confirmed at the biochemical level. The human α -chain differs much more from the β -, γ -, and δ -chains than the latter from each other. Nevertheless even in the case of the α -chain the remaining similarities with the others are striking. When the sequence of the first 30 amino acids of the β -chain became known from G. Braunitzer's laboratory, our own knowledge of the α -chain was limited to amino-acid sequences in peptides isolated from the chain by tryptic digestion by W. A. Schroeder, R. T. Jones, J. R. Shelton, and their collaborators. The succession of these peptides along the α -chain was unknown. By fitting them into the β -chain according to the principle of maximum homologies, a sequence of the first 31 residues of the α -chain was predicted that was later confirmed by Braunitzer *et al.* (1961). This showed for the first time that the homology principle could be put to work effectively, even between chains that differ considerably.

It is possible to evaluate very roughly and tentatively the time that has elapsed since any two of the hemoglobin chains present in a given species and controlled by non-allelic genes diverged from a common chain ancestor. The figures used in this evaluation are the number of differences between these chains, the number of differ-

ences between corresponding chains in different animal species, and the geological age at which the common ancestor of the different species in question may be considered to have lived. Braunitzer and Matsuda (1961) have recently found that there are a minimum of 15 differences in sequence between the horse and human α -chains. (Only one of the two main horse hemoglobin components was analyzed.) This number is not likely to be increased very much by subsequent work. If we estimate that the real number of differences in sequence is between 15 and 20, we may take 18 as a probable mean. From paleontological evidence it may be estimated that the common ancestor of man and horse lived in the Cretaceous or possibly in the Jurassic period, say between 100 and 160 millions of years ago (Piveteau, 1955; Dodson, 1960). For the sake of the calculation it is assumed that most effective mutations result in single amino-acid substitutions, as evidence from abnormal human hemoglobins indicates, and that the evolutionarily effective mutation rate, i.e., the rate of the mutations that have not been eliminated by natural selection, fluctuated during the time of evolution of hemoglobin around a mean without showing a predominant trend to increase or to decrease. Under these conditions the presence of 18 differences between the human and horse α -chains would indicate that each chain averages 9 evolutionarily effective mutations in 100 to 160 millions of years. This yields the figure of 11 to 18 million years per amino-acid substitution in a chain of about 150 amino acids, with a medium figure of 14.5 million years. Our results for the gorilla hemoglobin chains yield somewhat different figures. Because of considerable fluctuations that may be expected in cases where the number of evolutionarily effective mutations has been very small, it seems advisable to use the figure derived from the horse α -chain alone. As the amino-acid sequences of more animal hemoglobin chains become known and paleontological dating is improved, the calculation will have to be revised. Also the number of differences between the human chains is subject to moderate revision, especially the comparisons involving the γ - and δ -chains, based on the results of Schroeder *et al.* (1961), W. A. Schroeder and J. R. Shelton (personal communication, 1961), and Ingram and Stretton (1961).

As Table I shows, the evaluation of the time elapsed since the β - and δ -chains differentiated places their common ancestor at the time of origin of the Primates or somewhat earlier. This checks with the fact that so far δ -chains have been found only in Primates (Kunkel

et al., 1957) and furnishes evidence that, at least with respect to more recent evolution, the present evaluation is not unreasonable. Also the time of derivation of man and gorilla from their common ancestor as calculated on the basis of the figures derived from man and horse, 11 million years, falls on the lower limit of the range estimated on paleontological grounds, 11 to 35 million years.

TABLE I
THE APPROXIMATE TIME OF DERIVATION OF DIFFERENT HEMOGLOBIN CHAINS FROM THEIR COMMON ANCESTOR

Chains being compared	Number of differences ^a	Estimated time of derivation from common chain ancestor	Corresponding geological period
β and δ	~6	44×10^6 years	Eocene
β and γ	~36	260	Beginning of Carboniferous
α and β	78	565	Toward end of Pre-Cambrian
α and γ	~83	600	Toward end of Pre-Cambrian
Gorilla α and human α	2	14.5	Pliocene
Gorilla β and human β	1	7.3	Mean 11

^a The presence or absence of one to several contiguous amino-acid residues in one of the chains is counted as one mutational change.

Of course, the uncertainty increases as we go further back in time. The common ancestry of the β - and γ -chains is placed at the beginning of the Carboniferous period, that is, about at the time of the first amphibians. Differences between fetal and adult hemoglobins have however been found also in contemporary fish (Manwell, 1957, 1958a). It is conceivable that these have arisen from a gene duplication independent of the one that led to the differentiation of β - and γ -chains.

The α - and β -chains are so different that the present evaluation places their common chain ancestor in the Pre-Cambrian, before the apparent onset of vertebrate evolution. The differences between the α - and γ -chains check reasonably well with those between the α - and β -chains. If the figures were taken at face value, it would seem that vertebrate hemoglobin with its differentiation into he-

moglobin polypeptide chains derives from an invertebrate hemoglobin. Thus the ancestors of the present α - and β -chains would have to have been already present as differentiated chains in those primitive vertebrates in which the hemoglobin chains presumably did not associate into tetramers, as in the contemporary Cyclostomes, mentioned earlier. The finding of hemoglobin heterogeneity in the lamprey (Andinolfi *et al.*, 1959) is suggestive in this respect.

The figures in Table I also make it appear unlikely that corresponding chains, say α -chains and their homologs in animals, when the most distantly related vertebrates are compared, will be found to differ from each other more than human α - and β -chains differ. The figures also strengthen the presumption that hemoglobin has not been evolved independently more than once during vertebrate evolution and suggest, as stated, that even the most primitive among the ancestral vertebrates had already inherited their hemoglobin from other forms.

Polypeptide chains that are clearly not homologous, such as horse-heart cytochrome c (Margoliash and Tuppy, 1960) and mammalian hemoglobin chains, may still have a common molecular ancestor, in the sense in which all protein molecules of a given organism may conceivably have one, but such an ancestor would have existed so far back in Pre-Cambrian times that comparative studies on contemporary organisms have no significant chance of revealing a kinship. For all practical purposes it is therefore correct to say that horse-heart cytochrome c and horse hemoglobin chains have evolved independently.

Our best excuse for making the present evaluation is that it affords us the opportunity to point out why it is probably wrong. The sources of error involved are factors in gene evolution that deserve to be mentioned here.

We do not know whether the present "major" hemoglobin components have once been derived from "minor" components. The contribution of minor components as oxygen carriers is mostly negligible. Unless they have other unknown functions, natural selection will not be expected to act upon them. Thus all mutations will probably be preserved in a minor component until one of the three following possible events occurs: a mutation that makes it unrecognizable as a hemoglobin chain; a mutation that brings about a total inhibition of its synthesis, or a mutational change that transforms it into a major component. If the ancestors of the human hemoglobin chains

that are important quantitatively (the α -, β -, and γ -chains) have started out as minor components, they will during that remote period have retained many more mutations per unit time than we have assumed, and from this point of view the figures given in the table would be overestimates.

They tend to be underestimates for other reasons. In the comparison between the chains, possible back-mutations, of which we have no knowledge, had to be neglected, as well as successive different effective substitutions at the same amino-acid residue. The likelihood of these events increases with the increase in the number of amino acids affected by change in a given chain. Thus the number of effective mutational events that have actually occurred since the α - and β -chains have evolved from their common ancestor may be significantly greater than is presently apparent.

Furthermore, even if we assume the intrinsic mutation rate of the hemoglobin chain genes to have remained fairly constant throughout the geological periods, disregarding the probable effects of changes in temperature and in intensity of ionizing radiations, the effective mutation rate may have varied widely according to the "ecological" conditions of hemoglobin within the organism and of the organism within its environment. In particular, during evolutionary transition periods such as afforded by a change from aquatic to terrestrial habits the effective mutation rates may have been much higher than at other times. The size of the populations at every stage is also of paramount importance in determining the evolutionarily effective mutation rate. Some other factors, the impact of which is equally difficult to evaluate, should also be taken into account. Fortunately the over-all result of the interplay of all factors is expressed in the speed of evolution, which has been evaluated. The general finding is that in the course of time evolution has become accelerated (Rensch, 1954). The more recent terrestrial groups of animals, on the average, have evolved faster than the more ancient aquatic groups. We may expect that this generalization, based on morphological characteristics, has its counterpart in the speed of evolution of deoxyribonucleic acid (DNA) and of the proteins. On the average, a lesser number of evolutionarily effective mutations per unit time may thus have affected the hemoglobin molecules during the initial phases of vertebrate evolution than in later periods. Our guess is that the numbers given in Table I are more likely to be underestimates than overestimates.

In the preceding evaluations we have equated one mutational event to one amino-acid substitution in the polypeptide chain. As mentioned, present evidence tends to show that this is the most frequent type of evolutionarily effective mutation in hemoglobin genes. It is too early to generalize to structural genes at large. Work in progress in several laboratories has a direct bearing on this question, such as the work of Yanofsky's group on *Escherichia coli* tryptophan synthetase at Stanford University, that of Fraenkel-Conrat's group on tobacco mosaic virus (TMV) mutants at Berkeley, that of Levinthal's group on *E. coli* alkaline phosphatase at M.I.T., and that of Brenner's group on phage-head protein at the Cavendish laboratory. There is ample evidence that mutational events other than single amino-acid substitutions exist. Some such evidence is derived from the comparison of the hemoglobin α - and β -chains themselves. In several regions of both chains sequences of contiguous amino acids numbering from 1 to 5 (or 6, if we include in the comparison the C-terminal sequence of sperm whale myoglobin, to which no counterpart is as yet known in hemoglobin chains) are missing. Three types of mechanisms might account for such observations: terminal growth of the chains, in the case of terminal differences; deletions; or insertions. Insertions may be duplications of chain segments, associated or not with a reversion of the segments. Although the events at the level of the chromosome and of a single DNA molecule may be qualitatively quite different, one must not discount the possibility that events of the type first described by Sturtevant (1925)—the insertion into a chromosome of a duplicate of a chromosomal region—may have its counterpart within one structural gene.

There are two one-amino-acid "holes" in the α as compared to the β -chain and one two-amino-acid "hole" and one five-amino-acid "hole" in the β as compared to the α -chain. If each of these is tentatively considered to be attributable to a single mutational event, then of a total of 78 mutational events that have led to the present differentiation of the α - and β -chains four events, i.e., 5% of the total, represent deletions, insertions, or terminal chain growth. It is possible that the actual proportion of mutations that result in events other than the substitution of a single amino acid in the polypeptide chain is much higher, if such mutations are more often lethal than "substitution mutations," as seems likely indeed.

"Substitution mutations," such as have so far always been found in mutant alleles of hemoglobin chain genes, if occurring in a major

chain gene and advantageous either will be eliminated rather quickly or will eventually replace entirely the wild-type gene. Likewise, "substitution mutations" that cause molecular disease either will be eliminated before there is time for a second mutation to occur in the mutant or, if a selective advantage exists under special circumstances for the heterozygote, as in the case of sickle-cell hemoglobin, will be confined to a small enough number of individuals so that the appearance of a second mutation in the same gene will be improbable. Considering a given species, there are only two types of cases where the appearance, by repeated single substitutions, of more than one difference between originally identical genes would be favored: (1) when the heterozygote is universally favored over the wild-type homozygote; this would apply to sickle-cell hemoglobin, if humans were universally exposed to malaria; (2) when a gene has duplicated and the conformity of the duplicate gene to the original model is not selected for. Assuming that the duplicate gene contributes to protein production, this conformity will be selected for only if an increase in output of a given polypeptide chain is advantageous.

D. THE DESTINY OF DUPLICATE GENES AND THE FUNCTION OF GENIC MULTIPLICITY

If gene duplication is one of the means of increasing the output of a given protein, one may distinguish two phases in this respect. Up to an optimum number of duplications, the duplicate gene will be selectively retained with a structure identical to and a position rather near to that of the mother gene. Beyond this point, duplicate genes will be progressively more strongly selected against. During this latter phase they will be in part eliminated with their carriers, and in part subjected to progressive change. When they are preserved and changed, their destiny may be of three types. They might evolve new useful functional properties. In this case they will be retained as active genes, and to the extent to which polypeptide output depends on gene duplication their own duplicates will be kept unchanged by natural selection. Secondly, functionless or unfavorable duplicates will not maintain duplicates to their own likeness and may themselves be translocated to other chromosome parts and be reduced to minor-component genes by a position effect. Some such minor-component genes, more or less profoundly changed in the meantime, may be selected for later in evolution and be changed

into major-component genes. Thirdly, the activity of the duplicate may be reduced to zero.

This elimination of gene activity may again take three forms. The changed structural gene may be bodily eliminated through the loss of the part of the chromosome that carries it; or it may be modified to such an extent that its products, although significant in amount, are no longer recognizable in terms of the original protein; or it may be preserved in a modified state, but totally or subtotaly deprived of the power of expression.

The existence of such "dormant genes," although difficult to verify, is a plausible inference from two types of observations: firstly, that within a given tissue, say the hematopoietic tissue, major and minor structurally distinct components of a given type of polypeptide chain are found. Since at a given time the relative quantities of hemoglobin chains vary between 100 and 1%, other genetically distinct minor components may be present in such small amounts that they are practically undetectable. Secondly, there are numerous examples of proteins that are produced exclusively in one type of tissue, and of which no trace is found with presently available analytical means in other tissues. The nonproduction of hemoglobin in muscle cells and of myoglobin in reticulocytes is one example. This example shows that some among even relatively closely related structural genes may, within a given tissue, be the ones strongly expressed, the others unexpressed. We must assume that all the structural genes have during embryological development been communicated to all cell lineages. It is therefore quite likely that there exist in every organism numerous structural genes that do not find in any of the existing tissues conditions favorable to their expression and thus remain permanently dormant.

Furthermore the relative structural similarity of minor hemoglobin components to one of the major components affords yet another argument in favor of the existence of dormant genes. Indeed, as we have seen, the human δ - and β -chains are quite similar. Likewise structurally distinct components of orangutan hemoglobin have been found to be quite similar, and the same holds for pig hemoglobin components (E. Zuckerkandl, R. T. Jones, Y. Nishiwaki and L. Pauling, 1959-1961, unpublished). If duplicate genes remained usually expressed, one would expect to find a series of minor-component chains differing from all the other chains as much as the human α - from the human β -chain. This is, however, not

the case. One is led to think that, in the long run, duplicate minor-component genes most often cease to be expressed. There is no apparent reason why one should assume that they have most times been bodily eliminated. Other possibilities are that they have been implied in a further translocation, that a mutation or transposition of a controlling element (repressor) has occurred, or that the structurally modified mutated gene possesses specificity characteristics that fail to comply with the specificity requirements for polypeptide production under the conditions prevailing in the cell.

Dormant genes of course are conceived as dormant only as far as their expression, and not as far as their mutability goes. Mutations in dormant genes and in minor-component genes will never be lethal, unless the latter have some distinct specific function, which would then lead us to consider them as "major components" of another protein type. Minor-component genes and, mainly, dormant genes may thus furnish an important and perhaps the principal part of the genic raw material for macro-evolutionary experiments of nature. A new translocation, or the transposition of a controlling element such as those described by McClintock (1956), or some other genetic modification may reactivate the dormant gene after a very long period of time during which mutations have changed it enough so that it now controls the production of a new kind of protein. In this fashion new enzymes, new functions can arise without the corresponding loss of old enzymes, old functions. We have recalled earlier the importance to evolution of the loss of functions through the mutations of active genes. But it is evident that evolution, while it makes the best of such losses and of molecular disease, could not be based on them alone. There are a great many more different functions to be carried out by a great many more different types of enzymes than we are allowed to suppose can have existed in early evolutionary times. Primordial living matter must have been limited to a few simple functions. Therefore the notion of evolution by gain is a necessary complement to the notion of evolution by loss.

Horowitz (1945) made a lasting contribution to our thinking about evolutionary gain at the enzymatic level. He described how new reaction chains might arise in certain circumstances through the chance combination of the necessary genes and furthermore proposed a general mechanism for the stepwise building up of complex enzyme-systems, presenting us with a plausible scheme of macro-evolution at the molecular level. Obviously, as enzyme systems

become more complex, more different enzyme molecules are needed. There are reasons to think that the same molecules cannot usually be expected to carry out several different enzymatic functions. Therefore, new genes are needed for the building up of new functions. This is where the concept of the mutational reactivation of dormant genes complements Horowitz's picture. Minor-component genes are not to be excluded from a similar role, but may usually not yet be different enough from their parental genes to be fit for carrying out novel functions. Most minor-component genes, as stated above, are liable to be eventually turned into dormant genes because natural selection will not prevent their transfer to synthetically inactive chromosome regions, nor their coming otherwise under the influence of a repressor gene, nor structural changes that might place the genes outside the range of specificity requirements of the available macromolecules that collaborate with the structural gene in protein synthesis. As dormant genes become reactivated after periods of cryptic existence corresponding perhaps to geological ages, they may produce potential enzymes that do not disrupt existing chains of reactions, but are able to add new processes to the old ones, perhaps in the ways described by Horowitz. One of the possible mechanisms of the reactivation of dormant genes is the reactivation of the chromosomal region where it is located through a change in intracellular environment. The hope of demonstrating the existence of dormant genes rests on this possibility. Such a change in intracellular environment could result from the adaptation of the organism to changes in the external milieu. In adaptational changes, during an initial phase, gain and loss mutations may be balanced, or loss mutations only may occur, so that the total complexity of the organism either remains constant or tends to decrease. In the process, however, the environment of the chromosomes may be altered in some tissue and, on account of this alteration, genes be activated in some regions of the chromosomes, inactivated in others. Inactivations of this kind will be mostly lethal and genomes with corresponding inactivation-resistant chromosome regions will be selected for. The newly activated genes on the other hand will now be available to respond to further adaptational needs and will furnish a series of gain mutations without corresponding losses. Parts of this concept are supported by observation. Changes in conformation of the genic DNA molecules are presumably related to changes in the activity of the genes, and Schmitt (1956) has shown that the state of chromosomal DNA

depends on the chemical environment of the chromosomes. Furthermore, genes in mice that display the same activity characteristics during the individual's life time have been found to be closely linked on the same chromosome (Paigen and Noell, 1961).

Through the reactivation of dormant genes by a modification of the intracellular environment an initial adaptational stress of great magnitude would appear instrumental in producing a rise in complexity of the organism. Thus would be solved the old paradox expressed in this question: why should organisms get to be more complex, since simpler organisms are evidently adapted as well to their environment? Once more Biology will show that it can do without any "élan vital" or "entelechy."

The present concept leads one to predict that the fastest evolution toward more highly organized forms should take place after the occurrence of major environmental stresses. Evolutionary history bears out this expectation. For a long time paleontologists have noted that the initial phases in the development of new types of forms has an "explosive" character (consult, for instance, Rensch, 1954). According to the present theory, we assume that at the onset of such an explosive evolutionary phase a change in intracellular environment has brought about the reactivation of a number of previously dormant genes. Rensch believes that the phenomenon of explosive evolutionary phases is adequately accounted for by increased natural selection accompanying the conquest of new biotopes. He thus considers as instrumental a change in *external* environment only. This however does not explain the trend toward increased complexity of the forms.

In the sense that has been laid out here a marked change in environmental conditions may lead to what a layman would call a shake-up of the genome, and this may be the part of reality behind the poorly documented and falsely interpreted observations of the Soviet anti-geneticist A. Lysenko.

To sum up, mutations of active genes controlling "major" protein components suffice to explain how an organism can adapt to changes in its environment. Mutations of minor-component genes and dormant genes, however, seem to be able to furnish the organism with the genic raw materials that eventually allow it not only to adapt to a new environment but also, in the process, to become more highly organized. Minor-component genes and dormant genes may thus prepare the major steps in evolution.

Beside a minor-component multiplicity, there is also a major-component multiplicity in protein production, especially the one characterized by a succession of major components in time. Why are hemoglobin β -chains substituted for γ -chains? The reason may not be that β -chains would not be fit to meet the respiratory prerequisites of intrauterine life and the γ -chains those of adult life, but, rather, that the structural genes corresponding to the two chains are located in two distinct chromosomal regions, one of which is activated in the particular intracellular environment determined by the organism at early developmental stages, and will not be active under other conditions, while the second chromosomal region, on the contrary, will be activated only at the end of embryonic development due to the presence or absence of some particular factors in the cell at that time.

To carry out a given function it is thus not sufficient for a cell to possess a favorable structural gene. A further prerequisite is that this gene either be located in a chromosome region that remains active in protein synthesis in spite of the changes of intracellular environment that occur during development, as is the case of the hemoglobin α -chain gene; or that there exist several duplicates of the gene, each of which is located in a chromosomal region that is active during certain phases of development, and that these duplicates are distributed in such a way over the genome that at any time of development at least one of them remains "on duty." These different genes will never be identical because, although they have supposedly arisen by duplication of an original gene, translocation is apparently a rarer event than amino-acid substitution, so that translocation genes will be expected to differ from each other by more than one amino-acid substitution.

Constant vital functions thus frequently need to have at their disposal several editions of a given type of genes in several regions of the genome that are successively activated and inactivated, or vice versa, with respect to protein-synthesizing ability. This view is advanced as an explanation of the generality of the most important types of "major-component" multiplicity in proteins. These types are on the one hand the successive embryonic and adult editions of a protein; and on the other hand the different editions found at any one time in different tissues of the same animal. The latter type of protein heterogeneity has been referred to earlier and may be interpreted in the same terms. In each tissue the particular intracellular

environment provides different conditions for the distribution of active and inactive regions throughout the genome, and thus different duplicates of a given gene that have undergone different translocations will be "on duty" in different tissues.

It thus appears that the unavoidable change in intracellular environment during embryonic development is a great challenge to embryonic development itself, because of the obligation that many critically important proteins be produced throughout and in spite of this change. One may therefore venture to say that there could be no embryonic development without gene duplication followed by gene translocation. In this theory the same events that furnish the genic raw materials for evolution also furnish the genic raw materials for ontogeny.

Other types of genes, of course, are active in protein synthesis during one or the other phase of ontogenetic development only. Certain functions, for instance, are developed only in the adult organism. When their expression is delayed or hastened, paedomorphosis or palingenesis ensues (see for instance Rensch, 1954). Chromosomal events of the type described may be postulated to cause these phenomena also.

After having insisted on the cause and on the function of genic multiplicity, in particular in relation to hemoglobin structural heterogeneity, we must here point out that there is a limit to hemoglobin heterogeneity, at least so far as phenotypic expression goes. A universal intraspecific structural multiplicity of all types of proteins is on the way to being established and this finding may be an important step forward in the recognition of biological reality. Yet part of this step is the rejection of older speculations about protein heterogeneity. These postulated a continuous spectrum of structural variants in the case of each protein. As a result, the apparent amino-acid composition of a protein would be only a mean composition, and proteins would not exist as strictly defined chemical species. Work of recent years has shown that, except for a possible low percentage of "errors" in synthesis, the chemical formula of proteins is as rigorously defined as that of simple molecules. We can no longer agree with J. S. Haldane and J. G. Priestley, who wrote in 1935 "It does, in fact, appear to be fairly certain that each individual has a specific kind of hemoglobin, just as he has a specific nose." Concordant amino-acid sequence analyses in three laboratories (Braunitzer in the Max Planck Institute in Munich; Schroeder in the California

Institute of Technology; Hill and Konigsberg in the Rockefeller Institute) of human hemoglobin chains obtained from different individuals, as well as the comparison of tryptic peptide patterns (Ingram, 1958) obtained with different human hemoglobins in many more instances, shows that this is not so. Of course a greater number of structurally distinct human hemoglobin chains are probably produced than we presently know about. This applies in particular to the abnormal human hemoglobins, about thirty of which have so far been described, and whose inventory one will probably never be able to consider as complete. This also applies to the normal human hemoglobin chains controlled by non-allelic genes, of which a few more may eventually be found.

Some workers go further and believe that a number of undetected hemoglobin alleles may be present in the population. They may escape notice if they behave identically in electrophoresis, but differ in uncharged amino acids. One of us (L.P.) and Itano (1957) have proposed this idea as a possible explanation for the inhibition of hemoglobin synthesis in the molecular diseases known under the name of the thalassemias. Changes in structural genes may indeed lead to an inhibition and even to a total loss of the ability to synthesize a protein. Ingram and Stretton (1959) have developed this idea as one of several possible explanations of the thalassemias. It has ceased to be likely that such cryptic mutants are of very general occurrence in normal populations. The great similarity between gorilla and human hemoglobin chains is a piece of evidence against such a view. With the exception of one difference relating to a serine residue in one of the chains, the uncharged amino acids may all be the same. This similarity suggests that nonpathogenic undetected structural variants of hemoglobin chains must be rather rare, unless a human is more often like a gorilla than like a human. It is true that this hypothesis is supported by more observations than most biological theories.

III. Three Types of Molecular Diseases

After having considered molecular disease in its relation to evolution and genic variability in its relation to evolution, we may comment on genic variability in relation to molecular disease. Three types of molecular disease may be distinguished. Mutations may (1) interfere with molecular function, (2) interfere with the adaptation of the molecule to the intracellular environment, and (3) interfere

with the rate of synthesis of a molecule that is functionally fit. Hemoglobin mutants offer examples of all three types.

A. INTERFERENCE WITH FUNCTION

The only known alterations of amino-acid sequence that lead directly to an interference with the oxyphobic function of hemoglobin are those of the pigments collectively known under the name of HbM. They are all characterized by the formation of methemoglobin, in which the iron is oxidized to the tripositive state (Gerald, 1960). It was recently shown that several distinct mutations give rise to abnormal methemoglobin formation (Gerald, 1958, and personal communication, 1960). It is interesting that structural studies have revealed that the amino-acid substitutions leading to HbM formation all affect a certain region of the hemoglobin polypeptide chains. This region has been called the "basic center," because it comprises a relatively large amount of basic amino acids, and we may call it "basic center I" to distinguish it from a second basic center further along the chain. The two basic centers have in common the property of containing a histidine that seems to be in relation with the heme iron. While the main linkage of the heme iron to the globin is supposed to be to a histidine of basic center II, the histidine of basic center I is, according to present evidence (Watson and Kendrew, 1961), placed opposite the sixth coordination position of the iron, the one that binds oxygen in the oxygenated state and water in the deoxygenated state (Haurowitz, 1949). It may be that the second histidine is the one mentioned by Conant (1934) and Coryell and Pauling (1940). Table II shows the variability of some features of sequence and the constancy of others of the peptide region around basic center I in different animal species. Abnormal human mutants in which this region is affected are also listed. Corresponding results relative to basic center II have not yet been forthcoming.

The seven consecutive amino-acid residues shown in the table comprise residues numbers 56 to 62 in the α -chain and numbers 61 to 67 in the β -chain, counting from the amino end. Three of the seven residues are shown to be substituted in some normal respiratory pigments and the changes, as in the ovine α -chain and in one of the orangutan chains, may involve the substitution of an acid for a neutral amino acid without any fundamental interference with hemoglobin function. Of course such substitutions would be expected to affect one or several of the physical parameters of oxygenation.

TABLE II
"BASIC CENTER I"

LIST OF HEMOGLOBIN CHAINS DIFFERING IN THIS REGION
FROM THE HUMAN α - OR β -CHAIN

Chain	Sequence	Reference
Human α , normal	-lys-gly-his-gly-lys-lys-val-	Schroeder <i>et al.</i> (1961)
Human β , normal	-lys-ala-his-gly-lys-lys-val-	Braunitzer <i>et al.</i> (1961)
Human α , Norfolk	-lys-asp-his-gly-lys-lys-val-	Baglioni (1961)
Human β , Milwaukee	-lys-ala-his-gly-lys-lys-glu-	P. S. Gerald <i>et al.</i> (personal communication, 1960)
Human α , M _{Boston}	-lys-gly-tyr-gly-lys-lys-val-	P. S. Gerald <i>et al.</i> (personal communication, 1960)
Human β , M _{Memory}	-lys-ala-tyr-gly-lys-lys-val-	P. S. Gerald <i>et al.</i> (personal communication, 1960)
Human β , Zürich	-lys-ala-arg-gly-lys-lys-val-	C. J. Muller and Kingma (1961)
Bovine α	-lys-gly-his-gly-ala-lys- (or arg)	C. J. Muller (1961)
Ovine α	-lys-gly-his-gly-glu-lys- (or arg)	C. J. Muller (1961)
Goat α	-lys-gly-his-gly-glu-lys- (or arg)	C. J. Muller (1961)
Horse α	-lys-ala-his-gly-lys-lys-	Inferred from Braunitzer and Matsuda (1961)
Orangutan α or β	-lys-asp-his-gly-lys-lys- (or glu)	C. Baglioni (personal communication, 1961)
Sperm whale myoglobin	-lys-val-his-gly-ileu-glu-val- (or glu)	Watson and Kendrew (1961); Edmundson and Hirs (1961)

The reduction of the basicity of this peptide region is still more marked in sperm whale myoglobin. While both types of chains of human, gorilla, chimpanzee, beef, and horse hemoglobins, one of the chains of orangutan hemoglobin, and the β -chains of sheep and goat possess four basic amino acids in the stretch of seven shown in the table, sperm whale myoglobin probably retains only the histidine

opposite the heme group as well as the initial lysine. In various types of fish the corresponding peptide must also be less basic than in man, as tryptic peptide patterns indicate (Zuckerkandl *et al.*, 1960). However, four amino-acid residues out of the seven have so far not been shown to vary in any of the normal respiratory pigments, and may be essential to the oxyphobic function.

Methemoglobin formation in man has so far been shown to be caused by substitutions at two different residues, and one of these substitutions has been observed in both α - and β -chains. In HbM_{Milwaukee} there is a substitution at the valine in the fourth position C-terminally with respect to the histidine that is supposedly in relation with the heme iron. In the two other kinds of HbM that have been structurally analyzed, HbM_{Boston} and HbM_{Emory}, it is this histidine that is affected; in both cases it is replaced by tyrosine. A replacement of this histidine by another amino acid does however not necessarily lead to methemoglobinemia, since methemoglobinemia has not been reported as a feature normally observed in the family possessing Hb_{Zürich} (Hitzig *et al.*, 1960), although oxydation may be facilitated. In Hb_{Zürich} the histidine in question is replaced by arginine. Thus the basic character of histidine seems to be more essential in protecting the heme iron from oxidation than its particular configuration, and the affinity for heavy metals, which is much greater in histidine than in arginine (Albert, 1952), seems also not to be involved.

Since the "basic center I" region is probably part of an α -helix, as is the corresponding region in sperm whale myoglobin (helical region E, Watson and Kendrew, 1961), the fourth residue after the critical histidine should lie next to it on the helix. It is plausible that the acid residue found in this position in the case of HbM_{Milwaukee} interferes with function. Thus all types of HbM that have so far been analyzed seem characterized by a change in state or in kind of the critical histidine residue. On the other hand, the aspartic acid that replaces a neutral amino acid in Hb_{Norfolk} does not seem to interfere basically with the function of the neighboring histidine, and this is probably due to the fact that because of the different orientations of neighboring residues in a helix the side chains of neighboring residues can be further apart than the side chains of residues four removed along the helix.

A further interesting observation related to the basic center I is that the change in the orangutan peptide is apparently the same as

in the abnormal Hb_{Norfolk} (C. Baglioni, 1961, and personal communication, 1961). In this respect human carriers of Hb_{Norfolk} have orangutan hemoglobin. The disease is mild; but this hemoglobin has not yet been observed in the homozygous state. The case shows that what may appear as a "molecular disease" in one species may be the norm in another.

B. INTERFERENCE WITH NORMAL INTERMOLECULAR RELATIONS

Mutations that do not significantly affect the oxyphobic function of hemoglobin may nevertheless lead to severe molecular diseases if they alter the physicochemical properties of the molecules that are of importance in its relation to sister molecules and to other constituents of the red cell. Among the most important of these properties is solubility. Considering the great proportion of nonpolar groups usually found in proteins, the building of a soluble protein molecule appears to be a difficult accomplishment. The readiness with which solubility is lost upon changing spatial conformation (denaturation) demonstrates that the solubility of most proteins is very sensitive to the distribution of their side-chains in space. Mutations that result in the substitution of a polar by a nonpolar amino-acid residue, as well as other mutations that weaken the conformational stability, may thus be expected to interfere frequently with the functionally required solubility characteristics of a protein. One might surmise that many molecular diseases should involve such losses of solubility. If it is justified to consider aging as a multiple molecular disease arising through somatic mutations, then aging probably also expresses in part the loss of solubility of certain proteins.

Hemoglobin is expected to be particularly sensitive in this respect to mutational change because it is in solution in the red cell at a concentration of about 30%, a concentration that not many molecules are able to achieve. Even a slight change in the properties of molecular interaction may under such circumstances lead to a drastic effect. Under certain conditions the abnormal human hemoglobins H [composed of four β -chains (Jones *et al.*, 1959a)] and Zürich (Hitzig *et al.*, 1960) tend to precipitate in the red cell. Among the cases in point, the most well known, and the one that typified molecular diseases in general (Pauling *et al.*, 1949), is that of HbS, the hemoglobin of sickle-cell anemia chemically characterized by the substitution of a valyl for a glutamyl residue in the sixth position from the N-terminus of the β -chain (Ingram, 1959). This substitution does not lead to a

decrease in solubility of the oxyhemoglobin (Perutz and Mitchison, 1950), but upon deoxygenation the molecules, when in sufficiently concentrated solution, interact and align along fibers that seem to form a network reminiscent of gelification (Bessis *et al.*, 1958), and at the level of cellular dimensions the alignment of the fibers is expressed by the formation of tactoids (Harris, 1950), which deform the cell membrane and lead to an early destruction of the red cells and to interference with blood flow in capillaries.

The oxygen-dissociation curve, its position, shape, and dependence on pH, and thus the fitness of the hemoglobin with respect to a given set of circumstances may be affected indirectly by amino-acid substitutions that alter the relation between the respiratory pigment and its cellular environment. The oxygen affinity within the erythrocytes of HbS is lower than that of HbA (Becklake *et al.*, 1955), whereas very similar affinities have been reported for these two pigments in cell-free solution (Allen and Wyman, 1954). Recent evidence indicates however that also in cell-free solution the oxygen affinity of HbS is lower than that of HbA (Riggs and Wells, 1961). As to HbF, it has a greater oxygen affinity than HbA within the red cell, yet in cell-free dialysed solution the affinities of HbA and HbF appear to be nearly identical (Allen *et al.*, 1953), except at pH values below 7 (Manwell, 1960). Red cells containing fetal or adult hemoglobin, respectively, differ in certain parameters such as surface, volume, and thickness (Riegel *et al.*, 1959). This difference may possibly express the influence of structural characteristics of the hemoglobin molecule on structural characteristics of the red cell. Admittedly other factors, such as the relative rate of synthesis of different hemoglobins, will also intervene here. This leads us to consider briefly the third type of molecular diseases, associated with the hampering of protein synthesis.

C. INTERFERENCE WITH SYNTHESIS

Hemoglobin synthesis may be interfered with by various mechanisms. One of the most obvious is the absence of red-cell production. This apparently occurs as a normal feature in a family of antarctic bony fish, the *Chaenichthyidae* (Ruud, 1954). When red cells are present a decrease in the output per cell of individual hemoglobin chains is known as a heritable character in many human families. Most of the known abnormal human hemoglobins, while not unfit as oxygen carriers, are present in the red cells in subnormal amounts.

HbJ, which is present in higher quantities per red cell than HbA, is so far the only exception to this statement (Thorup *et al.*, 1956). The decreased ratio of abnormal to normal hemoglobins may be assumed to express a decreased relative rate of synthesis.

Itano (1953, 1957) was the first to point out that the structure of a gene may have a direct influence on its synthetic activity. The decrease in rate of synthesis of most abnormal hemoglobins may be visualized in at least two different ways. It may be due to an interaction between the structural gene and its substrate on the chromosome; or it may be due to an interaction between the gene and extra-chromosomal macromolecules, which might be called co-determinant factors. The collaboration of co-determinant factors with the genes or the primary gene products is assumed to be necessary for protein synthesis. An alteration of the gene structure would upset the balance of attractive and repulsive forces (of electrostatic, van der Waals and steric nature) between these molecules. A decreased rate of synthesis would then result from a decreased degree of fitting between the two entities. We have referred earlier to the probable change of state in the genes during various phases of the development of an organism. Co-determinant factors also may change in tertiary structure as a result of altered intracellular conditions, or they may combine with smaller molecules, perhaps hormones, that change their specificity and activity. The rate of protein synthesis as a function of the intracellular environment may thus be determined by changes of state of the genes on the chromosomes as well as by changes of state of other macromolecules in solution.

Evidence is accumulating to the effect that large parts of protein molecules are not connected with any of their known specific functions. Various hypotheses can be used as partial tentative explanations of this fact such as the stabilization of the tertiary structure in larger proteins, the random survival of structures that were useful during earlier evolutionary stages, and the resistance to diffusion of large molecules (Pauling, 1957b). Consideration should also be given to the possibility that many parts of the protein molecule that appear functionally neutral may have a function in connection with the rate of synthesis of the protein. Natural selection may well act not only at the level of the finished protein product, but also at the level of its production.

The decrease in the rate of synthesis of hemoglobin chains is often more considerable in thalassemias than in the case of abnormal

hemoglobins. While a single genetic event probably causes both the amino-acid substitution and the change in rate of synthesis in abnormal hemoglobins (Itano, 1959), there is evidence that more than one type of genetic event may be involved in thalassemias. The mechanism of these diseases is being investigated in many quarters and has been recently discussed by Ingram and Stretton (1959) as well as other authors. None of the theories so far proposed is completely satisfactory. Whatever the theory, it should be kept in mind that the inhibition of protein synthesis is a nonspecific effect that may have different causes and that there may be several types of thalassemia, not only, as is commonly recognized, with respect to the particular hemoglobin chain that is most severely affected, but also with respect to the underlying mechanism.

IV. Fighting Molecular Disease

In the more or less distant future an enzyme deficiency such as the one that causes phenylketonuria may be met by endowing the organism permanently with a certain quantity of a stable artificial enzyme (Pauling, 1956). Of course, such a solution could hardly be considered in the case of a respiratory pigment, where the quantities required with respect to the "substrate" are not catalytic, but stoichiometric.

Another conceivable means of fighting molecular disease that cannot at present be theoretically excluded is the activation in a certain tissue through drug action of a minor-component gene or a dormant gene, representing a functionally fit duplicate of the functionally unfit "major-component" mutant. Whether or not such a treatment is possible depends on the nature of the control of gene activity. If this activity is in part controlled by position effects in relation to factors present in the intracellular environment, medical research might take such a course, although of course the problems of cell permeability and of the toxicity to the activity of other important genes of the agents capable of modifying specifically the intracellular environment might create major obstacles. The fact that this question can be raised shows however that research on the mechanisms of control of gene activity is not only of fundamental importance to biology, but of great interest to medicine as well.

None of the means of meeting the challenge of molecular disease will, however, be as satisfying as the elimination of the disease-causing mutant genes from the human populations. This is theoretically

the best, and it is at present the only concrete efficient measure that can be proposed. H. J. Muller (1959, 1961) who has given considerable attention to eugenics, has recently proposed the creation of germ-cell banks, from which prospective parents could draw the choicest human genomes. Such discussions and proposals are of paramount importance, even though one might not share Muller's optimism, which leads him to believe that in a process of free choice of genomes on the part of the populations the greatest human values would all get their fair share. It may be anticipated that governments would advocate and propagandize the choice of the socially minded, the active, the efficient, at the expense of the contemplative person; and the "good fellow" who represents the majority of humans would tend to procreate by choice a "super good fellow," a super corporation man, more able in conforming than in intellectual accomplishments. Without the contemplative, endowed with refined perceptive abilities of the qualitative and of the significance of forms, the human race would deteriorate. They are indeed humanity's greatest asset, both in the realm of doing and in the realm of being. It is true that the present development of the world toward a huge hive where nothing can stand in the path of technology and mass production promises to individuals so endowed little in the future except frustration and unhappiness, and one might contend that to relieve their suffering it might be a good deed to eliminate them genetically.

On the other hand, no objection can be legitimately raised, it seems to us, against the ambition to eliminate from human heredity those genes that lead to clearly pathological manifestations and great human suffering. The means of achieving this goal need to be discussed. We know now that in the United States about 10% of the Negro population carry one HbS or HbC gene. Therefore about one 400th of the children born to Negro parents have the deadly disease sickle-cell anemia. A simple test permitting the detection of the heterozygous carriers of a sickle-cell-hemoglobin gene exists, and as a first protective step there should be a law requiring all persons within a population in which this gene is present to any significant extent to submit themselves to this test.

If people carrying the mutant gene were to refrain from marrying one another, but married normal individuals, the incidence of the gene would remain constant in the population, and the problem of eliminating the gene would not be solved. To eliminate the mutant

gene the following rules may be proposed. If two heterozygotes marry they should have no children of their own. If a heterozygote marries a normal person they should have a number of progeny smaller than the average. In this way the mutant gene would be eliminated in the course of time in a way not involving the suffering caused by the birth of the defective children. Similar measures should be taken in the case of phenylketonuria and other molecular diseases.

In a marriage of heterozygotes, who do not suffer from the diseases caused by recessive genes, the chance that each child is homozygous for the defective gene is 25%. This percentage is much too high to let private enterprise in love combined with ignorance take care of the matter. And although interference of law in questions that are to a great extent of a very personal nature is to be avoided whenever possible, it would be clearly unethical to oppose such an interference in the case of molecular disease, at the very least in those cases, such as sickle-cell anemia, where presently available palliative measures are inadequate.

We may accordingly have hope that the increase in knowledge about molecular disease will in the course of time lead to a significant decrease in the amount of human suffering in the world.

ACKNOWLEDGMENT

We thank Professors E. B. Lewis and N. H. Horowitz for stimulating discussions about some of the topics referred to in this article.

REFERENCES

- Albert, A. (1952). *Biochem. J.* **50**, 690.
- Allen, D. W., and Wyman, J., Jr. (1954). *Rev. hémato.* **9**, 155.
- Allen, D. W., Wyman, J., Jr., and Smith, C. A. (1953). *J. Biol. Chem.* **203**, 81.
- Allison, A. C., (1957). *Exptl. Parasitol.* **6**, 418.
- Andinolfi, M., Chieffi, G., and Siniscalco, M. (1959). *Nature* **184**, 1325.
- Baglioni, C. (1961). *Federation Proc.* **20**, 254.
- Becklake, M. R., Griffith, S. B., McGregor, M., Goldman, H. I., and Schreve, J. P. (1955). *J. Clin. Invest.* **34**, 751.
- Bessim, M., Nomarski, G., Thiery, J. P., Breton-Gorini, J. (1958). *Rev. hémato.* **13**, 249.
- Braunitzer, G., and Matsuda, G. (1961). *Z. physiol. Chem.* **324**, 91.
- Braunitzer, G., Rudloff, V., Hilse, K., Liebold, B., and Müller, R. (1960a). *Z. physiol. Chem.* **320**, 283.
- Braunitzer, G., Hilschmann, N., Hilse, K., Liebold, B., and Müller, R. (1960b). *Z. physiol. Chem.* **322**, 96.

- Braunitzer, G., Hilschmann, N., Rudloff, V., Hilse, K., Liebold, B., and Müller, R. (1961). *Nature* **190**, 480.
- Bridges, C. B. (1935). *J. Heredity* **26**, 60.
- Cepellini, R. (1959a). *Acta Genet. Med. et Gemellol.* **8**, Suppl. II, p. 47.
- Cepellini, R. (1959b). In "Biochemistry of Human Genetics" (Ciba Foundation Symposium), p. 133. Churchill, London.
- Conant, J. B. (1934). *Harvey Lectures Ser.* **28**, 159.
- Cori, G. T. (1954). *Harvey Lectures Ser.* **48**, 145.
- Coryell, C. D., and Pauling, L. (1940). *J. Biol. Chem.* **132**, 769.
- Crick, F. H. C. (1958). *Symposia Soc. Exptl. Biol.* **12**, 138.
- Dodson, E. O. (1960). "Evolution: Process and Product," 352 pp. Reinhold, New York.
- Edmundson, A. B., and Hirs, C. H. W. (1961). *Nature* **190**, 663.
- Gerald, P. S. (1958). *Blood* **13**, 936.
- Gerald, P. S. (1960). In "The Metabolic Basis of Inherited Disease" (J. B. Stanbury, J. B. Wyngaarden, and D. S. Fredrickson, eds.), pp. 1068-1085. McGraw-Hill, New York.
- Gratzer, W. B., and Allison, A. C. (1960). *Biol. Revs.* **35**, 459.
- Haldane, J. S., and Priestley, J. G. (1935). "Respiration," New ed., 493 pp. Clarendon Press, Oxford.
- Harris, J. W. (1950). *Proc. Soc. Exptl. Biol. Med.* **75**, 197.
- Haurowitz, F. (1949). In "Haemoglobin" (P. J. W. Roughton and J. C. Kendrew, eds.), pp. 53-56. Interscience, New York.
- Hitzig, W. H., Frick, P. G., Betke, K., and Huisman, T. H. J. (1960). *Helv. Paediat. Acta* **15**, 499.
- Horowitz, N. H. (1945). *Proc. Natl. Acad. Sci. U.S.* **31**, 153.
- Huisman, T. H. J., van de Brande, J., and Meyering, C. A. (1960). *Clin. Chim. Acta* **5**, 375.
- Hunt, J. A. (1959). *Nature* **183**, 1373.
- Ingram, V. M. (1958). *Biochim. et Biophys. Acta* **28**, 539.
- Ingram, V. M. (1959). *Biochim. et Biophys. Acta* **36**, 402.
- Ingram, V. M. (1961a). "Hemoglobin and its Abnormalities," 153 pp. Charles C Thomas, Springfield, Illinois.
- Ingram, V. M. (1961b). *Nature* **189**, 704.
- Ingram, V. M., and Stretton, A. O. W. (1959). *Nature* **184**, 1903.
- Ingram, V. M., and Stretton, A. O. W. (1961). *Nature* **190**, 1079.
- Itano, H. A. (1953). *Am. J. Human Genet.* **5**, 34.
- Itano, H. A. (1956). *Ann. Rev. Biochem.* **25**, 331.
- Itano, H. A. (1957). *Advances in Protein Chem.* **12**, 215.
- Itano, H. A. (1959). In "Abnormal Haemoglobins" (J. H. P. Jonxis and J. F. Delafresnaye, eds.), pp. 1-17. Charles C Thomas, Springfield, Illinois.
- Itano, H. A., and Pauling, L. (1957). *Svensk Kem. Tidskr.* **69**, 509.
- Jones, R. T. (1961). Chromatographic and Chemical Studies of Some Abnormal Human Hemoglobins and Some Minor Hemoglobin Components. Ph.D. Thesis, California Institute of Technology.
- Jones, R. T., Schroeder, W. A., Balog, J. E., and Vinograd, J. R. (1959a). *J. Am. Chem. Soc.* **81**, 3161.

- Jones, R. T., Schroeder, W. A., and Vinograd, J. R. (1959b). *J. Amer. Chem. Soc.* **81**, 4749.
- Kaplan, N. O., Ciotti, M. M., Hamolsky, M., and Bieber, R. E. (1960). *Science* **131**, 392.
- Kleihauer, E., Braun, H., and Betke, K. (1957). *Klin. Wochschr.* **35**, 637.
- Konigsberg, W., Guidotti, G., and Hill, R. J. (1961). *J. Biol. Chem.* **236**, PC55.
- Kunkel, H. G., and Wallenius, G. (1955). *Science* **122**, 288.
- Kunkel, H. G., Cappellini, R., Muller-Eberhard, U., and Wolf, J. (1957). *J. Clin. Invest.* **36**, 1615.
- Lehmann, H. (1959). *Nature* **184**, 872.
- Lenhert, P. G., Love, W. E., and Carlson, F. D. (1956). *Biol. Bull.* **111**, 293.
- Lewis, E. B. (1951). *Cold Spring Harbor Symposia Quant. Biol.* **16**, 159.
- Lwoff, A. (1943). "L'évolution physiologique. Etude des pertes de fonctions chez les microorganismes," 308 pp. Hermann, Paris.
- McClintock, B. (1956). *Cold Spring Harbor Symposia Quant. Biol.* **21**, 197.
- Manwell, C. (1957). *Science* **126**, 1175.
- Manwell, C. (1958a). *Physiol. Zool.* **31**, 93.
- Manwell, C. (1958b). *J. Cellular Comp. Physiol.* **52**, 341.
- Manwell, C. (1960). *Comp. Biochem. Physiol.* **1**, 267.
- Margoliash, E., and Tuppy, H. (1960). Presented at the 138th Annual Meeting of the American Chemical Society, New York, September, 1960.
- Markert, C. L., and Møller, F. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 753.
- Metz, C. W. (1947). *Am. Naturalist* **81**, 81.
- Mizushima, S. I., and Shimanouchi, T. (1961). *Advances in Enzymol.* **23**, 1.
- Muller, C. J. (1961). A Comparative Study on the Structure of Mammalian and Avian Hemoglobins. Ph.D. Thesis, Groningen, Netherlands.
- Muller, C. J., and Kingma, S. (1961). *Biochim. et Biophys. Acta* **50**, 595.
- Muller, H. J. (1956). *Acta genet. statist. med.* **6**, 157.
- Muller, H. J. (1959). *Perspectives in Biol. and Med.* **3**, 1.
- Muller, H. J. (1961). *Science* **134**, 643.
- Neel, J. V. (1959). In "Abnormal Haemoglobins" (J. H. P. Jonxis and J. F. Delafresnaye, eds.), p. 158. Charles C Thomas, Springfield, Illinois.
- Paigen, K., and Noell, W. K. (1961). *Nature* **190**, 148.
- Pauling, L. (1956). *Am. J. Psychiat.* **113**, 492.
- Pauling, L. (1957a). In "Arbeiten aus dem Gebiet der Naturstoffchemie. Festschrift Arthur Stoll," p. 597. Birkhäuser, Basel.
- Pauling, L. (1957b). *Am. Inst. Biol. Sci. Publ. No.* **2**, 186.
- Pauling, L., Itano, H. A., Singer, S., and Wells, I. C. (1949). *Science* **110**, 543.
- Perutz, M. F., and Mitchison, J. M. (1950). *Nature* **166**, 677.
- Piveteau, J. (1955). In "Traité de Zoologie" (P. P. Grassé, ed.), vol. 17, p. 1. Masson, Paris.
- Rensch, B. (1954). "Neuere Probleme der Abstammungslehre. Die transspezifische Evolution," 2nd ed., 346 pp. Ferdinand Enke, Stuttgart.
- Rhinesmith, H. S., Schroeder, W. A., and Pauling, L. (1957). *J. Am. Chem. Soc.* **79**, 4682.
- Rhinesmith, H. S., Schroeder, W. A., and Martin, N. (1958). *J. Am. Chem. Soc.* **80**, 3358.
- Riegel, K., Bartels, H., and Schneider, J. (1959). *Z. Kinderheilk.* **83**, 209.

- Riggs, A., and Wells, M. (1961). *Biochim. Biophys. Acta* **50**, 243.
- Roche, J., and Fontaine, M. (1940). *Ann. inst. oceanogr.* **20**, 77.
- Rossi-Fanelli, A., and Antonini, E. (1956). *Arch. Biochem. Biophys.* **65**, 587.
- Rossi-Fanelli, A., Antonini, E., and Giuffrè, R. (1960). *Nature* **186**, 896.
- Rumen, N. M. (1959). *Acta Chem. Scand.* **13**, 1542.
- Ruud, J. T. (1954). *Nature* **173**, 848.
- Schmitt, F. O. (1956). *Proc. Natl. Acad. Sci. U.S.* **42**, 806.
- Schroeder, W. A., and Matsuda, G. (1958). *J. Am. Chem. Soc.* **80**, 1521.
- Schroeder, W. A., Jones, R. T., Shelton, J. R., Shelton, J. B., McCormick, J., and McCalla, K. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 811.
- Shelton, J. R., and Schroeder, W. A. (1960). *J. Am. Chem. Soc.* **82**, 3342.
- Siniscalco, M., Bernini, L., Latte, B., and Motulsky, A. G. (1961). *Nature* **190**, 1179.
- Smith, E. W., and Thorbert, J. V. (1958). *Bull. Johns Hopkins Hosp.* **101**, 38.
- Sturtevant, A. H. (1925). *Genetics* **10**, 117.
- Svedberg, T. (1933). *J. Biol. Chem.* **103**, 311.
- Thorup, O. A., Itano, H. A., Wheby, M., and Leavell, B. S. (1956). *Science* **123**, 889.
- Watson, H. C., and Kendrew, J. C. (1961). *Nature* **190**, 670.
- Wilson, S., and Dixon, G. H. (1961). *Nature* **191**, 876.
- Zuckerkandl, E., and Schroeder, W. A. (1961). *Nature* **192**, 984.
- Zuckerkandl, E., Jones, R. T., and Pauling, L. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1349.

Biochemical Catalysis

The Significance of Absolute Configuration in Optical Rotation and in Catalysis¹

HENRY EYRING, LEON L. JONES,² AND JOHN D. SPIKES

*Departments of Chemistry and Experimental Biology,
University of Utah, Salt Lake City, Utah*

I. Introduction	229
II. The Nature of Optical Rotation	231
III. A Consideration of Optical Rotation Using a Harmonic Oscillator Model	232
IV. Optical Rotation and Molecular Structure	237
V. The Conformation of the Condensation Products of Naphthylphenylaminomethane	239
VI. Optical Activity and Life	240
A. Introduction	240
B. Prebiological Formation of Organic Compounds	242
C. Prebiological Evolution	243
D. Origin of Life	245
E. The Selective Step for L-Amino Acids in Protein Synthesis	245
F. Application of Rate Theory to the Origin of Life	247
G. Summary	248
References	249

I. Introduction

The importance of optically active molecules in biological processes was first shown by Pasteur, who noticed that only one of the isomers of tartaric acid was produced by common plants. An ordinary chemical synthesis, on the other hand, would produce equal numbers of the D- and L-forms of a given compound. Pasteur also showed that the mold *Penicillium glaucum* when fed on a racemic mixture of tartaric acid selected only the dextrorotating isomer for metabolism. Since that time virtually all biochemical processes have been shown to be specific for one of the two possible optical isomers.

¹ The preparation of this paper was supported in part by the National Science Foundation (Grant G-9648), and the Division of Biology and Medicine, U.S. Atomic Energy Commission, under Contract No. AT(11-1)-875.

² Present address: Department of Chemistry, University of Wisconsin, Madison, Wisconsin.

The specificity of enzymes for particular optical isomers of sugars led Emil Fischer in 1894 to propose a "lock and key" simile for enzyme reactions in which a reactant molecule must fit into a particular site on the enzyme. This specificity implies that there must be at least a three-point interaction between the molecule and the enzyme since it is necessary to locate the position of three different groups to specify the configuration of an optically active molecule. This three-point interaction may just be a hydrogen bonding of two groups and a space requirement for a third group, as in the selectivity of the L-amino acids where the —COOH and —NH₂ groups hydrogen bond to the enzyme surface and leave a space for a third group which is only large enough to contain a hydrogen. Such a scheme provides a complete exclusion of the corresponding D-amino acid. It is thus apparent that studies of optical activity help in the detailed understanding of the role of configuration in catalysis.

The measurement of optical activity can be used in kinetic studies. The first quantitative study of a chemical reaction was Wilhelmy's (1850) polarimetric investigation of the inversion of sucrose by acids. Complex reactions leading to the formation of a chromophore on an optically active product can be followed by using incident light at a wavelength near an absorption wavelength of the chromophore. The rotation for this product may change on the order of thousands of degrees while formation of other products not involving the chromophore are apt to produce changes of only a small fraction of this amount.

Perhaps the most interesting application of the optical rotatory dispersion technique has been in the examination of the conformation of polypeptides and proteins (Blout, 1960). A typical use is that of determining the secondary structure of helical polypeptides. The optical rotation for a helical polypeptide is composed of a residue rotation for the individual units along the chain and a rotation for the helix (for a particular sense of the helix). For a polypeptide which is entirely a right-handed helical conformation there is a value of the helical contribution which is independent of the nature of the groups which compose the helix. This value changes sign for a polypeptide in a left-handed helix. For a polypeptide which is only partially in the helical conformation, the value of the helical contribution is proportionally less, attaining a zero value when the polypeptide is completely in the random configuration. The importance of optical rotatory dispersion techniques is indicated by the

recent international symposium on their theory and application, the proceedings of which were published as Volume 13 of *Tetrahedron* (1961).

Optical rotatory dispersion can be used to determine the absolute configuration around a chromophoric group in such molecules as the steroids (Djerassi, 1960 and 1961).

In the case of substituted cyclohexanones or the decalones, where more than one ring conformation is possible, optical rotation coupled with the octant rule provides a rapid method of determining the ring conformation. Tartaric acid can exist in two intramolecularly hydrogen-bonded forms, one forming a five-membered ring and the other a six-membered ring structure which is dextrorotatory. The dependence of the optical rotation of diethyl tartrate on solvents and temperature shows that these rings are formed in different solvents. In some solvents the rotation is dextro and in others it is levorotatory. The rotation in all solvents approaches a limiting value at higher temperatures which may be identified as the optical rotation of the non-hydrogen-bonded structure with free internal molecular rotation.

The remainder of this paper will be devoted to a discussion of two major topics. First, we will examine the relation between the optical activity of molecules and their absolute configuration, using a harmonic oscillator model. Then we will discuss the intriguing problem of the almost universal preference in living organisms for one of the two possible optically active isomers of compounds. In order to do this it is necessary to discuss briefly the origin and evolution of living material.

II. The Nature of Optical Rotation

The question immediately arises as to the actual mechanism by which an asymmetric molecule rotates the plane of a beam of polarized light. We envision that when an electron in a molecule is raised to a higher energy level by the absorption of a quantum, a hole is left. Electrons from other parts of the molecule then flow toward the hole. If the molecule is asymmetric the electrons will be given a whirling motion as they move into the hole. The direction of rotation of the whirling electrons depends on the screw-sense of the molecular asymmetry, i.e., the direction of rotation will be one way with one isomer and in the opposite direction with the other isomer. The whirling motion of the charged particles (electrons) moving

into the hole adds an electromagnetic component to the light transmitted by the molecule. This results in a rotation of the plane of the transmitted light in one direction or the other, depending on the molecular asymmetry.

The magnitude of the optical rotation produced by a molecule is dependent on the wavelength of the incident light. This dependence can be expressed by the Drude equation

$$\phi = \sum_i \frac{R_i}{\lambda_i^2 - \lambda^2} \quad (1)$$

where λ is the wavelength of the incident light, λ_i is the wavelength corresponding to an electronic transition in the molecule, and R_i is a parameter, either positive or negative, whose magnitude depends on the electronic transition and the configuration of the molecule in the neighborhood of the group in which the transition occurs. The group in which the transition occurs is called the chromophoric group. There will also be an absorption band of the molecule at the wavelength λ_i . This absorption band is a characteristic of the chromophoric group and occurs whether the chromophore is in an asymmetric environment or not. Thus, the absorption band and the associated rotation band can be characterized as to the particular electronic process, either an electric dipole or a magnetic dipole transition, which exhibit strong and weak absorption, respectively.

In order to describe this phenomenon in modern physical-chemical terms, we are obliged to use the somewhat austere language of quantum mechanics. While recognizing that most biologists will not follow the argument in detail, we hope that the following brief treatment will indicate the lines of reasoning involved.

III. A Consideration of Optical Rotation Using a Harmonic Oscillator Model

A quantum mechanical expression for optical rotation was derived by Rosenfeld (1928). A convenient form of this equation is given by Eyring *et al.* (1944) where ϕ is the rotation in radians per centimeter and the transitions are from the ground state, designated by the eigenfunction ψ_0 , to the excited states, designated by the function ψ_i ,

$$\phi = \frac{16\pi^2 N_i}{3ch} \left(\frac{n^2 + 2}{3} \right) \sum_i \frac{R_{i0}\nu^2}{\nu_{i0}^2 - \nu^2}. \quad (2)$$

Here R_{i0} is the imaginary part of the scalar product of the electric and magnetic transition moments for the transition $\psi_0 \rightarrow \psi_i$

$$R_{i0} = \text{Im} \{ (\psi_0|R|\psi_i) \cdot (\psi_i|M|\psi_0) \}. \quad (3)$$

N_i is the number of active molecules per cubic centimeter, n the refractive index of the solution, ν_{i0} the frequency corresponding to the transition from state $\psi_0 \rightarrow \psi_i$ and ν the frequency of the incident light.

A straight-forward calculation of the rotation is impossible since the eigenfunctions describing the molecules are not known. It is necessary to employ a model for optically active molecules. The simplest model for optical rotation is the anisotropic, asymmetric, harmonic oscillator model of Condon *et al.* (1937). They used the potential

$$V = \frac{1}{2}k_1x^2 + \frac{1}{2}k_2y^2 + \frac{1}{2}k_3z^2 + Axzy. \quad (4)$$

The harmonic oscillator part was used to obtain the initial eigenfunctions, and the term $Axyz$ was treated as a perturbation. They obtained for the rotation when the harmonic oscillator is in an initial state characterized by the quantum numbers (n_1, n_2, n_3) for oscillators in the x , y , and z directions, respectively,

$$\begin{aligned} \phi = & [(n^2 + 2)NAhe^2/(144\pi^3\mu^3)] \\ & \{(1/\nu_2 - 1/\nu_1)(n_1 + n_2 + 1)\{[(\nu_1 + \nu_2)^2 - \nu^2](\nu_3^2 - \nu^2)\}^{-1} \\ & + (1/\nu_1 + 1/\nu_2)(n_2 - n_1)\{[(\nu_1 - \nu_2)^2 - \nu^2](\nu_3^2 - \nu^2)\}^{-1}\}. \end{aligned} \quad (5)$$

+ terms obtained by replacing subscripts (1,2,3) by (3,1,2) and by (2,3,1). Here μ represents the reduced mass of the oscillators, which are electrons, and

$$\nu_i = \frac{1}{2\pi} \left(\frac{k_i}{\mu} \right)^{1/2}$$

are the frequencies corresponding to the transitions in each of the principal directions. It can be seen that when the frequencies for the transitions in two of the principal directions are the same, the rotation vanishes.

This model is somewhat artificial as there are no chromophores which follow the potential of Eq. (4); however, one can write the actual eigenfunctions in terms of the harmonic oscillator eigenfunctions with a choice of the quantum numbers n_1, n_2, n_3 and the frequencies ν_1, ν_2 , and ν_3 such that the harmonic oscillator eigen-

functions give the correct energy levels of the chromophore. The interaction between the chromophore and the rest of the molecule can likewise be expanded in a series, in which case the first term which has the proper symmetry for optical rotation will be of the form $Axyz$, where A is a function of the distance to the neighboring groups and of the particular interaction.

To apply this model, it is convenient to choose the coordinates such that the z direction is the most polarizable, that is, such that the transition in the z direction has the lowest energy. The y axis will be chosen as the next most polarizable direction, and the x direction will be chosen as the least polarizable direction such that the coordinates x , y , and z form a right-handed coordinate system. This choice leads to the relation between the frequencies $\nu_1 > \nu_2 > \nu_3$. If we concern ourselves with the lowest frequency transition in the chromophore, we have two cases. When the transition is electric dipole-allowed we have a transition in the direction from a state which may be characterized by the harmonic oscillator in its ground state, $(0,0,0)$, to a state with an excitation in the z direction, or $(0,0,1)$. The other possible transition is a magnetic dipole-allowed transition and will be from a state which may be characterized by an excitation in the y direction, $(0,1,0)$, to a state with an excitation in the x direction $(0,0,1)$. This corresponds to a transition of frequency $(\nu_1 - \nu_2)$. This will be the most important magnetic transition since the chromophore will usually exhibit approximate cylindrical symmetry about the z axis and the transition frequencies ν_1 and ν_2 will be approximately equal, making the difference $(\nu_1 - \nu_2)$ small.

For these two cases we need only consider the terms which have the energy denominators $(\nu_3^2 - \nu^2)$ and $(\nu_1 - \nu_2)^2 - \nu^2$ or the terms shown explicitly in Eq. (5). Since the only values of the n 's which must be considered are $n_1 = 0$, $n_2 = 0$ or 1 , $n_3 = 0$, and with the choice of frequencies $\nu_1 > \nu_2 > \nu_3$ we can see from Eq. (5) that the rotation, when the frequency of the incident light is smaller than that of the smallest frequency transition, $\nu < \nu_3$ and $\nu < (\nu_1 - \nu_2)$, will follow the sign of the parameter A , the combination of terms on the right side of Eq. (5), with the exception of A , will be inherently positive.

The perturbing term $Axyz$ will be made up of contributions from all groups lying off the principal polarizability axes. The contribution to the term $Axyz$ for a neighboring group will be negative if the potential on the chromophoric electron is lower in the direction of

the group than it would be if the neighboring group were replaced in such a way as to introduce a plane of symmetry in the chromophoric group. Such a group will be more polarizable than the group which would make the chromophore symmetric or else will carry a greater net positive charge. If the neighboring group has a net negative charge, its contribution to $Axyz$ will be positive, corresponding to a high potential for the chromophoric electron to move in that direction.

As an example, we will consider the contribution to the optical rotation of substituted cyclohexanones from the carbonyl transition which occurs at about $290 \text{ m}\mu$ wavelength. Figure 1A shows a cyclohexanone ring and Fig. 1B shows the ring as it would appear when viewed from the oxygen along the $C=O$ axis. The figure is drawn with the assumption that all angles are tetrahedral. We consider the cyclohexanone skeleton locked in this configuration by side groups, as in bicyclic systems. The z direction is directed from the oxygen toward the adjacent carbon. The next most polarizable direction, the y direction, is perpendicular to z and is in the plane containing the four atoms oxygen, C1, CR2, and CL2 and is directed toward CR2. The x direction will then point toward the top of the page as drawn in Fig. 1B. The transition is magnetic dipole-allowed in the carbonyl group and is from a nonbonding state, primarily $2py$ on the oxygen atom, to the anti-bonding molecular orbital, π_x , located on the oxygen and carbon atoms. We can see that this satisfies the conditions for locating the polarizability ellipsoid as described above.

Consider next the contribution to the optical rotation which arises when substituents are placed at various equatorial and axial positions. It is convenient to assume that all these positions initially contained hydrogen in which case the molecule would show no optical rotation. The contribution to the parameter A will then arise from the difference in properties between the substituted group and

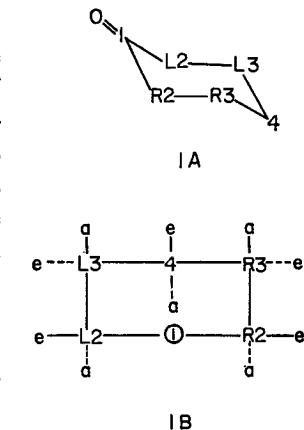


FIG. 1. A: Cyclohexanone ring; the z direction is directed from the oxygen atom toward the adjacent carbon atom C1. B: Cyclohexanone ring viewed along the z axis from the oxygen atom; e and a denote equatorial and axial substituents.

the hydrogen atom. A group substituted at R_{2e} will lie approximately on the plane defined by the *y* and *z* axes and will not contribute to the rotation. A group substituted at R_{2a} will lie in the octant characterized by the coordinate signs, *x* = −, *y* = +, *z* = +, so that the product of the three coordinate signs is negative. When the polarizability of the substituted group is larger than the polarizability of hydrogen, as in the case of a bromine, chlorine, or methyl group, electrons from the vicinal group will flow correspondingly

TABLE I
SIGNS OF THE CONTRIBUTIONS OF VICINAL GROUPS (MORE POLARIZABLE THAN HYDROGEN) TO ROTATION OF THE CHROMOPHORIC CARBONYL GROUP IN SUBSTITUTED CYCLOHEXANONES

Position	Substituent ^a	Octant			Sign of contribution to <i>A</i>
		<i>x</i>	<i>y</i>	<i>z</i>	
R ₂	a	−	+	+	+
	e	0	+	+	0
R ₃	a & e	+	+	+	−
	4	+	0	+	0
L ₃	a & e	+	−	+	+
	L ₂	−	−	+	−
L ₂	a	0	−	+	0

^a In the substituent column, a and e refer to the axial or equatorial position, respectively.

more into the positive hole left by the excited electron of the chromophore. This extends the resultant many-electron hole into the vicinal group just as though the original chromophoric electron had so extended because of a negative value for the potential term *Axyz* at the locus of the vicinal group. For these substituents at R_{2a} the term must accordingly be positive and the contribution to the rotation by this group will be positive. For the case of a substituted fluorine, since its polarizability is smaller than that of hydrogen, the term *Axyz* must be positive in that quadrant and hence, the value of *A* must be negative. A similar analysis of the other positions around the cyclohexanone ring leads to the information shown in Table I for the contribution to the rotation in cases where the substituent is more polarizable than hydrogen, such as bromine, chlo-

rine, or methyl. For substituents which are less polarizable than hydrogen, such as fluorine, the values in the table for the sign of contribution to *A* will all be of opposite sign. The magnitude of the contribution to the rotation will be approximately proportional to the difference in polarizability between the substituent and hydrogen. These findings provide the basis for the "octant rule" which was proposed by Djerassi and co-workers on the basis of a large number of rotatory dispersion measurements taken on substituted steroids (Djerassi, 1960).

This method may be further illustrated by its application to lactic acid. The chromophore in this case is the carboxyl group, and its rotation would be expected to determine the rotation in the visible spectrum. If the oxygen atom of the carboxyl group is hydrogen-bonded to the hydroxide as shown in Fig. 2, the *z* axis will lie along the C=O bond, chosen positive in the direction of the carbon atom. The second most polarizable axis, *y*, will lie in the plane of the ring, chosen positive in the direction of the hydroxyl hydrogen. The *x* axis then points out of the plane of the paper. The most polarizable group lying out of the plane of the ring is the methyl group, which is in the octant, *x* = +, *y* = +, *z* = +. Since *Axyz* must be negative, the rotation will then be negative as it is known to be for D-lactic acid.

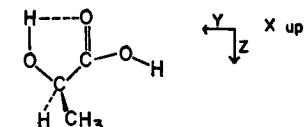


FIG. 2. Structural formula for D-lactic acid with the axis designations *x*, *y*, and *z* as discussed in the text.

IV. Optical Rotation and Molecular Structure

Perhaps even more impressive than the use of optical rotation in studying absolute configuration is its utility in studying the conformation of molecules. The one-electron theory was developed in a series of papers starting with Condon *et al.* (1937), while the latest in this series is a paper by Kwok and Eyring (1950). Except in unusual circumstances the many-electron or coupled oscillator effect on optical rotation overshadows the effect due to one electron so that it is important to discuss the estimation of the effective value of *A* in Eq. (5) for the case of many electrons.

As stated earlier, the virtual excitation of an electron by light falling on a chromophoric group causes vicinal electrons to surge into the positive hole in the same way that substituents feed elec-

trons into or remove them from a molecule by inductive or resonance effects as demonstrated by changes in ionization or reaction rate constants. Thus one expects for the multiple electron effective contribution, A'' , made by vicinal groups to $A = A' + A''$ of Eq. (5) the expression

$$A'' = \sum l_i m_i n_i \rho_i \sigma_i. \quad (6)$$

This same form of equation applies also for the one-electron part A' . In Eq. (6) the subscript i refers to the i th vicinal group and the summation is over all such groups; l_i , m_i , and n_i are the direction cosines of the vector joining the center of the vicinal group to the origin of coordinates which lies at the equilibrium position of the chromophoric electron. The minus sign in Eq. (6) applies to groups more polarizable than hydrogen with plus for less polarizable groups. As explained earlier, the right-hand coordinate system is chosen with z in the most polarizable direction, y the next most and x the least polarizable direction for the chromophoric electron; σ_i as in Hammett's (1940) relation is characteristic of the vicinal group and closely related to the polarizability; ρ_i has to do with the molecular bond structure which channels the electrons from the vicinal to the chromophoric group and with the effectiveness of the chromophoric group as a pump of vicinal electrons as the chromophore is excited. It will be convenient to express the energy, $\rho_i \sigma_i$, in absolute units instead of in multiples of kT as is usually done in the applications to ionization and to reaction kinetics (Hammett, 1940).

If a molecule changes its conformation because of change in the solvent or of the temperature there will be corresponding changes in the various products ($l_i m_i n_i$) of the direction cosines. If in addition the vicinal group moves nearer to the chromophore there will be a corresponding change in ρ_i due to the easier passage of electrons across the intervening space, but little change in the channeling of electrons along the bonds. Another important point which the present approach to optical rotation emphasizes is the advantages to be gained by breaking up the separate contributions of the various chromophoric groups through comparative studies of molecules with an orderly variation in structure and also by analyzing separately the individual Drude terms in their optical rotation. Properly used, optical rotation is one of the sharpest tools we have for establishing absolute configuration and likewise for delineating changes of conformation.

V. The Conformation of the Condensation Products of Naphthylphenylaminomethane

The theory just outlined makes it possible to rationalize the data in Table II from a paper by Betti (1930). In Fig. 3A the structure of substituted naphthylphenylaminomethane is given and in Fig. 3B the corresponding acid is drawn. In Table II in the first

TABLE II
MOLAR OPTICAL ROTATORY POWER, $[M]_D$, OF A SERIES OF ALDEHYDIC DERIVATIVES OF β -NAPHTHYLPHENYLAMINOMETHANE AS RELATED TO THE IONIZATION CONSTANTS OF THE ACIDS CORRESPONDING TO EACH SUBSTITUENT ALDEHYDE^a

$[M]_D$ of the aldehydic compounds	Aldehyde substituent	$K \times 10^5$ at 25° of the corre- sponding acids
+2676.0°	p-Dimethylaminobenzoic	0.94
+1049.5°	p-Oxybenzoic	2.9
+648.0°	3-Bromo-p-oxybenzoic	—
+588.8°	Protocatechuic	3.3
+559.6°	3-Nitroanisic	—
+504.5°	m-Tolue	5.6
+373.1°	Benzoe	6.6
+362.6°	m-Oxybenzoic	8.33
+311.8°	p-Chlorobenzoic	9.3
+280.9°	m-Bromobenzoic	13.7
+255.9°	m-Chlorobenzoic	15.5
+167.6°	m-Nitrobenzoic	34.8
−85.7°	Salicylic	106
−128.4°	o-Chlorobenzoic	132
−308.2°	o-Bromobenzoic	145
−990.7°	o-Nitrobenzoic	657

^a Data were taken from the paper of Betti (1930).

column is given the optical rotation of the substituted naphthylphenylaminomethane, $C_6H_5CH(C_{10}H_8OH)N=HC-R$, while the corresponding aldehyde, $RCOH$, is listed in the second column. The last column contains the dissociation constant, K , of the corresponding acid, $RCOOH$. Now if the group $N=C$ is the chromophore with the most polarizable direction of its polarizability ellipsoid lying about as indicated by the z axis then the *ortho* substituent X should

have a different sign for the product of its direction cosines (l, m, n_i) occurring in Eq. (6) than if X is in the *meta* or *para* positions indicated by *m* and *p*, respectively, in Fig. 3A. This fact explains the difference in the sign of rotation of the *ortho* as compared with the *meta* and *para* compounds. The expected absolute values of (l, m, n_i) should be largest for *para* and least for *ortho* which explains the observed slopes of $\log K$, plotted against optical rotation, $[M]_D$, as shown on page 395 of the review by Kauzmann *et al.* (1940) where one will find many other examples which can be understood in terms of Eq. (6). Rotation of the phenyl group, on which X is substituted, about the axis of the group $-\text{N}=\text{C}$ will leave all our qualitative conclusions unchanged but will explain deviations from the straight

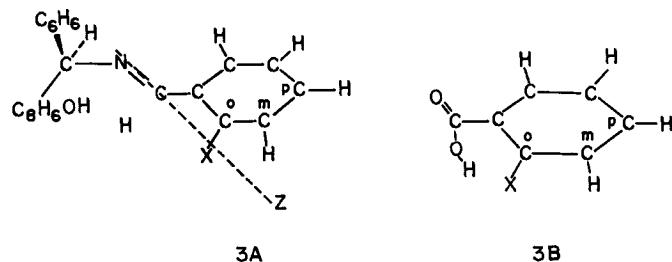


FIG. 3. A: Structural formula for a substituted β -naphthylphenylaminomethane showing the *z* axis as discussed in the text. B: Structural formula for the acid corresponding to the substituted groupings in Fig. 3A.

line relationship for optical rotation versus $\log K$. Various other changes in conformation can complicate this model. While the molecular conformation suggested is speculative, it or something very like it is required to explain the observed facts. The magnitude of ρ, σ_i in optical rotation can also be measured by appropriate dipole moments as has been emphasized by Rule (1930) and as explained by Ree and Eyring (1940) for rates of reaction.

VI. Optical Activity and Life

A. INTRODUCTION

In the preceding section we have discussed the atomic-molecular basis of optical activity. A large number of the organic molecules making up living material are optically active. One of the most fascinating problems related to the absolute configuration of molecules is the fact that, generally speaking, living organisms do not produce a

racemic mixture of a particular compound, but only one of the isomers, either the *D*-form or the *L*-form. For example, the proteins of all terrestrial life forms are made up of approximately the same twenty amino acids, and in all cases only the *L*-isomers are involved. It is apparent that if a single amino acid in the protein backbone were replaced with the *D*-isomer, the configuration of the protein molecule would be altered significantly. Certain *D*-amino acids are found in organisms, of course; however, these have never been found in proteins. For example, several of the polypeptide antibiotics (gramicidin, bacitracin, etc.) contain *D*-amino acids, and the capsular and cell wall materials of several kinds of microorganisms contain *D*-glutamic acid and *D*-alanine.

In the case of the sugars, most of which are optically active, only one of the isomers will generally occur in living material, although this case is not so clear-cut as that of the amino acids in protein. In the pentoses we find only *D*-ribose or *D*-deoxyribose and *D*-xylose; however, both *D*- and *L*-arabinose occur. With the hexoses, only *D*-glucose and mannose occur in nature, while galactose occurs in both forms.

Deoxyribonucleic acid (DNA) is the cellular material primarily concerned with the transfer of genetic information at cell division. The Watson-Crick model for DNA consists of two polynucleotide strands arranged in a helical structure. Each strand consists of a chain structure of alternating phosphate groups and *D*-deoxyribose with a basic (purine or pyrimidine) side group attached to each sugar molecule. The stability and symmetry of the double helix results from hydrogen bonding between the base pairs adenine-thymine and guanine-cytosine. It is apparent that a uniform spiral could not be produced if even a single *L*-deoxyribose were introduced into one of the chains. Thus, again we see the absolute requirement for isomers of a single type in biological systems. Ribonucleic acid similarly contains only *D*-ribose.

To us, one of the most intriguing problems in biology is the universality, from viruses to man, of the *L*-amino acids in proteins and of *D*-ribose and deoxyribose in the nucleic acids. Two major questions may be asked: (a) how were these particular isomers selected originally, and (b) what is the basic relationship, if any, between the protein helical conformation based on a chain structure composed of *L*-amino acids and the nucleic acid helical structure based on a chain composed of alternating *D*-pentose and phosphate residues. We usu-

ally assume that organisms adapt to the environment through the process of mutation followed by survival of the fittest. This mechanism does not account for the universal presence of L-amino acids in proteins and D-sugars in the nucleic acids since life based on the mirror images of these molecules should show equal fitness for survival. Oddly enough, with few exceptions (Rush, 1957; Blum, 1957), recent discussions of the origin and evolution of life devote little attention to this problem, which to us seems so basic (Oparin, 1957; Calvin, 1956, 1959, 1961a; Fox, 1956; Barghoorn, 1957; Sagan, 1961). In order to understand the selections that have been made for optical activity in living organisms, we must look back through evolutionary history to the origin of life. In present living material all optically active molecules appeared to be synthesized using corresponding optically active templates. We must assume, therefore, that the present templates had their remote origin in a single optically active template which was in some way sufficiently superior so that it eliminated all competitors and covered the world with replicas of itself. In order to examine this situation more carefully, we will now briefly discuss the prebiological formation of organic materials and the origin of life.

B. PREBIOLOGICAL FORMATION OF ORGANIC COMPOUNDS

It can be estimated (Sagan, 1961) that the earth's mantle became stabilized roughly 4.5 billion years ago, thus permitting the initial processes necessary for the origin of life. It probably required relatively little time for the earth's surface to become well stocked with organic compounds produced by nonbiological reactions. There are a number of ways in which these compounds could have been produced. As Lederberg and Hartline (1960) have pointed out, the primitive condensation of free atoms to form the stars produced tremendous quantities of organic materials, since hydrogen, carbon, oxygen and nitrogen have a high cosmic abundance. To what extent the planets still retain organic compounds formed in this way is not known. It may be pointed out, however, that at least some meteorites contain organic compounds and that the spectra of comets indicate the presence of organic groupings (Calvin, 1961b). Another postulated prebiological source of organic material was the production of olefins by the reactions of metallic carbides with water (see Barghoorn, 1957, however) and the subsequent reactions of the olefins with ammonia and water.

Most of the organic compounds probably resulted, however, from the interactions of ionizing radiation, ultraviolet radiation, electrical discharges, and heat with the primitive atmosphere of the earth. This field has been reviewed most recently by Sagan (1961). As early as 1950, Calvin subjected carbon dioxide-water mixtures to cyclotron irradiation and observed the production of formic acid, acetic acid, and other reduced carbon compounds. The primitive atmosphere itself was probably in a reduced state consisting of ammonia, methane, water vapor, etc. (probably becoming "oxidized" much later through the activities of photosynthetic organisms). By passing electric discharges through such a mixture, Miller (1957) produced a variety of organic compounds, including the biologically important amino acids glycine and alanine. Fox and co-workers (see Harada and Fox, 1960) have studied the thermal copolymerization of a wide variety of amino acids in mixtures. Linear peptides can be obtained from the products of such reactions. Fox and Harada (1961) demonstrated the thermal synthesis of uracil under conditions similar to the prebiological environment. Oró and Kamat (1961) found that heating HCN-ammonia mixtures even at temperatures below 100°C produced a variety of simpler organic compounds plus polymeric materials. In particular, they reported the production of the purine adenine plus two imidazoles. It appears reasonable, then, given sufficient time, that one might expect the prebiological formation of all of the amino acids required for protein synthesis, of protein-like materials, of the purine and pyrimidine bases necessary for nucleic acid formation, of the pyrroles used for porphyrin manufacture, and of a variety of carbohydrates. Further, the quantities of such compounds produced were probably large; Urey (1952) estimated that the primitive seas were perhaps 1% solutions of organic materials (also see Sagan, 1961).

It should be stressed that this primeval organic world was a racemic world, since none of the processes described above would be expected to favor the production of one isomeric form of a molecule over the other. This primeval world, however, permitted an essential period of what might be termed "prebiological evolution."

C. PREBIOLOGICAL EVOLUTION

As Rush (1957) has pointed out, biologists tend to think of evolution by means of natural selection solely in terms of living organisms. In a very real sense, however, even simple molecules can be

"selected for survival" in a given physical-chemical environment. For example, in an environment containing very low concentrations of chloride ions, the population of silver ions would be relatively large; in the presence of high concentrations of halide ions, the silver ion concentration would be very low. In the primeval environment the array of organic compounds became more and more complex since the activation energy provided by heat and radiation would permit interactions between the simpler molecules. Some of these molecules would be more stable in a particular environment than others and would thus accumulate. In particular, many types of molecules would collect at specialized sites such as at interfaces, on finely divided clay particles and in coacervates. Such sites may well have provided the environment necessary for the evolution of the more complex systems leading to life itself.

Calvin (1959) has recently discussed the importance of autocatalytic processes in the selective survival of particular molecules during this period of chemical evolution. If a given reaction in a series of competitive processes were autocatalytic, then in a short time it would probably dominate all of the simpler processes competing for the same pool of precursor molecules. Calvin (1959) has also pointed out that present enzymatic catalysts have their counterparts in simple systems. Ferric ions in water solution, for example, catalyze the dismutation of hydrogen peroxide. If the ferric ion is surrounded by a porphyrin ring, as in heme, the catalytic effectiveness increases a thousand times. If the iron porphyrin is combined with protein to form the enzyme catalase, the catalytic activity is further increased by a factor of 10^7 . Thus, in any system where peroxide breakdown is advantageous, iron porphyrins would have a much greater selective advantage than ferric ion, and catalase even more so.

In summary, then, it is reasonable to assume that a long period of prebiological evolution occurred during which a stupendous array of molecules and molecular aggregates were formed, altered, and reformed. This period was probably essential for the ultimate appearance of replicating structures. Gaffron (1957) has pointed out that the environmental conditions described above which originally permitted the appearance of life have probably been irreversibly lost. Environmental alterations include the disappearance of hydrogen gas from the atmosphere, the blocking of short-wavelength ultraviolet penetration to the earth's surface, the appearance of molecu-

lar oxygen in high concentration, and the oxidative and metabolic destruction of the primeval organic materials. Thus, if all present life on earth were wiped out (as by radioactive warfare), new life would have to evolve through entirely different developmental lines and against an initial strongly oxidizing environment.

D. ORIGIN OF LIFE

We usually regard "life" as originating with the first appearance in the primeval world of a reasonably stable, self-replicating, mutable molecular arrangement. Once this occurred we could have, in principle at least, natural selection and evolution. Since nucleic acids are involved in all of the replicating processes in present living material, it is customary to assume that the original "life" form was a nucleic acid of some type.

During the era of prebiological or chemical evolution described above, rudimentary enzymes and nucleic acids were probably continuously forming, breaking down, and reforming. Nucleic acid self-replication was so inefficient that possible template molecules would break down before reproducing themselves. Ever present in this situation, however, was the possibility of an explosive chain-like reaction which finally occurred at the instant a template molecule formed which could produce more than one replica of itself before it was degraded. This event signaled the end of inefficient, random chemical evolution and initiated a self-replicating world which would permit biological evolution.

This step would also mark the end of a world of racemic chemistry, since with the evolution of more and more efficient templates, more and more of the organic compounds of the world would be synthesized with a specific configuration. It is most likely, then, that the selection for L-amino acid proteins and D-pentose nucleic acids occurred at this time. In summary, then, we feel that once a really efficient, optically active template with the ability to direct the synthesis of cooperating enzymes first appeared, it would have such a great selective advantage that its optical isomer would never have the chance to evolve.

E. THE SELECTIVE STEP FOR L-AMINO ACIDS IN PROTEIN SYNTHESIS

Protein synthesis in contemporary organisms does not occur through simple polymerization of mixed amino acids, although

protein-like materials can be produced by this technique *in vitro* (Harada and Fox, 1960). The question naturally arises as to where the selective step or steps for L-amino acids in protein synthesis do occur in protoplasm. We can now speculate to some extent concerning this problem on the basis of the large amount of very elegant work on protein synthesis which has appeared during the past few years (see Webster, 1961).

In a living cell the synthesis of a single type of protein molecule requires a rather staggering array of subcellular machinery, including some twenty different specific amino acid-activating enzymes; approximately twenty different specific kinds of soluble, low molecular weight ribonucleic acid (sRNA); and a template of some sort to control the sequence of amino acids in the protein. The templates appear to be large RNA molecules present in the subcellular particles termed ribosomes, but originally synthesized in the nucleus under the direction of DNA. An additional array of enzymes is also probably required for peptide bond synthesis, for releasing the completed protein molecule from the ribosomal template, and for reactivation of the sRNA. Further, energy in the form of adenosine triphosphate (ATP) must be made available at appropriate steps in the sequence.

The synthetic process probably proceeds as follows: First an amino acid is "activated" by reaction with ATP to give the aminoacyl adenylate. This step requires an enzyme, and there seems to be a different enzyme for each kind of amino acid. The aminoacyl adenylate, which apparently remains attached to the activating enzyme, is then transferred to sRNA (variously called soluble RNA, transfer RNA, or acceptor RNA). In the next step, each sRNA-amino acid binds to a specific site on the ribosomal RNA template. The amino acids would thus be lined up in a sequence determined by the template. Presumably, then, an array of enzymes assembles the lined up amino acids into the corresponding protein and then separates it from the sRNA.

The selective process for L-amino acids thus appears to occur at the activation step, since the activating enzymes are specific for the L-amino acids. This particular point of selection has probably arisen subsequent to the origin of life. It is not known, of course, what kind of molecule made the first selection of a particular isomer in the origin of life. We tend to assume that it must have been something analogous to DNA because of the role this compound plays in the determination of enzyme configuration in modern organisms.

It could, however, have been a RNA-like or a protein-like material. Thus we are in no position to say definitely which came first, nucleic acids or proteins.

F. APPLICATION OF RATE THEORY TO THE ORIGIN OF LIFE

One of the questions which always arises in discussions of the origin of life is the time factor, i.e., was there enough time available between the stabilization of the earth's mantle and the appearance of life to have permitted its origin by the processes described above. Further, it has been argued that compounds of the complexity necessary to permit self-replication could not possibly arise by chance. These questions have been examined recently by Eyring and Johnson (1957) using the theory of absolute reaction rates (Johnson *et al.*, 1954). Rate theory permits writing the following equation for any chemical reaction

$$dc/dt = c_1 c_2 \cdots c_n \kappa (kT/h) \exp - (\Delta F^\ddagger / RT) \quad (7)$$

In the present case dc/dt would represent the most probable rate of formation of the critical complex leading to the production of the first successful self-replicating molecule, where c_1 to c_n are the concentrations of the various reactants involved and κ is the transmission coefficient with a value of one. The expression kT/h is the frequency of reaction of the activated complex with k , T , and h representing the Boltzmann constant, the absolute temperature, and Planck's constant, respectively; ΔF^\ddagger is the free energy of activation and R is the gas constant.

Equation (7) can be put in a more useful form for our discussion by converting from concentrations to the total number of reactants of the type n , as follows:

$$dn/dt = n_1 c_2 \cdots c_n (kT/h) \exp - (\Delta F^\ddagger / RT) \quad (8)$$

This equation gives the number of critical complexes formed per second, dn/dt . In order to apply Eq. (8) we must make a few somewhat arbitrary assumptions as follows: all concentrations will be estimated as 1/1000 molal; the temperature selected will be 300° absolute; in the production of the critical complex we assume that m reactants will join in a linear manner; $m - 2$ of the bonds thus formed will be assumed to contribute factors of 1/100th each to K^\ddagger , the equilibrium constant for the formation of the activated complex; the final bond will be assumed to contribute the factor $e^{-(30,000/RT)}$ to

$K \ddagger$ as it passes over the potential barrier corresponding to the activated complex. The factor of 1/100th assigned per bond provides a contribution of +2.6 kcal to $\Delta F \ddagger$, a value corresponding approximately to peptide bonds. We can now estimate the total number of reactant molecules of the type n available by assuming enough pools to average 10^6 liters per square mile over the whole surface of the earth. This comes out to be approximately 1.2×10^{35} molecules. If we now assume that our critical complex is formed only once per billion years, Eq. (8) gives us a value of $m = 10$ for the number of reactants combining to form the critical complex. This may be regarded as a reasonable minimal size for a primeval self-replicating molecule.

G. SUMMARY

The recent expansion of man's activities into space, and the reasonable possibility of obtaining first-hand information on other planets within the next decade or so has stimulated new interest in the possibilities of extraterrestrial life. Along with this there has been a considerable expansion of interest in the origin of life. We have summarized some of the current thinking in this latter field with particular reference to reaction rate theory and the significance of optically active molecules. In particular we have stressed the fascinating problem of the universality in living organisms of the L-amino acids in proteins and of D-ribose and deoxyribose in nucleic acids.

The question may well be asked as to where we go from here in the study of the origin of life. Certainly the experimental approaches of Miller, Calvin, Fox, etc., as described above must be expanded; we need to know more about the types of molecules formed and the mechanisms of synthesis under prebiological environmental conditions. Absolute rate theory should play an increasingly useful role here. For example, using Eq. (8), if we increase the concentration of reactants to 0.5 molal, and take n as only a mole of molecules, then the rate for the spontaneous generation of a critical complex with $m = 10$ would be speeded up from once per billion years to once every 9 days. This latter rate is in a range which might well be studied experimentally.

There are a number of other interesting problems which can be approached experimentally. For example, a broader survey should be made of the proteins of contemporary organisms to determine

just how universal the L-amino acids are in their occurrence. More detailed studies like those of Abelson (1957) should be made of the amino acids and other organic compounds in fossils to determine whether the isomeric distribution was the same in the past as it is now. It would be of interest to understand the role of those D-amino acids which do occur in certain polypeptide antibiotics in terms of whether or not the D-form is necessary for antibiotic activity. It appears that American coniferous trees manufacture l-pinene whereas the European conifers produce d-pinene (see Rush, 1957). These and other regional occurrences of particular isomers should be examined and evaluated from an evolutionary point of view. The work of Calvin (1961a) and others on the organic components of meteorites should be expanded with the aim, not only of determining what organic compounds are present, but also what isomers appear. The absolute configurations of L-amino acid protein helices and D-pentose nucleic acid helices should be compared to permit speculation on their origin and evolution, and of their compatibility as catalytic agents and code-containing entities. We feel that the coming era of space exploration, coupled with basic laboratory studies on the origin and evolution of life, although representing very difficult types of biology, cannot be other than meaningful and fruitful in the future.

REFERENCES

- Abelson, P. H. (1957). *Mem. Geol. Soc. Am.* **67**, 87.
- Barghoorn, E. S. (1957). *Mem. Geol. Soc. Am.* **67**, 75.
- Betti, M. (1930). *Trans. Faraday Soc.* **26**, 337.
- Blout, E. R. (1960). In "Optical Rotatory Dispersion: Applications to Organic Chemistry" (C. Djerassi, ed.), Chapter 17. McGraw-Hill, New York.
- Blum, H. F. (1957). In "Rhythmic and Synthetic Processes in Growth" (D. Rudnick, ed.), pp. 155-170. Princeton Univ. Press, Princeton, New Jersey.
- Calvin, M. (1956). *Am. Scientist* **44**, 248.
- Calvin, M. (1959). *Science* **130**, 1170.
- Calvin, M. (1961a). *Ann. Internal Med.* **54**, 954.
- Calvin, M. (1961b). *Chem. Eng. News* **39**, 96.
- Condon, E. U., Altar, W., and Eyring, H. (1937). *J. Chem. Phys.* **5**, 735.
- Djerassi, C. (ed.) (1960). "Optical Rotatory Dispersion: Applications to Organic Chemistry." McGraw-Hill, New York.
- Djerassi, C. (1961). *Science* **134**, 649.
- Eyring, H., and Johnson, F. H. (1957). In "The Influence of Temperature on Biological Systems" (F. H. Johnson, ed.), pp. 1-8. Ronald Press, New York.
- Eyring, H., Walter, J., and Kimball, G. E. (1944). "Quantum Chemistry." Wiley, New York.

- Fox, S. W. (1956). *Am. Scientist* **44**, 347.
- Fox, S. W., and Harada, K. (1961). *Science* **133**, 1923.
- Gaffron, H. (1957). In "Rhythmic and Synthetic Processes in Growth" (D. Rudnick, ed.), pp. 127-154. Princeton Univ. Press, Princeton, New Jersey.
- Hammett, L. P. (1940). "Physical Organic Chemistry." McGraw-Hill, New York.
- Harada, K., and Fox, S. W. (1960). *Arch. Biochem. Biophys.* **86**, 274.
- Johnson, F. H., Eyring, H., and Polissar, M. J. (1954). "The Kinetic Basis of Molecular Biology." Wiley, New York.
- Kauzmann, W. J., Walter, J. E., and Eyring, H. (1940). *Chem. Revs.* **26**, 339.
- Kwoh, T. C., and Eyring, H. (1950). *J. Chem. Phys.* **18**, 1186.
- Lederberg, J., and Hartline, H. K. (1960). In "Science and Space," Chapter IX. National Academy of Sciences—National Research Council, Washington, D.C.
- Miller, S. L. (1957). *Biochim. et Biophys. Acta* **23**, 480.
- Oparin, A. I. (1957). "The Origin of Life," 3rd ed. (translated by A. Syngle). Oliver and Boyd, London.
- Oró, J., and Kamat, S. S. (1961). *Nature* **190**, 442.
- Ree, T., and Eyring, H. (1940). *J. Chem. Phys.* **8**, 433.
- Rosenfeld, L. (1928). *Z. Physik* **52**, 161.
- Rule, H. G. (1930). *Trans. Faraday Soc.* **26**, 321.
- Rush, J. H. (1957). "The Dawn of Life." Hanover House, Garden City, New York.
- Sagan, C. (1961). *Radiation Research* **15**, 174.
- Urey, H. L. (1952). "The Planets." Yale Univ. Press, New Haven, Connecticut.
- Webster, G. (1961). *Ann. Rev. Plant Physiol.* **12**, 113.
- Wilhelmy, L. (1850). *Ann. Physik* **81**, 413, 499.

On the Metabolic Significance of Phosphorolytic and Pyrophosphorolytic Reactions

ARTHUR KORNBERG

*Department of Biochemistry,
Stanford University School of Medicine, Palo Alto, California*

I. Introduction	251
II. Biosynthetic Pathways Dependent on PP-Releasing Reactions	253
III. Degradative Reactions Utilizing Phosphorolytic Cleavages	258
IV. Distinct Biosynthetic and Degradative Pathways	260
V. Summary	261
References	261

I. Introduction

When enzymes were first studied in cell-free systems it became apparent that they could catalyze the reversible changes dictated by the thermodynamic potentialities of a reaction. It seemed possible therefore that in the cell too the same enzyme that was responsible for breakdown of a compound was also the catalyst for its synthesis. But it has become clear during the past decade that proteases are not responsible for protein synthesis, that nucleic acid synthesis is not a reversal of nuclease action, that the paths of fatty acid synthesis and breakdowns are quite distinct. And the same can be said for the biosynthesis and degradation of amino acids, nucleotides, coenzymes, phospholipids, and steroids. Finally, the dilemma of assigning glycogen synthesis to the reversal of phosphorylase action in the face of high intracellular phosphate concentration has been resolved by the discovery of a separate synthetic route from UDPG¹ (Leloir *et al.*, 1959). In fact, the only major two-way street that is still on the metabolic maps is the Embden-Meyerhof conversion of hexoses to pyruvate. Would it be unreasonable to suggest that in this case, too, there is a different avenue for glucose synthesis from pyruvate? This is a question to which we shall return later.

¹ UDPG, uridine diphosphate glucose; PP, inorganic pyrophosphate; PRPP, 5'-phosphoribosyl pyrophosphate; all other abbreviations used are as described in the *Journal of Biological Chemistry*.

As we proceed now to consider the metabolic significance of phosphorolytic and pyrophosphorolytic reactions we will look to see which are most readily identified with biosynthetic pathways and which with degradative routes. Phosphorylases utilize inorganic orthophosphate to cleave the glycosidic bonds of polysaccharides, nucleosides, and the phosphodiester bonds of ribonucleic acid (RNA). The ready reversibility of these reactions immediately suggested that they are feasible routes for cellular synthesis of glycogen, nucleosides, and RNA. Pyrophosphorylases were named by analogy with the phosphorylases (Kornberg, 1950) since they can utilize inorganic pyrophosphate for the cleavages of glycosidic and phosphate bonds. However, these enzymes are recognized primarily for their synthesis of coenzymes, deoxyribonucleic acid (DNA) and a variety of nucleotidyl derivatives in the course of which inorganic pyrophosphate is released. Are the *in vivo* functions of the phosphorylases and pyrophosphorylases chiefly degradative, as these names imply? When we reviewed this subject several years ago (Kornberg, 1957) we stressed instead their biosynthetic role. Yet there appeared to us to be a sharp distinction between the reactions grouped under each of these two headings. The synthesis of DNA, coenzymes, and proteins at some point in the reaction sequences involved the release of PP. In fact it appeared to us that these biosynthetic pathways with an associated release of PP were very likely driven and made "irreversibly anabolic" by the hydrolytic destruction of PP after its release. Hoffmann-Ostenhof and Slechta (1957) ventured a similar speculation in their review of pyrophosphate metabolism. On the other hand, phosphorolytic cleavages seemed to be identified with pathways designed for degradative utilization of compounds (i.e., glycogen) for the energy needs of the cell.

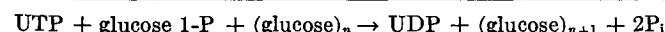
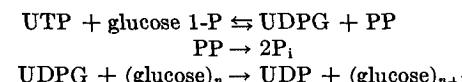
These early considerations of the physiologic significance of the phosphorylases and pyrophosphorylases may be amplified on the basis of two recent discoveries. One mentioned already is that in the synthesis of glycogen by UDPG, transglycosylation is preceded by the release of PP (Leloir *et al.*, 1959). For this and other reasons it now appears most likely that the glycogen phosphorylase serves mainly for the degradative utilization of polysaccharides. The other discovery is that DNA-directed RNA synthesis (Burma *et al.*, 1961; Chamberlin and Berg, 1962; Hurwitz *et al.*, 1960; Stevens, 1960; Weiss, 1960; Weiss and Nakamoto, 1961) and RNA-directed RNA synthesis (Reddi, 1961) and completion of the acceptor RNA chains

(Hecht *et al.*, 1958; Preiss *et al.*, 1961) occur by a condensation of nucleoside triphosphates with a concurrent release of PP; here again the polynucleotide phosphorylase may prove to be primarily responsible for the breakdown rather than the synthesis of RNA.

I propose, in this essay, to present and discuss the hypothesis that degradative pathways often involve phosphorolytic mechanisms and that biosynthetic sequences often rely on PP release. Barring exceptions to this proposal, phosphorylases are therefore appropriately named; but reactions in which PP is released, rather than pyrophosphorylases, are better regarded as "synthetases." In a recent and eloquent editorial, Stetten (1960) has also made a case for the synthetic function of these PP-releasing reactions.

II. Biosynthetic Pathways Dependent on PP-Releasing Reactions

It is evident from the tabulation of synthetases which release PP (Table I) that these are key enzymes in the synthesis of nucleic acids, coenzymes, proteins, carbohydrates, and lipids. In the case of the nucleic acids, nucleotides, and certain of the coenzymes and terpenes, the actual assembly of the constituent units is catalyzed by a PP-releasing reaction. In the synthesis of oligo- and polysaccharides, phospholipids, and probably of proteins, the PP-releasing synthetase is one or more steps removed. In almost all the reactions cited the free energy change associated with the release of PP is relatively small and the synthetic direction of the reaction is demonstrably reversible upon the addition of PP. However, an extremely active inorganic pyrophosphatase is found in all cell extracts examined for it (Schmidt, 1951; Nordlie and Lardy, 1961) and PP itself if detectable in tissues is present in extremely small amounts. Assuming then that the PP released is promptly hydrolyzed, reaction sequences, as for example transglycosylations, can be pulled far in the synthetic direction, as follows:



What evidence can be cited that is lacking for or is inconsistent with the concept that PP-releasing synthetases are the favored devices for biosynthetic pathways? Four considerations come to mind.

TABLE I
BIOSYNTHETIC PATHWAYS DEPENDENT ON PP-RELEASING REACTIONS^a

Biosynthetic pathway	Compound synthesized	Reactants	Ref. ^c	
<i>Coenzymes of electron transport</i>				
DPN, TPN, FAD, thiamine-PP	Nicotinic acid ribose 5'-P Deamido-DPN DPN FAD Thiamine-P	Nicotinic acid Nicotinic acid ribose 5'-P Deamido-DPN, glutamine Riboflavin 5'-P "Thiazole"-P	+PRPP +ATP +ATP +ATP "pyrimidine"-PP	77 77 77 85 11, 11a
<i>Nucleic acids</i>				
	DNA	DNA primer	+dATP, dCTP, dTTP, dGTP	46
	Acceptor RNA	RNA primer	+ATP, CTP	33, 78
	DNA-directed RNA	DNA primer	+ATP, CTP, UTP, GTP	10, 13, 36 93, 99, 100
	RNA-directed RNA	RNA primer	+ATP, GTP, UTP, GTP	81
<i>Nucleotides</i>				
"De novo": pyrimidine purine	Orotidylate 5-Phosphoribosylamine Guanylate	Orotate Glutamine Xanthylate, glutamine (or NH ₃)	+PRPP +PRPP +ATP	51a 28, 29 1, 49, 63
"Salvage": pyrimidine purine	Uridylate Adenylate, inosinate, guanylate	Uracil Adenine, hypoxanthine, guanine	+PRPP +PRPP	17 43, 82
<i>Proteins</i>	Amino-acyl RNA	Amino acid + acceptor RNA	+ATP	7, 8, 38

TABLE I (Continued)

Biosynthetic pathway	Compound synthesized	Reactants	Ref. ^c	
<i>Amino acids</i>				
Histidine	Phosphoribosyl-ATP	ATP	+PRPP	2, 64
Tryptophan (indole glycerol P)	P-ribosyl anthranilic acid	Anthranilic acid	+PRPP	20, 89
Arginine	Argininosuccinate	Citrulline + aspartate	+ATP	75
<i>Lipids</i>	Acyl CoA	Fatty acid + CoA	+ATP	6, 42, 58
<i>Phospholipids^b</i>				
	CDP-choline	P-Choline	+CTP	39, 40
	CDP-ethanolamine	P-Ethanolamine	+CTP	39, 40
	CDP-diglyceride	Phosphatidic acid	+CTP	73
<i>Polysaccharides^b</i> (pectins, chitins) disaccharides, bacterial cell walls, bacterial capsules, ascorbate, T-even bacteriophage DNA, glycolipids, gangliosides				
	UDP-glucose	Glucose 1-P	+UTP	26, 66
	UDP-N-acetylglucosamine	N-Acetylglucosamine 1-P	+UTP	27, 59, 87, 95, 96
	UDP-xylose	Xylose 1-P	+UTP	25
	GDP-mannose	Mannose 1-P	+GTP	66a, 67
	TDP-glucose	Glucose 1-P	+dTTP	48, 74
<i>Steroids, Terpenes</i>				
	Geranyl-PP	Isopentenyl-PP	+dimethylallyl-PP	55
	Farnesyl-PP	Isopentenyl-PP	+geranyl-PP	55
	Squalene	Farnesyl-PP	+farnesyl-PP	30, 54, 76
	Cholyl CoA	Cholic acid + CoA	+ATP	21

TABLE I (Continued)

Biosynthetic pathway	Compound synthesized		Reactants	Ref. ^c
<i>Coenzyme of acylation; CoA</i>	Pantothenate	Pantoic acid + β -alanine	+ATP	56
	Diphospho-CoA	Phosphopantetheine	+ATP	34
<i>Coenzyme of sulfurylation</i>				
Phosphoadenylylsulfate	Adenylylsulfate	Sulfate	+ATP	83, 88a, 101
<i>Polyamines, transmethylation</i>	Adenosylmethionine	Methionine	+ATP	65

^aThe term "biosynthetic pathway" includes those reactions which directly, or as part of a unified chain of reactions, lead to products molecularly more complex than the reactants.

^bCompounds found in nature which are most probably synthesized by reactions analogous to those cited here are CDP-ribitol (3a, 4, 4a, 96), CDP-glycerol (3, 5), and UDP-dihydroxyacetone (88).

^cThe numbers correspond to the numbers indicated in the Reference List.

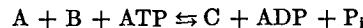
(1) Does inorganic pyrophosphatase dispose of PP promptly and efficiently in whatever compartment of the cell it is produced? There is at present no definitive answer. Sensitive measurements of the steady state levels of PP in actively growing cells are therefore needed; so are more kinetic data on the dissociation constants of the PP complexes of the various inorganic pyrophosphatases and synthetases. The demonstrated presence of PP in fungi (Mann, 1944) requires comment. PP has been identified in cultures which under conditions unfavorable for growth and biosynthesis accumulate inorganic polyphosphate. The most plausible interpretation of the presence of PP is the rapid and extensive degradation of the polyphosphate stores in cells which are in a phase of decline or autolysis (Kornberg *et al.*, 1956; S. R. Kornberg, 1957). There is no evidence from any of the studies on polyphosphate metabolism that the growing cells in the population build up measurable levels of PP.

Mention should also be made of the PP accumulation in sea urchin larvae exposed to high lithium concentrations (Lindahl and Kiessling, 1951); while such larvae fail to develop normally, it cannot be denied that extensive growth does take place.

(2) The PP-releasing synthetase, if inevitably coupled with inorganic pyrophosphatase, is an energetically wasteful process. While this is true the expense is trivial for the end achieved. The *de novo* synthesis of DNA consumes the equivalent of at least 60 ATP's (adenosine triphosphate) per nucleotide. At the cost of merely 2 additional ATP's (resulting from the hydrolysis of PP), the integrity of DNA is assured against the vicissitudes of PP concentration. Synthesis of glycogen via the UDPG pathway consumes 2 ATP's for every glucose unit polymerized, compared to 1 by the phosphorylase route. However the phosphorylase mechanism may not be available at all nor may there be any other alternatives to the UDPG pathway. In that case the glucose would be oxidized and "wasted" entirely as heat. Viewed in this light, the loss of 1 extra ATP of the potential 40 equivalents in glucose is hardly a high price to pay.

(3) Two PP-releasing reactions have been omitted from Table I because there is no apparent connection with biosynthetic events. These include (a) the formation of adenyl luciferin which with luciferase is responsible for light emission by the firefly (McElroy and Seliger, 1961), and (b) the formation of oxaloacetate by phosphorolytic cleavage of phosphopyruvate in propionic bacteria (Siu *et al.*, 1961).

(4) There is a group of reactions best described by the equation:



Reactions of this type are responsible for the synthesis of glutathione (Snoke and Bloch, 1952; Snoke *et al.*, 1953) and glutamine (Meister, 1956, 1960), the amination of purine and pyrimidine nucleotides (Carter and Cohen, 1956; Lieberman, 1955, 1956), and the formylation of tetrahydrofolic acid (Greenberg *et al.*, 1955; Rabinowitz and Pricer, 1956). In order to maintain a consistent argument it must be admitted that these reactions, excluding other reasons for kinetic irreversibility, are in fact susceptible to reversal by the influence of elevated P_i concentrations. Thus the compounds synthesized by this kind of route must be vulnerable to degradation by the same enzymes that formed them or there must be different and as yet undisclosed synthetases, perhaps of the PP-releasing variety.

None of the arguments raised seriously weakens the concept that PP-releasing reactions are designed for biosynthetic functions. On the other hand, no implication has been intended that this type of reaction may not also be used for driving nonsynthetic processes, such as the urea cycle or the examples already cited of light emission and the dicarboxylic acid step in the propionic acid fermentation. Finally it is also clear that there are hydrolytic and other reaction mechanisms outside the PP-releasing synthetase group that are effective in driving synthetic as well as degradative processes.

III. Degradative Reactions Utilizing Phosphorolytic Cleavages

The phosphorolytic cleavage of glycogen to form glucose 1-P, discovered by the Coris' was the first demonstration of how the energy contained in a typical glycoside is preserved (Table II). While not normally regarded as a phosphorylase, the phosphorolysis of a thiol ester intermediate by triose phosphate dehydrogenase qualifies it for inclusion in this group of enzymes. Other "phosphoroclastic" reactions (Table II) and the nucleoside phosphorylases are further examples of energy conservation by phosphorolytic cleavage. Polynucleotide phosphorylase has been disinherited by Cohn from the phosphorylase family because it splits a phosphodiester bond; but this enzyme still retains a kinship on the basis that it uses the phosphorolytic cleavage for efficient utilization of the energy in the polynucleotide chain.

The question to be asked about the thesis presented is whether the phosphorolytic reactions in Table II are properly identified

TABLE II
DEGRADATIVE REACTIONS UTILIZING PHOSPHOROLYTIC CLEAVAGES^a

Degradative pathway	Substrate degraded	Phosphorylated product	Ref. ^b
Polysaccharide → monosaccharide	Glycogen Sucrose Maltose 3-Phosphoglyceraldehyde (3-phosphoglyceryl-S enzyme)	α -Glucose 1-P α -Glucose 1-P β -Glucose 1-P 1,3-Diphosphoglycerate	16 18 19 80, 97
Monosaccharide → acyl P	Xylulose 5-P Pyruvate Aspartic- β -semialdehyde Intermediate in aerobic phosphorylation RNA Ribonucleosides Deoxyribonucleosides	Acetyl P Acetyl P Aspartyl P ATP (P _i -X) Nucleoside diphosphates Ribose 1-P Deoxyribose 1-P	32 69, 94, 102 9 14, 98 70 37, 71, 72 22, 23, 24
Theanine → acyl P "Acetyl" → CO ₂ + H ₂ O Nucleic acid			

^aThe term "degradative pathway" will include those reactions which directly lead to products molecularly less complex than the reactants and reactions which are part of a unified chain directed toward metabolic energy utilization.

^bThe numbers in this column correspond to the numbers indicated in the Reference List.

exclusively as steps in degradative pathways. This assessment of the physiologic functions of the polysaccharide, nucleoside, and polynucleotide phosphorylases is based largely on two points: one, that physiological concentrations of inorganic orthophosphate would favor breakdown by the phosphorylases rather than synthesis, and two, that there are established alternative routes (Table I) which appear well suited to synthetic purposes. However, the use of the term "degradative" may still be challenged since the products of all these phosphorolytic reactions (sugar, phosphates, purines, pyrimidines, and nucleotides), are ultimately used in the biosynthetic pathways. In the same vein the "degradation" of the monosaccharides provides the amino acids essential for nucleic acid and protein synthesis. It can be argued that all the reactions in the cell, including the hydrolytic ones, ultimately serve a biosynthetic function, so why separate anabolism from catabolism?

IV. Distinct Biosynthetic and Degradative Pathways

We might regard the cell as an efficient factory and anticipate that those mechanisms which provide for the supply of energy should be separate from those that utilize the energy for the various synthetic processes of the cell. While this analogy may appear reasonable and could be elaborated, it is less significant than the facts already available from analysis of biosynthetic and degradative pathways. In every case examined, whether it be protein, lipids, polysaccharides, or nucleic acids, the route of synthesis has been found to be distinct from the pathway of combustion. The only major exception, mentioned in the introduction to this essay, appears to be the lactic (or alcoholic) fermentation of glucose. Now we may question whether a cell or organ growing on acetate, lactate, succinate, or related carbon compounds utilizes the Embden-Meyerhof route for building the structural polysaccharide it requires. It is already clear that under some circumstances the conversion of pyruvate to phosphopyruvate may be shunted through a dicarboxylic acid rather than proceed by a direct reversal of the pyruvate phosphokinase reaction. In the same vein, is there also a bypass around triose phosphate dehydrogenase which may even sweep past hexose diphosphates directly to the hexosemonophosphates? Will the ubiquitous uridine diphosphate sugar compounds again be found as the device for reducing glycerates to trioses?

The breakdown of sugar to pyruvate and CO₂ provides the energy essential for biosynthetic reactions; it incidentally also pro-

vides building blocks along the way which can be salvaged for biosynthetic purposes. But in the growing cell the energy-yielding process and the synthetic processes must go on simultaneously! It seems not unreasonable to expect that as in the case of the proteins, nucleic acids, and lipids, the essential carbohydrate structures of the cell are built by reactions distinct from those involved in the degradation of these carbohydrates for the energy needs of the cell. With a functional division between anabolic and catabolic reactions established, it will then remain for subsequent investigations to describe the switching devices that control which of the metabolic pathways a given molecule is to follow.

V. Summary

A survey of enzymatic reactions involving inorganic ortho- and pyrophosphate leads us to the generalization that, in physiologic terms:

(1) Phosphorylases serve principally in degradative pathways. They catalyze phosphorolytic cleavages producing phosphorylated intermediates which conserve the energy of glycosidic and phosphodiester-linked compounds. But the main direction of the pathways involving phosphorylases is toward energy production.

(2) Pyrophosphorylases, by contrast, are better understood as synthetases which release inorganic pyrophosphate. Hydrolysis of the latter by inorganic pyrophosphatase promotes the irreversibility of the synthetic routes to coenzymes, nucleic acids, proteins, and structural carbohydrates and lipids.

On the basis of this generalization, certain predictions are ventured:

(1) Among the biosynthetic pathways to be discovered, some will contain a synthetase releasing inorganic pyrophosphate as a key reaction.

(2) Degradative pathways leading to energy production will utilize phosphorolytic or analogous group transfer reactions but these pathways will be distinct from the biosynthetic ones. Specifically, a pathway bypassing 3-phosphoglyceraledehyde dehydrogenase may be found to account for carbohydrate synthesis from small precursors.

REFERENCES

1. Abrams, R., and Bentley, M. (1959). *Arch. Biochem. Biophys.* **79**, 91.
2. Ames, B. N., Martin, R. G., and Garry, B. J. (1961). *J. Biol. Chem.* **236**, 2019.

3. Baddiley, J., Buchanan, J. G., Mathias, A. P., and Sanderson, A. R. (1956). *J. Chem. Soc.* p. 4186.
- 3a. Baddiley, J., Buchanan, J. G., Carss, B., and Mathias, A. P. (1956). *J. Chem. Soc.* p. 4583.
4. Baddiley, J., Buchanan, J. G., and Carss, B. (1957). *J. Chem. Soc.* p. 1869.
- 4a. Baddiley, J., Buchanan, J. G., and Carss, B. (1957). *J. Chem. Soc.* p. 4058.
5. Baddiley, J., Buchanan, J. G., and Sanderson, A. R. (1958). *J. Chem. Soc.* p. 3107.
6. Berg, P. (1956). *J. Biol. Chem.* **222**, 991.
7. Berg, P., and Ofengand, E. J. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 78.
8. Berg, P., Bergmann, F. H., Ofengand, E. J., and Dieckmann, M. (1961). *J. Biol. Chem.* **236**, 1726.
9. Black, S., and Wright, N. G. (1955). *J. Biol. Chem.* **213**, 39.
10. Burma, D. P., Kröger, H., Ochoa, S., Warner, R. C., and Weill, J. D. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 749.
11. Camiener, G. W., and Brown, G. M. (1960). *J. Biol. Chem.* **235**, 2404.
- 11a. Camiener, G. W., and Brown, G. M. (1960). *J. Biol. Chem.* **235**, 2411.
12. Carter, C. E., and Cohen, L. H. (1956). *J. Biol. Chem.* **222**, 17.
13. Chamberlin, M., and Berg, P. (1962). *Proc. Natl. Acad. Sci. U.S.* **48**, 81.
14. Chance, B. (1961). *J. Biol. Chem.* **236**, 1569.
15. Cohn, Mildred. Personal communication.
16. Cori, G. T., and Cori, C. F. (1943). *J. Biol. Chem.* **151**, 57.
17. Crawford, I., Kornberg, A., and Simms, E. S. (1957). *J. Biol. Chem.* **226**, 1093.
18. Doudoroff, M., Barker, H. A., and Hassid, W. Z. (1947). *J. Biol. Chem.* **168**, 725.
19. Doudoroff, M. (1955). In "Methods in Enzymology" (S. P. Colowick and N. O. Kaplan, eds.), Vol. I, p. 225. Academic Press, New York.
20. Doy, C. H., Rivera, A., Jr., and Srinivasan, P. R. (1961). *Biochem. Biophys. Research Commun.* **4**, 83.
21. Elliott, W. H. (1957). *Biochem. J.* **65**, 315.
22. Friedkin, M., and Kalckar, H. M. (1950). *J. Biol. Chem.* **184**, 437.
23. Friedkin, M. (1953). *J. Cellular Comp. Physiol.* **41**, Suppl. 1, 261.
24. Friedkin, M. (1954). *J. Biol. Chem.* **209**, 295.
25. Ginsburg, V., Neufeld, E. F., and Hassid, W. Z. (1956). *Proc. Natl. Acad. Sci. U.S.* **42**, 333.
26. Ginsburg, V. (1958). *J. Biol. Chem.* **232**, 55.
27. Glaser, L., and Brown, D. H. (1955). *Proc. Natl. Acad. Sci. U.S.* **41**, 253.
28. Goldthwait, D. A., Greenberg, G. R., and Peabody, R. A. (1955). *Biochim. et Biophys. Acta* **18**, 148.
29. Goldthwait, D. A. (1956). *J. Biol. Chem.* **222**, 1051.
30. Goodman, D. S., and Popjak, G. (1960). *J. Lipid Research* **1**, 286.
31. Greenberg, G. R., Jaenicke, L., and Silverman, M. (1955). *Biochim. et Biophys. Acta* **17**, 589.
32. Heath, E. C., Hurwitz, J., and Horecker, B. L. (1956). *J. Am. Chem. Soc.* **78**, 5449.
33. Hecht, L. I., Zamecnik, P. C., Stephenson, M. L., and Scott, J. F. (1958). *J. Biol. Chem.* **233**, 954.

34. Hoagland, M. B., and Novelli, G. D. (1954). *J. Biol. Chem.* **207**, 767.
35. Hoffmann-Ostenhof, O., and Slechta, L. (1957). *Proc. Intern. Symp. Enzyme Chem.*, Tokyo, p. 180.
36. Hurwitz, J., Bresler, A., and Diringer, R. (1960). *Biochem. Biophys. Research Commun.* **3**, 15.
37. Kalckar, H. M. (1947). *J. Biol. Chem.* **167**, 477.
38. Keller, E. B., and Zamecnik, P. C. (1956). *J. Biol. Chem.* **221**, 45.
39. Kennedy, E. P. (1957). *Ann. Rev. Biochem.* **26**, 119.
40. Kennedy, E. P. (1960). In "The Enzymes" (P. D. Boyer, H. Lardy, and K. Myrbäck, eds.), Vol. 2, p. 63. Academic Press, New York.
41. Kornberg, A. (1950). *J. Biol. Chem.* **182**, 779.
42. Kornberg, A., and Pricer, W. E., Jr. (1951). *J. Biol. Chem.* **193**, 481.
43. Kornberg, A., Lieberman, I., and Simms, E. S. (1955). *J. Biol. Chem.* **215**, 417.
44. Kornberg, A., Kornberg, S. R., and Simms, E. S. (1956). *Biochim. et Biophys. Acta* **20**, 215.
45. Kornberg, A. (1957). *Advances in Enzymol.* **18**, 191.
46. Kornberg, A. (1960). *Science* **131**, 1503.
47. Kornberg, S. R. (1957). *Biochim. et Biophys. Acta* **26**, 294.
48. Kornfeld, S., and Glaser, L. (1961). *J. Biol. Chem.* **236**, 1791.
49. Lagerkvist, U. (1958). *J. Biol. Chem.* **233**, 143.
50. Leloir, L. F., Olavarria, J. M., Goldemberg, S. H., and Carminatti, H. (1959). *Arch. Biochem. Biophys.* **81**, 508.
51. Lieberman, I. (1955). *J. Am. Chem. Soc.* **77**, 2661.
- 51a. Lieberman, I., Kornberg, A., and Simms, E. S. (1955). *J. Biol. Chem.* **215**, 403.
52. Lieberman, I. (1956). *J. Am. Chem. Soc.* **78**, 251.
53. Lindahl, P. E., and Kiessling, K. (1951). *Arkiv Kemi* **3**, 97.
54. Lynen, F., Eggerer, H., Henning, U., and Kessel, I. (1958). *Angew. Chem.* **70**, 738.
55. Lynen, F., Agranoff, B. W., Eggerer, H., Henning, U., and Mösllein, E. M. (1959). *Angew. Chem.* **71**, 657.
56. Maas, W. K. (1956). *Federation Proc.* **15**, 305.
57. McElroy, W. D., and Seliger, H. H. (1961). In "Light and Life" (W. D. McElroy and B. Glass, eds.), p. 219. Johns Hopkins Press, Baltimore, Maryland.
58. Mahler, H. R., and Wakil, S. J. (1953). *J. Biol. Chem.* **204**, 453.
59. Maley, F., Maley, G. F., and Lardy, H. A. (1956). *J. Am. Chem. Soc.* **78**, 5303.
60. Mann, T. (1944). *Biochem. J.* **38**, 345.
61. Meister, A. (1956). *Physiol. Revs.* **36**, 103.
62. Meister, A. (1960). In "Amino Acids, Proteins and Cancer Biochemistry" (J. T. Edsall, ed.), p. 85. Academic Press, New York.
63. Moyed, H. S., and Magasanik, B. (1957). *J. Biol. Chem.* **226**, 351.
64. Moyed, H. S., and Magasanik, B. (1960). *J. Biol. Chem.* **235**, 149.
65. Mudd, H. S., and Cantoni, G. L. (1958). *J. Biol. Chem.* **231**, 481.
66. Munch-Peterson, A. (1955). *Acta Chem. Scand.* **9**, 1523.
- 66a. Munch-Peterson, A. (1955). *Arch. Biochem. Biophys.* **55**, 592.

67. Munch-Peterson, A. (1956). *Acta Chem. Scand.* **10**, 928.
68. Nordlie, R. C., and Lardy, H. A. (1961). *Biochim. et Biophys. Acta* **50**, 189.
69. Novelli, G. D. (1955). *Biochim. et Biophys. Acta* **18**, 594.
70. Ochoa, S., and Heppel, L. A. (1957). In "The Chemical Basis of Heredity" (W. D. McElroy and B. Glass, eds.), p. 615. Johns Hopkins Press, Baltimore, Maryland.
71. Paege, L. M., and Schlenk, F. (1950). *Arch. Biochem. Biophys.* **28**, 348.
72. Paege, L. M., and Schlenk, F. (1952). *Arch. Biochem. Biophys.* **40**, 42.
73. Paulus, H., and Kennedy, E. P. (1960). *J. Biol. Chem.* **235**, 1303.
74. Pazur, J. H., and Shuey, E. W. (1961). *J. Biol. Chem.* **236**, 1780.
75. Petrack, B., and Ratner, S. (1958). *J. Biol. Chem.* **233**, 1494.
76. Popjak, G., Goodman, D. S., Cornforth, J. W., Cornforth, R. H., and Ryhage, R. (1961). *J. Biol. Chem.* **236**, 1934.
77. Preiss, J., and Handler, P. (1958). *J. Biol. Chem.* **233**, 493.
78. Preiss, J., Dieckmann, M., and Berg, P. (1961). *J. Biol. Chem.* **236**, 1748.
79. Rabinowitz, J. C., and Pricer, W. E., Jr. (1956). *J. Am. Chem. Soc.* **78**, 4176.
80. Racker, E., and Krimsky, I. (1952). *J. Biol. Chem.* **198**, 731.
81. Reddi, K. K. (1961). *Science* **133**, 1367.
82. Remy, C. N., Remy, W. T., and Buchanan, J. M. (1955). *J. Biol. Chem.* **217**, 885.
83. Robbins, P. W., and Lipmann, F. (1958). *J. Biol. Chem.* **233**, 681.
- 83a. Robbins, P. W., and Lipmann, F. (1958). *J. Biol. Chem.* **233**, 686.
84. Schmidt, G. (1951). In "Phosphorus Metabolism" (W. D. McElroy and B. Glass, eds.), Vol. 1, p. 443. Johns Hopkins Press, Baltimore, Maryland.
85. Schrecker, A. W., and Kornberg, A. (1950). *J. Biol. Chem.* **182**, 795.
86. Siu, P. M. L., Wood, H. G., and Stjernholm, R. L. (1961). *J. Biol. Chem.* **236**, 21P.
87. Smith, E. E. B., and Mills, G. T. (1954). *Biochim. et Biophys. Acta* **13**, 386.
88. Smith, E. E. B., Galloway, B., and Mills, G. T. (1961). *Biochem. Biophys. Research Commun.* **5**, 148.
89. Smith, O. H., and Yanofsky, C. (1960). *J. Biol. Chem.* **235**, 2051.
90. Snoke, J. E., and Bloch, K. (1952). *J. Biol. Chem.* **199**, 407.
91. Snoke, J. E., Yanari, J., and Bloch, K. (1953). *J. Biol. Chem.* **201**, 573.
92. Stetten, D., Jr. (1960). *Am. J. Med.* **28**, 867.
93. Stevens, A. (1960). *Biochem. Biophys. Research Commun.* **3**, 92.
94. Strecker, H. J. (1951). *J. Biol. Chem.* **189**, 815.
95. Strominger, J. L., and Smith, M. S. (1959). *J. Biol. Chem.* **234**, 1822.
96. Strominger, J. (1960). *Physiol. Revs.* **40**, 55.
97. Velick, S. F. (1954). In "The Mechanism of Enzyme Action" (W. D. McElroy and B. Glass, eds.), p. 491. Johns Hopkins Press, Baltimore, Maryland.
98. Wadkins, C. L., and Lehninger, A. L. (1958). *J. Biol. Chem.* **233**, 1589.
99. Weiss, S. B. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1020.
100. Weiss, S. B., and Nakamoto, T. (1961). *J. Biol. Chem.* **236**, 18P.
101. Wilson, L. G., and Bandurski, R. S. (1958). *J. Biol. Chem.* **233**, 975.
102. Wolfe, R. S., and O'Kane, D. J. (1955). *J. Biol. Chem.* **215**, 637.

Catalysis in Life and in the Test Tube¹

D. E. KOSHLAND, JR.

*Biology Department,
Brookhaven National Laboratory, Upton, New York*

I. Introduction	265
II. Catalysis in Life	266
III. Catalytic Power of Enzymes	268
IV. Enzyme Specificity	273
V. Ramifications from Studies on Enzyme Action	279
VI. Envoi	282
References	282

I. Introduction

A horizon might be defined as that point in the dimly seen future where the world of reality and the world of dreams meet. In scientific terms this might be translated as the point of intersection of the linear extrapolation of disciplined science with the hyperbolic fancies of science fiction. This chapter in "Horizons in Biochemistry" will, therefore, begin with the existing knowledge in enzymology and travel toward the unknown; whether the destinations reached are closer to reasoned extrapolation or to science fiction only the future can decide.

We have for some time now lost our awe of enzymes. They have been crystallized, "sequenced," and even X-rayed. They can be modified, denatured, and renatured. We are getting to know them better, but the secrets of their amazing properties are still undisclosed. Yet the ever accelerating knowledge suggests that the understanding of enzyme action may indeed be one of the achievements of the next decade of biochemistry. It is, therefore, probably not too early to cast our eyes toward that portion of the horizon which deals with the mechanisms of enzyme action.

¹ Research carried out at Brookhaven National Laboratory under the auspices of the U.S. Atomic Energy Commission.

II. Catalysis in Life

Catalysis in life is achieved by enzymes which are protein molecules. Some of these enzymes are stable and soluble and, therefore, can be isolated and purified with relative ease. In these cases it is fairly simple to establish that the *in vivo* properties of these catalysts are not changed by the isolation procedure. Enzymes, therefore, represent a class of biological materials which can be made to produce *in vitro* the precise action that they perform *in vivo*.

When enzymes are studied there are two properties which immediately stand out in comparison to the test tube catalysts of the chemist. The first is the ability of the enzymes to be excellent catalysts for chemical changes under mild conditions, i.e., at pH 7 in aqueous solution at room temperature. The second is their high specificity. Catalysis by chemical agents can be efficient but usually requires drastic conditions such as high temperature and high acidity. Such catalysis has very little specificity.

When the structure of a protein is examined, no immediate explanation for its superiority to man-made catalysts is apparent. Amino acid side chains contain groups whose chemical properties are familiar—sulphydryl, imidazole, indole, etc. The individual residues can act as catalysts in some cases, but they are pedestrian in their catalytic power. (The term "catalytic power" will be used to refer to the covalent bond-changing power of the enzyme in distinction to its specificity properties.) How can the relatively weak catalytic groups acquire these exceptional properties when they are included in a giant protein molecule? The search for the answer to this question has drawn increasing attention in recent times.

An important step in this search has already been taken. Test tube analogs have been found for a wide variety of enzymatic reaction. Some of the analogies were obvious at an early stage, e.g., both acid and invertase catalyze the hydrolysis of sucrose. Others took time and some very complex chemistry to unravel. Today test tube analogs have been demonstrated for such widely varied systems as the pyridoxal enzymes (Metzler *et al.*, 1954; Metzler and Snell, 1952), the DPNH enzymes (Mauzerall and Westheimer, 1955; Abeles *et al.*, 1957), the proteolytic enzymes (Bender and Neveu, 1958; Bender, 1957; Bruice and Schmir, 1959; Westhead and Morawetz, 1958; Jencks and Carriuolo, 1959), catalase (Wang, 1955), and thiamine-catalyzed condensations (Breslow, 1957). Moreover, the properties

of enzymes in general have been rationalized on the basis of known physical organic principles for the addition to double bonds (Gawron *et al.*, 1961; Speyer and Dickman, 1956; Alberty and Bender, 1959; Fisher *et al.*, 1955), the condensation reactions (Rieder and Rose, 1959; Bloom and Topper, 1956), the proteolytic enzymes (Gutfreund and Sturtevant, 1956; Wilson, 1960), and the transfer enzymes (Koshland, 1954). In fact it is probably safe to say that a nonenzymatic analog has been, or with little effort can be, found for every enzymatic reaction.

Before discussing the hopeful aspects of this conclusion, it may be well to mention its qualifications. The existence of an analogy does not establish that the enzyme is operating by the same mechanism as the analog. In many cases there are several analogs for a given enzymatic reaction. Hydrogen peroxide decomposition is catalyzed by a variety of agents from dust to ferric ion, adenosine triphosphate (ATP) hydrolysis is catalyzed by hydrogen ions and metal ions, ester hydrolysis can be catalyzed by acids and bases, and so on. The existence of an analogy does not, therefore, establish that it is a true model for the enzyme action. Moreover, even when a simple analogy exists, the extent of the similarity may be questioned. For example, both hydrogen ions and β -amylase catalyze the splitting of the glycosidic bonds of the starch molecule. The analogy does not explain, however, how the enzyme can selectively split the second glucosyl bond from the nonreducing end, whereas the acid acts essentially at random throughout the starch molecule. It does not explain the ability of the enzyme to hydrolyze the starch rapidly at room temperature at hydrogen ion concentration of $10^{-7} M$ when the acid-catalyzed reaction requires elevated temperatures at $1 M$ acid concentrations. We have not yet established a single model for a single enzymatic reaction or explained in detail the reasons for the efficiency and selectivity of the natural catalysts.

Nevertheless, the recurrent success at finding an analogy for each enzyme suggests a possible route to these explanations. Physical organic and physical inorganic chemistry have provided mechanisms in terms of electron densities and steric relationships. If enzymes are devoid of "black magic," if they can only carry out reactions which will proceed, albeit less rapidly and less selectively, without them, it seems probable that the search for new principles and the diligent examination of old principles in the light of these new problems may best begin in this area of mechanism studies. In the next two sec-

tions, mechanistic approaches to some phases of the problem of catalytic power and of specificity will be discussed.

III. Catalytic Power of Enzymes

As mentioned above, enzymes are able to accelerate a given reaction by many orders of magnitude under extremely mild conditions. The usual enzymatic situation involves two compounds which could be left together at room temperature in aqueous solution at neutral pH for years without detectable reaction. Addition of low concentrations of enzyme to these solutions accelerates the reaction so that the reactants are converted to products in minutes.

It is important to place a quantitative figure on this catalytic power but this is not a simple matter. The turnover number is used by many for this purpose. Catalase is thereby considered one of the most powerful enzymes since it has one of the highest turnover numbers known. Yet the reaction catalyzed by catalase—the decomposition of H_2O_2 —is one which goes with appreciable velocity in the absence of any catalyst, and can be catalyzed by such anemic catalysts as dust and glass walls.

A more meaningful comparison is obtained by quantitating the ratio of the velocities for a given reaction in the presence and absence of enzyme. When this is done some very large factors are found (Koshland, 1956). For example, hexokinase accelerates the reaction of glucose and ATP by a factor greater than 10^{11} , chymotrypsin accelerates the reaction of an ester and water by at least 10^6 and alcohol dehydrogenase speeds up the oxidation of alcohol by di-phosphopyridine nucleotide (DPN) by a factor of greater than 10^9 .

The ratios calculated this way are an improvement. They give a figure for the catalytic power of an enzyme, i.e., how much it accelerates the nonenzymatic reaction. They also give us a quantitative mark against which to evaluate our nonenzymatic models. They do not, however, point a pathway to the construction of such a model. To do that we would need a comparison based on theoretical principles which evaluated the contribution of various types of catalysis to the over-all rate. For example, the kinetics of the enzyme action in many cases is explained by the existence of a "ternary complex" between enzyme and two substrates. Presumably the substrates are absorbed to adjacent sites on the protein surface prior to reaction. It is important to have a quantitative estimate of the importance of this ternary complex on the catalysis by the enzyme.

Is it by itself, sufficient to explain the factors of 10^9 – 10^{12} mentioned above, or does it make only a minor contribution to this factor? The same question may be asked of each model or partial model for enzyme action. A quantitative evaluation based on theoretical rather than empirical principles is needed not only to guide us in the construction of models for enzyme action, but also to tell us when our quest has met with success.

A method for making such a comparison has recently been derived (Koshland, 1960, 1962) and it may be profitable to examine its nature and implications for a moment. Since the detailed mathematical manipulations have been developed elsewhere, only illustrative calculations will be repeated here. A somewhat simplified mathematical formula for calculating the maximum velocity V_E from the velocity in the absence of enzyme V_0 is given by Eq. (1) in which the θ 's represent orientation factors, (E_T) represents

$$\frac{V_E}{V_0} = \frac{(E_T)(55.5)^4 \theta_A \theta_B \theta_R \theta_S \theta_T}{(A)(B)(R)(S)(T)} \quad (1)$$

the enzyme concentration and R, S, and T represent the concentrations of the catalyst molecules during the nonenzymatic reaction. Let us assume that an enzyme catalyzes a reaction between two substrates A and B by the concerted action of three catalytic groups. The active site on the enzyme might in this case be represented by the schematic illustration of Fig. 1a. The arrows within the circles representing A, B, R, S, and T are used to indicate that these molecules are not spherically symmetrical and hence will have a preferred orientation for reaction. The catalytic groups R, S, and T are shown with covalent bonds (heavy lines) to the enzyme as would be the case if they were side chains such as imidazole or prosthetic groups like heme. However, as we shall see, a metal ion absorbed by noncovalent forces would not change the reasoning. The substrates are thus absorbed to a surface in which the catalytic groups are perfectly aligned and the alignment of the substrates themselves is accomplished by the specificity forces on the protein.

An infinitely sharp alignment of molecules is obviously not necessary and, therefore, a more realistic model is that of Fig. 1b in which the arrows are replaced by pie-shaped wedges to indicate the fraction of the solid surface ($1/\theta$) over which reaction can occur.

If the enzyme has as its only function the juxtaposition and

alignment of catalyst and substrate groups, the ratio of enzymatic and nonenzymatic velocities will depend on the concentration of the catalytic sites compared to the concentration of a 5-molecule constellation of the type shown in Fig. 2. In this figure, R, S, and T represent the same catalytic groups as before except that they are free in solution, e.g., R might represent a free imidazole molecule. It is this type of ratio which was

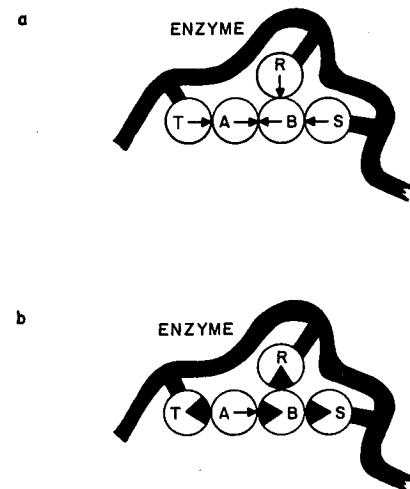


FIG. 1. Schematic illustration of a hypothetical reaction on an enzyme surface involving two substrates, A and B, and three groups on the enzyme, R, S, and T. The orientation for maximum reactivity is indicated by the arrows in (a), the enzyme groups being held in this position by the enzyme structure and the substrates being held by the specificity of the absorption process. In part (b), the same process is illustrated but that fraction of the solid surface over which reaction can occur is indicated by the pie-shaped wedges on each molecule.

used in the derivation of Eq. (1). In other words, the calculated V_E/V_0 ratio will agree with that observed experimentally if the assumed juxtaposition and orientation functions are the only ones that the enzyme performs. If these functions are inadequate to explain enzyme action, the ratio will be far less than the experimental factors of ca. 10^{10} and the extent of the difference will indicate the potential contribution of this type of catalysis. Application of this equation to specific cases will be published elsewhere, but certain general conclusions which have emerged will be discussed here.

In the first place, let us assume that reaction can occur for each of the substrates and catalysts over 10% of the surface of each of these molecules, i.e., $\theta_A = \theta_B = \theta_R = \theta_S = \theta_T = 10$. Values of concentrations for A, B, R, S, and T of $10^{-3} M$ and for E_T of $10^{-5} M$ are not unlike many experimentally realizable situations and hence may be used in this illustrative example. Substituting these values in Eq. (1) gives a V_E/V_0 ratio of 10^{21} . Stated in words, this says

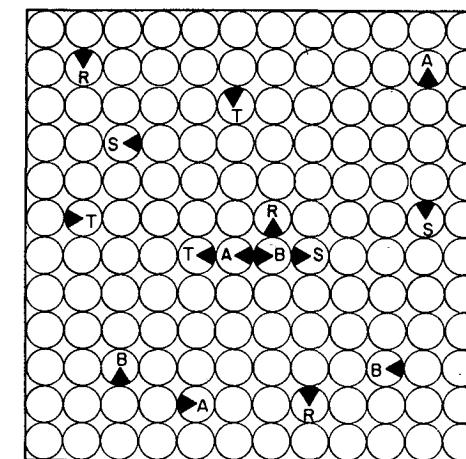


FIG. 2. Schematic illustration of the reaction of Fig. 1 occurring in solution in the absence of enzyme. Groups R, S, and T are the same reactive groups as before except that they are no longer attached to a protein structure. The combination of the five molecules in an orientation appropriate for reaction is indicated in the center of the figure. Intermediate associations, e.g., R and B, are not shown for simplicity in the diagram although it is clear they will occur. The concentration of the various combinations is calculable assuming random associations or from a detailed knowledge of the dissociation constants.

that a reaction involving 2 substrates and 3 catalysts with a relatively modest requirement for orientation will proceed 10^{21} times as rapidly on an enzyme surface even if the only enzyme function is to provide this oriented juxtaposition. This is considerably larger than the factors of 10^{12} given above and, therefore, offers the tantalizing possibility that such a mechanism might be sufficient to explain the high velocity of some enzyme reactions.

This large factor may account for enzyme action in some cases. Is it enough to explain all enzyme action? The answer to that ques-

tion depends on the values for the θ factors and whether or not several catalytic groups are involved. Assuming the same θ factors and concentrations, if one catalytic group and two substrates are involved, the V_E/V_0 ratio drops to 10^{10} . If only two substrates or one substrate and one catalytic group or two substrates (one of which is water) and a catalytic group are involved, the ratio drops to 10^3 . On the other hand, two substrates with θ factors of 10^3 each would give a V_E/V_0 ratio of 10^8 . Whether the oriented juxtaposition can account for enzyme action, therefore, will depend on the experimental evidence for the size of the θ factors and the number of catalytic groups.

Some minimum estimate of the size of θ factors can be made from reactions such as the Walden inversion reaction in which back-side attack occurs. The observed inversion of the asymmetric carbon during reaction demonstrates that the θ factor must be at least 4 and reactions in fused systems suggest that it may be considerably larger than this (Cristol and Arganbright, 1957). A factor of 10 does not, therefore, seem improbable as an initial estimate, but studies on specific cases will be needed before the contribution of this term can be determined precisely.

Some support for the idea of polyfunctional catalysis has been obtained (Swain and Brown, 1952), but the solidity of the evidence is in some question (Bell and Clunie, 1952; Bell and Jones, 1953). Nevertheless it is apparent that such catalysis will be difficult to establish in the absence of enzymes. Any reaction involving a change in covalent bonds involves two opposite electronic components. At one site electrons must be attracted and at another, repelled. Two reagents which can help in such processes would be an acid and a base, respectively. However, it is impossible to obtain high concentrations of an acid and a base in the same solution since they neutralize each other. The same arguments hold for oxidizing and reducing agents. The fixing of the catalytic groups in a three-dimensional protein matrix may help to resolve this basic incompatibility by preventing the direct neutralization reaction.² Moreover the orientation requirements of a multiple collision of free molecules are, as we have seen, formidable. The present scanty evidence for the concerted action of catalysts is, therefore, reasonable on the basis of

² The indirect neutralization via the buffer ions of the solution would still exist and hence geometrical contribution is necessarily oversimplified in this abbreviated discussion.

these considerations and not an argument against the existence of such catalysis on the enzyme surface.

Clearly the demands of the orientation factors and the number of catalytic groups are interdependent. If the θ factors are large, the number of catalytic groups needed for a given V_E/V_0 ratio will diminish. If the number of catalytic groups is 3 or more, rather moderate θ factors may suffice. The calculations discussed above, however, indicate that the oriented juxtaposition of substrates and catalysts probably play a significant role and may perhaps embody the answer to the enormous catalytic power of enzymes.

IV. Enzyme Specificity

The ability to distinguish between closely similar chemical structures, is one of the most important properties of biological systems. This specificity property is found not only in enzymes, but also in antibodies, permeable membranes, and systems synthesizing adaptive enzymes. It has also been observed in phenomena such as drug action and toxin action, but whether these specificities are caused by new specificity sites or by the interaction with the systems mentioned above is not yet known. However, a similarity in experimental data between the specificity exhibited by substrates to an enzyme, haptene to an antibody, or inducer to an induced enzyme synthesis, suggests that the specificity process, if not identical, is at least similar in all of these systems. Studies on enzyme specificity will therefore draw from and add to the studies on analogous systems.

The long-standing explanation for enzyme specificity is the classic "template" or "lock and key" model of Fischer, Ehrlich *et al.* In this theory a part of the enzyme surface is considered to be a three-dimensional negative of at least a portion of the substrate. Substrate and enzyme have, prior to interaction, complementary topographies which fit together in the manner of a jigsaw puzzle.

Recently, however, it has been suggested (Koshland, 1958, 1959), that such a fitting is not sufficient to explain enzyme specificity, but rather some, and perhaps all, enzymes undergo a significant change in shape on interaction with their substrates. Not all of the enzyme in its native state would be a simple negative of the substrate and enzyme and substrate could not be paired from simple external contours. If a rough mechanical analogy for this new theory is needed, a "glove" model might suffice. The glove is not a three-

dimensional negative of a hand before the hand is introduced. It may be in any of a number of shapes, but only after introduction of the hand is the close three-dimensional fit obtained. In the enzyme case a substrate is presumed to induce a proper alignment of catalytic groups so that enzyme action ensues. Thus the fitting feature of the template hypothesis is retained, but an added requirement of a necessary induced change in the protein structure is added. With this added feature, a number of phenomena which are, at the least, puzzling and, at the most, inexplicable by a template hypothesis can be understood.

This new hypothesis is far from proven, but the increasing number of phenomena uncovered in our own laboratory (Koshland, 1960; Thoma and Koshland, 1960) and in the laboratories of others (Grisolia and Joyce, 1959; Nirenberg and Jakoby, 1960; Tomkins *et al.*, 1961; Christensen, 1960) which can be readily explained by it, suggest that it may play an increasing role in the explanation of biological specificity. Some of its features and some of its consequences may, therefore, be worth discussing here.

In the first place, the fact that unfolding of the protein occurs does not mean that the unfolding is random or uncontrolled. Rather, it seems necessary that the unfolding is controlled and catalyzed by the substrate. For example, one group on the substrate may make initial contact with a complementary group on the enzyme. Because of this association, a partial unfolding of the protein may occur leading to the exposure of a previously concealed group on the enzyme which can now associate with a second group on the substrate, and so forth. A flexible enzyme theory does not, therefore, mean that the original enzyme exists as a random structure, but only that it can change upon interaction with its substrate. It goes without saying that a rigid substrate is not needed either. It is the final complex of enzyme and substrate that counts, and the changes leading to this complementary association may involve conformation changes of the substrate.

In the second place, the extent of the protein shape change is not, *a priori*, either large or small. The phenomenological requirement is that the change be enough to bring the catalytic groups into proper alignment. As we have seen in the previous section, such alignment may have to be very precise and hence a change from nonalignment to alignment may require only very small shape changes. Moreover, the active site itself is probably only a small

fraction of the total molecule. On comparing the molecular weights of proteins which are large, with the molecular weights of substrates which are usually small, it is obvious that only a small fraction of the amino acids of the protein can be in contact with the substrate at any one time. To obtain direct evidence that a protein change has occurred involves the frequently encountered problem of a small difference between two large quantities. It might not have been surprising, therefore, if such a difference had gone undetected. The indications that substrate-induced changes are sizable (Yankeelev and Koshland, *in press*), however, although preliminary, suggest that the changes may not always be small. Studies on ribonuclease and chymotrypsin have shown that amino acid residues which are distant from each other in the linear sequence of the protein chain, are probably involved in enzyme action (Dixon *et al.*, 1958; Stein, 1960; Hirs *et al.*, 1961). This does not mean that these residues must undergo large movements to become properly aligned, but it allows such a possibility. If these changes occur at diffusion velocities, they may be much more rapid than covalent bond changes, but the possibility also exists that the conformation changes themselves may be rate limiting in some cases of enzyme action.

A new theory causes a re-examination of old concepts and experimental procedures. It is amusing that some of these procedures which are clearly "evidence" for a given postulate may not be so clear-cut when examined from a different point of view. For example, if the presence of a substrate protects a sulphydryl group on an enzyme from reaction with iodoacetamide, it is usually deduced that the sulphydryl residue is at the active site and the protection results from the steric interposition of substrate between SH group and reagent. However, if the substrate induces conformation changes which cause unfolding of the protein, the unmasking and the masking of residues which are not in direct contact with the substrate may occur. The evidence cited above indicates just such an unmasking. It follows then, that some of the masking by substrates which has been, or will be, observed may also be explained by conformation changes and not by direct protection by substrate.

A second experimental criterion which has been used to identify a group at the active site is competitive inhibition. If a compound is a competitive inhibitor, it presumably competes with the true substrate for part or all of the active site. This is undoubtedly the

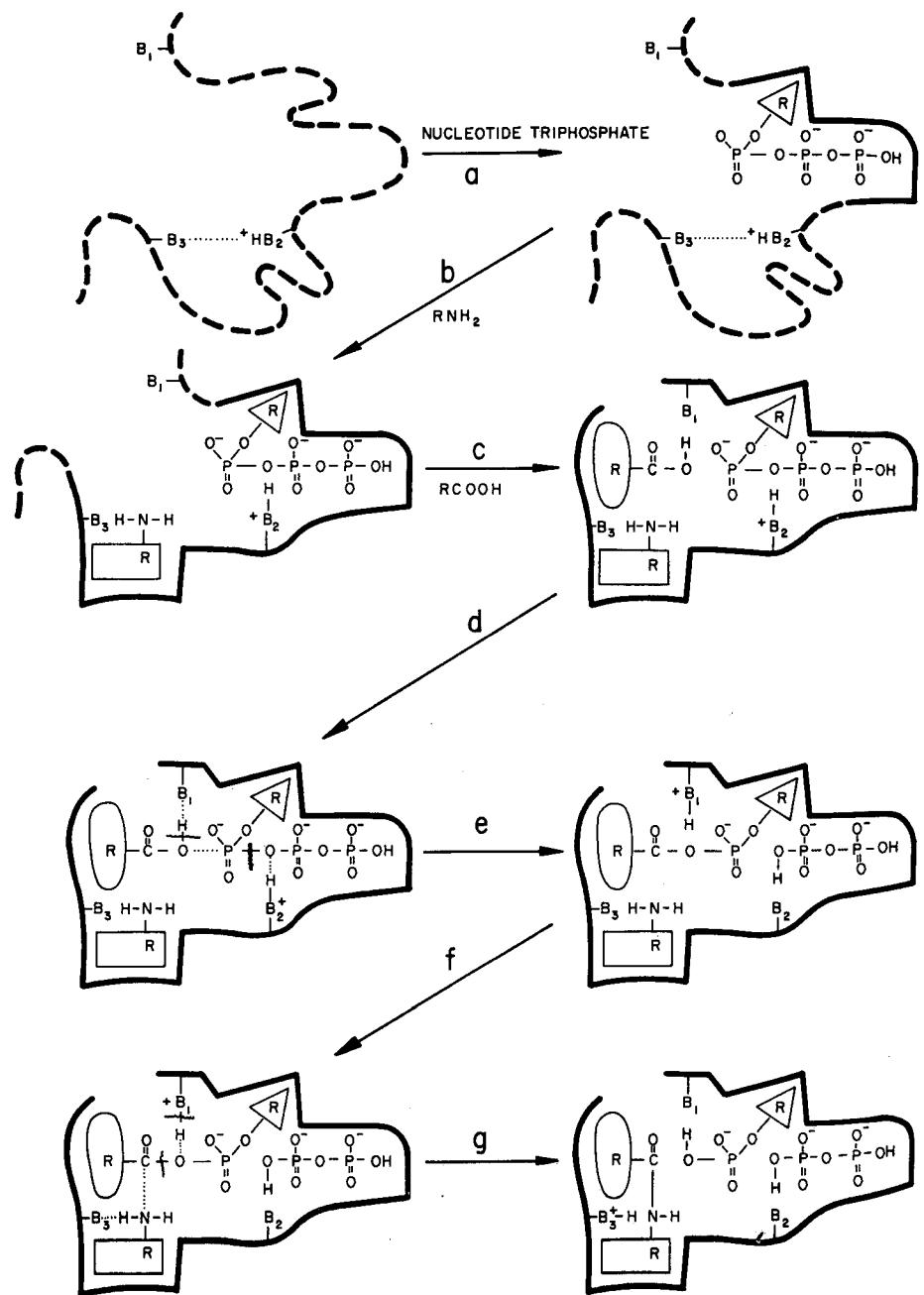


FIG. 3. For descriptive legend see opposite page.

correct explanation in most cases. However, if a change in protein shape accompanies the binding of compound X to an enzyme, the resulting molecule might not bind the true substrate. In this case, the observed kinetics would be that of competitive inhibition even though the two compounds were acting at different sites.

A flexible enzyme-substrate interaction provides an explanation for the sequential absorption and reaction of substrates on a single enzyme. In some cases this may have a pronounced biological value to the organism. An illustrative example which may be an explanation for the reaction of some synthetase-type enzymes is shown in Fig. 3. The reaction chosen in this case is a hypothetical one to explain the detailed steps of the over-all reaction of Eq. (2). Similar equations can be written for reactions like glutamine synthetase, pantothenic acid synthetase, acetyl coenzyme A synthetase, and amino acid activation. In these cases the acceptor molecule varies and the ATP split may produce adenosine diphosphate (ADP) and P_i instead of adenosine-5'-phosphate (AMP) and pyrophosphate, but the principles are the same,



In the figure it is seen that ATP, the amine, and the carboxylic acid are each absorbed in turn and each cause conformation changes in the protein. When all three molecules are absorbed, the catalytic groups B_1 , B_2 , and B_3 are aligned to allow reaction. Since the catalytic groups to activate the first reaction between carboxyl group and ATP are not properly aligned prior to the arrival of the amine, no initial reaction leading to the acyl adenylate derivative could occur until the amine is in place on the enzyme. Since the amine must be in place before the first step occurs, the second step will occur immediately thereafter because all the catalytic groups and substrates are properly aligned. This serves a desirable biological function since the acyl adenylate derivative is highly reactive and

FIG. 3. A schematic illustration of the formation of a highly reactive intermediate on a flexible enzyme surface. B_1 , B_2 , and B_3 represent the catalytic groups which must be induced into appropriate alignment prior to reaction. The heavy line represents a portion of the enzyme molecule. It is dotted to indicate the native structure in the absence of substrate and solid to indicate the changed conformation when substrate is absorbed. Steps showing desorption of the products to regenerate the free enzyme would occur after g, but are not detailed here.

might undergo numerous side reactions if it were to diffuse freely in the cell. Mere absorption to an enzyme might be less efficient. It is known for example, that DPN absorbed to an enzyme can be hydrolyzed by *Neurospora* DPNase (Astrachan, 1954) and hence absorption might not, *per se*, reduce side reactions sufficiently. A flexible enzyme can, therefore, perform a useful function in preserving and controlling the behavior of a highly reactive intermediate.

This mechanism suggests further ramifications. The molecule which leads to the proper alignment of the catalytic groups for the acyl adenylate does not have to be a reactant. It might be a group on the acyl substrate itself (Fig. 4a) or a chemically inert compound,

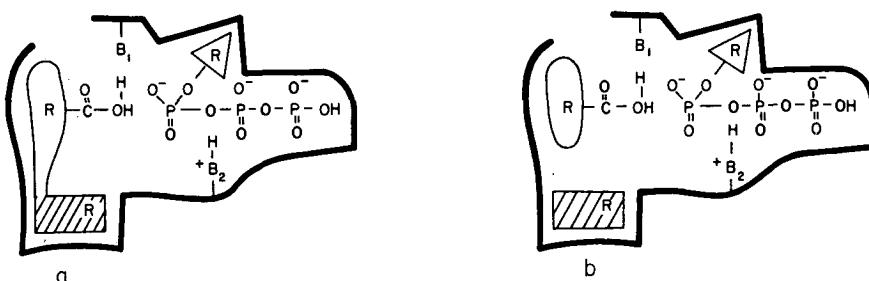


FIG. 4. Schematic illustration of a group R, either attached to the carboxylic acid (a) or as a separate molecule (b), which does not react itself but performs the function of inducing a proper protein shape in the carboxyl group-ATP reaction as does RNH₂ in Fig. 3.

R (Fig. 4b). The former explains why a minimum size is frequently needed in the substrate. An acyl molecule without the R group would not react because proper alignment of catalytic groups would not ensue. The latter suggests a role for hormone action since the compound R is necessary for reaction of the two substrates, but is not itself changed chemically in any way. A steroid or a polypeptide chain might well perform the function of the R group.

The model also allows an explanation for the confusing complexity of drug specificity. For example, if a drug replaced the hormone R, it might completely inhibit the enzyme action or, if it were a more effective activator than the natural hormone, it might accelerate the enzyme action.

Moreover, the model allows the possibility that a drug might be effective even in the case in which a natural hormone did not exist.

Enzymes do not always exist in their optimum forms, and some reagents which affect protein structure can enhance enzyme activity. For example, both dinitrophenol and *p*-chloromercuribenzoate enhance the ATPase activity of the muscle protein, myosin (Chappel and Perry, 1955; Kielley and Bradley, 1956). Hence it is reasonable to suppose that a synthetic molecule R might influence the structure of a protein and thus either accelerate or depress the normal enzyme action. If so, the design of drugs might be drastically affected by our developing knowledge of enzyme structure and the mechanism of enzyme specificity.

V. Ramifications from Studies on Enzyme Action

In the preceding sections some of the more specific developments in studies of enzyme action have been discussed. It may be worthwhile to conclude with some extrapolation to other areas of biochemistry.

To begin with the subject of metabolic control, it is clear that the kinetic properties of enzymes are of vital importance in the final understanding of the complex metabolic machinery of the cell. The variations with pH, mutation, reducing reagents, etc., must be known so that the effect of the actual cellular milieu and its variations on the velocity of the enzyme action can be calculated. With the aid of computers the data of the living system may ultimately be rationalized in terms of the *in vitro* studies. Indeed preliminary findings in this direction have already been started (Chance, 1960). For the soluble enzymes these data should not be too difficult to obtain. For the insoluble ones it may be quite difficult, if not impossible, to extract the enzyme with its native structure intact. For example, it might be covalently bonded to the particle, membrane, or other enzymes. Is this an insuperable obstacle? The answer seems to be "no." From our present knowledge of protein structure it seems quite possible to obtain an enzyme with its native activity even though its native structure has been altered. We know that large portions of an enzyme can be altered without effect on the enzyme activity. Changes in charge or in the nature of groups or even the deletion of much of the protein chain may have no detectable effect on the assay (Hill and Smith, 1956). In cases in which the enzyme is insoluble this knowledge may be of particular value since it would allow the removal of an enzyme from insoluble material even by breaking covalent bonds without a change in its

native activity. It is this native activity, rather than its native structure, which is of most importance in the kinetic analysis.

On the other hand, our structural knowledge may play a role in preventing hasty conclusions. Modification of some parts of a protein or even change in pH, metal ion concentration, etc., can change the specificity of a protein. The demonstration that an ATPase can be isolated from mitochondria on sonication does not establish that it is an ATPase *in situ*. The rather drastic conditions of isolation might convert a normal kinase into an ATPase. Thus, the increasing knowledge of protein structure may not only allow enzymes to be isolated with their native activity intact, but also may help in establishing when that activity is modified.

Another look at the horizon suggests that synthetic enzymes of much smaller molecular weight than the native enzyme may be made. The preponderance of evidence suggests that a small "active site" which is only a fraction of the total enzyme is responsible for enzyme catalysis and enzyme specificity. The rest of the structure appears to be a framework designed to hold the catalytic and specificity groups in the proper three-dimensioned orientation. Presumably some day methods may be found for designing only an active site with a low molecular weight superstructure, e.g., a phenanthrene molecule. These molecules would in the first place be far more stable than proteins and hence their usefulness in industrial processes would be greater. A column of synthetic enzyme over which reactants would be passed might last indefinitely. Katchalski has already constructed a column with an absorbed protein enzyme on it (Bar-Eli and Katchalski, 1960). A column packed with enzymes in a sequential array might place the advantages of the sequential processes of the living cell in the hands of laboratory and commercial producers of chemicals.

More important yet, however, is the potentiality of medical application of synthetic enzymes. The correction of "inborn errors of metabolism" and of the "aging process" would seem to be two areas where synthetic enzymes might be of value. As the infectious diseases steadily come under greater control by antimetabolites and antibiotics, the relative importance of the genetic diseases, of the degenerative diseases, and of the aging process will rise. How is therapy to be applied in these cases?

In the case of the "inborn errors of metabolism" it is conceivable to remedy the situation by injection of a synthetic enzyme or a syn-

thetic inhibitor depending on whether a deficiency or an excess is to be corrected. The insulin treatment of diabetes would be a somewhat distant analog of such a treatment. A more efficient process would be to correct the genetic defect at its source, i.e., on the chromosome. Our current knowledge of genetics suggests that a chemical defect in the DNA molecule at a specific locus becomes translated into a deficient enzyme and hence to a metabolic defect. Repairing this locus, particularly if a small change had occurred, in a DNA molecule of highly similar nucleotides would require a specific reagent and a synthetic enzyme would seem to be a possible solution to such a requirement.

The "aging process" is today still not understood and theories of its nature are diverse. A speculative hypothesis which is consistent with the theories of enzyme structure and hormone action mentioned above is that there is a "death hormone." Other hormones—growth hormone, pressor hormones, gonadotropic hormones—have clear-cut roles which aid the living organism. It may seem ridiculous to assume that the delicate machinery would also synthesize a deleterious hormone designed for its own destruction. On second thought, however, death has value to the species as a whole even if it is unpleasant for the individual. If natural selection is believed, death of the individual provides a mechanism for improvement of the species as a whole. It is nature's way of correcting errors and developing new and more efficient living systems. It allows living systems to adapt to environmental changes. As more adaptable, less environmentally sensitive forms evolved, complex growth patterns were controlled by hormonal messengers linked to built-in timetables. Why not, therefore, add an agent to ensure that death should occur naturally even in the absence of accidents?

Where would one look for such a hormone and how would it act? The symptoms of age, e.g., desiccation, decrease in metabolic rates, would seem to provide some clues for experimental approaches to the search of such a hormone. In view of the preceding discussion, the possibility that the "death hormone" is a general enzyme inhibitor rather than one blocking a specific enzyme, would also seem worthy of consideration. The juvenile hormone apparently acts by inhibiting hormones or enzymes which cause metamorphosis (Williams, 1956) and hence some analogy for such a mechanism can be found. At any rate, if a hormone for a time-controlled decay is

built into living systems, an enzyme or drug designed to counteract it would be of considerable interest.

VI. Envoi

A speculative article should not have a conclusion. Time will make its judgment of the evaluations and predictions. Scientists who well know the danger of predicting tomorrow's experiment, understand the hazard of describing the horizon. There is the danger of castles in the sand, but there is also the danger of timidity. Who, even in the 1930's, would have believed that by the early 1960's we would be synthesizing DNA by a cell-free system or discussing the side chains of a globular protein to 2 Å resolution? An author of such an article can only hope that some of the ideas are true, that some will serve as a stimulus for discoveries that might otherwise have taken longer to unravel, and that some are fantastic enough to keep up with the pace of modern science.

REFERENCES

- Abeles, R., Hutton, R., and Westheimer, F. H. (1957). *J. Am. Chem. Soc.* **79**, 712.
 Alberti, R. A., and Bender, P. (1959). *J. Am. Chem. Soc.* **81**, 542.
 Astrachan, L. (1954). *Federation Proc.* **13**, 177.
 Bar-Eli, A., and Katchalski, E. (1960). *Nature* **188**, 856.
 Bell, R. P., and Clunie, J. C. (1952). *Proc. Roy. Soc. A* **212**, 33.
 Bell, R. P., and Jones, P. (1953). *J. Chem. Soc.*, p. 88.
 Bender, M. L. (1957). *J. Am. Chem. Soc.* **79**, 1258.
 Bender, M. L., and Neveu, M. C. (1958). *J. Am. Chem. Soc.* **80**, 5388.
 Bloom, B., and Topper, Y. T. (1956). *Science* **124**, 982.
 Breslow, R. (1957a). *J. Am. Chem. Soc.* **79**, 1762.
 Breslow, R. (1957b). *Chem. & Ind. (London)*, p. 893.
 Bruice, T. C., and Schmir, G. L. (1959). *J. Am. Chem. Soc.* **81**, 4552.
 Chance, B. (1960). *J. Biol. Chem.* **235**, 2440.
 Chappel, J. B., and Perry, S. V. (1955). *Biochim. et Biophys. Acta* **16**, 285.
 Christensen, H. N. (1960). *Advances in Protein Chem.* **15**, 239.
 Cristol, S., and Arganbright, R. P. (1957). *J. Am. Chem. Soc.* **79**, 3441.
 Dixon, G. H., Kauffman, D. L., and Neurath, H. (1958). *J. Am. Chem. Soc.* **80**, 1260.
 Fisher, H. F., Frieden, C., McKee, J. S. M., and Alberti, R. A. (1955). *J. Am. Chem. Soc.* **77**, 4436.
 Gawron, O., Glaid, A. J., and Fondy, T. P. (1961). *J. Am. Chem. Soc.* **83**, 3634.
 Grisolia, S., and Joyce, B. B. (1959). *Biochem. Biophys. Research Commun.* **1**, 280.
 Gutfreund, H., and Sturtevant, J. M. (1956a). *Biochem. J.* **63**, 656.
 Gutfreund, H., and Sturtevant, J. M. (1956b). *Proc. Natl. Acad. Sci. U.S.* **42**, 719.
 Hill, R. L., and Smith, E. L. (1956). *Biochim. et Biophys. Acta* **19**, 376.
 Hirs, C. H. W., Halmann, M., and Kycia, J. (1961). In "Biological Structure and Function" (T. W. Goodwin and D. Lindberg, eds.), Vol. I, p. 41. Academic Press, New York.
 Jencks, W. P., and Carriuolo, J. (1959). *J. Biol. Chem.* **234**, 1272, 1280.
 Kielley, W. W., and Bradley, L. B. (1956). *J. Biol. Chem.* **218**, 653.
 Koshland, D. E., Jr. (1954). In "Mechanism of Enzyme Action" (W. D. McElroy and B. Glass, eds.), p. 608. The Johns Hopkins Press, Baltimore Maryland.
 Koshland, D. E., Jr. (1956). *J. Cellular Comp. Physiol. Suppl.* **1**, 47, 217.
 Koshland, D. E., Jr. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 98.
 Koshland, D. E., Jr. (1959). *J. Cellular Comp. Physiol. Suppl.* **1**, 54, 245.
 Koshland, D. E., Jr. (1960). *Advances in Enzymol.* **22**, 45.
 Koshland, D. E., Jr. (1962). *J. Theoret. Biol.* In press.
 Mauzerall, D., and Westheimer, F. H. (1955). *J. Am. Chem. Soc.* **77**, 2261.
 Metzler, D. E., and Snell, E. E. (1952). *J. Am. Chem. Soc.* **74**, 979.
 Metzler, D. E., Ikawa, M., and Snell, E. E. (1954). *J. Am. Chem. Soc.* **76**, 648.
 Nirenberg, M. W., and Jakoby, W. B. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 206.
 Rieder, S. V., and Rose, I. A. (1959). *J. Biol. Chem.* **234**, 1007.
 Speyer, J. F., and Dickman, S. R. (1956). *J. Biol. Chem.* **220**, 193.
 Stein, W. H. (1960). *Brookhaven Symposia in Biol.* No. **13**, 104.
 Swain, C. G., and Brown, J. F. (1952). *J. Am. Chem. Soc.* **74**, 2534, 2538.
 Thoma, J. A., and Koshland, D. E., Jr. (1960). *J. Biol. Chem.* **235**, 2511.
 Tomkins, G. M., Yielding, K. L., and Curran, J. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 270.
 Wang, J. H. (1955). *J. Am. Chem. Soc.* **54**, 822, 4715.
 Westhead, E. H., Jr., and Morawetz, H. (1958). *J. Am. Chem. Soc.* **80**, 237.
 Williams, C. R. (1956). *Nature* **178**, 212.
 Wilson, I. B. (1960). In "The Enzymes" (P. D. Boyer, H. A. Lardy, and K. Myrbäck, eds.), Rev. ed., Vol. IV, p. 501. Academic Press, New York.
 Yankelev, J., and Koshland, D. E., Jr. (In press). In "Biological Structure and Function." Pergamon Press, New York.

CATALYSIS

Enzyme Activity and Cellular Structure

H. A. KREBS

Medical Research Council Unit for Research in Cell Metabolism, Department of Biochemistry, University of Oxford, Oxford, England

One of the main efforts of biochemists is directed toward the study of isolated pure enzymes. This approach is based on the view that living cells are susceptible to the same type of chemical analysis as other chemical systems. As D. E. Green (1961) has put it, such systems "have to be taken apart into their components; the individual parts have to be characterized chemically; and the manner of association of the parts has to be regarded as no less a chemical problem than the study of the parts themselves." This traditional concept of chemistry is fully justified. Yet there can be no doubt that this must not be the only approach, because the properties of enzymes *in situ* can differ in many ways from those of the pure enzymes and such differences are often of great physiological interest, because they may throw light on the factors which regulate enzyme activity.

The effects of simple chemical or physical manipulations on the characteristics of an enzyme are strikingly demonstrated by the following experiments on the succinic oxidase system. The reaction investigated was the oxidation of succinate by molecular oxygen:



the rate of which depends on the activity of succinic dehydrogenase and of the electron transport carriers between succinic dehydrogenase and O₂. Most of the experiments were carried out on pigeon breast muscle minced in the Latapie mill and suspended in a saline medium, a material introduced into biochemical research by Albert Szent-Györgyi over 25 years ago (see Szent-Györgyi, 1935). The concentration of the tissue and the time of incubation was generally the same in the different experiments so that the rate values recorded in different tables are directly comparable.

When thoroughly washed pigeon breast muscle, which has been stored for at least 24 hours at -15°, is resuspended and incubated

with succinate aerobically the rates of succinate disappearance are relatively low—about one third or one quarter of those of fresh muscle—and the ratio (succinate removed/O₂ used) is near 2, indicating that reaction (1) is virtually the only oxidative process (Table I). The rates are not appreciably affected by 0.5 mM dinitrophenol or by 1.25 mM Amytal or by adenosine triphosphate (ATP).

TABLE I

EFFECTS OF AMYTAL, DINITROPHENOL, AND ATP ON THE OXIDATION OF SUCCINATE IN WASHED AND FROZEN PIGEON BREAST MUSCLE PARTICLES^a

Substances other than succinate (final concentrations)	O ₂ uptake (μ moles)	Succinate removed (μ moles)	Ratio Succinate/O ₂
None	16.7	34.7	2.08
Dinitrophenol (0.5 mM)	14.5	30.0	2.07
Amytal (1.25 mM)	16.8	33.6	2.10
ATP (2 mM)	15.8	31.9	2.02

^a The minced tissue was washed four times with 10 volumes of 0.01 M Na-phosphate buffer, pH 7.4 and stored for several days at -15°. The frozen tissue was suspended in the saline medium. The data refer to 4 ml suspension containing 400 mg (wet wt) tissue and 120 μ moles succinate, 30°; 30 minutes; O₂.

Different results are obtained when fresh tissue which has not been washed or frozen is used. In the example shown in Table II, the rate of removal of succinate is roughly three times, and that of the oxygen consumption more than four times, higher. The ratio (succinate removed/O₂ used) is smaller than 2 because substances other than succinate are oxidized. In this material dinitrophenol strongly inhibits the removal of succinate and to some extent also the O₂ uptake, and these inhibitions are largely removed by Amytal. ATP has no appreciable effects except in the presence of dinitrophenol when it slightly counteracts the inhibition of succinate removal by the latter.

Again very different reactivities are found in fresh breast muscle particles which have been thoroughly washed with the saline medium and are used immediately after washing (Table III). The rates of oxygen uptake and succinate removal in this material are less than one quarter of those found in unwashed material. Dinitrophenol and Amytal slightly accelerate the rate of succinate removal. A very large effect is shown by ATP. It stimulates by several

hundred per cent the rates of both O₂ uptake and succinate removal and restores them to almost the level of the unwashed material. This shows, *inter alia*, that the catalytic capacity of the electron transport carriers was not limiting the rate of oxygen uptake in the un-supplemented washed suspension. The ATP-stimulated oxidation of succinate, in contrast to the oxidation taking place before the addition of ATP, is inhibited by dinitrophenol and this inhibition is largely removed by Amytal. No further stimulation occurred on addition of diphosphopyridine nucleotide or coenzyme A (not recorded in Table III).

TABLE II

EFFECTS OF AMYTAL, DINITROPHENOL, AND ATP ON THE OXIDATION OF SUCCINATE IN FRESH MINCED PIGEON BREAST MUSCLE SUSPENSION^a

Substances other than succinate (final concentration)	O ₂ uptake (μ moles)	Succinate removed (μ moles)	Ratio Succinate/O ₂
None	51.8	65.0	1.25
Dinitrophenol (0.5 mM)	47.1	34.5	0.73
Amytal (1.25 mM)	57.6	80.6	1.40
Dinitrophenol + Amytal	50.3	55.5	1.10
ATP (2 mM)	58.7	64.0	1.09
ATP + dinitrophenol	49.0	44.9	0.92
ATP + Amytal	63.6	83.9	1.32
ATP + Amytal + dinitrophenol	53.4	56.0	1.05

^a The data refer to 4 ml suspension containing 400 mg (wet wt) tissue and 120 μ moles succinate, 30°; 30 minutes; O₂.

Some of these observations are not unexpected and have been made before on other materials (see Ernster, 1961; Slater and Hülsmann, 1961). The stimulation of the oxidation by ATP in the washed particles might be thought to be due to the circumstance that respiration is effectively coupled to phosphorylation in this system and that the level of adenosine diphosphate (ADP) is a controlling factor. ATP would act as a source of ADP. Other observations however do not fit in with the accepted concepts, e.g., the inhibition of succinate oxidation by dinitrophenol and the abolition of this inhibition by Amytal (Table II). Further, if it were correct that ATP increased the rate of oxidation of succinate because it served as a supply of ADP, then ADP should have the same effect as ATP, while adenosine monophosphate (AMP) should

be inactive. ADP is in fact as effective as ATP but AMP is not entirely inactive (Table IV); AMP has a marked though smaller effect than ATP. It cannot be excluded, though, that AMP is phosphorylated to ADP.

TABLE III
EFFECTS OF AMYTAZ, DINITROPHENOL, AND ATP ON THE OXIDATION OF SUCCINATE IN WASHED PIGEON BREAST MUSCLE PARTICLES^a

Substances other than succinate (final concentration)	O ₂ uptake (μ moles)	Succinate removed (μ moles)	Ratio Succinate/O ₂
None	11.3	9.5	0.84
Dinitrophenol (0.5 mM)	10.5	12.5	1.19
Amytal (1.25 mM)	15.4	18.4	1.19
Dinitrophenol + Amytal	12.3	16.7	1.36
ATP (2 mM)	38.5	45.7	1.19
ATP + dinitrophenol	24.0	20.7	0.86
ATP + Amytal	43.7	64.5	1.47
ATP + Amytal + dinitrophenol	36.5	40.5	1.11

^a The data refer to 4 ml of a suspension containing washed particles prepared as described in the text. Otherwise conditions are as in Table IV.

The main gap in the explanation of the findings thus concerns the inhibition of succinate oxidation by dinitrophenol, and the counteraction of Amytal. The only established action of dinitrophenol in this type of system is the inhibition of oxidative phosphorylation and the only established action of Amytal is the blocking of the electron transport between reduced diphosphopyridine nucleotide and the cytochromes. As these effects do not directly account for the inhibition of succinate oxidation by dinitrophenol and its abolition by Amytal, there must be either additional or secondary effects of the two inhibitors. It has been suggested that the level of oxaloacetate may be a key factor in controlling the rate of succinate oxidation (Schollmeyer and Klingenberg, 1960; Ernster, 1961; Slater and Hülsmann, 1961; Williams, 1961). It is true, as known from the work of Das (1937) working in Szent-Györgyi's laboratory, that oxaloacetate can act as a powerful inhibitor of succinic oxidase when added to washed pigeon breast muscle (see also Tyler, 1960). Under the conditions of the experiment recorded in Table I 0.06 mM oxaloacetate inhibited succinate oxidation by 84%. The assumption that oxaloacetate controls the rate of suc-

cinate oxidation is an attractive one, but some facts do not seem to fit in with this hypothesis. The addition of oxaloacetate to suspensions of fresh muscle, even 10 mM, has no appreciable inhibitory effect (Table V); probably because it is very rapidly metabolized by minced pigeon breast muscle (Krebs *et al.*, 1940). The rate of oxaloacetate removal by the tissue is not affected by dinitrophenol (Table VI). The non-inhibition of succinate oxidation by added oxaloacetate is difficult to reconcile with the assumption that oxaloacetate accumulation causes the inhibition of the oxidation of suc-

TABLE IV
EFFECTS OF ADENOSINE NUCLEOTIDES ON THE OXIDATION OF SUCCINATE IN WASHED PIGEON BREAST MUSCLE PARTICLES^a

Substances other than succinate (final concentration)	O ₂ uptake (μ moles)	Succinate removed (μ moles)	Ratio Succinate/O ₂
None	16.4	23.9	1.46
Dinitrophenol	9.3	13.5	1.45
AMP	22.5	35.3	1.57
AMP; dinitrophenol	10.9	18.4	1.68
ADP	33.2	48.7	1.38
ADP; dinitrophenol	17.1	22.0	1.29
ATP	38.1	47.5	1.25
ATP; dinitrophenol	20.1	22.0	1.10

^a The data refer to 4 ml washed particles prepared as described in text. Four milliliters contained 120 μ moles succinate. The concentration of dinitrophenol was 0.5 mM, of the adenosine nucleotides 2 mM.

cinate by dinitrophenol. One would have to postulate that the distribution of oxaloacetate within the mitochondria is uneven; that the locally formed oxaloacetate accumulates at the site of succinic dehydrogenase; that added oxaloacetate reacts at other sites. Although it is known that the permeability of mitochondria to carboxylic acids can be restricted, as in the case of citrate, the fact that oxaloacetate is readily metabolized on addition to mitochondria does not support the idea that its diffusion into, or within, mitochondria is blocked by permeability barriers.

An observation which may be of relevance is the inhibition of succinic oxidase by rather high concentrations of dinitrophenol. At 2.5 mM the inhibition was 40% (Table VII) and at 0.5 mM 13%. In fresh muscle (Table II) 0.5 mM inhibited succinate removal 47%

TABLE V
EFFECT OF OXALOACETATE ON THE OXIDATION OF SUCCINATE IN FRESH MINCED PIGEON BREAST MUSCLE^a

Substances other than succinate (final concentration)	O ₂ uptake (μ moles)	Succinate removed (μ moles)	Ratio Succinate/O ₂
None	25.2	36.2	1.44
Oxaloacetate (10 mM)	23.5	35.1	1.49
Oxaloacetate (5 mM)	23.9	33.6	1.41
Dinitrophenol (1.25 mM)	16.9	6.4	0.38

^a The data refer to 4 ml suspension incubated for 40 minutes at 20°. The lower temperature was chosen because of the greater stability of oxaloacetate at that temperature. It should be noted that the inhibition of succinate oxidation by dinitrophenol is greater at 20° than at 30°.

TABLE VI

EFFECT OF SUCCINATE AND DINITROPHENOL ON THE DISAPPEARANCE OF OXALOACETATE IN PIGEON BREAST MUSCLE SUSPENSION^a

Additions other than oxaloacetate	Oxaloacetate present in 4 ml suspension			
	0 min	5 min	10 min	20 min
(μ moles)				
None	39.2	15.6	8.1	2.4
Succinate (30 mM)	39.2	21.8	16.0	9.2
Succinate (30 mM) + dinitrophenol (0.5 mM)	39.2	21.4	15.4	8.9
Dinitrophenol (0.5 mM)	39.2	14.8	8.4	3.3

^a The data refer to 4 ml 10% suspension to which 39.2 μ moles oxaloacetate were added. 30°; O₂. Succinate inhibits oxaloacetate removal (see Tyler, 1960), but dinitrophenol has no effect.

and in fresh washed particles supplemented with ATP the inhibition was 55% (Table III). These findings suggest that the enzyme system may undergo modification in the process of washing and freezing. What kind of modification this is is a matter of conjecture. "Denaturation," conversion into an isoenzyme (see Markert and Møller, 1959) or formation of a less active fragment of the enzyme are possibilities.

No attempt is made here to elaborate explanations because they

are bound to be speculative. The object of the article is to illustrate the thesis that purification of enzymes may modify their properties and that it is therefore essential to investigate the behavior of enzymes not only in the purified state but also in their natural environment. The study of crude enzymes may bring to light important characteristics which may escape attention in the examination of the pure enzymes. This follows from the consideration that living cells are systems where the whole is more than the sum of its components. The integration of the parts to a unit involves an arrangement whereby the component parts influence each other's behavior. In the terms of chemistry this interlocking implies that cell constituents modify each other's chemical reactivity, an interplay which is

TABLE VII
INHIBITION OF SUCCINIC OXIDASE BY HIGH CONCENTRATIONS OF DINITROPHENOL^a

Final concentration of dinitrophenol added (M)	O ₂ uptake (μ moles)	Succinate removed (μ moles)	Ratio Succinate/O ₂	% Inhibition of succinate removal
0	13.0	-27.1	2.08	—
2.5 mM	8.3	-16.4	1.98	39.5%

^a The succinic oxidase preparation and the experimental conditions were as in Table I.

an essential part of the regulatory mechanisms. It distinguishes a complex unit from a complex mixture. The feedback inhibition of the synthesis of amino acids and pyrimidines, the repression of enzyme formation by metabolites, and the inhibition of enzymes by metabolites, are examples of such interplay.

Experimental Details

Pigeon breast muscle was minced in the Latapie mill and suspended in 6.5 volumes of a saline medium consisting of 90 ml 0.155 M KCl, 20 ml 0.1 M K-phosphate buffer, pH 7.4, and 1 ml 0.1 M MgCl₂. Three milliliters of this suspension were pipetted into the main compartment of a conical Warburg cup and 1 ml of various additions were made so that the final volume was 4 ml and the tissue concentration 10%. Usually substrates, inhibitors, and

other substances were added from the side-arm after 10 minutes' shaking in the water bath at 30°. The gas space contained O₂, the center well 0.2 ml 2N NaOH. The data given refer to a 30-minute period of incubation (unless otherwise stated) beginning at the end of the equilibration time.

"Washed" tissue particles were prepared by suspending the tissue in 6.5 volumes of the above medium and centrifuging in a high-speed head (24,000 g) for 5 minutes. This was repeated twice more and the tissue was finally suspended in the same volume.

The oxygen uptake was measured manometrically. When succinate was to be determined, 1 ml 15% (w/v) trichloroacetic acid was added as quickly as possible at the end of the incubation. The precipitate was centrifuged off and the supernatant was heated for 90 minutes in boiling water to remove the trichloroacetic acid, care being taken that no water was lost. It proved unnecessary to extract succinic acid with ether as the quantities of succinate were relatively high and interfering substances were absent. The succinate was determined manometrically in 1 ml (see Krebs, 1937). When oxaloacetate was to be determined, the enzymatic reactions were stopped by the addition of 1 ml N HCl to the cup contents and deproteinization was carried out with HClO₄. Oxaloacetate was determined either enzymatically (Hohorst *et al.*, 1959) or with aniline citrate (Edson, 1935).

REFERENCES

- Das, N. B. (1937). *Biochem. J.* **31**, 1124.
Edson, N. L. (1935). *Biochem. J.* **29**, 2082.
Ernster, L. (1961). In "Biological Structure and Function" First IUB/IUBS Joint Symposium, Stockholm, September 12-17, 1960 (T. W. Goodwin and O. Lindberg, eds.), Vol. II, p. 139. Academic Press, New York.
Green, D. E. (1961). Plenary Lecture, V Intern. Congr. Biochem., Moscow.
Hohorst, H. J., Kreutz, F. H., and Bücher, T. (1959). *Biochem. Z.* **332**, 18.
Krebs, H. A. (1937). *Biochem. J.* **31**, 2095.
Krebs, H. A., Eggleston, L. V., Kleinzeller, A., and Smyth, D. H. (1940). *Biochem. J.* **34**, 1234.
Krebs, H. A., Eggleston, L. V., and d'Alessandro, A. (1961). *Biochem. J.* **79**, 537.
Markert, C. L., and Møller, F. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 753.
Schollmeyer, P., and Klingenberg, M. (1961). *Biochem. Biophys. Research Commununs.* **4**, 43.
Slater, E. C., and Hülsmann, W. C. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1109.
Szent-Györgyi, A. (1935). *Z. physiol. Chem.* **236**, 1.
Tyler, D. B. (1960). *Biochem. J.* **76**, 293.
Williams, G. R. (1961). *Can. J. Biochem. and Physiol.* **39**, 103.

Molecular Organization

Giant Molecules and Semiconductors¹

L. BRILLOUIN

Columbia University, New York, New York

I. Introduction	295
II. Electronic Waves in Periodic Structures; Band Structure	296
III. Pauli Exclusion Principle. Distinction between Conductors and Insulators	300
IV. The Intermediate Case of Semiconductors; Donor and Acceptor Impurities	304
V. Some Physical Problems Concerning Giant Molecules in Biology	305
VI. A Simplified Model of Semiconducting Molecule	307
VII. Donors and Acceptors—Calculations by Pullman and Pullman	308
VIII. Example of an Application—DNA	313
IX. The Bases as Possible Donors and Acceptors	316
References	318

I. Introduction

The brilliant idea of comparing the gigantic macromolecules of biology with the semiconductors in crystal physics was first presented by A. Szent-Györgyi in 1941 (see also Szent-Györgyi, 1946). It immediately raised a great deal of interest. Many experimental results did require a mechanism for energy transfer along the molecules, but theoretical discussions were not too successful, and for some years this suggestion seemed to be almost abandoned. New investigations recently have revived the whole problem, and many convincing results show the possibility of overcoming the previous theoretical obstacles; once this roadblock is removed, a broad new field will be opened for investigation which may lead to a variety of discoveries.

The purpose of the present report is to give the viewpoint of a theoretician, who has been at the origin of the whole theory of semiconductors, and to discuss in very general terms the conditions required for a possible application of the theory of semiconductors to macromolecules. The similarity of these two problems (from the theoretical point of view) brings forward some suggestions about

¹ This investigation was supported by Contract N.o.n.r. 266(56), Washington, D.C.

things to look for: what, for instance, may be needed to make a macromolecule play the role of a transistor, or of a "tunnel" diode? This report is theoretical and will not touch upon the problems of experimental biology.

II. Electronic Waves in Periodic Structures; Band Structure

We shall start with a short summary of the theory, upon which the distinction between conductors and dielectrics has been based. This theory has been developed by the present author since 1930 and has been discussed in various papers (Brillouin, 1930, 1931, 1932, 1933) and books (Brillouin, 1946, 1953; Brillouin and Parodi, 1958).

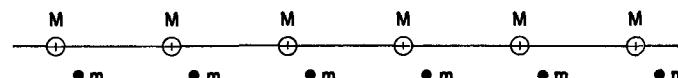


FIG. 1. Kelvin's model for optical dispersion.

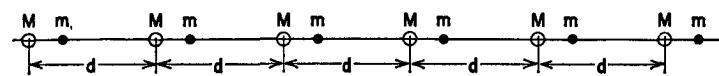


FIG. 2. A row of diatomic molecules.

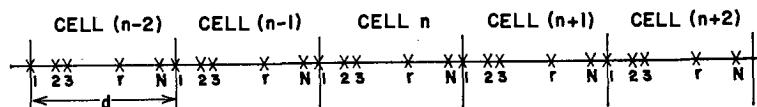


FIG. 3. A row of polyatomic molecules.

The most important part of the theory, when physical problems are considered, is related to wave propagation in crystals, that represent typical three-dimensional periodic structures. The three dimensions, however, introduce only some additional complications, and the main features of the theory may be more easily explained on periodic structures in one dimension. These results are the ones we may use for a comparison with long macromolecules.

We thus consider a long chain of atoms, which extends in one direction and repeats the same structure at regular intervals d . Figures 1, 2, and 3 represent three different examples of structures of this general type; these figures are self-explanatory.

The structure does not have to be along a straight line; it may just as well proceed in zigzag provided the complete pattern is repeated at regular intervals d .

Let us now study the propagation of waves along such a structure. Since we will be especially interested in electrons moving through the chain of atoms, we shall investigate the properties of de Broglie ψ waves, but the theory is general and applies to any kind of waves. In the books of the present author the discussion starts with elastic waves, where the physical model is easier to visualize.

Periodic structure of a system directly leads to band structure of the spectrum. This is the main idea developed in the author's books; it is best shown on a simple graph (Fig. 4).

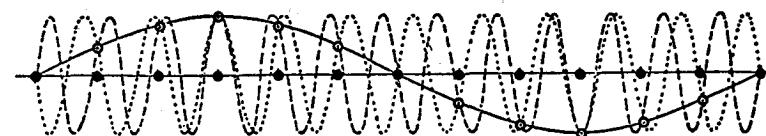


FIG. 4. Different sine curves passing through the position of the particles.

— sine wave for smallest value $a = a_0 < 1/d$

- - - sine wave for value $a_1 = a_0 + 1/d$

.... sine wave for value $a_{-1} = a_0 - (1/d)$

It is easily seen in this figure that the first two waves propagate to the right, while the third wave propagates backwards to the left; a_0 and a_1 are positive, but a_{-1} is negative.

Consider a chain of particles, with period d along the x axis; we can measure a certain quantity (displacement, electric polarization, etc.) only at the points where there is a certain particle, say, at points

$$x_0 = 0, x_1 = d, x_2 = 2d, \dots, x_n = nd \quad (1)$$

and we represent the motion by a wave

$$\psi = A \cos 2\pi(\nu t - ax) \quad a = \frac{1}{\lambda}; \lambda = \text{wave length.} \quad (2)$$

Remember ψ is defined only at the points belonging to the set (1). It is obvious that we can replace a by another quantity

$$a' = a + \frac{m}{d} \quad m, \text{ positive or negative integer} \quad (3)$$

without changing anything in the wave motion under consideration. This is easy to see, since a measurement at points x_n gives, with the waves (2) or (3)

$$\begin{aligned}\cos 2\pi(\nu t - a'x_n) &= \cos 2\pi\left(\nu t - ax_n - \frac{mx_n}{d}\right) \\ &= \cos 2\pi(\nu t - and - mn) = \cos 2\pi(\nu t - and)\end{aligned}\quad (4)$$

the quantity $2\pi mn$, in the phase, does not change the cosine.

Both waves in a or a' are equivalent for the representation of the propagating motion. Figure 4 illustrates the situation.

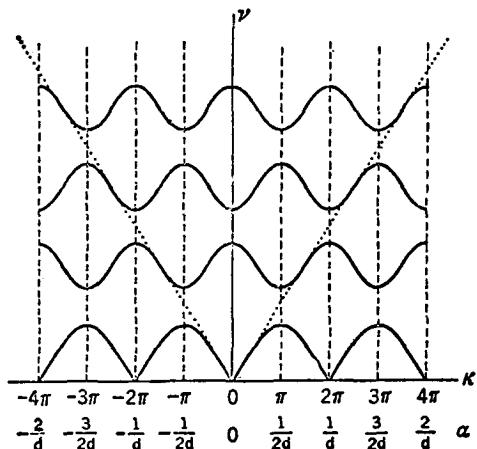


FIG. 5. ν as a function of k ($= 2\pi a$) or of a in a continuous medium;
— ν as a function of k or a in a periodic structure of period d along the x coordinate.

The condition (3) is expressed mathematically by the statement that the quantity a is defined only "modulus $1/d$." As a direct consequence, the frequency ν of the wave must take the same value for a or a' .

In a periodic structure (period d) the frequency is a periodic function of a ($= 1/\lambda$) with period $1/d$. This general theorem leads directly to frequency bands. Some plots and drawings will represent the best explanation of this relation. Figure 5 represents a typical example, where the frequency ν is plotted as a function of a (or of $k = 2\pi a$). For a homogeneous uniform medium, propagating waves with a constant velocity $W = \nu/|a|$, we obtain the straight (dotted)

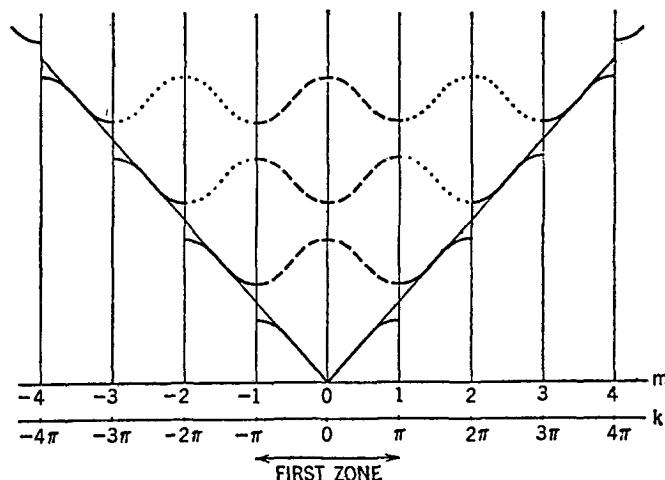


FIG. 6.

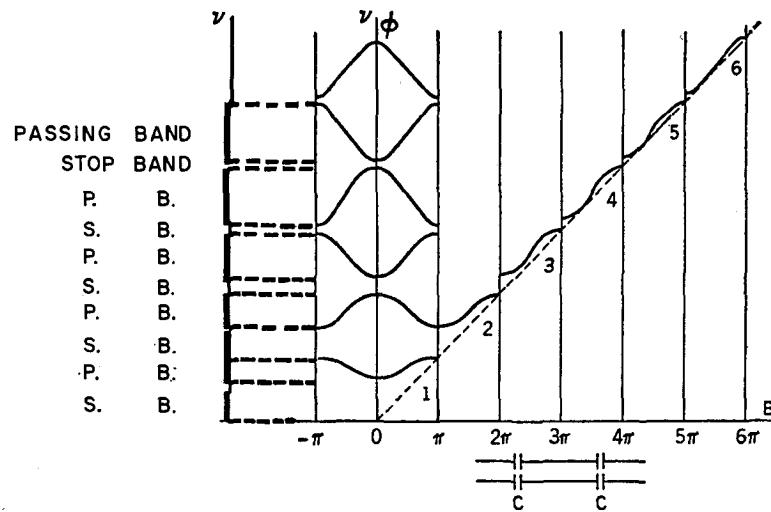


FIG. 7.

lines. If instead of a homogeneous medium we consider a periodic structure with period d we compute v as a periodic function of a with period $1/d$, represented by the solid curves. These curves fit rather closely to the dotted lines in the middle of the intervals. The solid curve corresponds to a system of frequency bands where some frequencies can be propagated, while some others do not correspond to any solid curve and cannot propagate. Figure 6 is another representation of the same situation. On account of the periodicity of the curves, we need not draw them completely, for a extending from $-\infty$ to $+\infty$, but it suffices to give the curve in one single period, for instance for $-(1/2d) \leq a \leq (1/2d)$. (This also means $-\pi \leq k \leq +\pi$.)

Another similar example is given in Fig. 7 and shows only the reduction of the curve to the fundamental interval $-\pi \leq k \leq +\pi$. The periodicity in a (or k) results in breaking the curve into a succession of periodic branches. There are *passing bands* corresponding to frequencies that may be propagated and *stop bands* for nonpropagating frequencies.

III. Pauli Exclusion Principle. Distinction between Conductors and Insulators

We may consider the same problem from another angle. Instead of discussing directly the properties of an infinitely long periodic structure, we may examine what happens when we start with one cell (considered as an elementary building block) and add progressively a second, a third cell . . . any number of similar cells, thus building up the whole chain link by link.

The results of the calculations are presented in Fig. 8, where a typical distribution of energy levels has been sketched. On the left we plot the discrete energy levels for one isolated cell and we number them I, II, III, IV, etc. Next, we assume two similar cells coupled together. Each of the preceding levels is split into 2 separate levels. For a chain of 3 coupled cells, we find groups of 3 levels; 4 coupled cells yield groups of 4 levels; n coupled cells give groups I, II, III, IV, each containing n close-packed levels. When the number n becomes very large, these groups of levels build up almost continuous passing bands which we may still number I, II, III, IV, etc., each passing band corresponding to one original level. Passing bands are separated by stop bands.

Such is the situation when the coupling between successive cells

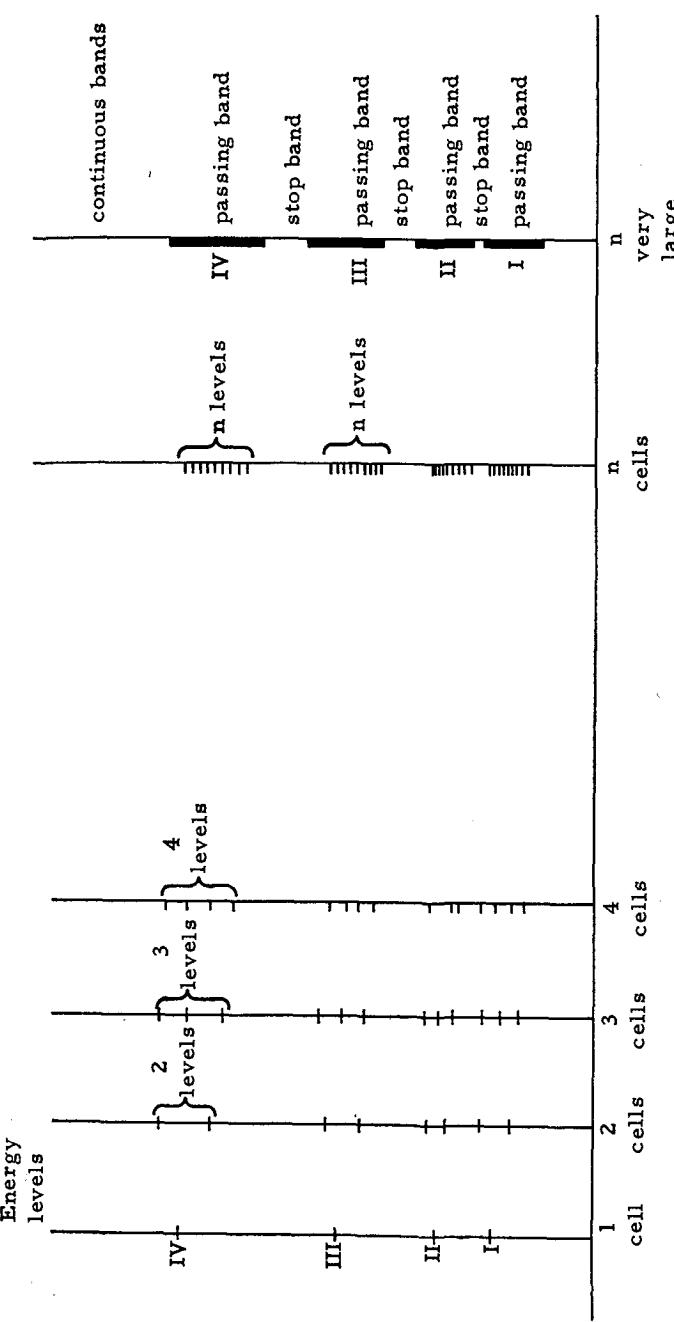


FIG. 8.

is not too strong. For very strong coupling, the successive bands I, II, III, IV, . . . may partly overlap and the description becomes more complex, but the main feature of the band structure is preserved.

We have in Fig. 9 another diagram visualizing the general results. Let us assume electrons to move through a succession of hills and valleys of potentials. We may have a few energy levels E_1' , E_2' , E_3' within the valleys. Their coupling from valley to valley is very small since it results only from tunnel effects through the hills. These levels will remain sharply defined, and the corresponding bands are very narrow. The higher level E_3' , lying near the top of

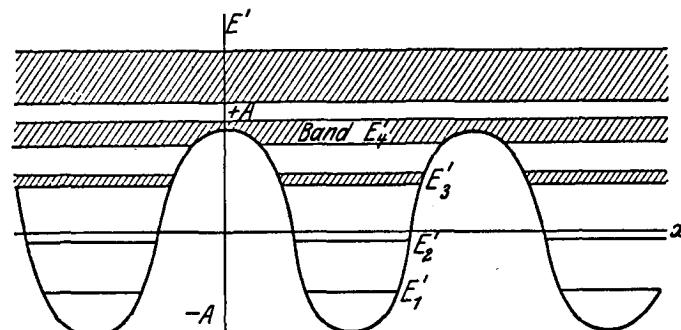


FIG. 9.

the hill yields already a slightly broader band structure. Levels E_4' and E_5' above the hills give very broad passing bands separated by narrow stop bands.

A similar band structure is represented in Fig. 10, where the frequency ν is plotted against the wave number a , running from $-(1/2d)$ to $(1/2d)$.

When we discuss the problem of electronic ψ waves, we must remember the *Pauli exclusion principle*. It states that each individual energy level may accommodate 2 electrons only (one with spin up, the other one with spin down). This means that each band I, II, III, IV, has room for $2n$ electrons on its n interval levels, when n is the number of cells in the chain. But how many electrons do we have to distribute on all these passing bands, and how are they distributed? We assume that the energy levels of the bands cover a total range of electron volts much greater than κT , the thermal

energy at temperature T . If we work at room temperature, κT corresponds to approximately 0.03 ev and our band structure may extend over many volts. Thermal agitation will only slightly perturb the most stable distribution, the one that would obtain at temper-

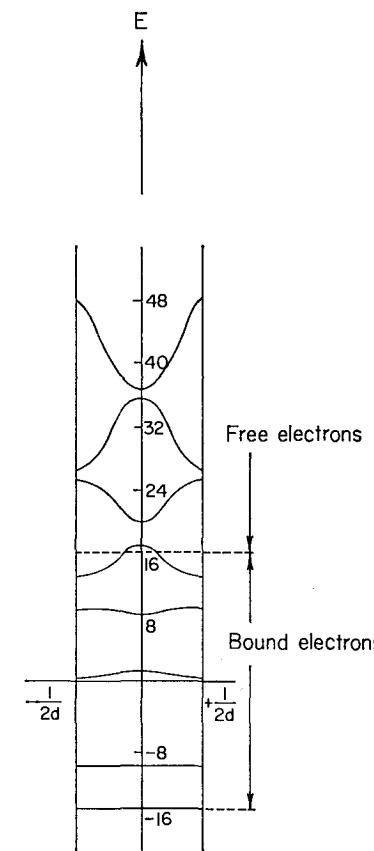


FIG. 10.

ature zero. Electrons accumulate on the lower levels, and fill up all the lower passing bands. We may assume bands I and II to be filled up with their full number of $2n$ electrons, and the highest electrons to accumulate on band III (for instance). We then have 2 cases to distinguish.

(A) The number of electrons per cell N is even ($N = 2p$). We

thus have a total of $2pn$ electrons to distribute and they will completely fill up a total of p passing bands. The preceding example corresponds to $p = 3$. The general result in this case is that all occupied bands, including the highest band are completely filled up.

(B) The number of electrons per cell N is odd ($N = 2p + 1$). Here we obtain $2p$ bands completely filled up, while the highest band is half filled and half empty.

Case A corresponds to an *insulator* (dielectric), since electrons jamming all bands have no possibility of moving. An electric field applied to the system has no effect whatsoever.

Case B represents a *conductor*; electrons in the half-filled upper band may move to empty levels and be set in motion by an electric field.

Note that we have assumed an ideal perfectly periodic structure, each cell being identical to any other cell. This ideal case corresponds to an ideal crystal structure, a perfection which is never realized. There are always some impurities in a crystal, or some irregularities in the structure (the crystal physicist calls it "mosaic-structure" and sees it on X-ray diagrams). The role of impurities will be discussed later and shall be proven very important. Here we may assume that it would yield an average number N of electrons per cell that would be neither exactly odd nor even but fractional.

The *Fermi level* is the highest energy level reached by electrons: in the ideal case A it reaches up to the top of band p ; in ideal case B it reaches to the middle of band $p + 1$.

IV. The Intermediate Case of Semiconductors; Donor and Acceptor Impurities

The preceding explanations make it clear that the conditions needed for an insulator are rather critical, and that the situation reveals a real instability. The upper passing band must be exactly filled to capacity: *no overflow*, that would bring some surplus electrons to an upper band; *no incomplete filling*, that would leave some holes in the band and thus allow some freedom of motion for the electrons. Two main causes of perturbation must be discussed: thermal agitation and impurities.

Thermal agitation may yield additional energies of the order of 0.03 ev at room temperature. This is very small compared to the width of the stop-band above the conduction band. Stop-bands usually amount to some electron volts, and are much too high for

thermal agitation to bypass them. Much higher temperature would be needed for this effect to play a role.

The situation is different with *impurities*. In a germanium crystal (valence 4) we have 4 free electrons per atom, just enough to fill exactly 2 passing bands. If we add in the crystal a few atoms of arsenic (valence 5) without disturbing the crystal lattice, these atoms will yield 5 free electrons each. The extra electron cannot be kept in the passing band, already filled to capacity; this electron must find room in a higher passing band where it will be free to move. The impurity atom is a *donor*.

We may, instead, introduce an atom of valence 3, boron for instance. This atom has one free electron less than germanium; it leaves a *hole* in the originally filled up band. This means a certain freedom of motion for the rest of the electrons. It has been proven that a *hole* behaves like a *positive charge* with a mass different from that of an electron. The foreign atom (boron) is called an *acceptor*.

With *donor* impurities, a small conductivity is obtained and is due to some free electrons in an upper band, this is a *n* (for negative)-type semiconductor. With acceptors, the small conductivity results from the motion of holes in the originally filled band. Holes behave like positive charges and the semiconductor is *p-type*.

These general ideas are the basis of the whole theory of semiconductors, and serve as the explanation of crystal-rectifiers, transistors, Esaki-tunnel-diodes, etc. Many textbooks have been published on these matters.

We discussed in the preceding section one-dimensional models, which will suffice for our further investigations. The technical semiconductors, which have taken on great importance in modern radio techniques, are crystals with three-dimensional periodic structures.

The extension of our preceding remarks to three-dimensional lattices reveals a curious complexity of geometric situations, which I characterized as "zones." These zones (now called Brillouin zones) were first discovered in 1930 (see Brillouin, 1930, 1931, 1932, 1933, 1946; Brillouin and Parodi, 1958) and have played a central role in all theories and applications of semiconductors.

V. Some Physical Problems Concerning Giant Molecules in Biology

Biochemists have been very busy investigating the structures and properties of giant molecules, such as proteins, ribonucleic acid

(RNA), deoxyribonucleic acid (DNA). Many of these molecules are built up like long chains of similar elements: they are almost periodic structures, and we may attempt to start a discussion from a periodic model that would represent a first approximation to the actual molecule. Thus we may establish a similarity between bio-molecules and periodic structures in one dimension. This assumption enables us to use many of the general results of the preceding discussion.

This was the point of departure of Szent-Györgyi in 1941 when he first assumed that giant bio-molecules might be semiconductors. He was especially interested in the need for an explanation of many experimental results.

Photosynthesis reveals that thousands of chlorophyll molecules react as one single functional unit, and may absorb, at distant points, the four quanta needed for the reduction of CO₂. This strongly suggests that excited electrons should move freely through the whole structure of chlorophyll molecules. Many experimental results prove that common energy levels exist within protein molecules. Energy absorbed at one point may be used to break a far distant link. In fibrous structures, it is quite reasonable to assume a quasi-periodic structure of many molecules in series.

A very strong argument is found in the *contraction of muscles*. The energy is derived from the splitting of ATP (adenosine triphosphate) at one point and is communicated to a great many myosin molecules along the muscle fiber.

Many experimental facts have been steadily accumulating since Szent-Györgyi's first remarks, and point to the need for some mechanism capable of transmitting energy throughout protein molecules. (See, for instance, a paper by R. Wurmser (1958) on the problem of transfers of chemical energies along protein structures, where many such examples are discussed.)

The same questions were recently discussed in Szent-Györgyi's excellent book (1957, 1960), where the author re-emphasizes the importance of his original suggestion, and strongly restates the similarity between bio-molecules and semiconductors.

The idea, first received with great favor, was almost abandoned when Evans and Gergely, in 1949, made a theoretical discussion of bands in proteins, and computed a width of about 3 ev for the forbidden band. This meant that thermal agitation would be perfectly unable to raise an electron from the fully occupied band to the upper

free passing band; the thermal energy would be $\frac{1}{100}$ of the energy needed. These first computations were rather crude, and will require corrections. Nevertheless, the order of magnitude of electron volts makes thermal excitation impossible.

This difficulty seems to have discouraged biochemists but we noted in Sections III and IV that the situation is exactly the same for crystalline semiconductors. Thermal excitation is not the main cause of excitation, but impurities, acting as donors or acceptors, represent the actual mechanism. Can we discover any similar situation in biochemistry?

VI. A Simplified Model of Semiconducting Molecule

In order to explain clearly our point of view, we will use the oversimplified model of Fig. 11. We consider a chain of C atoms, each one of them bearing the same number of H atoms. This represents

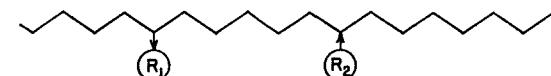


FIG. 11. The arrow shows the direction of the force acting on electrons, along the junction. R₁ is an acceptor, while R₂ is a donor.

a perfect periodic structure, and its electronic spectrum must exhibit the typical passing and stop-bands previously described. The lower passing bands will be filled up with electrons and the system will be a typical insulator. Now we may add a few side groups R₁, R₂ . . . replacing some of the H atoms. This is now a distorted periodic lattice, with impurities R₁, R₂ . . . The point is to show how such side-groups may play the role of donors or acceptors.

It is common knowledge that free molecules, branched upon a main chain, can be divided into 2 groups: electropositive and electronegative. Theoretically, this corresponds to the fact that there is an electric field along the junction between the molecule and the main chain, even when the molecule has no total electric charge. This field along the junction results from: (A) dipole, quadripole, . . . multipole structures within the molecule; and (B) uncompensated exchange effects.

Electropositive and electronegative types correspond to two opposite directions of the electric field between molecule and chain.

One type (R₁) will attract and fix (on the average) a negative

charge, in the chain, near the junction; the other type (R_2) repels (on the average) the negative electrons of the chain. The first case corresponds to *acceptors*, and the second one to *donors*.

There is a difference, however, with the impurities in a crystal semiconductor: each foreign atom introduced in the crystal may donate or accept one (or many) full electrons. In our new problem, the effects are not so sharply defined. One junction may, in average, fix or liberate a fraction of electron charge, and a few of them, working in parallel, will be needed to make up a full electron charge.

The whole character is, nevertheless, very similar in both problems: the groups may play the role of the impurities in usual semiconductors, and create along the chain a succession of regions of the "*n*"- or "*p*"-type.

Thus, a long bio-molecule must have properties very similar to those of a semiconductor with *n* and *p* regions in a certain order. Acceptors draw electrons from the filled band and create holes. Donors push electrons into a previously empty band just above the normal Fermi level. The new point is that it requires a certain number of electropositive or negative groups to do the job.

VII. Donors and Acceptors—Calculations by Pullman and Pullman

A great step ahead was made by Bernard and Alberte Pullman (1958) when they computed the properties of "donors" or "acceptors" of various molecules. This enabled them to state more precisely the role of these radicals in biochemical structures. The computation method is that of the molecular orbitals, which is described in their book (1952). The energy levels of the radicals are given by a formula

$$E = \alpha + k_i \beta \text{ with } \alpha, \text{ Coulomb integral and } \beta, \text{ resonance integral} \quad (5)$$

$k_i > 0$ corresponds to occupied levels;

$k_i < 0$ corresponds to empty levels.

Table I sums up part of the results.

These numbers represent values of k_i . However, one calculates $\beta \approx 3.26$ ev so that the values of k_i , multiplied by 3.26 yield electron volts. The differences in energy between the two columns yield a quantum of absorption $h\nu$, which can be observed in the optical spectrum, since this spectrum corresponds to raising one electron from an occupied level to an empty one. The energy of the highest

TABLE I

Compound	Energy of highest occupied molecular orbital	Energy of lowest empty molecular orbital
Adenine	0.486	-0.865
Guanine	0.307	-1.050
Hypoxanthine	0.402	-0.882
Xanthine (IV)	0.397	-1.197
Xanthine (V)	0.442	-1.005
Uric acid	0.172	-1.194
1-Methyl guanine	0.303	-1.064
9-Methyl guanine	0.302	-1.074
1-Methyl xanthine (IV)	0.397	-1.198
3-Methyl xanthine (IV)	0.354	-1.197
9-Methyl xanthine (IV)	0.394	-1.213
1-Methyl xanthine (V)	0.442	-1.009
3-Methyl xanthine (V)	0.395	-1.009
7-Methyl xanthine (V)	0.429	-1.049
1-Methyl uric acid	0.172	-1.201
3-Methyl uric acid	0.153	-1.204
7-Methyl uric acid	0.133	-1.200
9-Methyl uric acid	0.161	-1.204
Uracil	0.517	-0.960
Thymine	0.510	-0.958
Cytosine	0.595	-0.795
5-Methylcytosine	0.530	-0.796
Barbituric acid	1.033	-1.295
Alloxan	1.033	-0.757
Phenylalanine ^a	0.908	-0.993
Tyrosine ^b	0.792	-1.000
Histidine ^c	0.660	-1.160
Tryptophan ^d	0.534	-0.863
Riboflavin ^e	0.500	-0.344
Pteridine	0.864	-0.386
2-Amino-4-hydroxypyridine	0.489	-0.650
2,4-Diaminopteridine	0.544	-0.508
2,4-Dihydroxypyridine	0.653	-0.663
Folic acid	0.526	-0.647

^a The π -electron system of this molecule is assumed to be that of toluene.

^b The π -electron system of this molecule is assumed to be that of phenol.

^c The π -electron system of this molecule is assumed to be that of imidazole.

^d The π -electron system of this molecule is assumed to be that of indole.

^e The π -electron system of this molecule is assumed to be that of isoalloxazine.

The authors are indebted for this result to Dr. Karreman.

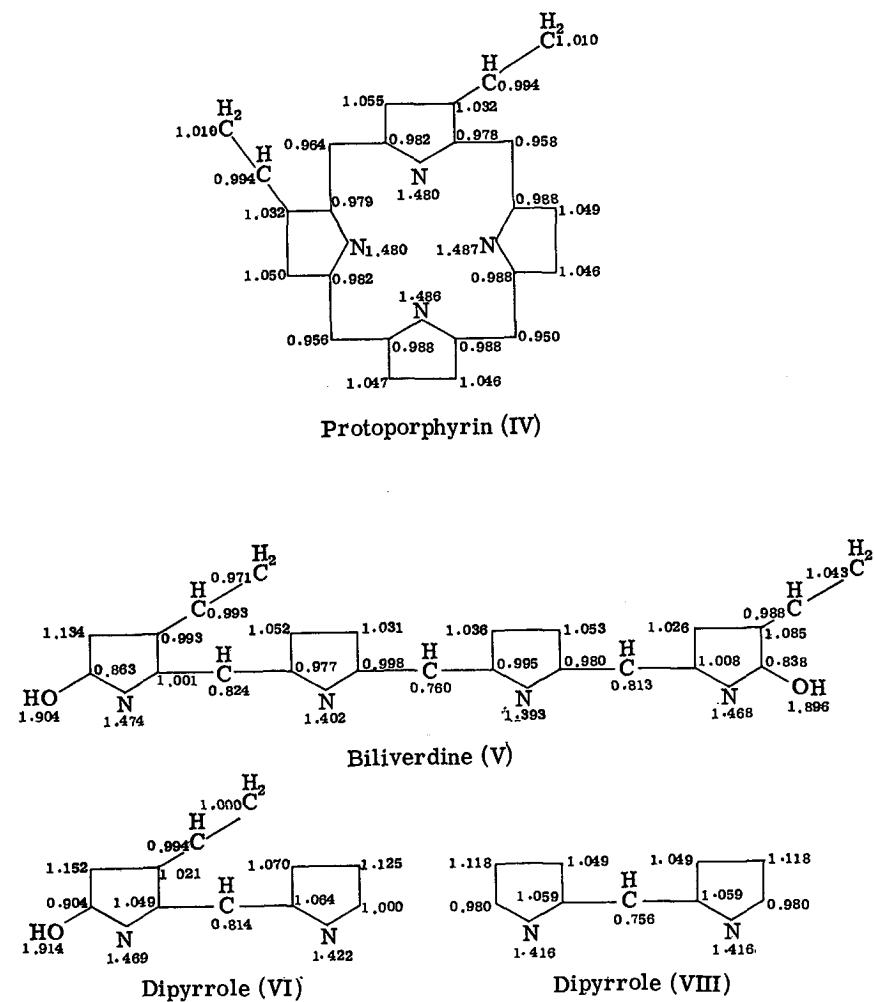
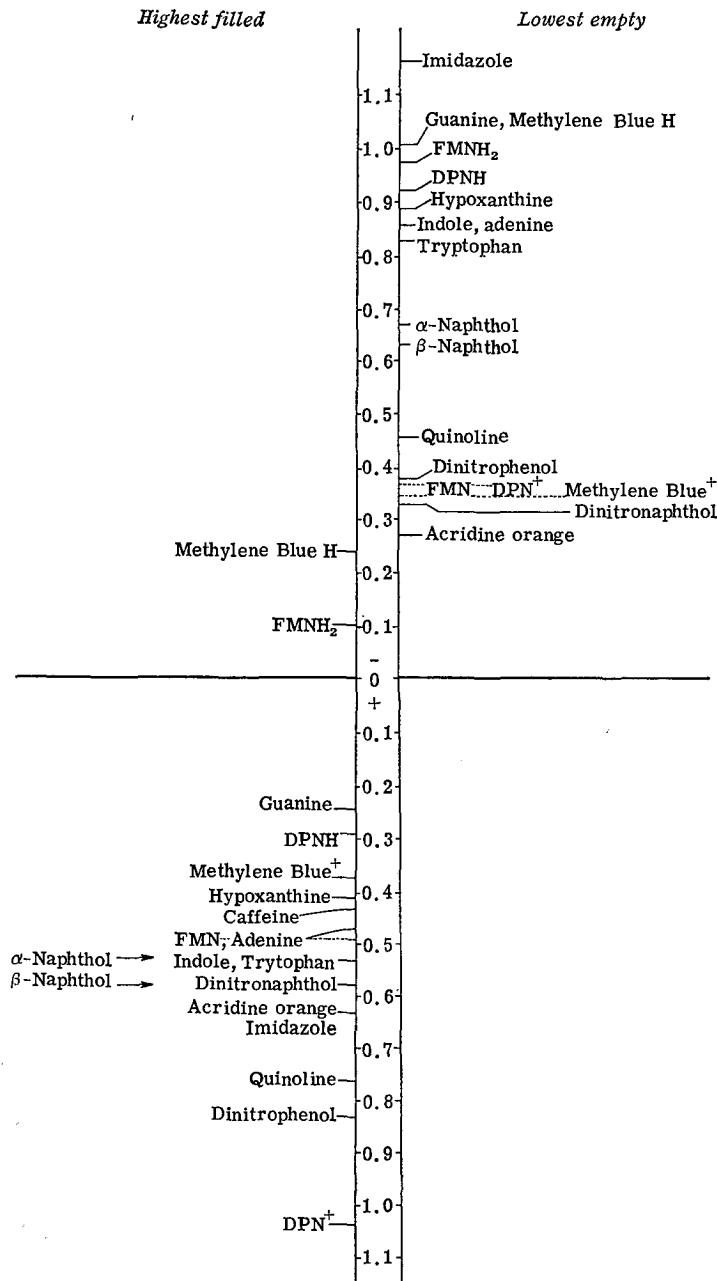


FIG. 12. Distribution of electronic charges.

occupied level represents ionization energy. The easier the ionization, the easier for the compound to play the role of electron donor. The energy of the lowest empty level corresponds to electroaffinity: the lower the level, the easier for the compound to "accept" an extra electron.

These authors have thus calculated the distribution of electronic charges in a series of structures (see Fig. 12 for various interesting

ENERGY OF MOLECULAR ORBITALS

FIG. 13. The k values of the highest filled and lowest empty molecular orbitals. From Szent-Györgyi (1960).

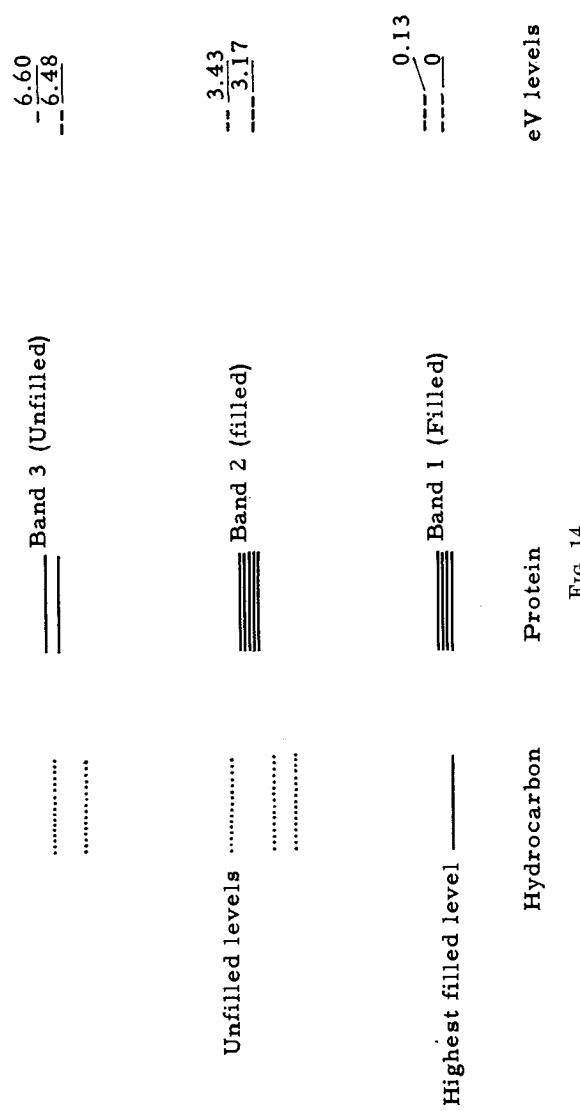


FIG. 14.

examples). From these distributions, it is easy to see that positive and negative charges are not distributed along exactly similar patterns. A structure, which is neutral, on the average, thus represents a very complicated electric multipole. This is one of the effects we mentioned above (Section VI) which, when combined with exchange effects about the junction show clearly that the molecule acts as donor or acceptor of electrons. Figure 13 shows another series of results presented by Szent-Györgyi (1960).

Isenberg and Szent-Györgyi (1959) have made many applications of the results found by Pullman and Pullman.

Other interesting results can be found in the thorough articles of R. Mason (1958a,b, 1959). Figure 14, taken from this author, gives an example in which a hydrocarbon (energy levels in the left column) will act as an "electron acceptor" with respect to a protein molecule (levels in the central column). Indeed, the electrons in band 2 (occupied) of the protein may fall into the empty levels of the hydrocarbon, which is practically at the same height. Figure 14 corresponds to a practical example discussed by Mason.

This diagram indicates the essential element of the transfer mechanism of an electron from a molecule to a protein structure. The reverse situation, in which the molecule would act as "donor" would require this molecule to have occupied levels at the very top of the diagram, facing band 3 (empty) of the protein.

VIII. Example of an Application—DNA

Let us try to explain, with a typical example, how the preceding remarks may be used for practical applications. The DNA looks like a sort of spiraled ladder. The sides of the ladder are phosphate sugars, while the rungs (or steps) represent organic bases (Fig. 15) (purine or pyrimidine of 4 different kinds) interconnected by links called H-bonds.

This structure principle of a ladder, twisted into a helix, was first discovered by Pauling and was later modified by Watson and Crick. Along this helical backbone, the arrangement of the 4 bases, up and down the ladder, constitutes a sort of 4-letter code. These four letters spell out long words in a very large vocabulary, and carry the information used in the biological reactions. We are, unfortunately, unable as yet to read out this coded information. Figures 16 and 17 give more details about the structure and come from an interesting paper by Perutz (1958).

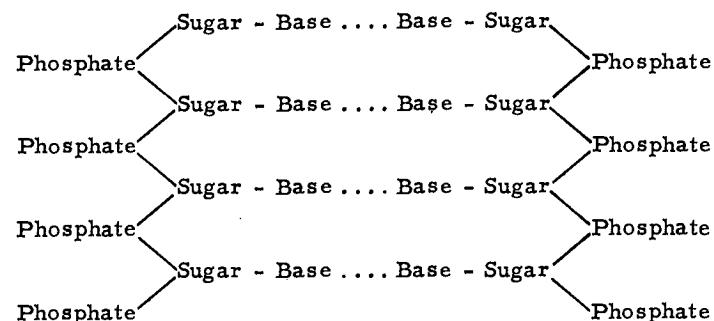


FIG. 15. Chemical structure of DNA. The compound consisting of one phosphate, sugar, and base is called a nucleotide. The dots, ..., represent a hydrogen bond.

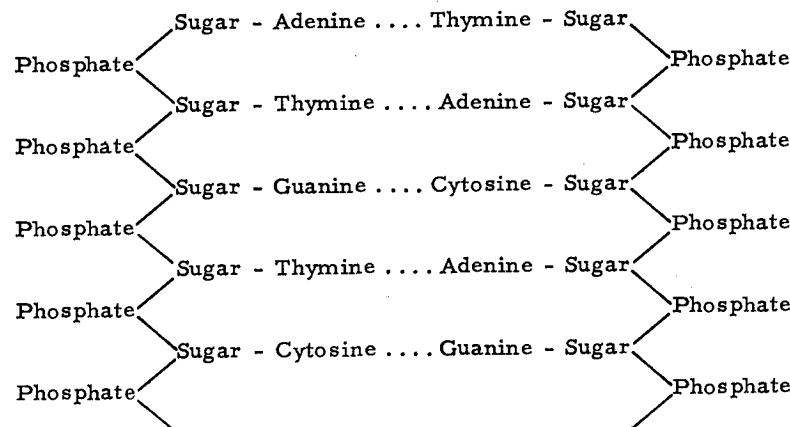


FIG. 16. Base pairing between two chains of the double helix of DNA. The dots, ..., represent a hydrogen bond.

Let us look again at the structure of DNA: it is a long periodic structure if we first ignore the bases and consider them as equivalent. We have:

I—a short period, L_1 , given by the height of each step of the ladder,

II—a longer period, L_2 , given by the height of one turn of the helix (one floor up on the spiraled staircase).

This doubly periodic structure goes on for a long vertical distance (the height of the total building) where it breaks, makes a

sharp angle, and starts in a different direction. Altogether it looks like a sort of pellet of crushed steel wire.

Let us discuss the problem of one long helical staircase. This is a one-dimensional periodic structure, with two periods L_1 , L_2 , and we may use the results of Section II, that enable us to predict a band spectrum. The wave number a [$a = (1/\lambda)$] exhibits two periods, a long one equal to $1/L_1$ and a short one equal to $1/L_2$.

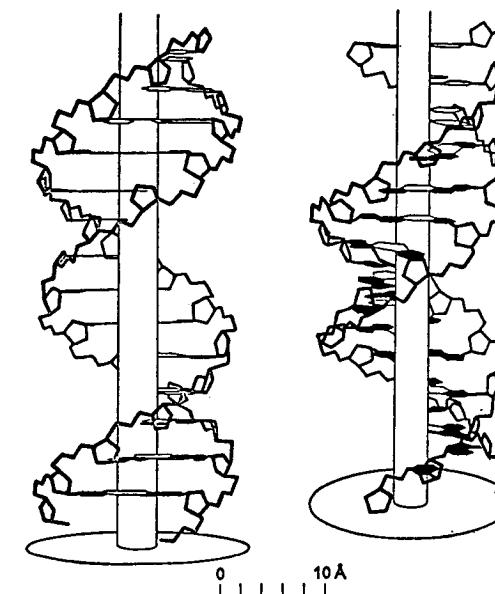


FIG. 17. The double helix of DNA. The rings and wire links signify the phosphate ester chain. The plates between them represent the bases. (Reproduced from F. H. C. Crick and J. D. Watson, *Proc. Roy. Soc.* **223**, 80, 1954.)

Propagation of the ψ waves (de Broglie electron waves) occurs up and down the staircase, along either side. If we did not have any connection between both sides (if all the steps of the ladder were broken), we would obtain the same average velocity of propagation for either side. But here the coupling occurs through the hydrogen bonds that build up the horizontal steps. These bonds have the following structure:



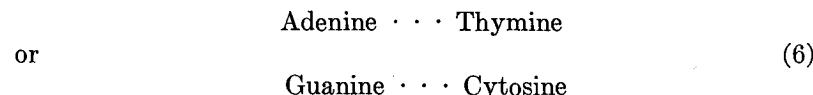
They are much weaker than the chemical bonds on the side of the

ladder. They provide a weak coupling between both sides, and result in doubling the bands; a first type of coupling yields bands due to ψ waves propagating parallelwise along both sides. A second band resulting from the coupling corresponds to ψ waves propagating in opposite directions along both sides. If we had no coupling, these two bands would be superimposed, and exhibit similar properties. The coupling splits these two bands from one another so that they yield two different frequencies for each wave number a .

Next comes the role of spins, which will again double the number of elementary bands according to spin orientation. Altogether, we may expect bands with a 4-fold structure; depending on the type of coupling this may yield two doublets or one singlet with one triplet. This type of singlet-triplet structure seems to appear in many biochemical examples.

IX. The Bases as Possible Donors and Acceptors

All these remarks are based upon the assumption of identical "bases," since in this first approximation we systematically ignored the differences between the 4 bases appearing in Fig. 16. These bases are grouped in pairs:



A special investigation should show whether these bases play the role of donors or acceptors. From the few data contained in Table I and Fig. 13, we suggest the possibility that the bases on the left (adenine and guanine) might be donors while thymine and cytosine would play the role of acceptors. This is just a crude assumption, selected for the only purpose of enabling a precise discussion. If this assumption would not prove correct, the resulting discussion could be easily modified accordingly.

Each of the two groups, in (6) would act as a sort of pump taking a fraction of an electron from the left side and pushing it to the right side. How big could this fraction be? We may look for information to the schemes of Fig. 12 where we find computations of electron numbers for a variety of atoms. Most N atoms exhibit fractions of electrons as high as 0.5 while C atoms have 0.1 to 0.2 electrons either in excess or in default. We remember that the H bond contains one very active N atom, and we assume that the fraction of

electron pushed from left to right in (6) may be about 0.5 electron. This is again a rough guess, and shall certainly be revised. If this order of magnitude obtains, it would need two hydrogen bonds acting in parallel to punch a complete hole in the filled band on the left side and push one full electron to an empty conduction band on the right. This would mean *p*-type semiconduction on the left and *n*-type on the right.

If the structure contains, on a certain step of the ladder, just one isolated junction of a certain type, with half a hole on the left and half an electron on the right, we would have an "exciton," where the hole and the electron should move together and could not be

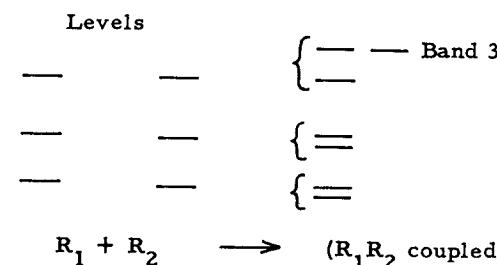


FIG. 18.

separated. The possible role of excitons in biomolecules has been repeatedly emphasized by Kasha. Here is a place where it might again be of importance.

Such are the general ideas suggested by a reading of the most recent articles. Many aspects remain to be discussed, from the theoretical as well as from the experimental point of view. Let us point out a possibility which does not seem to have drawn anyone's attention: a molecule may have rather high occupied levels, though they are still below the first empty levels of the protein. Such a molecule, if it is by itself, cannot be a "donor." However, the "donor" characteristic may eventually be produced by coupling between this molecule and a neighboring molecule (which cannot be a donor by itself either). Since originally they had levels close to each other, the molecules, when occupied, will form double levels whose higher one may be pushed up to band 3 (empty) of the protein. This is schematically indicated in Fig. 18. Many similar examples may occur.

REFERENCES

- Brillouin, L. (1930). *J. phys.* **1**, 377.
 Brillouin, L. (1931). "Quanten Statistik." Springer, Berlin.
 Brillouin, L. (1932). *J. phys.* **3**, 565.
 Brillouin, L. (1933). *J. phys.* **4**, 1, 333.
 Brillouin, L. (1946). "Wave Propagation in Periodic Structures," McGraw-Hill, New York (1953). Dover, New York.
 Brillouin, L., and Parodi, M. (1958). "Propagation des Ondes dans les milieux périodiques." Masson, Paris.
 Evans, M. G., and Gergely, J. (1949). *Biochim. et Biophys. Acta* **3**, 188.
 Isenberg, I., and Szent-Györgyi, A. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 1232.
 Mason, R. (1958a). *Brit. J. Cancer* **12**, 469.
 Mason, R. (1958b). *Nature* **181**, 820.
 Mason, R. (1959). *Discussions Faraday Soc.* **27**, 129.
 Perutz, M. F. (1958). *Endeavour* 190.
 Pullman, B., and Pullman, A. (1952). "Theories Electroniques de Chimie Organique." Masson, Paris.
 Pullman, B., and Pullman, A. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 1197.
 Szent-Györgyi, A. (1941). *Nature* **148**, 157.
 Szent-Györgyi, A. (1946). *Nature* **157**, 875.
 Szent-Györgyi, A. (1957). "Bioenergetics," pp. 9, 15, 22, 60. Academic Press, New York.
 Szent-Györgyi, A. (1960). "Introduction to a Submolecular Biology," pp. 3, 14, 48-75. Academic Press, New York.
 Wurmser, R. (1958). *Bull. soc. chim. biol.* **40**, 2091.

Is DNA a Self-Duplicating Molecule?

Some Observations on the Present Relationships of Biology to Chemistry and Physics

BARRY COMMONER

The Henry Shaw School of Botany and The Adolphus Busch III Laboratory of Molecular Biology, Washington University, St. Louis, Missouri

One of Albert Szent-Györgyi's greatest gifts to science has been his long and persuasive campaign to relate the phenomena of life to the behavior of the molecules, atoms, and subatomic particles which comprise the substance of living things. "We will really approach the understanding of life," he states, "when all structures and functions, all levels, from the electronic to the supramolecular, will merge into one single unit" (Szent-Györgyi, 1960). It is appropriate, then, to honor him by devoting ourselves to a consideration of how well we are learning to follow this exhortation.

What is to be gained by attempting to bridge the gap between biology and modern physics? The advantages are considerable:

Living things are exceedingly complex, inhomogeneous systems. Their properties vary with time. Certain of them are exquisitely sensitive to environmental changes; others resist all but the most brutal intrusions. Although the living system is highly organized—and is therefore a very improbable state—it not only manages to persist in all its improbability, but can expand its domain, and reproduce. Analysis of such a system sufficient to support prediction and control of its behavior is exceedingly difficult. That our present success is only moderate is evident from the rather little that can be done to alter the outcome of life. As it has for thousands of years, agriculture still consists of attempting to provide for animals and plants, only slightly different from those found free in nature, an environment suitable for growth. In medicine, our greatest successes derive from some knowledge of how to kill or remove harmful or diseased cells without unduly disturbing the human host. Transformation of life, changing the course of its development—as distinct

from simply killing, nourishing, or inhibiting it—is a largely unfulfilled goal.

What attracts us to physics is that it has succeeded where biology has seemingly failed. In the last 50 years, physical analyses of nature have yielded knowledge about the properties of matter that is so profound and so general as to disclose systems wholly new to human experience. As a direct result, nuclear and molecular transformations of matter hitherto absent from this planet have been produced on a massive scale. Moreover, the resultant body of basic theory—quantum physics—is inherently applicable to all forms of matter, and no evidence has arisen to dispute this claim. In that case we should expect modern physics to apply to life as well as to atomic and molecular systems and that its powerful theoretical and practical weapons can be used to breach the walls which have thus far protected the basic phenomena of life from experimental attack.

How well have these expectations been fulfilled? In the past 20 or 30 years there have been increasingly intensive efforts to develop experimental links between modern physics (and its related areas in the field of chemistry) and the problems of biology. The accomplishments are very great and well known, and it would appear that there is little to report but success. However, most of this work has related to those properties of life—such as energetic transformations—which it has in common with many nonliving systems. It can be argued, therefore, that these successes do not adequately test the usefulness of modern physics and chemistry in elucidating the *unique* attributes of life—reproduction and inheritance.

But now this wall too appears to be breached, and we have before us a growing body of physicochemical explanations of reproduction and inheritance. These enable a significant test of the idea that life—in all its uniqueness—can be closely correlated with what is known about the physicochemical properties of the components which comprise its substance. It is my purpose in this paper to examine some aspects of this relationship in the hope that it will help to improve our understanding of the lesson that Szent-Györgyi teaches.

In the last few years the application of physics and physical chemistry to this problem has given rise to a fairly simple and direct concept which proposes to explain the most fundamental property of living things—self-duplication. The basic features of this proposal are well known. It holds that reproduction and inheritance are in

the last analysis governed by a chemical substance, deoxyribonucleic acid (DNA). This substance has the capacity, according to the theory, to duplicate itself and to govern the specific chemical structures of proteins (for example, enzymes) which are responsible for establishing the pattern of chemical events characteristic of a given organism. A detailed physical structure has been proposed for DNA and its behavior in biological replication (the separation and complementary copying of complementary strands) is supposed to account for the biology of inheritance.

Various principles of physics have been employed to describe the structure of DNA, to account for its physical properties, and to calculate the probability requirements for the translation of its structure into a replica or into a protein analog. On the whole, these accomplishments appear to represent a quite successful example of the interaction between physics and biology.

But in one respect the picture is not so bright. From the purely biological behavior of living things (such as inheritance and development) biologists have, of course, already developed a series of principles which describe how living things behave and interact with their environment. Indeed, what little we can do to regulate the course of life—for example, induction of new organs in embryos—is due to the knowledge embodied in these principles. Biologists have placed considerable reliance on these principles, and it comes as a disappointment to discover that some of the most important of them have now been apparently contradicted by conclusions drawn from the new physical discoveries about DNA.

For example, the proposal that DNA is a self-duplicating molecule conflicts with one of the basic principles of biology—the cell theory—which holds that living substance cannot be reduced to any unit smaller than a single cell and yet exhibit the properties (such as self-duplication) characteristic of life.

Perhaps the most illuminating arena of conflict is that between current proposals regarding the properties of DNA and certain principles of development, a process uniquely associated with life. A major question at issue in the study of development is to determine how the numerous specific inherited features that appear in the adult arise during its formation from a highly undifferentiated precursor—the fertilized egg. The science of embryology was founded in the midst of a great controversy regarding the explanation of this remarkable phenomenon. There were those who, with

the aid of the then primitive microscopes, reported that a tiny fetus lay within the human egg or sperm. They believed that the features of the adult were already preformed, in miniature, within the germ cell and needed only to grow and expand to produce an adult. When microscopes were improved, and this view of the matter was overthrown by the actual facts, the concept of preformation was replaced by epigenesis, which holds that the characteristics of the adult are not present in the germ cell but are gradually elicited by the interaction of specific stimuli (either internal or external) with the pre-existing structure of the developing egg. Since then detailed experimental studies of development have proven, at least to the satisfaction of the embryologists, that a great many specific developmental processes are epigenetic rather than preformationist in character.

Current proposals regarding the genetic function of DNA are in conflict with the concept of epigenesis. The proposal that the inheritable features of the adult are encoded in the DNA of the germ cell substitutes for the miniature features of the preformationist's homunculus, *representations* of them in the form of nucleotide sequences. The theory holds that in development the pre-existing code is translated, without loss or gain of information content into the overt characters of the adult. It calls for a one-to-one correspondence between the final features, and their encoded representation in the zygote. It is in effect what Elsasser (1958) calls a "modernized version" of the preformationist principle. In the language of information theory, the DNA-code hypothesis holds that the information content of the adult is wholly derived from the information content of the germ cells' DNA, while epigenesis means that some other sources of information must arise during the course of development. The preformationist basis of the DNA theory has been put quite directly by Quastler who states ". . . there is a very real homunculus in every sperm and in every ovum; but it does not look at all like a little man any more than a description looks like the object described" (Quastler, 1959).

There is, therefore, an unmistakable disagreement between the conclusions about the mechanism of replication that are derived from biological evidence, and those which have been suggested by recent applications of physics to this question. Does this disagreement represent a denial of Szent-Györgyi's hope for unity between biology and modern physics? Or does the disagreement mean that

one of the two conflicting conclusions is in error? Possibly epigenesis is only a crude approximation, derived from the necessarily inexact data of embryology, which must now give way before the more refined and precise evidence of DNA structure. Alternatively, the DNA-code theory might be wrong.

In what follows I shall examine briefly the sources of this conflict, and possible means of resolving it.

To begin with, it will be useful to define the area of disagreement as precisely as possible. This requires that we understand quite explicitly what is meant by the general term *self-duplication*, and by the particular idea of a "self-duplicating molecule."

In biology, the term self-duplication describes a special feature of the living cell's capacity to replicate—autonomy. Although cellular replication occurs only in a suitable environment, which provides the substances necessary for growth, the *outcome* of replication—that is the kind of inheritable characteristics which appear in the daughter cell—is not appreciably affected by the composition of the environment. According to the biological evidence, self-duplication is a process of replication in which the inherited features of the progeny are controlled *solely* by the features of the parent, even though the process will occur only in a specified environment. There are cases in which the environment may exert some influence (e.g., "dauermodifications"), but these appear to be due to the effects on the relative rates of multiplication of intracellular structures, such as chloroplasts.

Accordingly, to establish that DNA is a "self-duplicating molecule," it must be shown that in the process of DNA replication within the cell (or in an *in vitro* system), the newly formed molecule acquires its specific nucleotide sequence *only* from the nucleotide sequence of the parental molecule. It must be shown, as well, that although the cell provides an environment which is specifically required by the process, no feature of the intracellular environment contributes to the structural specificity of the new polynucleotide strand. That this moderately rigorous requirement is a justifiable test of the proposal that DNA is the bearer of all genetic information has been tacitly admitted by one of the founders of that theory. Crick states that a major requirement of the DNA hypothesis is the principle that while information which is acquired by nucleic acid may be recovered from it, information transferred from nucleic acid to protein can never be retrieved from the latter (Crick,

1961). Hence, according to this fiat, which Crick terms "the central dogma,"¹ protein cannot contribute any information in DNA synthesis, thereby foreclosing the chief source of possible non-DNA information in DNA replication.

Our problem can now be stated more precisely: What support can be found for the proposal that in DNA replication, the genetic specificity (nucleotide sequence) of the newly formed polynucleotide is wholly derived from the nucleotide sequence of the parental molecule, i.e., that no other source of specificity is essential to the process?

Clearly the most satisfactory way to answer this question is the direct one: An experimental description of the physical processes which occur in DNA replication.

Unfortunately, the available experimental evidence on DNA replication, while highly informative in general, does not bear decisively on this particular question. None of the very interesting *in vivo* experiments (e.g., the appearance of 1:1 DNA isotope hybrids after a few cycles of replication in microorganisms) are relevant. Such results are indicative of some rather precise mechanism of DNA replication and may show that the specific composition of the pre-existing DNA participates in the processes that determine the structure of the new DNA molecules. But there is in these data no basis for ruling out the possibility that some of the other structurally specific and biochemically active agents of the cell—enzymes for example—also contribute to the determination of the new DNA molecule's nucleotide pattern. Nor can this difficulty be eluded by asserting that the specificity of any enzyme which participates in DNA replication is itself the product of a specific DNA. Such a proposal presupposes that only DNA serves as the vehicle of biological continuity, so that the specificity of all other cellular components must be ultimately derived from that of DNA. But there is, of course, no empirical basis for such a supposition. All cells are derived from other cells which contain relatively large amounts of cytoplasmic substance other than DNA and no cell which possesses a full complement of DNA but lacks a considerable

¹ Crick states that this term, rather unexpected in scientific discourse, has been introduced to emphasize the "speculative nature" of the proposal. But the fact is, of course, that dogmas serve the opposite purpose, and are commonly employed to discourage speculation.

cytoplasm—a sperm cell, for example—can divide and serve self-sufficiently, as a vehicle of biological continuity.

The DNA-primed synthesis of DNA in *in vitro* enzyme systems provides a hope—as yet unrealized—of decisive evidence on the physical origin of specificity in DNA replication. The properties of such systems do show that the nucleotide composition of the newly synthesized DNA is under the influence of the composition of the primer; but again there is no evidence that the specificity of the primer DNA, while a necessary factor, is a *sufficient* one. Several specific enzymes are also present, and in their absence no new DNA is formed. The possibility that the nucleotide composition is *jointly* determined by the DNA primer, and by the enzymes, cannot be excluded as yet. Some evidence suggests that polynucleotides with simple but non-random nucleotide sequences may be formed in the absence of primer. This would indicate that the enzymes must exert some degree of control over sequence, and that DNA replication is not achieved by self-duplication.

In the absence of direct physical evidence regarding the mode of replication of DNA, we must fall back on less satisfactory tests for the validity of the theory that DNA is a "self-duplicating molecule." Indeed, at this stage in the discussion one can argue that since such direct evidence is lacking the proposal should be regarded as thus far failing of proof, and conclude that it does not yet merit the right to challenge the concept of epigenesis—which is, after all, a well-established and hitherto undisputed principle of biology.

But there are, of course, important sources of support for the theory of the self-duplicating DNA molecule. In the main these derive from the fact that a structure of the DNA molecule has been developed which conforms with considerable physical evidence, and which also provides a *model* of how self-duplication *might* occur. The model is reasonable because it resembles a process—template copying—known to us in certain macroscopic realms of the physical world. Furthermore, certain elementary aspects of information theory—which is firmly based on well-established principles of physics—may be employed to describe possible mechanisms whereby information encoded in DNA can be transferred to a daughter DNA molecule and translated into a protein or RNA analog. In general, the strength of the hypothesis that DNA is a self-duplicating molecule derives from the physical and chemical

plausibility of its structure, and of the models for the process of self-duplication which are derived from this structure.

Thus, to continue the search for a means of resolving the apparent contradiction between the biologically established principle of epigenesis, and the hypothesis of DNA-encoded preformationist determination, we need to examine the degree to which the latter finds support in the principles of physics and chemistry.

We owe to Elsasser (1958), the initial development of considerations which provide rigorous and meaningful physical tests of the DNA-code model. He points out that the physical requirements for a successful code library can be specified from the now considerable experience with information storage in modern computers. The chief problem is "noise," i.e., the intrusion of random, thermal fluctuations into the code structure, which irreversibly destroy its information content. In computer practice codes are imprinted on mechanically stable structures such as tapes, cards, or wires. Elsasser shows that in a code library constructed on the molecular level, the problem of avoiding loss of information through thermal interactions is virtually insuperable. This conclusion is supported by current engineering experience with the smallest known code libraries (cryotron memory devices)—which are of course still vastly larger and more thermally stable than a molecular structure. Even in these systems thermal loss of imprinted information is so serious, that it must be combatted by using high levels of redundancy and by linking the device to a circuit which periodically regenerates its original information content from the remaining redundant elements of the code (Ittner & Kraus, 1961). In other words, on relatively simple thermodynamic grounds, we cannot expect DNA, or any other molecule to serve as an effective store of information.

Wigner has recently considered this problem from the viewpoint of quantum mechanics (Wigner, 1961) by evaluating the conditions required for the existence of a self-reproducing structure. The procedures used involve quantum mechanical descriptions of the relevant physical states. The term N is defined as the number of components which must be specified to describe *any* possible state of the substance present in a living system, i.e., the "space" of the organism is N -dimensional. The space of the nutrient not incorporated by the organism has R dimensions. Of the many possible state vectors in N -dimensional space, only a very small number (n) corre-

spond to living states that are capable of self-duplication. Wigner examines the dependence of self-duplication on the attributes of a quantity, S , which represents the interaction between the states of the substance that makes up the organism, and the states of the substance that makes up the nutrient.

In ordinary quantum mechanical relationships, S would be related to a random matrix with no specific properties. If this condition obtains, it can be shown that if reproduction is to occur N and n must be related by an equality in which the terms N^2 and n are on opposite sides of the equation. Since N must be very much larger than n , the equality cannot be satisfied under this condition. Wigner shows that the necessary equality can be obtained only if the interaction term S is not random, but highly specific. He concludes, "We arrive at the result that, if the interaction S is not 'tailored' so as to permit reproduction, it is infinitely unlikely that there be *any* state of the nutrient which would permit the multiplication of any set of states which is much smaller than all the possible states of the system." Since S represents the history or the change in the state of the reproducing system with time, this conclusion indicates that the outcome of reproduction is specified not only by the initial state of the organism, but by a *particular* time-course of interaction with the environment in which it duplicates. In biological terms this means that the specificity of reproduction is inherently dependent on the specificity of *development*, a condition which is of course met by ordinary living things. However, this condition does not play a part in the template model of molecular self-duplication, for in this case the specificity (information) is wholly contained in the structure of the initial DNA molecule and is simply transferred from molecule to molecule by a process (complementarity of structure) which does not permit the entry of specificity into the system through developmental interaction with other agents. Wigner's argument suggests that "molecular self-duplication"—in which all the specificity is autonomously contained in molecular structure—is an "infinitely unlikely" event.

These considerations indicate that the proposed model for DNA self-duplication comes into conflict not only with some principles of biology, but also with certain relevant principles of physics.

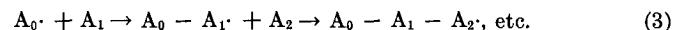
We can also look for confirmation—or contradiction—of the theory of DNA self-duplication in the results of the extensive experience with the *chemical synthesis* of polymers. The nonbio-

logical substances that most closely resemble polynucleotides and other biological macromolecules, are the artificial linear polymers that can now be synthesized in such profusion. It will illuminate our problem to summarize certain generalizations about the processes involved in the synthesis of such polymers.

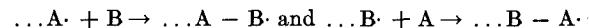
Most pertinent are the mechanisms which govern the formation of a linear copolymer, i.e., a polymer composed of two different types of monomer units (A and B). Typically such polymerizations are initiated by adding a free radical catalyst ($X\cdot$) to a mixture of A and B (1:1, let us say). Two reactions then occur:



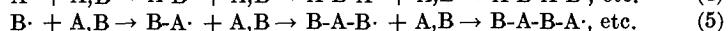
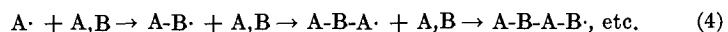
The two products ($A\cdot$ and $B\cdot$) are highly reactive free radical forms of the monomers, each of which can propagate a chain reaction, and thereby form a linear polymer. The simplest result is the formation of an homologous polymer from:



In the 2-monomer system, four different reactions are possible, the two homologous chains such as (3) above, and the two heterologous chains:



The kind of polymers that are actually formed in such a system depends on the relative concentrations of A and B and on the relative values of the rate constants for the 4 possible reactions. There are well-known examples in which the rate constants for the heterologous events are very much greater than the constants for the homologous events. In this case, only two of the 4 types of processes will proceed at significant rates, and the chain initiators, $A\cdot$ and $B\cdot$ formed by the initial events (1) and (2) will react as follows:



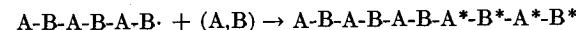
In this system the propagation of chain reactions will produce from a mixture of monomers A and B two specifically structured polymers:



and



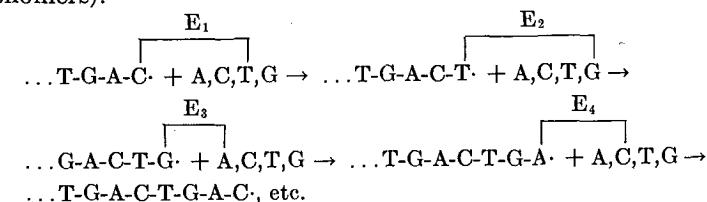
To what extent does such a system resemble a process of replication, as we understand it in biochemistry? The process is autocatalytic, in the sense that the free radical terminals catalyze the incorporation of monomers into the growing polymer. To demonstrate the inherent replicative character of the system we need only consider the consequence of, let us say, adding a short-growing chain (which terminates in a free radical) to a fresh mixture of monomers:



where the starred segment of the polymer is a newly synthesized replica of the older part.

These examples are cited only to show that linear polymer synthesis exhibits the rudiments of replication. The system is readily elaborated into something which resembles the empirical observations of polymer replication in biological systems, if we add to the rudimentary events (a) 4 or 5 monomers rather than only 2, and (b) catalysts which regulate the rates of specific types of additions.

In such a system, processes of this type might occur: (here we shall, for emphasis, employ the designations A, C, T, G for the monomers).



The enzymes involved in this system have been given certain arbitrary but plausible specificities. E_1 is an enzyme which specifically catalyzes the addition of T to a terminal C; E_2 catalyzes the addition of G to a terminal T; E_3 catalyzes the addition of A to a terminal G; and E_4 catalyzes the addition of C to a terminal A.

This system, in the oversimplified form illustrated above will accomplish the linear replication of the sequence $\dots T-G-A-C \dots$. At this level of complexity a system will of course yield only very limited patterns. But this is readily overcome by introducing catalysts for monomer addition which depend for their specificities, not only on the terminal monomer of the chain, but on the couple which includes the penultimate monomer. (Such effects have actually been

observed in polymer chemistry.) Since there are 16 possible couples, the complexity of the resultant polymer increases quite sharply, and may become adequate to produce chains of the type found among natural polynucleotides.

The chief purpose of the foregoing exercise has been to suggest that what is known about the synthesis of linear polymers in non-biological systems leads to the interesting possibility of replication by linear extension of a monomer pattern which is determined in part by the specificities of a complex system of reactions. As presented, the system has obvious and numerous inadequacies as a model for biological replication of polynucleotides: How are molecule lengths determined? What mechanism dictates the apparent biological inactivity of shorter-than-normal molecules? And many others. But just as in the case of the template model of DNA replication, difficulties of this type are readily overcome—at least in theory—by ingenious speculation.

But it is not my purpose here to urge the adoption of a new model of DNA replication. We are simply engaged in a search for points of contact between the problem of DNA replication and modern physics and chemistry. What we learn from the foregoing considerations is that the mechanisms which govern the synthesis of artificial polymers have the following properties relevant to our problem: (a) The macromolecule is formed by the *terminal* addition of monomers, and is synthesized *linearly*. (b) The processes are fundamentally autocatalytic and therefore possess the rudimentary requirements for a replicative system; the self-propagating element is represented by the terminal activated monomer, which is frequently in a free radical form. (c) In complex copolymers, the sequence of monomer residues in the resultant molecule is determined, in part, by the rate characteristics of the reactions which the polymer terminal can undergo with the available free monomer residues. (d) The specific structure of the resultant polymer may be governed *jointly* by the nature of the primer, and by the specific properties (especially rate-determining characteristics) of the catalysts present in the system.

The linear model differs from the conventional DNA template model in certain important respects. In the latter, the forces which specify the polymer structure are exerted laterally, and the template hypothesis requires no other restraints on linear sequence. Replication is direct, depending only on the parental structure; it is therefore basically preformationist. All of the information content

of the parental molecule is transferred directly, and essentially simultaneously, to the progeny.

In the linear model, replication is indirect. It is regulated, not by the total structure of the parental molecule, but by a relatively limited terminal segment. That this small amount of information suffices to specify the whole, is due to the fact that the primer interacts with a highly complex synthetic system, which contributes a great deal to the specificity of the final product. The active end of the primer molecule can be regarded as a stimulus (the specificity of which is determined by the entire polymer pattern of which it is the terminal result), which in turn sets off in a pre-existing system a series of chain reactions, so constructed as to yield in the end a replica of the original polymer molecule. In this model the parental molecule can not be regarded as capable of self-duplication, because part of the requisite information is contained in the specific catalytic system, and part in the structure of the parent molecule itself. In sum, the linear model of polymer replication is basically epigenetic rather than preformationist.

The position then, is as follows: There is no direct experimental evidence to support the concept that DNA is replicated by a process of self-duplication, i.e., that the nucleotide sequence of the product is governed *solely* by the nucleotide sequence of the parental molecule. The concept that DNA governs biological inheritance by virtue of a process of self-duplication is in conflict with certain well-established principles of biology, epigenesis in particular. However, the concept is a plausible extension of what is known about the physical structure of the DNA molecule, providing that it can be shown that a lateral, template-like alignment of parent polymer and newly acquired nucleotides occurs. But rigorous examination of the physical problems involved in template-like self-duplication shows that this process is exceedingly improbable, as is any process involving information storage on a molecular level. Furthermore, the mechanisms which govern the synthesis of artificial polymers, which are autocatalytic and in a rudimentary sense replicative, are linear rather than lateral, and derive their specificity in part from the properties of the reactive system, rather than from the structure of the parental molecule alone.

Thus, if we seek to adjudicate the conflict between the concept of DNA self-duplication and the biological principle of epigenesis by an appeal to the experiences of physics and chemistry—we discover only support for the nonself-duplicating models of DNA

replication. Indeed, the specific evidence from polymer chemistry suggests a particular nonself-duplicating mode of replication—linear extension—which conforms quite well to the requirements jointly established by the evidence from biology, physics and chemistry.

It remains only to be said that the concept of epigenetic replication of biological macromolecules is by no means new. Such schemes have been suggested by several authors, Claude (1949) and Hinshelwood (1956) in particular, on the basis of biochemical evidence. Hinshelwood, for example, points out that the complex cyclical chemical systems which are so characteristic of life, are basically similar to chain-reacting processes, which are autocatalytic and replicative, but in which *self*-duplication of separate molecules does not occur. In such systems, replication of a new molecule is partly under the control of the specificity of a pre-existing molecule, but the relationship is indirect, and also requires a series of chemically specific steps. A very simple example is the formation of trypsin by means of trypsin-catalyzed conversion of trypsinogen. A more complex example is the adenosine triphosphate (ATP)-dependent synthesis of ATP which arises from the fact that in the mechanisms of glycolysis, each glucose molecule is oxidized only after it has been phosphorylated at the expense of 2 ATP molecules. But once activated in this way, the glucose molecule gives rise to twice that many ATP molecules.

Such indirect cyclical processes are, of course, quite characteristic of cellular chemistry and as Hinshelwood points out: "A system of mutually dependent parts, each of which performs something like enzymatic functions in relation to another, will as can easily be shown, in the steady-state appear as a whole to be autosynthetic. No individual part need be credited with a new and mysterious virtue by which to duplicate itself" (Hinshelwood, 1956).

In effect, if we are to generalize from the rather extensive experience regarding the self-dependent synthesis of specific substances in biological systems, the operative mechanisms are indirect, and the specificity of the final product depends on the specificity of a series of interconnected reactions. Such a system is epigenetic in character rather than preformationist.

It is pertinent that recent investigations of the biosynthesis of tobacco mosaic virus (TMV) provide direct evidence for an epigenetic mode of replication (Commoner, 1959). In brief, it can be shown that the TMV rod is synthesized lengthwise, the ribonucleic acid (RNA) fiber and the surrounding protein subunits in a particu-

lar linear segment of the virus rod being formed simultaneously. The only plausible explanation for these empirical observations is that a relatively short terminal segment of RNA helps to determine the structure of a protein subunit (which is laid down at the terminal step-dislocation in the helical linear crystal in which the subunits occur) and that the newly deposited protein subunit in turn helps to determine the choice of new nucleotides in the terminal extension of the RNA fiber. Such a process is epigenetic, in the sense that the replication of a specific RNA nucleotide sequence is determined in part by the structure of a short terminal segment of the RNA fiber, and in part by the chemical specificity involved in the enzymatic processes of protein synthesis. That the virus RNA fiber is formed linearly, seems to preclude its replication by means of a lateral template-like process of self-duplication.

These considerations lead me to doubt the validity of the hypothesis that reproduction and inheritance of living things is governed by the self-duplication of DNA (or for that matter, of any other molecule). The evidence arises from a broad range of sources: (a) There is no direct experimental evidence to show that the nucleotide sequence of DNA produced during biological replication, the supposed structural source of its genetic specificity, is determined *solely* by the nucleotide sequence of a parental DNA chain. (b) Essential features of the proposal that DNA is a "self-duplicating code" conflict with basic principles of physics. (c) The chemical mechanism which is supposed to account for DNA self-duplication—i.e., lateral template-like orientation of new nucleotides—is unknown in the chemistry of nonbiological polymers; indeed the latter suggests an alternative, linear, nonself-duplicating mechanism for replication of biological polymers. (d) The proposal that inheritance and development is governed by DNA self-duplication conflicts directly with certain biological principles, epigenesis in particular, that have been firmly established by observations and experiments with living things.

Are we to generalize from the outcome of this inquiry and conclude, as well, that the hope—so appealingly expressed by Szent-Györgyi—of a unified understanding of the living thing and its molecular, atomic, and subatomic components, is not to be realized? Such a conclusion would I believe be premature, for it presupposes that the theory of DNA self-duplication is in fact a good example of what modern physics can tell us about biological replication. But from what has already been said, it is apparent that this proposal

does not come off with particularly good marks in chemistry or physics, and it is therefore not a fair test of the idea that physics and chemistry can be unified with biology.

As I have pointed out elsewhere (Commoner, 1961), one of the most penetrating attacks on the hypothesis of DNA self-duplication is derived from Bohr's views of the relationship between living systems and the principle of complementarity—which is a most basic view of the nature of matter. There appears to be some danger of error in the practice of exploiting the marvelous experimental tools which have been produced by modern physics, without at the same time accepting the intellectual guidance of the principles which are themselves the source of these tools.

The foundation of the concept of DNA self-duplication in biological theory is also faulty. Certain features of Mendelian genetics have an apparent preformationist character, and therefore suggest that they could be governed by a molecular "code." To this extent there is, of course, biological support for the DNA hypothesis. However this type of inheritance is a limited, and probably dependent, facet of biological replication, which includes as well the epigenetic phenomena of development, of cytoplasmic inheritance, and of non-Mendelian nuclear inheritance. Part of the weakness of the DNA hypothesis results from its failure to consider *all* of the principles of inheritance and development.

These deficiencies suggest a remedy. To build a successful bridge between biology and physics the structure must be equally based on two essential foundations: The basic principles of physics, and the basic principles of biology. This task still lies before us.

REFERENCES

- Claude, A. (1949). *Advances in Protein Chem.* **5**, 423.
- Commoner, B. (1959). *Nature* **184**, 1998.
- Commoner, B. (1961). *Science* **133**, 1745.
- Crick, F. H. C. (1961). *Symposia Soc. Exptl. Biol.* **12**, 138.
- Elsasser, W. M. (1958). "The Physical Foundation of Biology." Pergamon Press, New York; (1961). *J. Theoret. Biol.* **1**, 27.
- Hinshelwood, C. (1956). *Proc. Roy. Soc. B* **146**, 155.
- Ittner, W. B., III, and Kraus, C. J. (1961). *Sci. American* **205**, 124.
- Quastler, H. (1959). *Lab. Invest.* **8**, 480.
- Szent-Györgyi, A. (1960). "Introduction to a Submolecular Biology." Academic Press, New York.
- Wigner, E. (1961). In "The Logic of Personal Knowledge, Essays Presented to Michael Polanyi," p. 231. The Free Press, Glencoe, Illinois.

Remarks on Some Physical Properties of Nucleic Acids

JULES DUCHESNE

University of Liège, Liège, Belgium

Some years ago a method was developed (Duchesne and Monfils, 1955a,b) which made it possible to identify weakly piezoelectric substances. The technique consisted essentially in placing the substances in the alternating field of a condenser inserted in a circuit working as a superregenerative oscillator analogous to that used in nuclear quadrupole resonance. Several organic substances which were thus examined were found to give rise to many resonance lines distributed over a large radiofrequency range. These were mainly studied between 20 and 40 Mc per second (Duchesne and Monfils, 1955a,c). It was shown in particular that the lines are reversibly shifted by temperature and that the displacements called Δ are given by the simple relation

$$\Delta = -\frac{1}{\nu} \frac{d\nu}{dT} = \alpha\gamma.$$

In this equation, α is the thermal expansion coefficient (in volume) of the substance, and γ the Grüneisen constant characterizing the dependence of compressibility upon pressure. The main interest of the method involved lay in its high sensitivity, due to superregeneration, which permitted the detecting of even very weak piezoelectricity. On the whole the new technique had the advantage of being simple and of providing us with important information on the electrical, thermal, and mechanical properties of substances.

At the time the method was evolved, it appeared to us that one-dimensional large systems such as deoxyribonucleic acids (DNA) clearly characterized by long-range order and interactions might be shown to possess quite unexpected physical properties. Research performed with our "piezoelectric spectrograph" was started with DNA substances kindly put at our disposal by Professor Welsch of the Faculty of Medicine at the University of Liège, and the presence of a few radiofrequency lines within the region investigated was soon revealed. Since not only nucleic acids and nucleoproteins, but also

proteins, when submitted to similar tests (Duchesne *et al.*, 1960), were seen to give rise to a series of lines, it would seem that a new radiofrequency spectroscopy of macromolecules had thus come into being.

Our attention was immediately drawn to the very special fact that whereas, in general, radiofrequency lines of organic compounds were greatly dependent on temperature, those characterizing DNA's were exceedingly stable. Indeed, the DNA sodium salts studied (yeast, calf thymus, herring sperm) are characterized by values of the Δ 's as low as about 3×10^{-5} , and this compares remarkably with the corresponding quantities obtained for quartz, and metals such as platinum and iron, for which Δ approximates to 5×10^{-5} . In the case of the three latter substances the γ 's used in the calculation amount, respectively, to about 0.7, 2.5, and 1.5 (Slater, 1939). In contrast, Δ was in fact found to be a hundred times larger for Rochelle salt and antibiotics (Duchesne and Monfils, 1955a,c). Unfortunately, the method involved does not permit the disentangling of α and γ . However, the dispersion in the values of the γ 's hitherto established is not very great. It therefore seems possible to anticipate that the anomalously low values of the Δ 's mainly reflect the low thermal expansion coefficients of DNA. It is perhaps worthwhile noticing here that in the case of a ribonucleic acid of highly polymerized yeast which behaves in a similar fashion to the DNA, shifts of different signs were found, the exact physical meaning of which still remains to be determined. Experiments are now being performed in our laboratory to go further with these questions.

The main result emerging from all this certainly lies in the illuminating analogy between nucleic acids and metals as regards thermal expansion. Further, from empirical relations, according to which thermal conductivity in an insulator should vary inversely with Δ (Dugdale and MacDonald, 1955), it is to be expected that DNA's are good thermal conductors, and on this property also experiments are being carried out in our laboratory.

As regards piezoelectricity it is to be added that this was identified in our experiments by discontinuity in the reactance of the circuit at a certain resonant frequency. It is still difficult to state the nature and the scale of the polarization produced. It is however not at all excluded that piezoelectricity should be restricted to small domains in the material and even to the macromolecule itself. This possibility is suggested by significant experiments made on different compounds by Japanese authors (Kojima *et al.*, 1953, 1955). Very

recently it was shown (Polonsky *et al.*, 1960; Douzou *et al.*, 1960) that the relation between polarization and field was such that some kind of ferroelectric or antiferroelectric phenomenon seemed to be present. It is nevertheless most important to point out that the phenomena observed only appear for water contents at 20% and reach their maximum for 40%. In this respect, it is to be recalled that in our experiments also piezoelectricity was found to depend in a not yet well-defined way on the water content.

Let us now turn our interest to electrical conductivity. It was found that the variation of the dc electrical conductivity with temperature followed the well-known equation for semiconductors (Duchesne *et al.*, 1960). The activation energies deduced in the range 273°–313°K were about 1.8 ev in the case of the herring sperm and calf thymus DNA's as compared to 3.0 ev, a value which is commonly found in the case of proteins (Cardew and Eley, 1959). This lessening in the activation energy found for DNA's might, however, not be a characteristic of these substances as such, but might be due to their association with water.

Parallel with these investigations, magnetic properties have been the subject of an important series of research (Blumenfeld, 1959; Blumenfeld *et al.*, 1959). These authors would seem to have demonstrated that the DNA's are characterized by the presence of unpaired electrons, in a hitherto undetermined number, corresponding to each nucleotide and interacting over the whole molecular structure. These electrons are thought to be π -electrons belonging to the bases themselves, and the pseudoferromagnetic character is considered to be the result of interactions between them. The corresponding exchange integrals between the bases must however attain certain rather high values in order to give rise to the property referred to, and these are reached, according to the authors, on account of the considerable length of the molecular systems.

If we assume that all this is the case, it is clear that the ferromagnetic property is most favored when the bases are oriented perpendicularly to the molecular axis as, in these conditions, the overlap of the π -orbitals is maximum. If the bases should deviate from their original position, a decrease of the overlap occurs, with a subsequent decrease in the corresponding exchange integrals. A crude calculation which we have made shows that a tilting of 15° produces a decrease of about 15% in the exchange integral. On the other hand, recent experimental infrared studies (Bradbury *et al.*, 1961) have confirmed in a decisive way that in the case of the sodium salt of

DNA the base lie either perpendicular to the helix axis, for a water content higher than 40% (B form), or are tilted by at least 13° to the normal to the helix axis, for a lower water content (A form). Taking into account the dependence of the exchange integrals upon such structural changes, it may be anticipated that the pseudoferromagnetic property will not remain constant in intensity and would in fact reach its maximum in the B form. This prediction, which might undergo some variations in its quantitative aspect according to the alkali salt considered, ought to be the object of an urgent verification. This subtle behavior is not restricted only to pseudoferromagnetism, but extends, as we already mentioned, to piezoelectricity, ferroelectricity, and even semiconductivity. The latter case is perhaps particularly significant and it would be most revelant to know whether for a given water content one cannot reach values for the activation energies still lower than 1.8 ev found by us. Should this be the case, then the important problem of energy transfers for DNA's would again be revived.

As a general conclusion, it would perhaps not be too speculative to consider that even within the cell DNA's are subject, during mitosis, to discontinuous changes of their physical properties governed by discontinuities in their molecular structure produced by different degrees of hydration. This is perhaps the most extraordinary feature of cell dynamics, and it makes difficult the understanding of the mechanism of mitotic processes, which must nevertheless, one feels sure, depend at least partially, as regards the behavior of the forces involved, on the picture here suggested.

This would seem to open a new approach in this fundamental field. The kind of discontinuity here described brings to mind the possibility of the existence of some sort of universal constant ruling such biological phenomena at the cellular level.

Summary

Thermal, electrical and magnetic properties of DNA's have been considered. The possible effect of the water content on these properties is emphasized, and tentative suggestions are made in regard to their evolution during the mitotic processes.

REFERENCES

- Blumenfeld, L. A. (1959). *Biophysica (U.S.S.R.)* **4**, 515.
 Blumenfeld, L. A., Kalmanson, A. E., and Shen Pei-Gen (1959). *Doklady Akad. Nauk S.S.R.* **124**, 1144.

- Bradbury, E. M., Price, W. C., and Wilkinson, G. R. (1961). *J. Mol. Biol.* **3**, 301.
 Cardew, M. H., and Eley, D. D. (1959). *Discussions Faraday Soc.* **27**, 115.
 Douzou, P., Francq, J., Polonsky, J., and Sadron, Ch. (1960). *Compt. rend. acad. sci.* **251**, 976.
 Duchesne, J., and Monfils, A. (1955a). *Bull. classe Sci., Acad. roy. Belg.* **41**, 165.
 Duchesne, J., and Monfils, A. (1955b). *J. Chem. Phys.* **23**, 762.
 Duchesne, J., and Monfils, A. (1955c). *Compt. rend. acad. sci.* **241**, 749.
 Duchesne, J., Depireux, J., Bertinchamps, A., Cornet, N., and van der Kaa, J. M. (1960). *Nature* **188**, 405.
 Dugdale, J. S., and MacDonald, D. K. C. (1955). *Phys. Rev.* **98**, 1751.
 Kojima, K., Tsukada, K., Ogawa, S., and Shimauchi, A. (1953). *Phys. Rev.* **92**, 1571.
 Kojima, K., Tsukada, K., Ogawa, S., Shimauchi, A., and Matsumiya, N. (1955). *J. Phys. Soc. Japan* **10**, 265.
 Polonsky, J., Douzou, P., and Sadron, Ch. (1960). *Compt. rend. acad. sci.* **250**, 3414.
 Slater, J. C. (1939). "Introduction to Chemical Physics." McGraw-Hill, New York.

Semiconductivity in Biological Molecules

D. D. ELEY

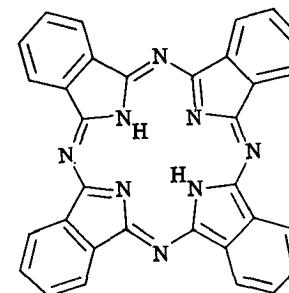
*Department of Chemistry, Nottingham University,
Nottingham, England*

I. Introduction	341
II. Semiconductor Behavior	343
A. The Specific Conductivity	343
B. Electron Transfer in Crystals	346
C. The Potential Box (Electron Gas) Model	348
D. The Energy Gap	350
E. Ohm's Law and κ_0	353
F. Mobility	354
III. Some Molecular Types	356
A. Solid Free Radicals	356
B. Charge-Transfer Complexes	357
C. The Porphyrins	360
D. The Proteins	361
E. Nucleic Acids and Nucleoproteins	368
IV. Photoconduction	369
V. Biological Systems	371
A. Cytochrome Systems	371
B. Chloroplasts	372
C. Carcinogenesis	373
D. Energy Conversion in Retinal Rods	375
E. Irradiation of Proteins	375
VI. Conclusion	376
References	377

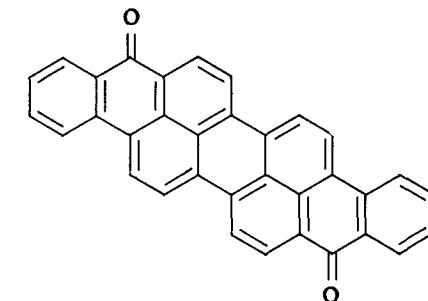
I. Introduction

Some twenty years ago Professor Szent-Györgyi (1941) gave the Koranyi lecture, in which he drew attention to the study of energy levels in biochemistry. There were a number of problems, such as the photosynthetic unit of the chloroplast (Möglich and Schön, 1938) and the inactivation spectrum of urease (Jordan, 1938) which could be interpreted in terms of the photoexcitation of an electron into an energy-band characteristic of a macromolecule or group of molecules, thus allowing for action at a point remote from the site of quantum absorption. Szent-Györgyi suggested at that time that

fibrous protein molecules might join together to form a continuum through which an excited electron might travel, and that such a system might link the insoluble oxidation enzymes of the cell. Baxter and Cassie (1941) brought forward results for the wool-water system (> 8% water) which showed a semiconductivity obeying Ohm's law with an activation energy of 1.3 ev, similar results being given by the wool-methyl alcohol and the glass-water systems. They suggested the energy band for conduction was associated with the adsorbed water on molecules or cell interfaces. These publications came at a time when the present author was interested in the role of globular protein molecules in promoting and directing the action of prosthetic groups in enzymes. It seemed possible that this interaction, for example between heme and globin in hemoglobin might be associated with a degree of electron mobility between the two parts. When opportunity offered, a program of experiments was set in hand, first at Bristol University and later at Nottingham to establish the electrical conducting properties of organic molecules, related to biologically active substances. Phthalocyanine (I) and its copper derivative were chosen for investigation because of their thermal stability and a definite semiconductivity established from the variation of their resistivity with temperature (Eley, 1948). In this work it was suggested that the π -orbitals overlap between the molecules to give energy bands common to the whole crystal, and that the observed activation energy was one-half the energy gap between the full ground band and the empty first-excited band. This idea has persisted with little change up to the present day. About the same date Vartanyan (1948) established a similar result, also for phthalocyanine. Vartanyan's result was the end product of a series of investigations of photoconductivity of ionic dyestuffs carried out in Terenin's laboratory, which led to a consideration of dark conductivity in non-ionic substances. Putzeiko (1948) in this laboratory showed the majority charge carriers in photoconducting phthalocyanine to be positive, and the work has recently been reviewed by Terenin (1961). Szent-Györgyi (1946) reported photoconductive effects in gelatin dyed with various ionic dyes, and further work on ionic dyestuffs was reported by Nelson (1951; see also Weigl, 1956). The work on phthalocyanine and the concept of semiconduction arising from the conjugated π -electron system were supported by the results of Akamatu and Inokuchi (1950) on isodibenzanthrone (II).



(I)



(II)

Several further years of effort were required to establish beyond doubt that electron semiconduction is an intrinsic property of many organic substances, and there are still many points of discussion, as shown at the first International Conference on the subject, at Duke University, North Carolina in 1960. This Conference paid much attention to electrode effects, and while these may be dominant for chemically reactive iodine electrodes (Kallmann and Pope, 1960), they may be regarded as of secondary importance for most of the cases discussed here.¹ The account below will necessarily follow the thread of the author's own work and omit many important contributions for reasons of space. An attempt will be made to put forward a coherent philosophy of organic semiconductors.

II. Semiconductor Behavior

A. THE SPECIFIC CONDUCTIVITY

The specific resistivity ρ is related to the measured resistance R by $\rho = RAV/l$ where A is the area and l the thickness of the specimen, and V a filling factor, less than unity for powders. The specific conductivity κ is then $1/\rho$.

By semiconductors we mean substances which show a specific conductivity κ vs. temperature, or resistance R vs. temperature relation of the type.

$$\kappa = \kappa_0 e^{-\Delta\epsilon/2kT}, \quad R = R_0 e^{\Delta\epsilon/2kT} \quad (1)$$

¹ Added in proof: In support of this view, see P. J. Reucroft, *J. Chem. Phys.* **36**, 1114 (1962).

where $\Delta\epsilon$ is usually referred to as the energy gap. The division between semiconductors and insulators is an arbitrary one, and in this field of research relates to the measuring instruments used which in practice are limited to $\kappa > 10^{-16}$ ohm $^{-1}$ cm $^{-1}$. Most measurements are made with direct current on samples held compressed between electrodes, so as to remove the resistance contribution due to interparticle contacts. These contacts may also be eliminated by

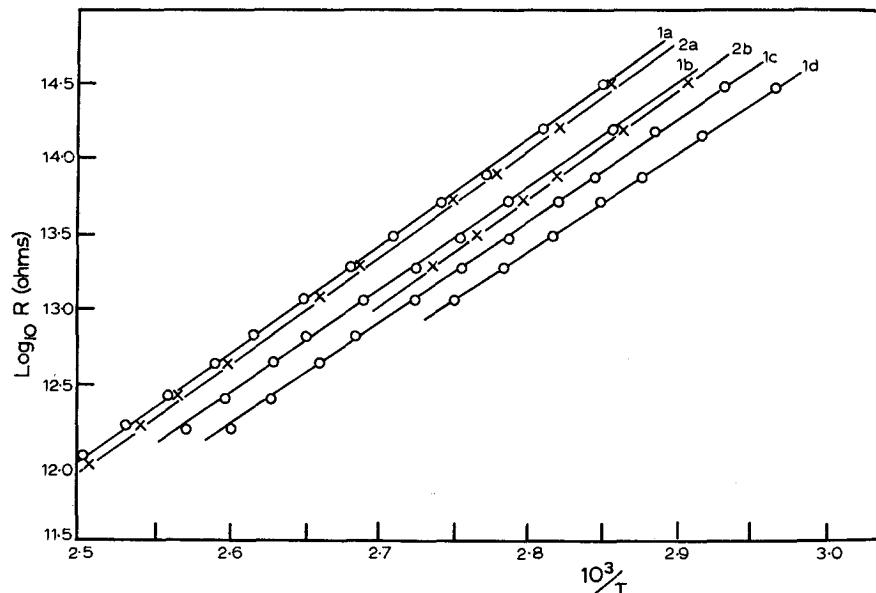


FIG. 1. This shows the resistance-temperature relation for the typical case of native hemoglobin (1d). The samples have been denatured to varying extents by initial heating to (1a) 150°C, (1b) 125°C, (1c) 115°C, (1d) 95°C, and alcohol-denatured specimens (2a) 145°C, (2b) 95°C.

high frequency a.c. measurements, which may be carried out down to $\kappa \geq 10^{-9}$ ohm $^{-1}$ cm $^{-1}$. Where a comparison has been made, the two methods give comparable $\Delta\epsilon$ values, but discrepancies have occasionally been found in κ_0 values, which are under investigation. Single crystal measurements have given similar $\Delta\epsilon$ values to powders measured under pressure. The conductivities of proteins and nucleic acids in the dry state are too small to be investigated by the a.c. technique; the d.c. technique has been used.

An example of the above relation is shown in Fig. 1 which shows

hemoglobin in different stages of denaturation (Eley and Spivey, 1960).

To obtain $\Delta\epsilon$ values characteristic of natural proteins in the dry state, it is necessary never to heat the specimen above 100°C, but the effects of denaturation on $\Delta\epsilon$ are seen to be small, raising $\Delta\epsilon$ from 2.63 to 2.88 ev.

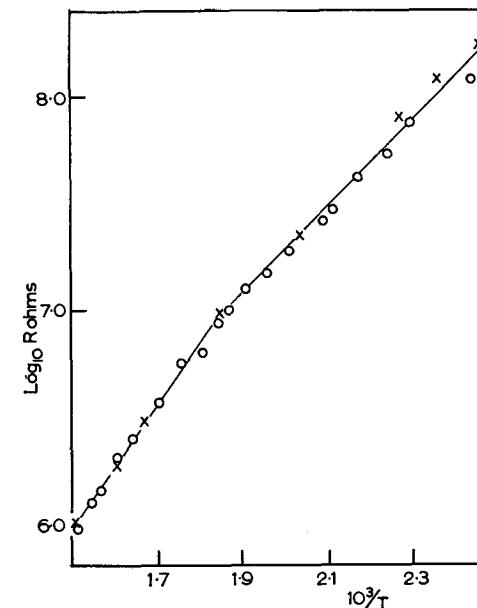


FIG. 2. To show the change from impurity to intrinsic conduction in phthalocyanine.

Equation (1) tacitly assumes that Ohm's law is obeyed, and this is an approximation which holds quite well up to 100 volts cm $^{-1}$ for many substances.

In the case where conduction arises from impurities, a change-over from a low $\Delta\epsilon$ value at low temperatures (impurity value) to a high $\Delta\epsilon$ value at higher temperatures (intrinsic value) has been found, as in inorganic crystals. The example in Fig. 2 is for supposedly metallic impurities in phthalocyanine (Eley and Parfitt, 1955) and the effect has been particularly well established for organic impurities in polycyclic hydrocarbons (Northrop and Simpson, 1954).

A collection of published $\Delta\epsilon$ and κ_0 values for all work to 1959 has been prepared by Garrett (1959). The values given in Table I have, with the exception of anthracene, been chosen from work at Nottingham so as to illustrate the main features found over the main groups of substances (Eley, 1959a). There is still some doubt about the value for anthracene. The $\Delta\epsilon$ values are in many cases accurate to ± 0.05 ev but further work will be necessary to establish $\log_{10} \kappa_0$ with similar accuracy.

TABLE I
A RANGE OF ORGANIC SEMICONDUCTORS
(All Substances Measured in the Dry State in Vacuo)

Substance	$\Delta\epsilon$ (ev)	$\log_{10} \kappa_0$ (ohm $^{-1}$ cm $^{-1}$)
Diphenyl picrylhydrazyl	0.26	6.1
Dimethylaniline-bromanil complex	0.45	6.7
Isodibenzanthrone	0.96	2.3
Phthalocyanine	1.49	0.9
Heme	1.74	2.8
Thymus nucleic acid	2.44	3.45
Anthracene ^a	2.7	2.0
Hemoglobin	2.75	4.6
Polyglycine	2.99	6.3

^a Inokuchi (1956).

It is proposed to present a model for the semiconduction process, in terms of which to discuss the experimental data.

B. ELECTRON TRANSFER IN CRYSTALS

Riehl (1955, 1956) has considered two energies, $\Delta\epsilon_1$ to transfer an electron from a molecule to its next neighbor in the lattice, and $\Delta\epsilon_2$ to transfer an electron from a molecule to a point in the lattice very far removed. Lyons (1957) has discussed these processes rather fully giving the energy quantities as:

$$\begin{aligned}\Delta\epsilon_1 &= I - A - e^2/r - P\mu \\ \Delta\epsilon_2 &= I - A - 2P_i\end{aligned}$$

Here I is the ionization potential and A the electron affinity of the molecule, *in vacuo*, e^2/r is the Coulombic energy of the ion pair,

$P\mu$ the polarization of the crystal due to the ion pair, and $2P_i$ the polarization of the crystal due to the two separated ions. Taking anthracene, Lyons calculated $\Delta\epsilon_1 = 2.2$ ev, $\Delta\epsilon_2 = 3.8$ ev. The latter energy should correspond to the observed $\Delta\epsilon$, but in fact smaller values have been observed, e.g., 2.7 ev (Inokuchi, 1956) and 1.95 ev (Northrop, 1959). There have been at least three attempts to reconcile the calculated $\Delta\epsilon_2$ and experimental $\Delta\epsilon$ values. Lyons suggested that the electron is trapped, adding a negative energy term to the expression for $\Delta\epsilon_2$, and accounting for the frequent finding of *p*-type behavior in photoconducting anthracene. Fox (1959) in a brief note considers that Lyons has underestimated the polarization energies, and concludes that a narrow energy band will occur at $\Delta\epsilon_2$, the value of which when corrected by quantum considerations, comes out as of the right order. Northrop (1959) pointed out that if $I - A = ^1E_1$, the energy of the singlet-singlet transition giving rise to the fundamental band, the predicted gaps agree with those found. However, this is a poor approximation in error by 2.5 ev for anthracene, though no doubt it holds better for the larger molecules. The author has made the same assumption in his own model.

In his own papers the author has consistently adopted a potential box model which has led to fairly successful correlations, but which has ignored Coulombic (except in the case of donor-acceptor complexes) and polarization terms. Before describing this model, it is worth considering these approximations in relation to the ion-pair calculation, $\Delta\epsilon_1$ above. We consider a molecule containing n π -electrons and, writing D as the dielectric constant,

- (a) identify A with the energy $E(n/2 + 1)$ of the lowest unoccupied level of the neutral molecule in the ground state (really this should be the negative ion)
- (b) neglect Coulombic e^2/r (or e^2/Dr) and polarization $P\mu$ terms.

For anthracene the error due to (a) is -2.5 ev, while for (b) it is +3.4 ev ($D = 1$) or 1.8 ev ($D = 2.29$). It is not clear what value of D should be adopted, but along the *c* axis it seems likely that it should be 2.29, since the charges will very probably be separated by 2 benzene rings. The above calculations refer to an interionic distance of 5.23 Å (Lyons, 1957) but this will probably be much larger along the *c* axis. In any case the point is that the errors tend to cancel, and for larger molecules (phthalocyanine, isodibenzan-

throne) the approximation

$$\Delta\epsilon_2 = I - A = -[E(n/2) - E(n/2 + 1)]$$

may not be too much in error.

C. THE POTENTIAL BOX (ELECTRON GAS) MODEL

This model, developed by Bayliss, Simpson, *et al.* has had a great success in discussing the spectroscopic behavior of conjugated molecules (cf. Hall, 1959). We consider a molecule containing an even number n of π -electrons, constrained to move in a one-dimensional box of length l . In the case of polyacenes, the $n \pi$ -electrons are spread out along a "perimeter path" of length $l = nd$, where d is the C—C aromatic bond length. This path may be "open" or bent on itself to form a closed ring, sometimes called the "doughnut model." In either case, the problem is to find the energy levels of an electron within the potential box, as indicated diagrammatically in Fig. 3 for the case of $n = 6$ (benzene or hexatriene). By treating the electron as a gas, the energy levels may be calculated and the energy $\Delta\epsilon$ to promote an electron from the highest filled $n/2$ th level to the lowest unfilled $(n/2 + 1)$ th level may be found. This is

Open Path:

$$\Delta\epsilon = \frac{\hbar^2}{8m_e l^2} (n + 1) = \frac{\hbar^2}{8m_e d^2} \frac{(n + 1)}{n^2}$$

Closed Path:

$$\Delta\epsilon = \frac{\hbar^2}{4m_e l^2} n = \frac{\hbar^2}{4m_e d^2} \frac{1}{n}$$

where m_e is the effective mass of the electron.

If the electrons have opposed spins then $\Delta\epsilon = {}^1E_1$ the singlet-singlet excitation energy, corresponding to the fundamental absorption band of the molecule. If the electrons have parallel spin then $\Delta\epsilon = {}^3E_1$ the (optically forbidden) singlet-triplet excitation energy. The electron gas model does not allow these two states to be energetically distinguished, but in practice the triplet energy is always appreciably lower than the singlet. Such an excited electron may tunnel through the intermolecular barrier to the equivalent empty level in the next molecule, in which case we generate an ion pair, and the energy $\Delta\epsilon$ is equal to $\Delta\epsilon_1$ of the previous section, to the approximation discussed. The electron transfer will tend to occur

in the direction of the applied field, V volt cm^{-1} , which will cause a potential energy drop per electron of $a\text{eV}$ from one molecule to the next. The effect of the applied potential, greatly exaggerated is shown in Fig. 3b. At the same time the vacancy in the central molecule may be filled by an electron from the left-hand molecule which is equivalent to a positive hole moving in the opposite direction.

In the first account of this theory (Eley and Parfitt, 1955) it

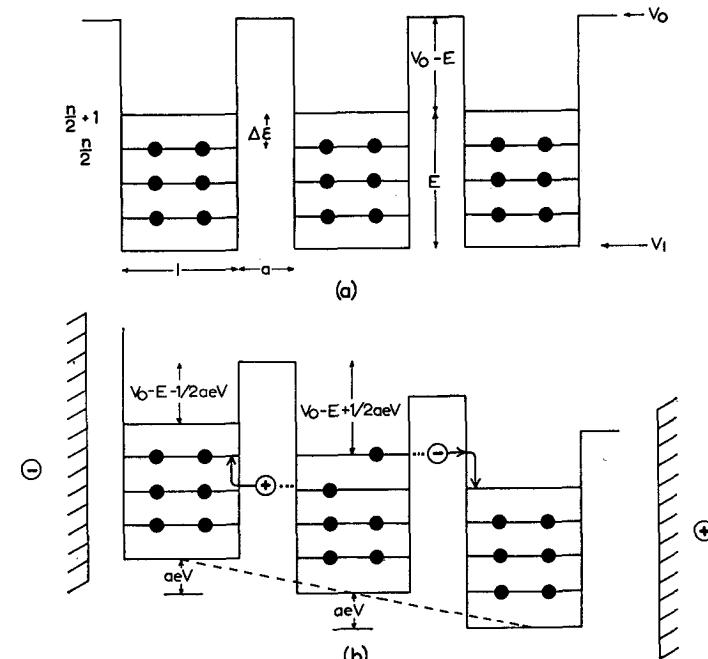


FIG. 3. The potential box or electron tunneling model. Barrier parameters are width a , height $V_0 - E$.

was suggested that the interaction between the excited $(n/2 + 1)$ levels was sufficient to form an energy band. It is now recognized that while this may formally be so, the band must be very narrow, and an appropriate way to treat the problem of movement of charge is in terms of the tunneling of electrons (and holes) through the intermolecular barriers (Eley and Willis, 1960). The total current is then

$$i = e(n_e v_{de} + n_h v_{dh})$$

where n_e is the number of excited electrons per cubic centimeter, v_{de} their drift velocity in the applied field, and n_h , v_{dh} for the holes, with e the electronic charge. Assuming an equilibrium of holes and electrons given by band theory

$$n_e = n_h = AT^{3/2}(m_e m_h)^{3/4} \exp(-\Delta\epsilon/2kT)$$

and equating v_{de} and v_{dh}

$$i = 2ev_{de}AT^{3/2}(m_e m_h)^{3/4} \exp(-\Delta\epsilon/2kT) \text{ amp cm}^{-2}$$

The drift velocity v_{de} is now calculated as the frequency with which electrons strike the barrier, times the difference in probability of passing forward with or backward against the field. The result is

$$i = A \sinh bV \exp(-\Delta\epsilon/2kT)$$

where A involves the barrier parameters indicated on the diagram, the effective masses of the charge carriers, and temperature. At low voltages, the voltage dependence becomes linear, and Ohm's law is obeyed at constant temperature.

$$i = AbV \exp(-\Delta\epsilon/2kT)$$

so $\kappa = \frac{i}{V} = Ab \exp(-\Delta\epsilon/2kT)$

and $\kappa_0 = Ab \text{ ohm}^{-1} \text{ cm}^{-1}$

D. THE ENERGY GAP

In Fig. 4 we show a plot of observed $\Delta\epsilon$ vs. the total number of π -electrons in the molecule except in the case of phthalocyanine where we take $n = 16$ in accordance with the obvious conjugation path (cf. Lonsdale, 1937). This is a revised version of an earlier plot (Eley and Parfitt, 1955), and contains additional data. Most points lie between the two lines for closed and open conjugation paths.

The points lying appreciably outside the two lines have been labeled, and there is little doubt that coronene (three authors) and ovalene (two authors) do not fit. The other points are unconfirmed points and further work is desirable on the compounds in question. The following points may need attention in any re-examination of energy gap values.

(1) Impurities. As an example, traces of anthraquinone on the surface increases the photoconductivity of anthracene (Bree and Lyons, 1960). Energy gaps of 2.7 ev (Inokuchi, 1956) and 1.95 ev

(Northrop, 1959) have been reported for single crystals of this material, and this observation may be relevant to these differences.

(2) Defect structure. Aftergut and Brown (1961) found $\Delta\epsilon = 2.1$ ev for zone-refined and $\Delta\epsilon = 3.8$ ev for sublimed samples of phenazine which they attribute to this cause.

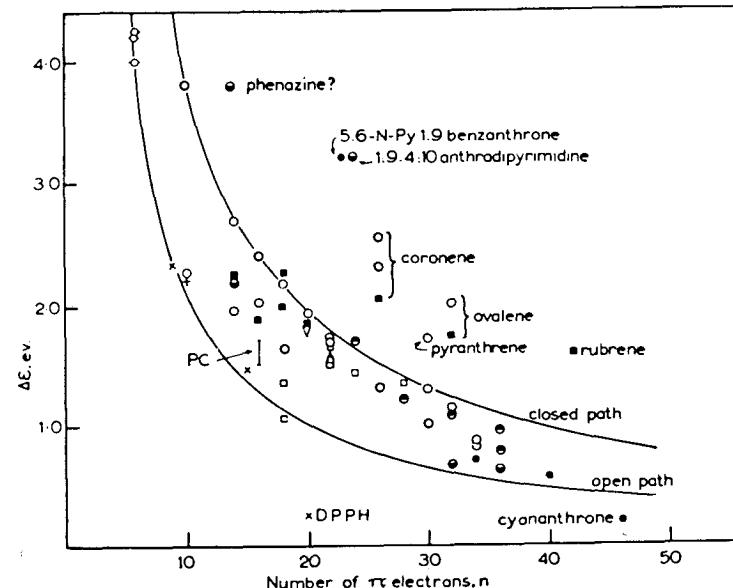


FIG. 4. Energy gap $\Delta\epsilon$ as a function of number of conjugated π -electrons in the molecule. The lines are calculated for open and closed conjugation paths. Points are experimental data, \circ polyacenes (mainly from Inokuchi and Northrop), \bullet contain N, \bullet contain O, \bullet contain N and O, \square cationic dyes, \blacksquare data from Wilk, X three solid free radicals, \triangle carotene (Rosenberg), ∇ porphyrins, + a dipyromethene (contains Br), \ominus phthalic anhydride, p -dihydroxybenzene and β -resorcinol [see T. Heydel-Zyczkowa, *Roczniki Chem.* **34**, 601 (1960); K. Pigoń, *ibid.* **29**, 239 (1955) and A. Golebiowski, *ibid.* **31**, 241 (1957)]. P.C. gives the spread of phthalocyanine data (see Eley, 1959a).

(3) Lattice plane factor. Conductivity behavior of anthracene is different along the three directions (Mette and Picke, 1953; Kepler, 1960). In this connection polymorphic forms of phthalocyanine behave differently (Eley and Parfitt, 1955; Wihksne and Newkirk, 1961).

(4) Form of the specimen, viz., single crystals, polycrystalline, evaporated film. So far as energy gap is concerned, agreement is

usually good, though differences in κ_0 have been noted, for example by Northrop (1959).

(5) Method of measurement. Energy gaps usually agree whether obtained by the d.c. method using a specimen compressed between electrodes, or by the a.c. method (Eley and Parfitt, 1955). The a.c. method often gives somewhat lower specific resistivity than the d.c. method at 20°C, e.g., by a factor 1.5 to 150 for charge-transfer complexes (Eley *et al.*, 1959) and 100 for 2,2-diphenyl-1-picrylhydrazyl (DPPH) (Eley and Inokuchi, 1959). This matter is under investigation, and may denote a conduction path localized over a limited number of molecules in the case of a.c. loss. Vartanyan has found very good agreement between thermal and photoconductivity $\Delta\epsilon$ values (Terenin, 1961).

Several attempts have been made to identify precise molecular energy states with energy gaps. The first such (Eley *et al.*, 1953) compared optical transitions in solution with energy gaps in the crystalline state for anthracene 3.33 ev ($\Delta\epsilon$ 1.65 ev) and phthalocyanine 1.75 ev ($\Delta\epsilon$ 0.86 to 1.13 ev). It was suggested that the conducting state was the first excited singlet state broadened by interactions in the solid, but in fact, absorption spectra in solid state and solution are little different and these postulated interactions are, therefore, small (e.g., Kasha, 1959). However, in many cases the conducting state is lower than the first singlet, and several authors have suggested that the first triplet is involved (Carswell *et al.*, 1954; Terenin, 1957; Northrop and Simpson, 1956). The correlation of $\Delta\epsilon$ with 3E_1 is quite close over a series of polyacenes (Lyons, 1957; Wilk, 1960) (cf. Fig. 5). Alternatively, there is a correlation between $\Delta\epsilon$ and 1E_1 , 1.5 ± 0.5 ev (Northrop, 1959; Wilk, 1960). The correlation is good between $\Delta\epsilon$ and 1E_1 in the case of the phthalocyanines, when the correct $\Delta\epsilon$ values are used (Fielding and Gutmann, 1957). The difference noted above was based on a too low $\Delta\epsilon$ value for phthalocyanine, the correct value being 1.5–1.7 ev (Eley, 1959a). A good correlation between $\Delta\epsilon$ and 1E_1 has also been found for porphyrins (Eley and Spivey, 1962). Another case where $\Delta\epsilon$ maybe equal to 1E_1 , is diphenyloctatetraene, 1E_1 3.3 ev, for which a dark conductivity was not measurable in 1953. It is difficult to understand this polyene result at present, since Fig. 4 would predict $\Delta\epsilon = 1.5 \pm 0.5$ ev for 20 π -electrons, and β -carotene with 22 π -electrons has $\Delta\epsilon = 1.5$ ev as expected (Rosenberg, 1961a). In fact, all compounds with pendant phenyl groups except rubrene

have been excluded from Fig. 4, since it seems unlikely that the phenyl groups are fully conjugated with either the polyene chain (as in diphenyloctatetraene) or the polyacene, as in rubrene or diphenyl-anthracene [cf. results in Wilk (1960)]. Based on their total number of π -electrons we predict energy gaps in these cases much smaller than those observed. We conclude that the relevant conjugation path in diphenyloctatetraene probably involves an open path of 8 π -electrons, for which we should expect a $\Delta\epsilon$ of 2.8 ev, which should be measurable with the present-day technique.

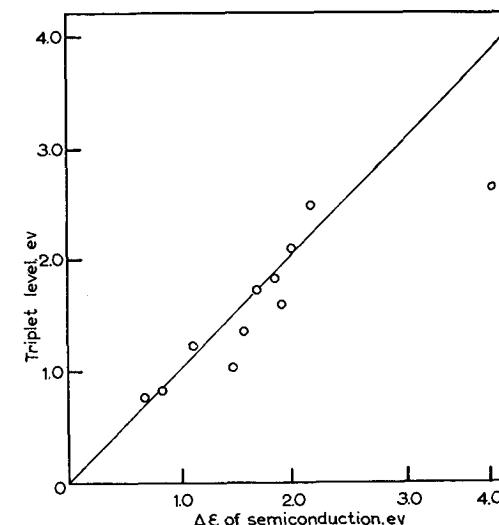


FIG. 5. A plot of observed energy gap $\Delta\epsilon$ against first triplet excitation energy for eleven polyacenes. After L. E. Lyons.

Referring again to Fig. 4, Epstein and Wildi (1960) found a polymerized copper phthalocyanine containing 140 π -electrons to show $\Delta\epsilon = 0.25$ ev in fair agreement with 0.20 ev predicted by the graph.

E. OHM'S LAW AND κ_0

Ohm's law deviations and κ_0 are clearly connected since the former are determined by bV and the latter by Ab . We expect $\sin h bV$ to approximate to bV up to $bV = 0.3$, where the deviation between the two will be only 1.3%. The tunnel theory determines the

parameter b , and using typical molecular magnitudes it was predicted that Ohm's law should be obeyed to this accuracy up to $V \sim 5 \times 10^7 \text{ V cm}^{-1}$. However, most of the substances examined in the author's laboratory, as films or in polycrystalline form show deviations at 10^3 V cm^{-1} so Eley and Willis (1960) suggested that these deviations in fact arose from electron tunneling through *inter-crystalline* rather than intermolecular barriers. Against this at the time was the fact that single crystals and powders of phthalocyanine showed similar κ_0 values. In fact, Ohm's law deviations show a systematic correlation with the chemical nature of the substances concerned in that proteins with high $\Delta\epsilon$'s have low deviations, and the deviations systematically increase as $\Delta\epsilon$ decreases, being very marked for charge-transfer complexes and for DPPH. However, the prediction has been definitely fulfilled in the case of anthracene, where single crystals have high κ_0 values and no deviations measurable at 10^5 V cm^{-1} (Northrop, 1959). To achieve this result the single crystals were grown onto the electrodes.

F. MOBILITY

An attempt to discuss mobilities was made by Eley and Parfitt (1955) using the band theory equation for an intrinsic semiconductor (Moss, 1952).

$$\kappa_0 = \frac{2(2\pi mkT)^{3/2}}{\hbar^3} e\mu = 4\mu \text{ (20°C)}$$

where μ is the mean mobility of electron and hole. This gave values of $10^{-6} \text{ cm}^2 \text{ volt}^{-1} \text{ sec}^{-1}$ for DPPH, 10^{-2} for isodibenzanthrone, and 2 for phthalocyanine. Many, Harnik, and Gerlich (1955) derived μ by making similar (actually slightly different) assumptions for 29 compounds, mainly polyacene data of Akamatu and Inokuchi (cf. a recent review, 1959), and found μ values varying from 10^{-12} to 10^{11} in order roughly of increasing $\Delta\epsilon$. They associated this result with the tunneling of electrons through potential energy barriers.

Mobility values calculated by theoretical equations for κ_0 on the assumption that m is the *free-electron* mass are best labeled as free electron values μ_{fe} . It is not at all clear that this assumption is justified since band theory gives much larger effective masses in the case of very narrow bands, as are bound to be the case for molecular solids. The barrier tunneling theory previously described relates μ to barrier dimensions and in cases using typical molecular magni-

tudes we obtained values of κ_0 of 10^{-3} to 10^{-1} (Eley and Willis, 1960). It is satisfactory that mobilities in anthracene established independent of m by a pulsed photoconductivity technique are of similar order, 0.3 to $3 \text{ cm}^2 \text{ volt}^{-1} \text{ sec}^{-1}$ (Kepler, 1960; Le Blanc, 1960).

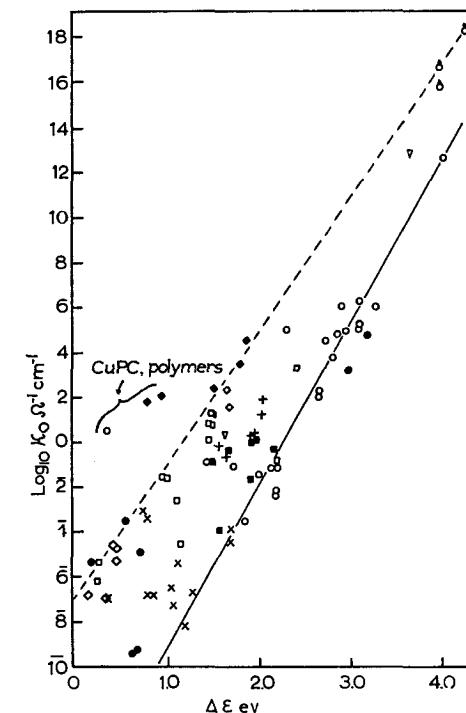


FIG. 6. A plot of $\log_{10} \kappa_0$ vs. $\Delta\epsilon$ for available data. Left-hand line, defined by charge-transfer complexes and phthalocyanines; right-hand line, proteins and amino acids.

In polymeric copper phthalocyanine, the Hall mobility has been determined as $10 \text{ cm}^2 \text{ volt}^{-1} \text{ sec}^{-1}$. It is very likely that in polymers higher values of the mobility still will be found, since the intermolecular barriers will occur less frequently than in polyacenes. There is a great need for further work on Hall mobilities.

To conclude this section we reproduce in Fig. 6 a revised version of a figure in an earlier paper (Eley and Parfitt, 1955). This considers all the available data and demonstrates a trend between $\log \kappa_0$ and $\Delta\epsilon$. The right-hand edge of this band is defined by amino

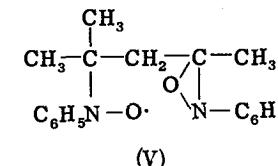
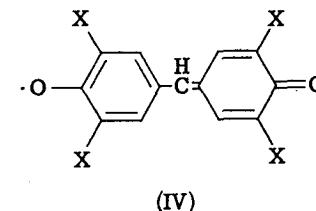
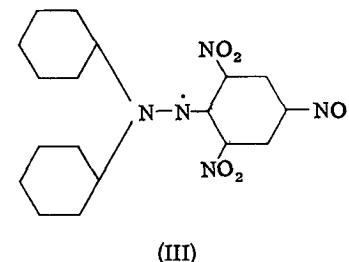
acid data. The left-hand edge is largely defined by phthalocyanines and charge-transfer complexes. A part of the scatter may arise from differences in degree of crystallinity and defect structure of specimens, which is known to effect κ_0 , but the two extreme lines may well correspond to different intermolecular barrier shapes, changing the probability of tunneling. The three points at $\Delta\epsilon < 1.0$ ev well above the dotted line are for phthalocyanine polymers (Epstein and Wildi, 1960; Felmayer and Wolf, 1958). They show that even in a case where $\Delta\epsilon$ is small, π -electrons excited into an extensive conjugated system will have a high mobility, reflected in a high κ_0 . This observation supports our explanation of the high κ_0 values found in proteins, which on the other hand have high $\Delta\epsilon$. Here the electrons are mobile through an extensive system of conjugated hydrogen bonds (Eley and Spivey, 1960).

III. Some Molecular Types

A. SOLID FREE RADICALS

The potential box theory predicts that a solid free radical, in which the uppermost filled molecular orbital contains only one electron, should have a zero or small energy gap. It should be possible for the electron to tunnel from one molecule to the next without excitation (apart from a possible Coulombic energy of separation of charges). This prediction was tested and found to be true for diphenyl picrylhydrazyl (III). To secure the true energy gap of 0.26 ± 0.10 ev (Eley and Parfitt, 1955; Eley and Inokuchi, 1959), it was necessary to recrystallize the DPPH many times from diethyl ether. The energy gap predicted by Fig. 4 for 20 π -electrons is 1.5 ± 0.5 ev and the very much lower observed value may simply be a residual Coulombic energy. The work has been recently extended to two further free radicals, but these behave very differently (Eley and Willis, 1960). Radical (IV) gives an energy gap of 1.45 ev and radical (V) one of 2.31 ev, and these values approximate to the optical threshold determined spectrophotometrically on the solids. Clearly, it is necessary to excite the electron to the next higher molecular level, ground doublet state to first excited doublet state presumably to increase the probability of electron tunneling through an unfavorable intermolecular barrier. This work, like the work on α - and β -phthalocyanine, focuses attention on the possible effect of molecular packing on intermolecular barriers. The molec-

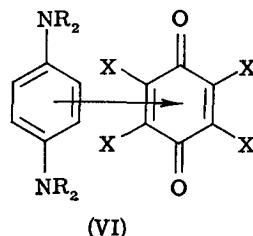
ular packing must be very favorable in DPPH for electron tunneling in the ground state, but less favorable in the two other free radicals.



B. CHARGE-TRANSFER COMPLEXES

Charge-transfer complexes show semiconductivity properties and broadly one may distinguish two classes, inorganic-organic and organic-organic complexes. In the first class we have the alkali-metal anthracene complexes described by Holmes-Walker and Ubbelohde (1954) and Slough and Ubbelohde (1957). These possess variable compositions in the range M_1An to M_2An and typical energy gaps are $Li_{1.16}An$ 2.68 ev, $Na_{1.08}An$ 2.40 ev, and $K_{1.18}An$ 2.20 ev. If we assume the intrinsic gap for anthracene is 2.70 ev (Inokuchi, 1956), then the alkali atom may be regarded as forming donor levels

just above the top of the "valence band." Similar results were obtained with benzoquinoline complexes. Density measurements indicate a degree of electrostriction which suggests a high degree of ionization of the Na atoms in anthracene. Striking results are also obtained in the study of complexes between polyacenes and halogen molecules (Akamatu *et al.*, 1956). Very high conductances and low energy gaps of 0.1 to 0.2 ev were obtained, but a certain chemical instability was present. Kommandeur and Singer (1960) in a detailed examination of pyrene- $2I_2$ and 2 perylene- $3I_2$, established by Hall effect an upper value for charge carrier mobility of $10^{-2} \text{ cm}^2 \text{ volt}^{-1} \text{ sec}^{-1}$. They found the spin concentrations determined by electron spin resonance (ESR) had the same temperature dependence as the conductivity, suggesting that the electron resonance arises from

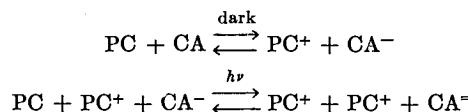


charge carriers. In the systems indole- I_2 (Smaller *et al.*, 1961) and carotene- I_3 (Huggins and Le Blanc, 1960), however, there was no such correlation, so the matter is still open.

There is a very wide range of organic donor-organic acceptor complexes possible, considered to range from the mainly covalent to the mainly ionic (Mulliken, 1952). The possibility of ionization in such complexes was first considered by Weiss (1942). Electron spin concentrations vary over a wide range, the more ionic complexes showing the larger spin concentrations in general (Kainer, Bijl, and Rose-Innes, 1960). Eley, Inokuchi, and Willis (1959) examined the six 1:1 complexes formed by the donors, dimethylaniline (DMA) and tetramethyl paraphenylenediamine (TMPD) with the three tetrahalogenated paraquinones (Cl, Br, and I) [see (VI)]. These complexes are considered to possess alternate donor and acceptor molecules stacked one above the other with molecular planes parallel, as in a pile of coins (Wallwork, 1961).

The complexes with the stronger donor, TMPD showed an electron resonance signal and much larger conductivity than the DMA

complexes, which showed no ESR signal. However, over the series of TMPD complexes the order of increasing conductivity was TMPD-iodanil (20) < TMPD-bromanil (2) < TMPD-chloranil (0.2) clearly the reverse of the free radical concentrations shown in parentheses. The less conducting DMA complexes had the lower energy gap, which suggests the intermolecular barriers were less transparent than in the DMA complexes. The TMPD complexes were also markedly the more non-ohmic which also suggests a difference in barrier dimensions. Furthermore, while temperature had a marked effect on conductivity, with $\Delta\epsilon \approx 0.5$ ev, it had no effect on the ESR signal. Data on twenty-three such complexes are now available (Labes *et al.*, 1960). Calvin and his colleagues (Tollin, Kearns, and Calvin 1960; Kearns and Calvin, 1961) have adapted the donor-acceptor idea to the preparation of laminated layers by evaporating *o*-chloranil solutions onto phthalocyanine layers, for example. They find a 10^7 -fold increase in dark conductivity and a 10^5 increase in steady state light conductivity, which they attribute to the following electron transfers at the interface:

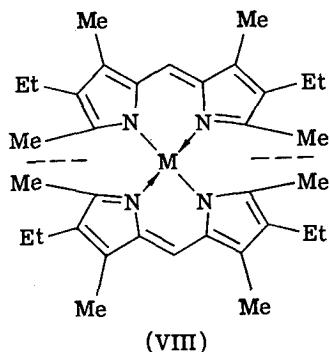
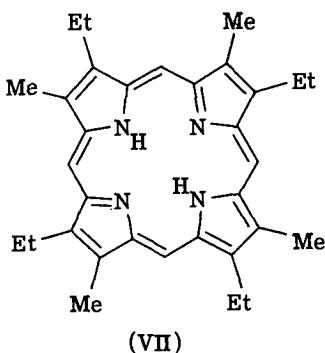


In each case the conductivity is attributed to positive holes in the phthalocyanine, the first reaction leading to an increased ESR signal (associated with CA^-), and the second to a decreased ESR signal (formation of $\text{CA}^=$). The results are of importance in connection with the possible function of laminated biological systems.

One may summarize by saying that the charge-transfer complexes furnish some of the most highly conducting of organic crystals. While something may be done to adapt the potential box model to the new situation (Eley, Inokuchi, and Willis, 1959), there are difficulties in correlating ESR signal with conductivity which may turn on the transparency of potential energy barriers between adjacent donor and acceptor molecules in the molecular stacks, i.e., on mobility differences residing in crystal structure considerations. The smaller interplane spacing may be responsible for the better conductance of TMPD complexes over DMA complexes. There is a need for more X-ray work on structure before generalizations are taken too far in this field.

C. THE PORPHYRINS

Formulas (I), (VII), and (VIII) give examples of phthalocyanine, porphyrin (actually etioporphyrin I), and the metal M complex of dipyrromethene I (Johnson *et al.*, 1959). As is well known,



both (I) and (VII) give metal derivatives by substitution of the two H atoms on the central N atoms, and formation of two coordinate links from the other N atoms to the metal to give a 4-coordinate structure. The energy gap for phthalocyanine determined in four laboratories by four different methods all lie in the range

1.44–1.87 ev (Eley, 1959a). The most recent result is 1.82 ev for phthalocyanine in the β modification (Wihksne and Newkirk, 1961). The Cu, Ni, Co, and Pt derivatives have almost identical $\Delta\epsilon$ values, in the range 1.5–1.8 ev (Fielding and Gutmann, 1957; Felmayer and Wolf, 1958). It is one of the most striking and important results that the metal has so little effect on the conductivity. Cardew and Eley (1959) found that ferric heme isolated from hemoglobin, viz., ferric protoporphyrin had an energy gap of 1.74 ev, recently re-measured as 1.80 ev. Eley and Spivey (1961b) have completed a series, viz., etioporphyrin I and its Co, Cu, Ni, and Mg derivatives which have $\Delta\epsilon$ in the range 1.81–1.99 ev, and coproporphyrin III 1.99 ev. These values agree with those for the structurally related phthalocyanines and confirm the absence of any marked effect due to the central metal atom, which at the most lowers $\Delta\epsilon$ by 0.2 ev.

The dipyrromethene complexes are closely similar in structure to the porphyrins, except that for steric reasons the two half-molecules cannot be exactly in the same plane. However, the copper and cobalt complexes of the dipyrromethene shown (VIII), have energy gaps of 1.88 and 1.85 ev, closely similar to the porphyrins. Presumably, the π -electron conjugation is linked through the metal atom in these two cases. However, in a second group of dipyrromethene complexes, where the substituents on the two pyrrole rings are Br, Me, Et, Me Et, Me, in place of the Me Et Me Me Et Me of (VIII), we obtained energy gaps of 2.3 ev, identical in fact, for the HBr salt of the dipyrromethene itself [viz., the part above the dotted line in (VIII)]. Presumably the substituent bromine atom by its electron attractive inductive effect has reduced the electron density in the nitrogen-metal bond and broken the conjugation, which would otherwise occur through the metal atom.

D. THE PROTEINS

1. *Conductivity and Current Carriers*

Baxter (1943) investigated the fibrous proteins collagen, wool, and silk, and all three had $\Delta\epsilon$ values of about 2.7 ev, compared with cotton, 1.92 ev, for water contents in the range 3–28.5%. For dry wool (in a current of dry nitrogen) a lower $\Delta\epsilon$ of 2.2 ev was found. Baxter suggested the current carriers were electrons tunneling from one adsorbed water molecule to the next. On the other hand, King and Medley (1949) found that a current passed through keratin

containing 15% water gave rise to evolution of 0.88 equivalents of hydrogen. They associated conduction with the movement of unspecified ions, and explained the effects of temperature and dielectric constant in terms of Bjerrum's ion-dissociation theory.

Eley, Parfitt, Perry, and Taysum (1953) examined the *globular* proteins plasma albumin, fibrinogen, and edestin in the dry state under dry nitrogen, finding energy gaps of 2.2, 2.7, and 2.5 ev, respectively. Adsorbed water lowered the energy gap in plasma albumin to 1.7 ev. Cardew and Eley (1959) improved the technique, by examining the proteins in the very dry state in a high vacuum apparatus. Under these conditions it is now known that most proteins have specific conductivities at 30°C of less than 10^{-18} ohm cm $^{-1}$ and energy gaps of about 2.7 ev (compare Fig. 1) (Eley and Spivey, 1960). Eley and Cardew found an energy gap for salt-free powdered hemoglobin of 2.75 ev. The absence of any observable polarization in this protein supported the idea that the current carriers were electrons and holes. Addition of about 1% NaCl gave rise to a markedly curved log κ vs. T^{-1} plots, which suggests that the straight lines observed in dialyzed purified proteins are not to be associated with conduction by fortuitously present ions. Eley and Cardew's previously unpublished graph demonstrating the salt effect is shown in Fig. 7. A number of amino acids and polyglycine gave similar results for dialyzed dry hemoglobin, and for this reason the authors associated conductivity with π -electrons moving in the CO \cdots HN bridges, which is common to all these systems. Evans and Gergely (1949) some years ago applied molecular orbital theory to the hydrogen bridges in the β -protein structure, as in Fig. 8. They showed that the energy levels fall into three narrow bands, width about 0.2 ev, the lower two filled and the upper one empty. A model in which the N atoms are assumed to possess a trigonal sp^2 arrangement of bonds gave an energy gap of 3.05 ev between the top of the highest filled and the bottom of the empty band, in good agreement with our observed values. The conductivity in a single crystal of glycine was also found to be greater *along* than *across* the layers of hydrogen bridges. The thermoelectric powers were small but observable, as would be expected for intrinsic electronic semiconductors, hemoglobin being *p*-type and glycine *n*-type, the hemoglobin result possibly arising from electron-trapping effects.

Riehl (1956, 1957) on the other hand, observed an energy gap for gelatin containing some water of 1.8 ev. Since he found an iden-

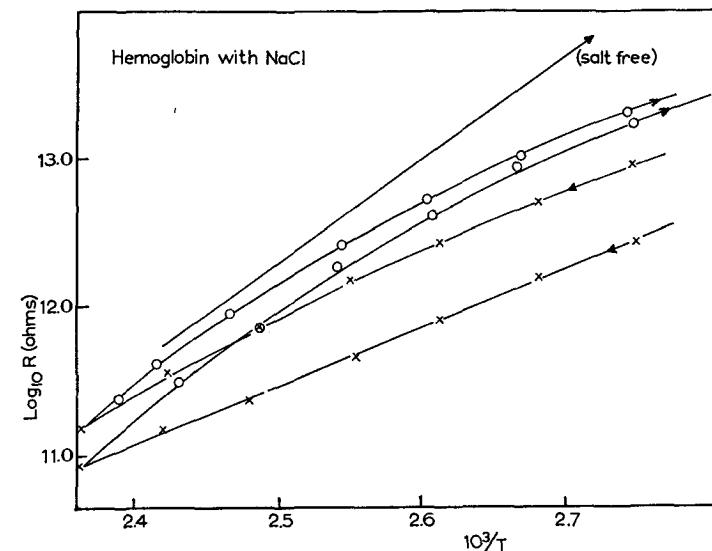


FIG. 7. The effect of 1% NaCl on dry hemoglobin, crosses: heating; circles: cooling (M. H. Cardew).

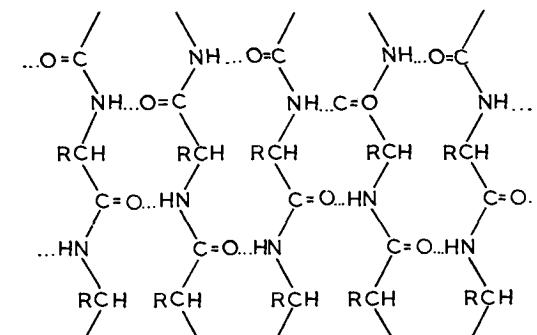


FIG. 8. The hydrogen-bridge system in a β -protein. After Evans and Gergely (1949).

tical value for ice he concluded the charge carriers in gelatin were protons associated with immobile water molecules structurally bound to the protein. However, Bradley (1957) has recently found a much lower $\Delta\epsilon$ value for electrical conduction in ice of 1.07 ev, which agrees with another value of about 1.2 ev (Gränicher *et al.*, 1957). Riehl's agreement is, therefore, not very certain.

Eley and Spivey (1960) carefully studied the effect of predenaturation by heat or alcohol on the $\Delta\epsilon$ value for solid dry hemoglobin and insulin. In the first case denaturation raised $\Delta\epsilon$ from 2.63 to 2.89 ev, and Eley and Cardew's hemoglobin with a value of 2.75 ev, therefore, was probably partly denatured. In the absence of moisture hemoglobin will survive heating at 100°C without denaturation. Spivey's data for predenaturation of hemoglobin is shown in Fig. 1, and similar results were obtained for insulin. A wide range of proteins showed energy gaps in the dry state around 2.7 ev, and very high mobilities were calculated. The results were discussed in terms of an extension of the potential barrier theory, shown in Fig. 9. In this model, the CO \cdots HN band system is assumed to extend

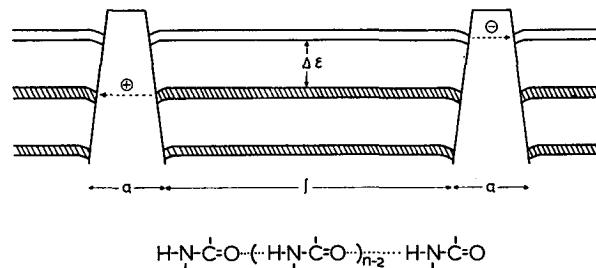


FIG. 9. Application of the potential box model to proteins (or other polymers).

to the end of the α -keratin spiral, or to a discontinuity in the β -protein sheet. An excited electron may tunnel through the barrier into such a system of H-bonds and then continue right through it, giving a rather large net displacement, hence a large mobility.

C. P. S. Taylor (1961) has investigated conductivity in cytochrome c. To explain the effect of moisture content on conductivity, he has invoked King and Medley's view, that the main effect of water is to increase the dielectric constant of the protein system, and so assist the separation of the ions. He considers the mobile ions to be protons. However, Rosenberg (1962) has shown that the same equations are applicable to the separation of electrons and positive holes.

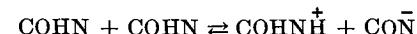
So we see that ions, protons, and electrons and positive holes have been postulated as current carriers in proteins. It would seem very probable, in the light of King and Medley's experiment, that

protons are conductors in very highly hydrated proteins. In dry proteins the electronic charge carrier is much more likely however. Davis, Eley, and Snart (1960) have found that small amounts of the electron acceptor chloranil incorporated in proteins such as serum albumin raise the conductivity of the dry protein by a factor 10⁶. It seems fairly certain that the chloranil has conferred *p*-type semiconduction on the protein, and in this state addition of moisture *lowers* the conductivity by a factor 10. Since water is an electron donor, presumably the first effect is to fill the positive holes (lower κ by 10⁶) and then to donate electrons into the protein making it *n*-type and raising κ by 10⁶, to give the net effect of a lowering of κ by a factor 10.

Assuming a band theory equation for κ_0 and a free electron (fe) mass for the charge carrier one may estimate an average mobility μ_{fe} as 10⁴ cm² V⁻¹ sec⁻¹, a similar value being found for all proteins. An explanation for this large mobility has already been given.

2. Comparison with Polyamides

In earlier work Baker and Yager (1942) demonstrated a d.c. conductivity in polyamides 66 and 610 and suggested that amide protons acted as charge carriers. Recent work in our laboratory has confirmed this point of view (Eley and Spivey, 1961). Nine polyamides were examined and it was found that the conductivity was 10³ times that for proteins and amino acids, the latter assumed to be electronic semiconductors. The polyamide conductivity showed a strong time-dependent polarization, and it was markedly non-ohmic. The log conductivity vs. T^{-1} graphs consisted of two straight lines of different slope intersecting at about 100°C, which other evidence associates the onset of rotation of the CONH groups in the chains. Below this temperature the activation energy for conductivity E_1 in $\kappa_0 \exp(-E/kT)$, is about 2.5 ev, falling to $E_2 = 1.0$ ev above the transition (we reserve the term energy gap $\Delta\epsilon = 2E$ for electronic conduction). The activation energy E_1 above 100°C is probably the energy needed for self-ionization,



where the chain segments are freely rotating a proton may be passed from one NH group to another on a neighboring chain and so on. Below the transition temperature the rate-determining step becomes that of rotation of the chain segments, and an activation energy

of 2.5 ev corresponds to the need to simultaneously break some eleven hydrogen bonds each of energy 0.217 ev.

Thus chain rotation is essential for proton conduction, and this is assisted in polyamides by the $(\text{CH}_2)_n$ sections between CONH groups, which function as internal plasticizers. In protein there is a much greater frequency of hydrogen bonds between adjacent chains, which effectively suppresses any chain rotation (the transition temperature for proteins will be above the temperature of onset of chemical decomposition) and thus one is left only with the electronic conduction, which can now be detected. Electronic conduction in the polyamides will be effectively masked by proton conduction.

3. Prosthetic Groups

The effect of heme on the electrical conductivity of globin is only to lower its energy gap from 2.97 to 2.75 ev and to raise its specific conductivity at 20°C by a factor 25, in the dry state. Crystalline heme itself has a $\Delta\epsilon$ of 1.74–1.80 ev and we conclude that in hemoglobin the semiconducting heme molecules are embedded in a relatively insulating matrix of protein (Cardew and Eley, 1959). Cytochrome c, with a $\Delta\epsilon$ of 2.60 ev is clearly similar in behavior (Eley and Spivey, 1960). Thymus nucleoprotein has $\Delta\epsilon = 2.57$ ev and contains 37% of DNA with a $\Delta\epsilon$ of 2.44 ev. If we may assume the protein moiety has a $\Delta\epsilon$ of about 2.8 ev, the observed $\Delta\epsilon$ for the nucleoprotein is intermediate between the two components (Eley and Spivey, 1960). The visual pigment rhodopsin, made up of retinene coupled to the protein opsin has $\Delta\epsilon = 2.2$ ev (Rosenberg, Orlando, and Orlando, 1961). Since retinene has an open conjugation of 10 π -electrons, from Fig. 4 we should predict $\Delta\epsilon = 2.2$ ev for this molecule, so that in this case it may be that retinene and opsin are coupled so that excited π -electrons in the retinene may enter the conduction band of the protein, in contrast to the situation just outlined for hemoglobin.

4. Hydration

The adsorption of water by protein crystals obeys the Brunauer-Emmett-Teller isotherm in most cases, which leads to a value for the number of adsorbed water molecules in a "monolayer," V_m . Pauling (1945) showed that this number corresponded roughly to the number of polar side chains in the protein, suggesting that water

molecules penetrated throughout the whole crystal to reach these adsorption sites. Pauling's view was that the relatively small water adsorption on nylon indicated that peptide links do not adsorb water, but the question is still open as there is evidence for such adsorption on diglycylglycine (Frey and Moore, 1948) and poly-glycine (Mellon *et al.*, 1948). Cardew and Eley (1958) have studied water adsorption on hemoglobin, which is known from the X-ray

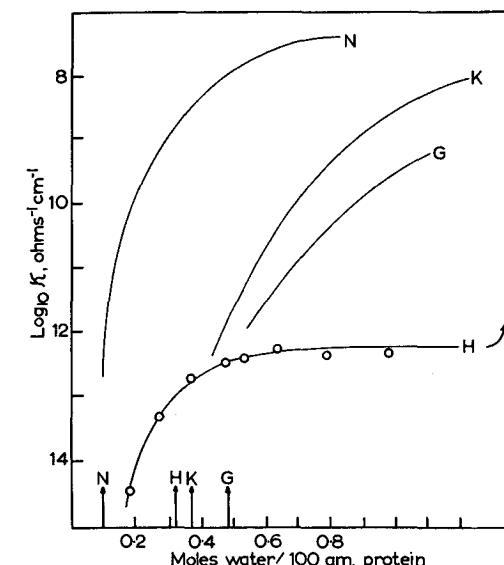


FIG. 10. Adsorbed water vapor and electrical conductivity in proteins; N-nylon, K-keratin, G-gelatin, H-hemoglobin. The arrows mark BET monolayer coverage.

studies of Perutz to take up water only on the molecular surfaces. The molecular dimensions indicate a molecular surface area corresponding to a saturated film of 1160 moles water per 10^5 gm protein and the chemical analysis indicates a content of 435 moles of polar side chains, which are probably mostly on the molecular surface. The observed BET monolayer was 319 moles of water, corresponding to 73% of the total polar side chains present.

The effect of adsorbed water on proteins is to increase the electrical conductivity. Our own experiments (Eley and Spivey, 1961a) on hemoglobin indicate that a definite saturation value of the con-

ductivity is reached at a water content rather above the BET V_m value. This data, together with other published results, are shown in Fig. 10, and it seems clear that water has a fairly specific action in increasing conductivity in different proteins. The electron donor action of water at low water contents was established by its action in *decreasing* the conductivity of *p*-type proteins, i.e., proteins containing traces of an electron acceptor. We therefore suppose the first action of water is to donate electrons into the conduction band of the protein and this suggests some of the adsorption sites are either CONH groups or are at least adjacent to them. This conductivity will saturate when all the surface-available polar groups are occupied, i.e., around V_m , which is about one-quarter of the available surface in hemoglobin. As the total available surface becomes covered a proton conductivity will become possible, by a Grotthus mechanism involving neighboring adsorbed water molecules, as suggested by Riehl. This is the mechanism to be expected in keratin with 15% adsorbed water which corresponds to 2.5 V_m . The very large effect of water on the conductivity in nylon is probably due to a third action, i.e., a plasticizing action permitting chain rotation and proton transfer between NH groups on adjacent chains, as suggested by Baker and Yager (1942).

E. NUCLEIC ACIDS AND NUCLEOPROTEINS

Deoxyribonucleic acid has a structure of two intertwined helical chain molecules built up of phosphate-ribose base repeating units. The bases point toward the middle of the spiral and are paired by hydrogen bonds, adenine-thymine and guanine-cytosine, so as to hold the two spirals together (Crick and Watson, 1954). Each base contains 10 π -electrons, thus each base pair 20 π -electrons, and the base pairs are arranged one on top of the other like a pile of coins, with an interplane spacing of 3.4 Å close to that found in graphite. Figure 4 would suggest that an arrangement of 10 π -electron molecules in a crystal should have an energy gap of 3.0 ± 1.0 ev. If conjugation were perfect a 20 π -electron molecule crystal might have $\Delta\epsilon = 1.5 \pm 0.5$ ev. A number of samples of DNA in the dry state have been examined and found to have a specific resistivity of 5×10^{11} ohm cm at 400°K and $\Delta\epsilon = 2.42$ ev (Eley and Spivey, 1962). This suggests that the conjugation due to hydrogen bonds between the base pairs is relatively weak, so that the individual molecules function separately as 10 π -electron units, which agrees

with the rather small stabilizing energy calculated by molecular orbital theory (Pullman and Pullman, 1959). The conclusion is that π -electron conduction occurs down the axis of the molecule, which agrees with the orbital overlap calculations of Ladik (1960). A sample of yeast RNA gave very similar results to DNA, which perhaps is surprising, in view of structural differences. It is very likely that all the DNA specimens were denatured to some extent. Duchesne *et al.* (1960) have reported energy gaps of 1.80 ev for DNA, but they indicate that adsorbed water may be present, and this, in our view, explains the difference between the two values.

Tobacco mosaic virus (6% RNA) has been found to have in the dry state $\Delta\epsilon = 2.92$ ev, while thymus nucleoprotein (37% DNA) has $\Delta\epsilon = 2.57$ ev. These values are what one might expect for a simple mixture effect for nucleic acid and protein.

Using band theory equations and a free electron mass, a mobility of $10^4 \text{ cm}^2 \text{ V}^{-1} \text{ sec}^{-1}$ may be calculated from κ_0 , similar to that for proteins. In the same way we may assume that thermal excitation is necessary for an electron to tunnel from one DNA spiral to the next, but that once in the activated state it may pass along the whole length of the molecule.

IV. Photoconductivity

The energy to generate an electron and a hole may be provided by a light quantum as well as thermally but the exact relation between the two is in doubt. The photocurrent-wavelength graph in many cases follows the absorption spectrum of the organic solid (Lyons and Morris, 1957). This is so where the two electrodes are disposed on the illuminated surface of the specimen, the so-called surface cell (Chynoweth and Schneider, 1954). In a second type, the sandwich cell, the specimen is placed between one electrode and a transparent electrode, on which the light is incident, and different behavior is found here. Values of the energy gap have been equated to the quantum energy for the first detectable photoconduction (Inokuchi), and to the energy corresponding to half the maximum in the photoconduction curve, the $\lambda_{1/2}$ method (Vartanyan). The latter method is based on band theory as applied to inorganic substances (Moss, 1952). Agreement with the thermal energy gap has been reported for both methods, in the first case for polyacenes (Akamatsu and Inokuchi, 1952; Inokuchi, 1954) and in the second case for phthalocyanines (Vartanyan and Karpovich, 1958). Var-

Vartanyan has also extensively used the method of "photoelectric lines" (Moss, 1952) to determine the threshold quantum which agreed with the $\lambda_{1/2}$ method for the phthalocyanines to about 0.05 ev. Vartanyan has found agreement between the thermal energy gap and this threshold quantum for cationic, anionic, and non-ionic dyestuffs (Terenin, 1961).

Earlier we mentioned the views of Lyons (1957) and Terenin (1957), that semiconduction occurs via the triplet state, viz., by a narrow energy band formed from this molecular state and lying slightly below it in energy. Photoconduction can only occur via the optically permitted singlet-singlet transition, so that the observed activation energy should be *greater* than that for semiconduction, if the latter involved the triplet state. Terenin (1961) now concludes that semiconduction involves the singlet state because of the agreement in activation energies mentioned above.

Rosenberg (1958, 1961a; Rosenberg and Camiscoli, 1961) has pointed out that the temperature dependence of the photocurrent provides definite evidence for the triplet state in certain instances. For anthracene, Compton, Schneider, and Waddington (1957) found that $i = i_0 \exp(-E/kT)$ with $E = 0.17$ ev, a result confirmed by Rosenberg. If we refer to Fig. 11, we see that photoexcitation to the first excited singlet may be followed by a thermally activated transition to the triplet state regarded as responsible for photoconduction. The thermal activation may involve one or more vibrational quanta and the above E value corresponds to one such quantum. The rise and decay of photoconductivity in anthracene followed monomolecular kinetics with a time constant of 0.6 millisecond, and such behavior is expected for the triplet state model.

Adsorbed gases may increase or decrease surface photoconduction currents (Vartanyan, 1948), and Schneider and Waddington (1956) have classified the effect of gases on anthracene according to their tendency to accept or donate electrons, electron acceptors leading to increased currents, suggesting that most of the conduction occurs by positive charge carriers. Rosenberg (1961) found a temperature-photocurrent activation energy of 0.37 ev for *trans*- β -carotene, which is unaffected by the presence of oxygen although the photocurrent is raised 25-fold. However, oxygen lowered the semiconduction energy gap of β -carotene from 1.52 to 1.29 ev and raised the dark conduction at 25°C by 1100-fold. Rosenberg suggested that the fact that the triplet state of chlorophyll at 1.43 ev

lies midway between these values gives support to the theory of a donor-carotene-acceptor complex in photosynthesis, due to Platt (1959).

The triplet state mechanism has been applied to explain the dark observed for chlorophyll by Rosenberg and Camiscoli (1961) and Snart and the present author (Snart, 1961).

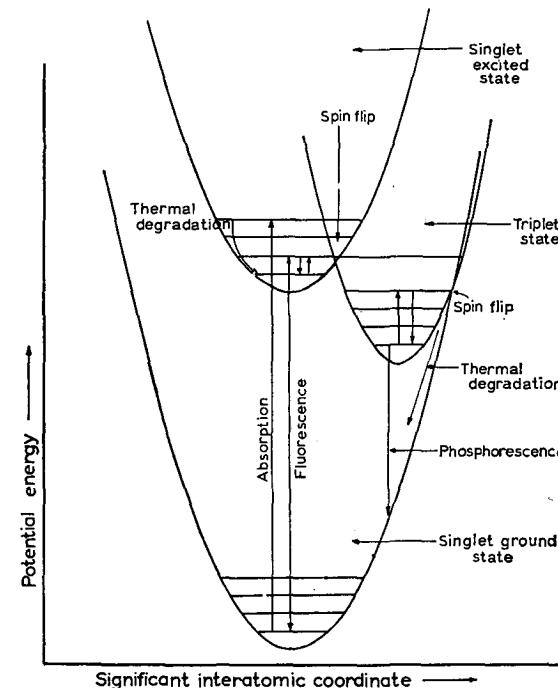


FIG. 11. The triplet state in photoconduction. After B. Rosenberg (1958).

V. Biological Systems

Semiconduction mechanisms will be discussed in relationship to cytochrome systems, chloroplasts, retinal rods, radiation effects, and carcinogenesis. In addition, semiconduction data may throw light on prosthetic group-protein interactions.

A. CYTOCHROME SYSTEMS

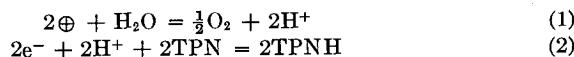
Electron transfer between cytochromes localized in mitochondria may occur by a semiconduction mechanism. However, Cardew and

Eley (1959) have estimated the resistivity of dry protein to be too high by a factor 10^{16} to pass the electron current equivalent to the observed respiration rate in a sea urchin egg. Hydration of the proteins concerned may lower the resistivity by 10^8 but a formidable factor still remains. It is possible that electron acceptors such as quinones may induce highly conducting *p*-type protein systems (in hydrophobic environments). Experiments on mitochondria themselves will be needed to give a final answer, but the semiconduction mechanism seems not very probable at present.

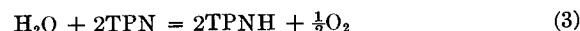
B. CHLOROPLASTS

Katz (1949) and Bradley and Calvin (1955) have discussed how the chloroplast may function as a semiconductor-like unit [cf. a recent review by Calvin (1959)]. Thus, the absorption of light may produce electrons and positive holes in the chlorophyll. If this is sandwiched between a *p*-type layer and an *n*-type layer of lipid or lipoprotein (the chloroplast has a laminar structure) then the electrons and holes may diffuse away in opposite directions to start off the chemical cycles of reduction and oxidation.

An absorbed quantum of red light would produce an electron and a positive hole in the chlorophyll with free energy 40 kcal per mole of which 5 kcal might be lost in transfer leaving 35 kcal. This chemical potential is then available to drive the spatially separated oxidation and reduction reactions.



giving the over-all hydrogen transfer



In this primary reaction hydrogen, provided by photolysis of water, gives reduced triphosphopyridine nucleotide. Calvin has suggested that the semiconducting layers might be carotene molecules embedded in an insulating matrix of lipid. Another possibility suggested by our own recent work is that the *p*-type layer could be a protein-vitamin K complex, and the *n*-type layer a protein-water complex. Vitamin K is a quinone that is known to be an essential chloroplast component for the Hill reaction (Lynch and French, 1957). Spin resonance spectra have been observed in wet chloro-

plasts and tentatively attributed to the trapped electron responsible for reaction (2) (Sogo, Pon, and Calvin, 1957).

Photoconductivity and semiconductivity in chlorophyll has been investigated by Terenin *et al.* (1959), by Rosenberg and Camiscoli (1961), and by Snart and Eley (1961). For crystalline chlorophyll Rosenberg found a semiconduction activation energy of 1.44 ev, and Snart found an identical value for an amorphous layer of the mixed chlorophylls. Rosenberg found a temperature coefficient of photoconduction corresponding to 0.33 ev whereas Snart found 0.15 ev. Rosenberg refers to Rabinowitch's observation that the primary photophysical process in photosynthesis is temperature independent, in contrast to the above result. However, he concludes that until the chlorophyll-protein complex has been examined we cannot rule out photoconduction as a determining process in photosynthesis.

C. CARCINOGENESIS

There have been several attempts to explain the relative carcinogenic activity of polyaromatic compounds. Thus A. and B. Pullman (1955) associated this action with a high electron density in the "K region" of the hydrocarbon. Mason (1958, 1959) considered that the primary process is one of electron transfer from the highest filled band of the protein (Evans-Gergely model) to the empty levels in the hydrocarbon as in Fig. 12. This would create a positive hole in the valence band of the protein. He supposes that for electron transfer a close matching of electron energy levels are necessary, and from his matching procedure concludes that the excitation energy in the hydrocarbon must lie in the range 3.23 ± 0.19 ev calculated for the separation $E_2 - E_1$ of first and second valence bands of the protein. While the Evans-Gergely calculations are very approximate, the close agreement between the thermal energy gap ($E_3 - E_2$) calculated by them and that found by Eley and Cardew supports the use of their calculation in this way. It is a matter of some interest that the five most active hydrocarbons out of a total of thirty-four appear to have an excitation energy in this range, while most of the others lie outside it.

The model is open to criticism in that

- (a) The energy level schemes for the two molecules should be adjusted with respect to the free electron as energy zero, and not by equating the valence bands as shown in Fig. 12.
- (b) The exact matching of energy levels is a requirement for

single electron bond formation, but not for electron transfer, from protein to hydrocarbon, which simply requires that the highest filled level of the protein is higher than the lowest unfilled level of the hydrocarbon (Eley, 1959b).

This kind of complex electron transfer is found from a protein to a very strong acceptor such as chloranil, and gives a positive hole in the protein and a large rise in conductivity. However, hydrocarbons are relatively weak acceptors, so, accepting Mason's correlation as

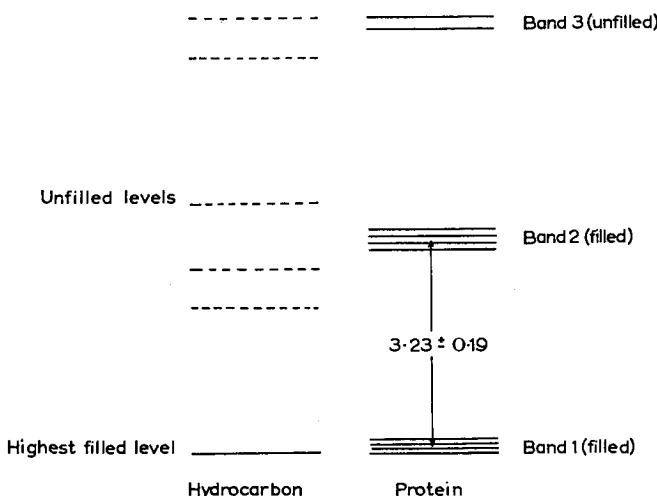


FIG. 12. Electron transfer from protein to hydrocarbon as postulated by R. Mason (1958).

correct, its basis may well be single-electron bond formation rather than complete electron transfer. The conductivity method would appear adequate to test Mason's hypothesis and some work has already been started at Nottingham. More recently Hoffmann and Ladik (1961) have calculated the energy band structure for DNA using the molecular orbital method. They favor the idea of electron transfer from the DNA to the carcinogenic hydrocarbon. They further postulate the presence of an electric field to cause the movement of positive charge to the end of the DNA particle, a distance of perhaps 15μ , where its Coulombic repulsion energy effect may initiate separation of the two chains, the first stage in DNA duplication. The

charges could equally well arise from a radiation effect as from a carcinogenic agent.

D. ENERGY CONVERSION IN RETINAL RODS

If a photon of green light is absorbed by any one of millions of rhodopsin molecules in a dark-adapted rod of a human retina it has been shown there is a 30% chance of generating a nervous impulse (Hecht, Shlaer, and Pirenne, 1942). Haggins and Jennings (1959) have, therefore, pointed out the need for a mechanism for energy migration in the system of rods. These authors found the retinal rods to be free of all photodichroism, which suggests an efficient process of intermolecular radiationless energy transfer. However, quite concentrated solutions of vitamin A, a molecule closely related to retinene, the prosthetic group of rhodopsin, show a strongly polarized fluorescence which indicates a relatively inefficient radiationless energy transfer process. They conclude that the photosensitizing action of rhodopsin must depend upon some other process of signal transmission such as photoconduction, electrolytic conduction, or molecular diffusion. Rosenberg, Orlando, and Orlando (1961) found a semiconduction activation energy of 2.29 ± 0.3 ev for both water-washed and sucrose-washed dried sheep rods, but only the second specimen showed photoconductivity which they attributed to the presence of retained water in the specimen. The action spectrum was not measured accurately, but the indication was that it fitted in with the known optical absorption of rhodopsin. With continued exposure to light the photoconduction decreased, which may be associated with the conversion of hindered *cis*-retinene to all-*trans*-retinene, still coupled to opsin, a kind of light adaptation phenomenon.

E. IRRADIATION OF PROTEINS

It was pointed out by Cardew and Eley (1959) that a quantum of energy 3 ev (4112 Å) should suffice to photoexcite an electron into the conductivity band of a protein, leaving a positive hole in the valence band. The effect of hydration may well be to lower this energy, as determined from semiconduction investigations in the first instance for plasma albumin (Eley *et al.*, 1953). Rosenberg (1962) has found a linear relation between lowering of energy gap and water content in hemoglobin. Since hemoglobin adsorbs water only on the molecular surfaces, other proteins may well show much more marked effects of this kind, and it has been emphasized that

effects of hydration vary from protein to protein (Eley and Spivey, 1961a and Fig. 10). Thus the photoinactivation of urease, and the photodissociation of carbon monoxyhemoglobin may result from excitons (electron-hole pairs) produced in the protein and migrating to the reaction center. Very recently Allen and Ingram (1961) have found that light of 3600 Å (3.4 ev) produces strong free radical ESR signals in egg albumin and bovine serum albumin held at liquid nitrogen temperature. Three amino acids in aqueous solution gave no signal while the fourth, leucine, gave only a small one. For this reason the authors associated the signal with the excitation of an electron into the conductivity band of the solid protein. The signal was retained on warming the powdered proteins in vacuum up to room temperature, but disappeared on warming in the case of the aqueous solutions. In the case of egg albumin the crystals after irradiation, on raising to room temperature in vacuo showed a well-defined doublet, suggesting that the electron concerned is trapped at a proton, perhaps in the hydrogen bonding system. Irradiation of powders at liquid nitrogen temperature was also carried out by light of 2537 Å wavelength. This produced a more asymmetrical signal which did not disappear on warming to room temperature even in the presence of moisture.

It is interesting also that Rajewsky and Redhardt (1961) found a strong transient line after X-irradiation of fibrous proteins and amino acids, with a *g*-value in the neighborhood of the free spin value. A rise in electrical conductivity was observed, which fell off after the end of irradiation.

VI. Conclusion

Plenty of evidence is available for the semiconducting properties of a wide range of molecules of biological interest in the dry crystalline or amorphous state. There is also evidence for semiconduction in certain organized systems such as chloroplasts and retinal rods, and the laminar structure of such systems may well have electrical significance as discussed by Calvin (cf. a review, Fernández-Morán, 1959). At various times A. Szent-Györgyi has raised the topics of semiconductivity (1941), hydration (1957), and charge transfer (1960). The present author would suggest that an important function of adsorbed water is to transfer electrons to biological molecules such as proteins and nucleic acids, so as to form *n*-type laminae. The interaction of quinones and other electron acceptors, such as

a free radical source with proteins in lipoidal environments then forms *p*-type laminae. The juxtaposition of these structures, as in transistors, will allow a wide range of electron conduction behavior, and permit the localization of oxidation and reduction behavior in chloroplasts, mitochondria, etc. This localization may be coupled with stereochemical considerations to give a high degree of specificity. This new hypothesis may serve to unify the above three concepts of Szent-Györgyi, and is offered as a subject for further experimentation.

REFERENCES

- Aftergut, S., and Brown, G. (1961). *Nature* **189**, 827.
- Akamatu, H., and Inokuchi, H. (1950). *J. Chem. Phys.* **18**, 810.
- Akamatu, H., and Inokuchi, H. (1952). *J. Chem. Phys.* **20**, 1481.
- Akamatu, H., and Inokuchi, H. (1959). In "Carbon: Proceedings of the Third Biennial Conference" (S. Mrozwoski, M. Studebaker, and P. L. Walker, eds.), p. 51. Pergamon Press, New York.
- Akamatu, H., Inokuchi, H., and Matsunaga, Y. (1956). *Bull. Chem. Soc., Japan* **29**, 213.
- Allen, B. T., and Ingram, D. J. E. (1961). "Free Radicals in Biological Systems," p. 215. Academic Press, New York.
- Baker, W. O., and Yager, W. A. (1942). *J. Am. Chem. Soc.* **64**, 2171.
- Baxter, S. (1943). *Trans. Faraday Soc.* **34**, 207.
- Baxter, S., and Cassie, A. B. D. (1941). *Nature* **148**, 408.
- Bradley, D. F., and Calvin, M. (1955). *Proc. Natl. Acad. Sci. U.S.* **41**, 563.
- Bradley, R. S. (1957). *Trans. Faraday Soc.* **53**, 687.
- Bree, A., and Lyons, L. E. (1960). *J. Chem. Soc.* p. 5719.
- Calvin, M. (1959). *Rev. Modern Phys.* **31**, 147, 157.
- Cardew, M. H., and Eley, D. D. (1958). "Fundamental Aspects of the Dehydration of Foodstuffs" (Society of Chemical Industry), p. 24. Macmillan, New York.
- Cardew, M. H., and Eley, D. D. (1959). *Discussions Faraday Soc.* **27**, 115.
- Carswell, D. J., Ferguson, J., and Lyons, L. E. (1954). *Nature* **173**, 736.
- Chynoweth, A. G., and Schneider, W. G. (1954). *J. Chem. Phys.* **22**, 1021.
- Compton, D. M. J., Schneider, W. G., and Waddington, T. C. (1957). *J. Chem. Phys.* **27**, 160.
- Crick, F. H. C., and Watson, J. D. (1954). *Proc. Roy. Soc. A* **223**, 80.
- Davis, K. M. C., Eley, D. D., and Snart, R. S. (1960). *Nature* **188**, 724.
- Duchesne, J., Depireux, J., Bertinchamps, A., Cornet, N., van der Kaa, J. M. (1960). *Nature* **188**, 405.
- Eley, D. D. (1948). *Nature* **162**, 819.
- Eley, D. D. (1959a). *Research* **12**, 1293.
- Eley, D. D. (1959b). *Discussions Faraday Soc.* **27**, 242.
- Eley, D. D., and Inokuchi, H. (1959). *Z. Electrochem.* **63**, 29.
- Eley, D. D., and Parfitt, G. D. (1955). *Trans. Faraday Soc.* **51**, 1529.
- Eley, D. D., and Spivey, D. (1960). *Trans. Faraday Soc.* **56**, 1432.

- Eley, D. D., and Spivey, D. (1961a). *Nature* **188**, 725.
- Eley, D. D., and Spivey, D. (1961b). *Trans. Faraday Soc.* **57**, 2280.
- Eley, D. D., and Spivey, D. (1962). *Trans. Faraday Soc.* **58**, 405, 411.
- Eley, D. D., and Willis, M. R. (1960). In "Symposium on Electronic Conductivity in Organic Solids," Duke Univ. Conf. (B. H. Kallmann, ed.). Interscience, New York. In preparation.
- Eley, D. D., Parfitt, G. D., Perry, M. J., and Taysum, D. H. (1953). *Trans. Faraday Soc.* **49**, 79.
- Eley, D. D., Inokuchi, H., and Willis, M. R. (1959). *Discussions Faraday Soc.* **28**, 54.
- Epstein, A., and Wildi, B. S. (1960). *J. Chem. Phys.* **32**, 324.
- Evans, M. G., and Gergely, J. (1949). *Biochim. et Biophys. Acta* **3**, 188.
- Felmayer, W., and Wolf, I. (1958). *J. Electrochem. Soc.* **105**, 141.
- Fernández-Morán, H. (1959). *Rev. Modern Phys.* **31**, 319.
- Fielding, P. E., and Gutmann F. (1957). *J. Chem. Phys.* **26**, 411.
- Fox, D. (1959). *J. Phys. Chem. Solids* **8**, 439.
- Frey, H. J., and Moore, W. J. (1948). *J. Am. Chem. Soc.* **70**, 3644.
- Garrett, C. G. B. (1959). In "Semiconductors" (N. B. Hannay, ed.), p. 634. Reinhold, New York.
- Gränicher, H., Jaccard, C., Scherrer, P., and Steinmann, A. (1957). *Discussions Faraday Soc.* **23**, 50.
- Hagins, W. A., and Jennings, W. H. (1959). *Discussions Faraday Soc.* **27**, 180.
- Hall, G. G. (1959). *Rev. Progr. Phys.* **22**, 1.
- Hecht, S., Shlaer, S., and Pirenne, M. H. (1942). *J. Gen. Physiol.* **25**, 819.
- Hoffmann, T. A., and Ladik, J. (1961). *Cancer Research* **21**, 474.
- Holmes-Walker, W. A., and Ubbelohde, A. R. (1954). *J. Chem. Soc.* p. 720.
- Huggins, C. M., and Le Blanc, O. H. (1960). *Nature* **186**, 552.
- Inokuchi, H. (1954). *Bull. Chem. Soc. Japan* **27**, 1.
- Inokuchi, H. (1956). *Bull. Chem. Soc. Japan* **29**, 131.
- Johnson, A. W., Kay, I. T., Markham, E., Price, R., and Shaw, K. B. (1959). *J. Chem. Soc.* p. 3416.
- Jordan, P. (1938). *Naturwissenschaften* **26**, 693.
- Kainer, H., Bijl, D., and Rose-Innes, A. C. (1960). *J. Chem. Phys.* **30**, 765.
- Kallmann, H., and Pope, M. (1960). *Nature* **186**, 31.
- Kasha, M. (1959). *Rev. Modern Phys.* **31**, 162.
- Katz, E. (1949). In "Photosynthesis in Plants" (J. Franck and W. E. Loomis, eds.), Iowa State College Press, Ames, Iowa.
- Kearns, D. R., and Calvin, M. (1961). *J. Am. Chem. Soc.* **83**, 2110.
- Kepler, R. G. (1960). *Phys. Rev.* **119**, 1226.
- King, G., and Medley, J. A. (1949). *J. Colloid Sci.* **4**, 1, 9.
- Kommandeur, J., and Singer, L. S. (1960). In "Symposium on Electronic Conductivity in Organic Solids" (B. H. Kallmann, ed.). Duke Univ. Conf. Interscience, New York. In preparation.
- Labes, M. L., Sehr, R., and Bose, M. (1960). *J. Chem. Phys.* **33**, 868.
- Ladik, J. (1960). *Acta Phys. Acad. Sci. Hung.* **11**, 239.
- Le Blanc, O. H. (1960). *J. Chem. Phys.* **33**, 626.
- Lonsdale, K. (1937). *Proc. Roy. Soc. A* **159**, 149.
- Lynch, V. H., and French, C. S. (1957). *Arch. Biochem. Biophys.* **70**, 382.

- Lyons, L. E. (1957). *J. Chem. Soc.* p. 5001.
- Lyons, L. E., and Morris, G. C. (1957). *J. Chem. Soc.* p. 3648.
- Many, A., Harnik, E., and Gerlich, D. (1955). *J. Chem. Phys.* **23**, 1733.
- Mason, R. (1958). *Nature* **181**, 820.
- Mason, R. (1959). *Discussions Faraday Soc.* **27**, 129.
- Mellon, E. F., Korn, A. H., and Hoover, S. R. (1948). *J. Am. Chem. Soc.* **70**, 3040.
- Mette, H., and Pick, H. (1953). *Z. Physik.* **134**, 566.
- Möglich, F., and Schön, M. (1938). *Naturwissenschaften* **26**, 199.
- Moss, T. S. (1952). "Photoconductivity in the Elements," p. 15. Academic Press, New York.
- Mulliken, R. S. (1952). *J. Am. Chem. Soc.* **74**, 811.
- Nelson, R. C. (1951). *J. Chem. Phys.* **19**, 798.
- Northrop, D. C. (1959). *Proc. Phys. Soc.* **74**, 756.
- Northrop, D. C., and Simpson, O. (1954). *Proc. Phys. Soc.* **67**, 892.
- Northrop, D. C., and Simpson, O. (1956). *Proc. Roy. Soc. A* **234**, 123.
- Pauling, L. (1945). *J. Am. Chem. Soc.* **67**, 555.
- Platt, J. R. (1959). *Science* **129**, 372.
- Pullman, A., and Pullman, B. (1955). *Advances in Cancer Research* **3**, 117.
- Pullman, B., and Pullman, A. (1959). *Biochim. et Biophys. Acta* **36**, 343.
- Putzeiko, E. K. (1948). *Doklady Akad. Nauk S.S.R.* **59**, 471.
- Rajewsky, B., and Redhardt, A. (1961). *Intern. Biophys. Congr., Stockholm, Abstracts*, p. 98.
- Riehl, N. V. (1955). *J. Phys. Chem. (U.S.S.R.)* **29**, 1152.
- Riehl, N. V. (1956). *Naturwissenschaften* **43**, 145.
- Riehl, N. V. (1957). *Kolloid-Z.* **151**, 66.
- Rosenberg, B. (1958). *J. Chem. Phys.* **29**, 1108.
- Rosenberg, B. (1961). *J. Chem. Phys.* **34**, 812.
- Rosenberg, B. (1962). *J. Chem. Phys.* **36**, 816.
- Rosenberg, B., and Camiscoli, J. F. (1961). *J. Chem. Phys.* **35**, 982.
- Rosenberg, B., Orlando, R. A., and Orlando, J. M. (1961). *Arch. Biochem. Biophys.* **93**, 395.
- Schneider, W. G., and Waddington, T. C. (1956). *J. Chem. Phys.* **25**, 358.
- Slough, W., and Ubbelohde, A. R. (1957). *J. Chem. Soc.* p. 983.
- Smaller, B., Isenberg, I., and Baird, S. L. (1961). *Nature* **191**, 168.
- Snart, R. S., and Eley, D. D. To be published.
- Sogo, P. B., Pon, N. G., and Calvin, M. (1957). *Proc. Natl. Acad. Sci. U.S.* **43**, 387.
- Szent-Györgyi, A. (1941). *Nature* **148**, 157.
- Szent-Györgyi, A. (1946). *Nature* **157**, 875.
- Szent-Györgyi, A. (1957). "Bioenergetics." Academic Press, New York.
- Szent-Györgyi, A. (1960). "Introduction to a Submolecular Biology." Academic Press, New York.
- Taylor, C. P. S. (1961). *Intern. Biophys. Congr., Stockholm, Abstracts*, p. 214.
- Terenin, A. N. (1957). *Radiotech. i. Electron. (U.S.S.R.)* **1**, 1127.
- Terenin, A. N. (1961). *Proc. Chem. Soc.* p. 321.
- Terenin, A. N., Putzeiko, E., and Akimov, I. (1959). *Discussions Faraday Soc.* **27**, 83.

- Tollin, G., Kearns, D. R., and Calvin, M. (1960). *J. Chem. Phys.* **32**, 1013, 1020.
 Vartanyan, A. T. (1948). *J. Phys. Chem. (U.S.S.R.)* **22**, 769.
 Vartanyan, A. T., and Karpovich, I. A. (1958). *J. Phys. Chem. (U.S.S.R.)* **32**, 274.
 Wallwork, S. C. (1961). *J. Chem. Soc.* p. 494.
 Weigl, J. W. (1956). *J. Chem. Phys.* **24**, 364.
 Weiss, J. (1942). *J. Chem. Soc.* p. 245.
 Wihksne, K., and Newkirk, A. E. (1961). *J. Chem. Phys.* **34**, 2184.
 Wilk, M. (1960). *Z. Elektrochem.* **64**, 930.

On the Molecular Organization of Biological Transducing Systems

D. E. GREEN AND S. FLEISCHER

*Institute for Enzyme Research, University of Wisconsin,
Madison, Wisconsin*

I. Introduction	381
II. The Particulate State and Biological Organization	383
III. Experimental Approaches to the Study of the Mitochondrion	384
IV. Unit of Mitochondrial Action	386
V. Electron Transport Chain	389
VI. Fragmentation and Reconstruction of the Electron Transport Chain	391
VII. Structural Protein	395
VIII. Lipid	396
IX. Concept of the Elementary Particle	402
X. Mechanism of Electron Transport	406
XI. Oxidative Phosphorylation	410
XII. The Assembly of Mitochondria <i>in Vivo</i>	414
XIII. Pharmacology and Transducing Systems	415
XIV. Universality of the Molecular Principles of Mitochondrial Structure and Function	415
References	418

I. Introduction

Living systems contain a variety of structured devices or machines that transform energy from one form to another—radiant to chemical energy, chemical to mechanical energy, chemical to radiant energy, etc. These devices known as transducing systems are fundamental to the work performance of living cells and without such devices all cellular activity would be impossible (Green and Fleischer, 1960a). Any physiological process such as transmission of a nerve impulse, muscular contraction or vision may involve the close collaboration of several machines, e.g., the mitochondrion and the membrane apparatus in nerve transmission; the mitochondrion and the myofibril in muscular contraction; and the mitochondrion with elements of the retinal rods, retinal cones, and nerve cells in vision.

Transducing systems as a class have several properties in common. They are all contained within subcellular particles or mem-

branes of highly intricate structure. They do not occur singly but rather in clusters or large numbers of repeating units that are embedded in some structured medium. In general, these systems contain a high proportion of lipid and exhibit double-membraned (lamellar) structure.

The transduction is a molecular process that is consummated by the cyclical performance of specialized transducing compounds, and the organized structure of the transducing systems is designed or hand-tailored as it were for the performance of these specialized molecules. Chlorophyll in the photosynthesizing chloroplasts (Gaffron *et al.*, 1957), luciferin in the luminescent organ of the firefly (Rhodes and McElroy, 1958), rhodopsin in the retinal cones and rods (Wald, 1956), and myosin and actin in the myofibrils (Weber, 1958) are among the well-documented transducing molecules. The key to transduction is clearly the mode of action of these transducing molecules but it must be remembered that only in the context of the organized system can the action of these molecules be properly studied. Structure and function are so inextricably bound together that the loss of one inevitably leads to the loss of the other. The cycle of events involving the absorption of photons by chlorophyll and the transmission of electrons to other components of the electron transport chain requires the structured milieu of the chloroplast and when this structured system is modified beyond the point of no return, chlorophyll no longer can function in precisely the same catalytic fashion as it does under physiological conditions.

The complexity of the structure-function relationship, the difficulties of working with particulates, and various other technical factors have raised formidable barriers to the study of transducing systems. Until very recently these basic systems were assiduously avoided by biochemists. The mitochondrion, however, was the one transducing system, the study of which could not be avoided or postponed by biochemists and this for two reasons. Firstly, the mitochondrion is the seat of the Krebs citric acid cycle and the prime source of oxidative energy (Lehninger and Kennedy, 1949; Schneider and Potter, 1949). With all key biochemical processes being directly or indirectly energized by adenosine triphosphate (ATP) it was impossible to neglect the key system for synthesis of ATP. Secondly, the transduction process implemented by the mitochondrion is chemical in nature, and a problem of this type is not too far removed from chemical experience.

II. The Particulate State and Biological Organization

When we say that the mitochondrion is a particle, it is important to know what is meant by such a term. A particle is a water-insoluble complex that is readily sedimented in a relatively low gravitational field. There is a continuum between particles and soluble proteins. One and the same complex may behave as a particle under one set of conditions and as soluble protein under another. The distinction between particles and soluble proteins is rather arbitrary and when made should have no more than operational or descriptive value. Thus, there are particles the monomers of which may have relatively low molecular weights, and soluble proteins that may have relatively high molecular weights. The particulate state need not necessarily be an expression of molecular size, and may be a reflection of a preponderance of water-repelling groups in the molecular unit.

Just as we can talk about pure soluble proteins so there are analogous criteria that can be developed to describe what we may call pure particles. Pure is used in the sense that each individual particle has exactly the same chemical composition as every other particle in the population but not necessarily the same size. There are many ways of separating and purifying particles, and the isolation of a pure virus particle or a pure mitochondrial particle or a pure submitochondrial particle is now a relatively straight-forward procedure.

A particle could be a single protein or a polymer thereof or a complex of a number of proteins linked together in a precise fashion. In theory, cell particulates could represent random statistical aggregates; in practice, the evidence is all to the contrary. When properly isolated, cellular particulates are highly reproducible molecules or complexes (by chemical criteria) and they can be described with a degree of precision comparable to the precision that can be attained in describing any soluble protein in homogeneous state.

Cell particulates are held together by various types of bonds—covalent, hydrophobic, hydrogen, electrostatic, etc. If we accept the thesis that particulates are either molecules or molecular complexes of a high degree of reproducibility then it is possible and feasible to determine the structure of particulates by the special application of the methods of classic chemistry. These elementary and obvious truisms have to be emphasized to dispel the widely held notion that particulates lie beyond the framework of chemical reasoning and

study. On the contrary, there is no reason why either insolubility in water or relatively large molecular size should bar the application of chemical methodology.

The particulate state is no happenstance and indeed offers logistic advantages for introducing structure into a molecular array. Nature usually, if not invariably, selects the particulate state for the arrangement of proteins and other components in a patterned sequence. The solid state provides a degree of rigidity and tensile strength that is more difficult to achieve in solution. Thus the particulate state and biological organization are usually two sides of the same coin. Water insolubility and the supramolecular state are the hallmarks of biological transducers and there is no alternative but to accept these properties and to find solutions for dealing with these special giant molecules.

Now we may define the problem of transducing systems somewhat more precisely. The units of transducing systems are water-insoluble particles which have an exact composition. These contain special proteins that are linked to one another and to lipid in a definite pattern and stoichiometry. The biochemical objectives are threefold: (1) to isolate the particle; (2) to characterize its components and the pattern and sequence of their arrangement; and (3) to determine the mechanism of the transduction process. The problems are largely chemical in nature and it is of interest that the direct chemical approach to the mitochondrion was in the end the most decisive approach yet attempted (Green, 1958).

III. Experimental Approaches to the Study of the Mitochondrion

There are two rather different ways of studying the mitochondrion. The classic approach is that of studying the intact system (Chance and Williams, 1956). The chromophoric components are monitored spectrophotometrically when the mitochondrion is exposed to a variety of reagents, conditions, or inhibitors. The enzymatic activity is measured by a variety of methods and under a variety of conditions. Gross behavioral properties are observed such as swelling, shrinking, and deformation in shape. From studies of this kind, conclusions are drawn about the properties of the system. This may be described as the slot machine approach. Something is added to the system and then the investigator studies what happens. In our view the shortcoming of this strategy is that there is no way of getting behind the scenes, and thus relatively little can be learned

directly about the system itself. The direct chemical approach involves more drastic methods of experimentation (Green, 1958). The mitochondrion is fragmented into less complex parts—each of which is isolated, characterized, and documented. The over-all enzymatic activity is resolved into its component activities by fragmenting the mitochondrion into subunits that carry out only segments of the total activity. The organizational features of the system are analyzed in chemical terms, and the relation of organization and composition to function is studied.

The chemical approach entails large-scale isolation of the mitochondrion since adequate amounts of material are needed for chemical analysis (Crane *et al.*, 1957). The mitochondrion is in effect treated as a chemical entity that can be stepwise disassembled into its component parts and the parts eventually reassembled into the original unit.

The mitochondrion can be examined from the standpoint of chemical composition, structural and physical characteristics, enzymatic function and kinetics, and mechanism of various reactions. All these aspects are pertinent to the problem of how the mitochondrion ticks. It is necessary to be concerned with all these aspects if real progress in understanding is to be made. Nothing short of a total examination can fill the bill. Admittedly this will inevitably lead at first to superficiality but by successive approximations the whole picture can in time be seen even though this may be marred by blank spots here and there. A total picture notwithstanding its inevitable imperfections has great tactical advantages over a detailed picture of a small segment of the problem with the rest in total darkness.

The stepwise approach to the study of mitochondrial structure and function has involved finding the answers to the following basic questions: (1) What is the smallest unit of mitochondrial function? (2) How many parts make up this unit? (3) How are these parts held together? (4) How do the parts function individually and in concert? The search for the answers has led to phenomena at the borderlines of physics, chemistry, and biology. Practically every aspect of biochemistry—from intermediary metabolism to protein and lipid chemistry, from macromolecules to metals and cofactors has had to be dealt with. Oddly enough there appears to be more simplicity at the level where all the parts are assembled. It is precisely when the parts are disassembled that the real complexity

becomes apparent. The multiple chemical and physical principles that underlie mitochondrial structure and function can at present be recognized only very dimly.

IV. Unit of Mitochondrial Action

The mitochondrion is primarily a device for combusting pyruvic acid to CO_2 and water (by way of the citric acid cycle) and conserving the energy thus liberated in the form of a bond established by the interaction of adenosine diphosphate (ADP) with inorganic phosphate (synthesis of ATP). The oxidation of pyruvic acid to CO_2 and water can go on in complete absence of oxygen; for formal purposes the cycle can be looked upon as a mechanism for supplying "high energy" electrons for the combustion chamber which is, in reality, an electron transport chain, ending with molecular oxygen as the final acceptor (Keilin, 1933). During the passage of a pair of electrons through this chain to oxygen three molecules of inorganic phosphate are esterified to ADP and three molecules of ATP are synthesized (Belitzer and Tsibikova, 1939; Lardy, 1956; Kalekar, 1939). Thus, there are three ingredients of mitochondrial function: (1) the oxidative release of electrons from citric cycle substrates or other substrates; (2) the transfer of these released electrons through the combustion chain; and (3) the coupling of electron flow to synthesis of ATP. The intact mitochondrion carries out all these functions. When mitochondria are broken up by sonic irradiation the resulting particles no longer have the capacity to carry out the citric cycle but may still retain the other two functions (Linnane and Ziegler, 1958). The sources of electrons now are either succinate or dihydriodiphosphopyridine nucleotide (DPNH) rather than the citric cycle intermediates. If certain precautions are not taken during sonic treatment the coupling function as well as the capacity for carrying out the complete citric cycle is lost. Electrons from succinate or DPNH move through the chain but this movement is no longer coupled to synthesis of ATP. It is thus possible to simplify the problem of function enormously. The mitochondrion can be comminuted by sonic irradiation to a much smaller particle that retains only an intact electron transport chain and this particle is stripped of the complement of enzymes and cofactors required for the complete citric cycle oxidations and for the coupling function. In this way, electron transport can be studied independent of coupling; or coupling independent of the citric cycle oxidations;

or the citric cycle oxidations independent of the other two activities (cf. Table I). The sonically derived particle (electron transport particle) is the hard structural core of mitochondrial function—stripped of peripheral and ancillary elements, and this hard core is an organized mosaic of lipid and protein components.

The one danger that is courted is that the particle derived by sonic treatment of mitochondria, which no longer can couple electron flow to synthesis of ATP, may have an altered electron transport chain. There is, however, much evidence to suggest that the basic features of the electron transport process are not altered when the coupling function is lost since it is now possible to restore the coupling capacity of this particle by adding back specific proteins (Linnane and Titchener, 1960; Pullman *et al.*, 1958; Webster, 1962).

TABLE I
THE PROPERTIES OF MITOCHONDRIA, ETP_H AND ETP

	Citric cycle oxidations	Electron transport	Oxidative phosphorylation (coupling function)
Mitochondria	+	+	+
ETP_H	0	+	+
ETP	0	+	0

NOTE: The plus sign denotes activity and the symbol 0 denotes loss of activity. ETP represents the electron transport particle; the subscript "H" signifies the form of this particle that retains the coupling function.

The particle that is obtained by sonic irradiation of mitochondria is minute in size compared with the parent mitochondrion (Ziegler *et al.*, 1958). The ratio of sizes is about one to fifty thousand (cf. Fig. 1). In other words, the individual mitochondrion consists of several thousand units each of which contains a complete electron transport chain and the capacity for coupling (Green and Oda, 1961). When the mitochondrion is sonicated and the cluster of enzymes implicated in citric cycle oxidations is detached from this unit, the latter remains particulate while the citric cycle enzymes become water soluble.

Particles prepared by sonic irradiation of mitochondria which retain their magnesium content also retain their coupling capacity

(Linnane and Titchener, 1960). Under conditions leading to the release of magnesium from the particle the coupling function is lost. The particle obtained by sonic irradiation is known as the electron transport particle, and the form of this particle, still endowed with coupling capacity, is distinguished from the noncoupling form by the subscript "H" (cf. Table I).

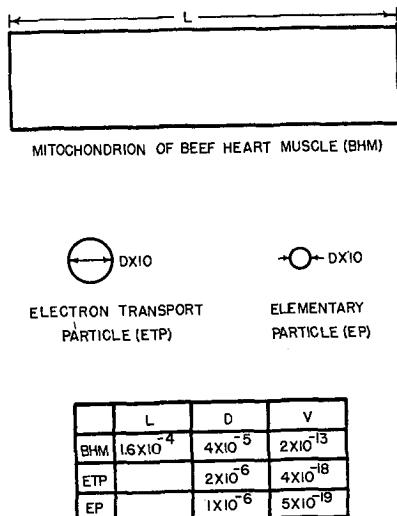


FIG. 1. Dimensions of a beef heart mitochondrion, an electron transport particle, and an elementary particle. The data for the mitochondrion and the electron transport particle were compiled by Green and Oda (1961). The calculations of the volume of the electron transport particle and of the elementary particle were made on the assumption that the particles were perfect spheres of diameter, respectively, 2×10^{-6} and 1×10^{-6} cm.

Apart from soluble protein that is released (about 20% of the total in the case of heart mitochondria) the conversion of mitochondria to the electron transport particle is essentially quantitative. This means that, at least in the case of heart mitochondria, the electron transport particle accounts for about 80% of the dry weight of the mitochondrion and is, in effect, the main mitochondrial constituent.

The electron transport particle represented the first milestone in the emancipation of the investigator from the shackles of mitochondrial complexity; it provided a stable system in which one basic

mitochondrial function, viz., electron transport, could be studied with comparative ease.

V. Electron Transport Chain

The electron transport chain may be defined as the structured mosaic of proteins and lipid, which contains an array of multiple oxidation-reduction components through which electrons from either succinate or DPNH are transferred to molecular oxygen (Green and Fleischer, 1960b). The sequence and the stoichiometry of these components are invariant features of the chain. In other words, the electron transport chain must be looked upon as a reproducible, precise entity. It is only complexity that distinguishes such an integrated unit from a classic protein. The chain is made up of at least six proteins having groups capable of oxidation-reduction and one protein lacking such groups (to which reference will be made in a later section). Two flavoprotein enzymes and four cytochromes comprise the total of proteins with oxidation-reduction groups (cf. Fig. 2). The molecular proportions of these proteins are as follows. For each molecule of succinic dehydrogenase flavoprotein (F_s), there are, respectively, one molecule of DPNH dehydrogenase flavoprotein (F_D), one molecule of cytochrome c_1 , three molecules of cytochrome b , and six molecules of cytochrome a (Oda, Blair, and Green, unpublished studies; Rieske, Tisdale, and Green, unpublished studies). There is at least one molecule of cytochrome c per molecule of F_s but for reasons that will be apparent later it is difficult to assign the exact number of molecules of cytochrome c in a given chain.

Besides the flavin and heme groups, associated with the six proteins (one flavin or heme group per protein molecule), there are two additional protein-bound oxidation-reduction groups present. Iron in a form other than iron porphyrin is linked to F_s and F_D and to both cytochromes b and c_1 . The links are between iron and some group in each of these proteins. Copper is linked to the protein of cytochrome a . For each molecule of F_s in a single chain there are some eleven atoms of non-heme iron and four atoms of copper.

There is yet one additional oxidation-reduction component that should be mentioned to complete the roster. This is coenzyme Q (ubiquinone)—a fully substituted benzoquinone with a long side chain of ten isoprenoid units on carbon atom five in the ring. For each molecule of F_s in a complete chain there are some fifteen molecules of coenzyme Q.

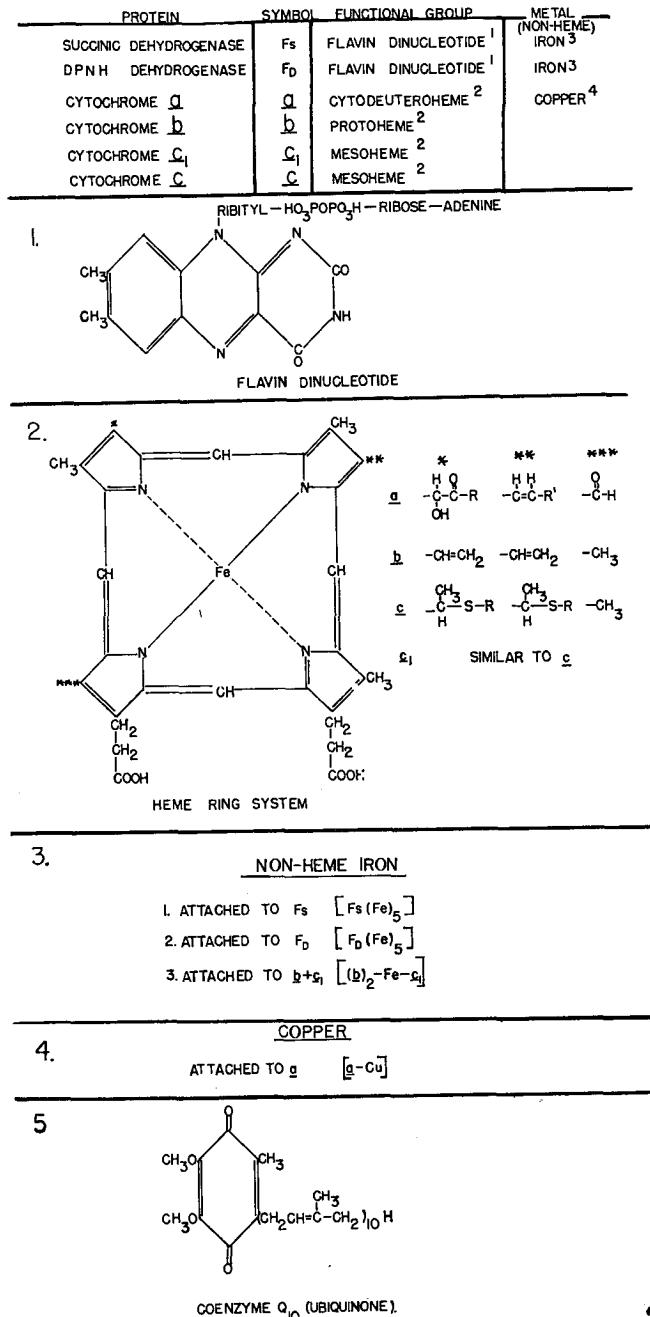


FIG. 2. The six oxidation-reduction proteins of the electron transport chain and coenzyme Q.

One must distinguish between components like coenzyme Q and cytochrome *c*, on the one hand, both of which are readily extracted from, and reinserted into, the chain and, on the other hand, the rest of the oxidation-reduction components that are not readily separable from the chain. The former will be referred to as the mobile components of the chain, the latter as the fixed components. During the preparation of submitochondrial particles extraction of some of the cytochrome *c* takes place and this component must be added back to restore full enzymatic activity.

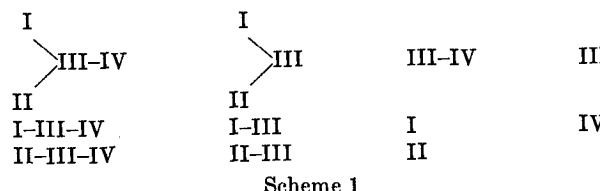
VI. Fragmentation and Reconstruction of the Electron Transport Chain

The electron transport chain has the character of a jigsaw puzzle that can be resolved into its component functional parts in one way only, i.e., by separation along the natural cleavage lines. Lipid is one of the main ingredients in the cement that holds together the parts, and only those reagents that can cleave protein-lipid and/or lipid-lipid bonds are effective in resolution of the chain. In point of fact, there are precious few reagents that have proved useful for this purpose. Cholate and deoxycholate among the bile acids, *tert*-amyl alcohol and butanol among the lower alcohols, and petroleum ether and cyclohexane among the hydrocarbons represent the full gamut of useful fragmenting reagents. The action of bile acids requires the presence of salts while the action of hydrocarbons requires the presence of bile acids.

It has been demonstrated in our laboratory that, under conditions leading to fragmentation of the chain, the lipids of the particles can rapidly exchange with externally added lipid (Fleischer and Brierley, 1961). It is probably the weakening of the protein-lipid or lipid-lipid bonds that controls the fragmentation of the electron transport chain.

Reagents such as detergents, which modify protein-lipid or lipid-lipid bonds, may not necessarily be effective in fragmentation. Quite apart from inhibitory effects there are probably critical steric and electrostatic requirements that have to be satisfied and these factors set a severe limit on the possible number of useful reagents.

The electron transport chain can be resolved into four complexes (I, II, III, and IV) and various intermediary stages in the fragmentation process can be isolated as shown in Scheme 1 (Green and Hatefi, 1961):



In the above scheme the line between complexes indicates the links holding the complexes together.

Each of the four complexes has now been isolated in a state of relatively high purity. All but one is of molecular weight 250,000 or less. Each complex contains unique components as summarized below:

I: trypsin-extractable flavin and non-heme iron (Ziegler and Doeg, 1959).

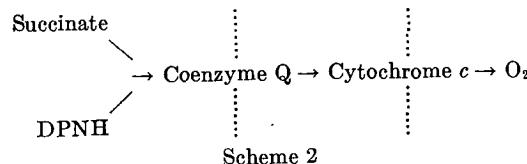
II: acid-extractable flavin and non-heme iron (Hatefi *et al.*, 1961a).

III: cytochromes *b* and *c₁*; and non-heme iron (Hatefi *et al.*, 1961b).

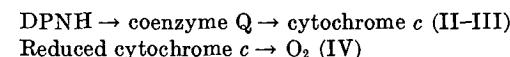
IV: cytochrome *a* and copper (Griffiths and Wharton, 1961a; Okunuki *et al.*, 1958).

The non-heme iron groups of complexes I, II, and III have quite different properties such as differential susceptibility to inhibitors and distinctive electron spin resonance spectra (Beinert and Lee, 1961; Beinert and Sands, 1959; Doeg and Ziegler, 1962b; and Ziegler 1961). These three species of non-heme iron have to be looked upon as highly specific in both their chemistry and function. They are lumped together under one heading because as yet we cannot describe in chemical terms the differences among them.

The guidelines for fragmentation of the electron transport chain were functional criteria. If the chain is represented by Scheme 2,



then at any cleavage point (denoted by dotted line) it should be possible to divide the chain into two segments—one that reduces the component in question and one that oxidizes the component. Thus, DPNH oxidase (which can be represented in terms of the complexes referred to above as II, III, and IV linked together or II-III-IV) can be fragmented into DPNH-cytochrome *c* reductase (II-III) (Hatefi *et al.*, 1961c) and reduced cytochrome *c* oxidase (IV):



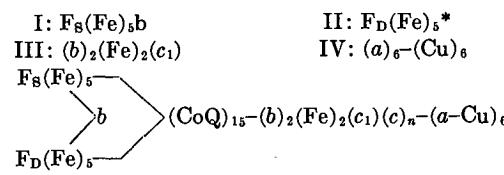
In the same way succinic oxidase can be fragmented into succinic-cytochrome *c* reductase (I-III) (Green and Burkhard, 1961) and reduced cytochrome *c* oxidase (IV). By the same token succinic-DPNH oxidase [I-(II)-III-IV] can be fragmented into succinic-DPNH-cytochrome *c* reductase [I-(II)-III] and cytochrome oxidase (IV).

The same tactics can be applied to coenzyme Q. The chain can be fragmented into a segment that reduces coenzyme Q (I or II) (Hatefi *et al.*, 1961a; Ziegler and Doeg, 1959) and a segment that oxidizes coenzyme Q (III or III-IV) (Hatefi *et al.*, 1961a; Hatefi, 1959). Thus there are two complexes that reduce coenzyme Q [succinic-Q reductase (I) and DPNH-Q reductase (II)] and one complex that oxidizes the reduced form of coenzyme Q (III). If III is tested alone cytochrome *c* is the only natural acceptor that can be used; if III-IV is tested then either O₂ or cytochrome *c* can be used.

The two natural cleavage points of the chain correspond to the lines that separate one complex from the next in sequence. Coenzyme Q lies between complexes I and II on the reducing side and complex III on the oxidizing side while cytochrome *c* lies between complex III on the reducing side and complex IV on the oxidizing side. Both coenzyme Q and cytochrome *c* serve as electron bridges or shuttles between complexes in the chain. It is probably not a coincidence that these are the only two mobile oxidation-reduction components in the electron transport chain.

By this type of stepwise degradation of the chain it has been possible to isolate each of the four complexes and on the basis of both the composition and position in the chain of each complex

the sequence of components in the chain could be assigned in a straightforward fashion (Scheme 3).



Scheme 3

The amount of cytochrome *c* is difficult to determine because of the continuous loss of this cytochrome during isolation of the electron transport particle. There appears to be one molecule of cytochrome *c*₁ per mitochondrial chain.

The electron transport chain as isolated is a composite of four component complexes into which it can be resolved by various procedures. In general, a combination of bile acids (cholate and deoxycholate) and salts (ammonium sulfate, ammonium acetate) can achieve separation of these four complexes. Isolated complexes, freed from bile acids and inorganic salts, can readily recombine to form a unit which now shows integrated electron transport activity (Fowler and Hatefi, 1961; Hatefi *et al.*, 1961b).

When the four isolated complexes are mixed together under appropriate conditions they reassemble to form a close facsimile of the original electron transport system and the product formed is capable of integrated electron transport activity. There are, of course, factors that can complicate or even obscure the specificity of the recombination reactions but precise realignment of the component complexes into an integrated whole must have taken place. Like the pieces of a jigsaw puzzle the four complexes stand in a precise relation one to the other and the recombination reactions must re-establish this complementarity. The time required for recombination of the complexes can be a matter of seconds at 0°. But at least one of the reacting components must be present in relatively high concentration for efficient combination. With dilution of the components prior to mixing both the rate and extent of recombination falls off logarithmically. Once the complexes have interacted

* Complex II as isolated contains Fe:flavin ratios of at least 15 but there is reason to believe that only five iron atoms of the total were originally associated with the complex. Some translocation of iron during isolation is indicated.

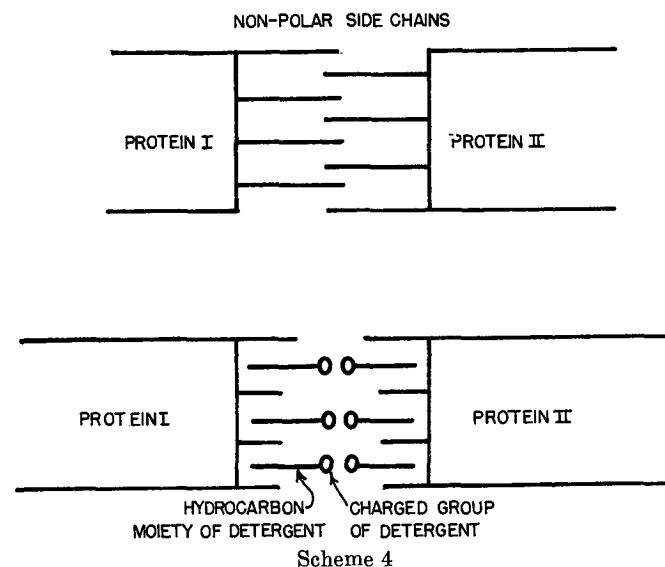
dilution has no effect (Fowler and Hatefi, 1961; Hatefi *et al.*, 1961b; Fowler and Richardson, 1962).

VII. Structural Protein

The oxidation-reduction proteins of the electron transport chain and of the associated complexes comprise at most 50% of the total protein of mitochondria and probably a minimum of 40%. The balance is made up of a colorless protein lacking oxidation-reduction groups (Green *et al.*, 1961a). This protein has been designated structural protein. At neutral pH it exists only in a polymeric form that is utterly water insoluble. The polymeric form can be solubilized by dilute alkali, by 60% acetic acid, and by cationic detergents such as dodecylsulfate. Depolymerization and solubilization are two sides of the same coin since it is only the monomer (M.Wt. of about 25,000) that is water soluble.

The polymer-monomer transition also applies to each of the three fixed cytochromes (*a*, *b*, and *c*₁) (Ambe and Venkataraman, 1959; Goldberger *et al.*, 1961; Criddle and Bock, 1959). At neutral pH these exist only as polymers that can be either water-soluble (cytochrome *c*₁) or water-insoluble (cytochrome *b* and *a*). The polymeric forms can be depolymerized to the monomeric forms by dilute alkali, detergents, or strong acetic acid. The monomeric forms are relatively small molecules (M.Wt. of 30–40 × 10³).

The monomerization of the polymeric species by detergents at low concentration is a diagnostic of the hydrophobic bond. The structural protein and the three cytochromes are presumed to contain regions in which the side chains of the amino acids are all non-polar. These localized lipid-like or hydrophobic areas underlie the tendency of the protein to polymerize. The coalescing of hydrophobic regions in two interacting protein molecules is the essence of the polymerization process. As a rule of thumb one may say that proteins containing multiple regions with a pronounced hydrophobic character will spontaneously polymerize within the physiological range of pH. These hydrophobic regions must lie close to the periphery of the molecule to make protein-protein interaction possible. Charge repulsion induced by alkali and concentrated acid leads to depolymerization by overcoming the hydrophobic forces holding the polymer together. Detergents probably act as monomerizing agents in a manner such as is represented by Scheme 4.



The detergent forms a hydrophobic complex with the proteins by a displacement reaction. Protein-detergent complex replaces the protein-protein complex; in this manner the polymer is monomerized (Green *et al.*, 1961b). It is assumed that each protein molecule in the polymer (except the terminal molecules) is linked to at least two other molecules by hydrophobic bonds.

If a molecule of structural protein can readily form a hydrophobic bond with a fellow molecule it should also show the same tendency to combine with the monomeric form of any of the cytochromes. Indeed, 1:1 molecular complexes of structural protein with each of the three cytochromes have been prepared and characterized (Criddle *et al.*, 1961). These are all stable complexes and are readily formed. These complexes can be split into their component protein molecules by the same reagents that depolymerize the polymeric form of structural protein and of the cytochromes.

VIII. Lipid

The mitochondrion contains about 30% by weight (dry) of lipid (Fleischer *et al.*, 1961). Phospholipid accounts for more than 90% of this complement of lipid. In beef heart mitochondria four phospholipids [phosphatidylcholine, phosphatidylethanolamine, phos-

phatidylinositol, and cardiolipin (polyglycerol phosphatide)] account for essentially all the phospholipid (ca. 98%).

When mitochondria are fragmented randomly, i.e., without regard to the preservation of integrated activities, the particles or fractions so derived show variable lipid composition—from 0–96% by weight of lipid. However, when the fragmentation is so conducted as to preserve integrated activities, the lipid content of the derived particles remains at about the same level as in the original mitochondrion (about 30%) (Fleischer *et al.*, 1961).

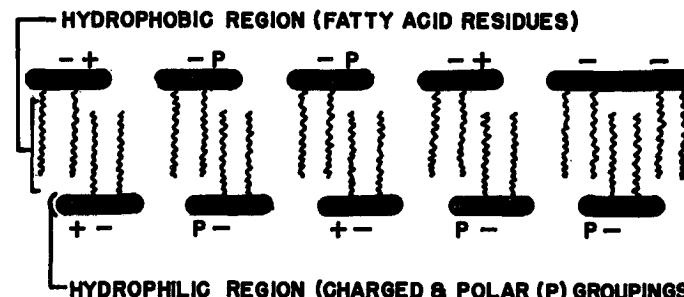


FIG. 3. Diagrammatic representation of a segment of a "solubilized" phospholipid micelle of mixed phospholipid composition. In such micelles there could be many hundreds of paired arrays and these arrays could be coiled to form spirals or other complex lamellar arrangements. The important idea is that of a molecularly dispersed array in which hydrophobic and hydrophilic areas alternate.

Lipids are normally characterized by their solubility in organic solvents and by their insolubility in water. When a triglyceride is shaken up with water it separates out as a bulk phase and no measurable amount is demonstrable in the aqueous phase. Phospholipids, however, differ fundamentally from the neutral lipids in that they have a polar end; both classes of lipids have non-polar side chains. In principle, the bimodal molecules can be oriented in water to form micelles. Optically clear or at worst faintly opalescent preparations in water of the individual mitochondrial phospholipids and of mixtures of these phospholipids have been made in our laboratory (Fleischer and Klouwen, 1961). These clear "solutions" of the phospholipids contain micelles in which the polar groups are oriented at the periphery and the non-polar, hydrocarbon chains are directed toward the interior of the micelle (Hartley, 1955; Rouser, 1958) (cf. Fig. 3). There is a variety of factors that deter-

mine the stability of a phospholipid micelle—the nature of the charges on the charged groups, the degree of unsaturation of the fatty acid residues, the presence of impurities in the micelle, etc. Micelles with mixed phospholipids have quite different properties from those of micelles formed by pure phospholipids. The degree of unsaturation of the fatty acid residues in the phospholipid has a profound influence on the shape, dimensions, and properties of the micelle and particularly on the capacity of the micelle to incorporate non-polar, water-insoluble molecules. When phospholipids of natural origin (the fatty acids of which have a high degree of unsaturation) are hydrogenated the resulting saturated phospholipids do not give rise to the near water-clear, molecularly dispersed micelles (Fleischer and Klouwen, 1961; Rouser, 1958). The important point at issue is that the properties of micelles are profoundly influenced by a variety of factors and there is a wide assortment of micelle types. Nature has acquired great proficiency in use of the physical and chemical properties of lipids oriented in micelles.

Germane to our discussion of micelles is the fact that many compounds, normally water insoluble, can be solubilized within the aqueous micelles of the phospholipids. Thus, coenzyme Q and cholesterol, both of which are utterly water insoluble, can be readily solubilized in aqueous medium by phospholipid micelles—the latter to the extent of 20% by weight of the phospholipid (Fleischer and Brierley, 1961a; Basford and Green, 1959). There can be little doubt that the transport of water-insoluble materials in the blood is accomplished in part by the stratagem of a phospholipid micelle. (The phospholipid micelle is further modified by association with a specific protein, the lipid-protein system being referred to as a lipoprotein.)

In view of these remarkable molecular properties of phospholipids it cannot be a mere accident of evolution that the lipid of the mitochondrion is predominantly phospholipid and that its fatty acids enjoy a high degree of unsaturation. The possibility that mitochondrial phospholipid may exist as a bulk phase cannot be seriously entertained in view of the concatenation of factors ideal for micelle formation, and in view of the clear-cut electron microscopic evidence that lipid in bulk does not exist in mitochondria or submitochondrial particles. In order to understand the function of phospholipid we must orient our thinking in terms of the capacity of phospholipids to form micelles, and in terms of the

potentiality of these micelles to serve as bridges between aqueous and non-aqueous phases. The high concentration of phospholipids in cell membranes argues for phospholipid structures serving as the determinants of cell permeability.

Until very recently coenzyme Q was the only lipid for which an absolute requirement could be demonstrated in mitochondria (Lester and Fleischer, 1959). This demonstration encouraged us to inquire whether there might be additional lipid requirements for the activities of the electron transport chain. There is a vast literature dealing with activating effects of fatty acids, alcohols, and aldehydes and of detergents. Most, if not all, of these effects are irrelevant to the problem of lipid involvement. We made the decision that a *sine qua non* for proof of lipid involvement was: (1) to remove the lipid from particles; (2) to show a correlation between removal and loss of activity; and (3) to show a correlation between restoration of activity and uptake of lipid. Consequently methods were devised to remove the bulk of the lipid from mitochondria (Lester and Fleischer, 1961). Indeed it was found that extraction does lead to loss of activity, and that rebinding of lipid is a necessary prerequisite for restoration of activity (Fleischer, Brierley *et al.*, 1962; Fleischer and Klouwen, 1961). The degree of restoration and the degree of rebinding run parallel courses. It is not lipids, in general, that are implicated in the loss of activity and in its restoration but phospholipid specifically (Fleischer, Brierley *et al.*, 1962; Fleischer and Klouwen, 1961). Some lipid-like molecules that are not of mitochondrial origin can mimic, in small measure, the behavior of phospholipids but this mimicry does not gainsay the fact that the phospholipids are essential for enzymatic activity. These molecules share some of the special physical properties that underlie the role of phospholipid in the electron transport system just as certain synthetics can replace the natural steroid hormones.

One of the mitochondrial pyridinoprotein enzymes, the β -hydroxybutyric dehydrogenase, shows an absolute requirement for a specific phospholipid, viz., lecithin. The available evidence suggests that lecithin combines with the protein and that the enzyme protein achieves the proper conformation for activity only when combined with lecithin (Jurtshuk *et al.*, 1961; Sekuzu *et al.*, 1961).

When lipid is extracted from mitochondria there is no evidence that the mitochondrion falls apart. On the contrary it is more difficult to fragment the mitochondrion after lipid extraction than

before. The same reagents that are effective with normal mitochondria are ineffective with lipid-extracted mitochondria. The electron micrograph, shown in Fig. 4, of a mitochondrion divested of 75–80% of its phospholipid complement shows that the over-all integrity of structure of the extracted mitochondrion persists. Indeed it is not grossly different from that of the normal.

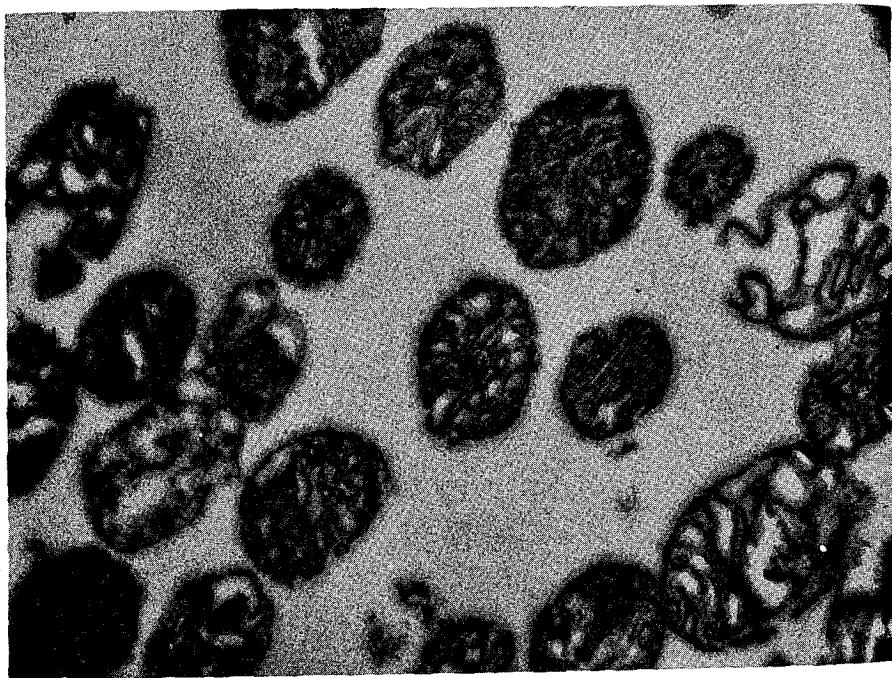


FIG. 4. An electron micrograph of aqueous acetone-treated mitochondria (Lester and Fleischer, 1961). The mitochondria have been divested of 75–80% of their lipid content (Fleischer *et al.*, 1962) by acetone extraction. The electron micrograph was prepared and supplied by Dr. David Slatterback, Department of Anatomy, University of Wisconsin, Madison, Wisconsin.

We have recently observed that mitochondrial-bound phospholipids exchange with externally added solubilized phospholipid (labeled with P^{32}); exchange, however, occurs only in the presence of those reagents that have been found effective in the fragmentation of the mitochondrion (Fleischer and Brierley, 1961b). Furthermore, with removal of the fragmenting reagent, the phospholipid molecules

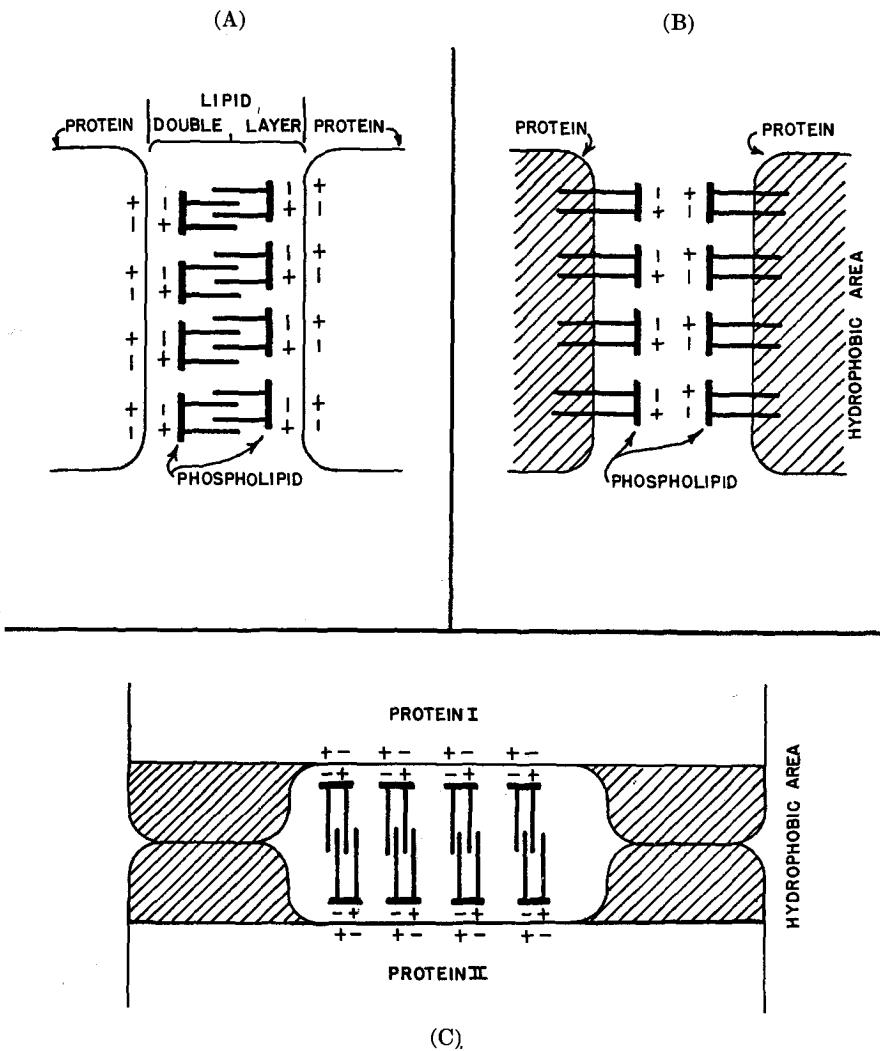


FIG. 5. Diagrammatic representations of the modes of binding of phospholipid to protein. (A) Electrostatic binding of protein to lipid. (B) Binding of lipid to hydrophobic areas of a protein. (C) Electrostatic binding of lipid to protein in an enclosed cage formed by the interaction of equally spaced hydrophobic areas in two proteins.

once again become firmly bound and no longer exchange with externally added phospholipid.

At present it is not possible to evaluate the relative importance of the bonds linking protein to lipid—electrostatic, hydrophobic, and the combination of the two (cf. Fig. 5 for a diagrammatic representation of these possibilities). The fact that the net charge on the phospholipid appears to determine the tightness of binding to protein (phospholipids carrying the negative charge only being more reactive than those with both positive and negative charges) would argue for electrostatic bonds. If true, then the action of bile acids would be difficult to explain in terms of electrostatic bonds.

Model systems for studying protein-phospholipid interaction are fortunately now available and these may in time provide the yardsticks by which a decision as to bond types can be reached. Cytochrome c forms a hydrocarbon-soluble micelle with phospholipids (Widmer and Crane, 1958). In this model system the phospholipid must necessarily be bound to the protein by electrostatic bonds since the micelle is soluble in hydrocarbons. Furthermore, certain mixtures of phospholipids lead to micelle formation whereas the individual phospholipids may not give rise to micelles under the same experimental conditions (Green and Fleischer, unpublished observations). The distribution of charges on the protein probably explains why the mixture is effective whereas the individual phospholipids are not.

These model systems can provide the biochemist with the means to eliminate guesswork about the mode of protein-lipid interactions. Theory has fallen far behind experiment and it is high time that the gap be narrowed.

IX. Concept of the Elementary Particle

From the relative volumes of the mitochondrion and the electron transport particle a computation can be made of the number of such particles per mitochondrion. The mitochondrion of beef heart muscle has the shape of a (right circular) cylinder with rounded ends at both extremities (cf. Fig. 1) (Green and Oda, 1961). The average dimensions are a length of 1.6×10^{-4} cm and a radius of 2×10^{-5} cm. The volume of such a cylinder is 1.8×10^{-13} cm³. When the mitochondrion is disrupted by sonic oscillations the monomeric form of the resulting particles (known as

the electron transfer particles) has a molecular weight of 4.6×10^6 , based on composition data. On the assumption of a spherical particle, the volume corresponding to this molecular weight is computed to be 6×10^{-18} cm³. Each mitochondrion could contain $(1.8 \times 10^{-13})/(6 \times 10^{-18})$ electron transport particles, i.e., 30,000 if there were no spaces. Since the structured elements in the mitochondrion account for no more than 50% of the total volume the more probable estimate is about 15,000 electron transport particles per mitochondrion. The protein would contribute about 3×10^6 to the molecular weight of such a particle; when allowance is made for lipid the molecular weight is calculated to be about 4×10^6 (Green and Oda, 1961). Each such particle contains one molecule of succinic- and DPNH dehydrogenase and the appropriate number of molecules of the other members of the electron transport chain as discussed above. By all chemical criteria the electron transport particle appeared to be the smallest common denominator of mitochondrial function.

Two developments, however, have led to a re-evaluation of this interpretation: (1) the discovery in our laboratory (Green *et al.*, 1961a) that structural protein, a species having no oxidation-reduction group, accounts for 50 to 60% of the mitochondrial protein; and (2) the discovery by H. Fernández-Morán (unpublished experiments), using an electron-microscopic technique, that mitochondria contain a repeating unit, with diameter of about 125 Å and a probable molecular weight of 1.2×10^6 . This demonstration of a unit smaller than the electron transport particle at once suggested that the latter is a composite of a yet smaller particle and structural protein and that it should be possible to isolate the smaller particle free from its coat of structural protein. A particle of the correct size has indeed been isolated in our laboratory (Oda, Blair, and Green, unpublished studies) and this particle contains all the components of the electron transport chain in concentrations some three times those found in the mitochondrion (cf. Table II). The preparation of this particle which has been called the elementary particle by Fernández-Morán involves the separation of the mitochondrion into three fractions: (1) the colorless structural protein; (2) the intensely colored elementary particle; and (3) a supernatant medium containing soluble proteins and substances released during the fractionation procedure (Green *et al.*, unpublished experiments). The structural protein fraction accounts for about 60% of the total protein, the elementary particle fraction for about 20% of the protein, and the

supernatant fraction for the rest. The entire succinic oxidase and DPNH oxidase activity of the original mitochondrion is localized in the elementary particle. We have tentatively equated the elementary particles with the particles of 125 Å seen in the electron micrographs.

TABLE II

THE CONCENTRATION OF COMPONENTS AND ACTIVITY IN THE ELEMENTARY PARTICLE AS COMPARED TO THE MITOCHONDRION^a

	Mitochondria (m μ moles/mg protein)	Elementary Particle (m μ moles/mg protein)
F _S	0.23	0.81
F _D	0.26	0.81
Cytochrome <i>a</i>	1.53	5.00
Cytochrome <i>b</i>	0.72	2.62
Cytochrome <i>c</i> + <i>c</i> ₁	0.43	1.12
Copper	1.3	3.8
Non-heme iron	3.3	—
Succinoxidase activity (μ moles succinate oxidized/min/mg protein)	1.0	3.0

^a Data of Oda, Blair, and Green (unpublished experiments).

The elementary particle catalyzes the oxidation of succinate and DPNH by molecular oxygen through the electron transport chain and this process is inhibited by the same specific reagents that suppress oxidation in the intact mitochondrion. Two processes no longer are demonstrable in preparations of the elementary particle: (1) the citric acid cycle; and (2) oxidative phosphorylation. It can be shown that the enzymes of the citric acid cycle are in fact released during the preparation of the elementary particle; these enzymes occur in complexes that are normally closely associated with the electron transport chain. These complexes are readily released by treatment of the mitochondria with sonic irradiation or by exposure of the mitochondria to bile acids (as in the preparation of the elementary particle). The loss of oxidation phosphorylation is, in large measure, also a reflection of the release of certain enzymes required for coupling electron flow to phosphate esterification.

In the mitochondrion the elementary particles are arranged in

paired tiers in both the outer membrane and the cristae (Fernández-Morán, unpublished experiments). The ease with which the enzymes of the citric acid cycle are lost during the isolation of the elementary particle suggests that the various complexes containing the enzymes of the citric acid cycle may nest between the paired elementary particles in a manner such as is shown in Fig. 6. According to this interpretation the double-membraned structure provides a type of

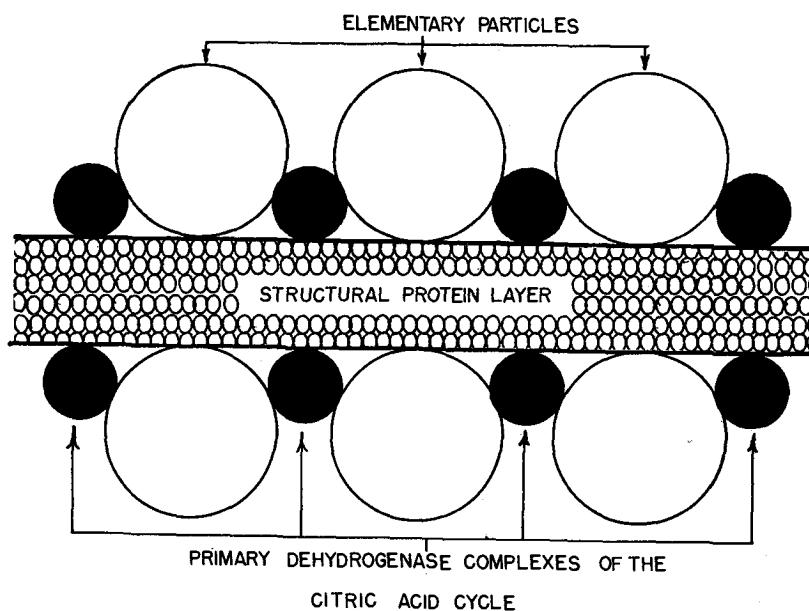


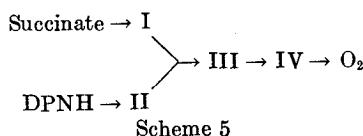
FIG. 6. Diagrammatic representation of the nesting of primary dehydrogenase complexes between arrays of elementary particles.

molecular cage in which various complexes can be fixed in position. The forces that bind these complexes to the elementary particle and to the structural protein matrix may be relatively weak but the stability of the over-all system is considerable as long as the paired arrangement obtains. Once this arrangement is destroyed then the forces binding the complexes to the elementary particle are too weak to prevent separation one from the other. The importance of positioning as a factor in holding various components together in the intact mitochondrion should not be underestimated. It is of interest to point out that the complexes containing the enzymes of the citric

acid cycle end up in the fraction containing structural protein when the mitochondrion is subjected to the procedure for the isolation of the elementary particle.

X. Mechanism of Electron Transport

The four complexes, when appropriately linked together, form a forked chain in which electrons from succinate and DPNH are ultimately transferred to molecular oxygen (Scheme 5).



Scheme 5

Between successive complexes (I and III, II and III, III and IV) there are special mobile coenzymes that can shuttle electrons. For example, coenzyme Q is the link between I and III, and II and III; and cytochrome *c* is the link between III and IV. These mobile coenzymes serve as electron acceptor for one of the paired complexes and as electron donor for the other. Each complex has at least two bound oxidation-reduction groups and, in addition, two dissociable oxidation-reduction groups (cf. Table III). The meaning of this

TABLE III
BOUND AND DISSOCIABLE GROUPS OF THE FOUR COMPLEXES
OF THE ELECTRON TRANSPORT CHAIN

Complex	Bound Groups	Dissociable Groups	Sequence
I ^a	F _S ; Fe; <i>b</i>	Succinate; Q	Succinate → F _S → Fe → Q
II ^b	F _D ; Fe	DPNH; Q	DPNH → F _D → Fe → Q
III ^c	<i>b</i> ; <i>c</i> ₁ ; Fe	QH ₂ ; <i>c</i> ...	QH ₂ → <i>b</i> → Fe → <i>c</i> ₁ → <i>c</i>
IV ^{d,e}	<i>a</i> ; Cu	<i>c</i> ...; O ₂	<i>c</i> ... → <i>a</i> → Cu → O ₂

NOTE: The dissociable groups are at the beginning and end of the electron transport sequence catalyzed by the individual complexes. Arrows in the equation represent direction of flow of electrons.

^a Ziegler and Doeg (1959).

^b Hatefi *et al.* (1961a).

^c Hatefi *et al.* (1961a).

^d Griffiths and Wharton (1961b); Okunuki *et al.* (1958).

^e The reduced *c* cytochrome oxidase (IV) is a polymeric species containing six molecules of cytochrome *a* and six atoms of copper.

apparently complicated chain may be found in the following interpretation. The oxidation of both succinate and DPNH takes place in three stages:

Stage	Succinate	DPNH
1	Succinate → Q	DPNH → Q
2	QH ₂ → <i>c</i> ...	QH ₂ → <i>c</i> ...
3	<i>c</i> ... → O ₂	<i>c</i> ... → O ₂

Each stage is catalyzed by a separate complex. The energy liberated in the over-all oxidation of DPNH by oxygen is divided, so to speak, into three more or less equal parts—one for each of the complexes. In the case of the oxidation of succinate the first segment of the total energy drop is smaller than the other two segments and, as we shall see later, this is probably the basis of the fact that succinate oxidation involves a P/O ratio of 2 while that for DPNH is 3.

The complex is the smallest common denominator of the electron transport chain. Consider for example the complex that catalyzes the oxidation of reduced coenzyme Q by cytochrome *c* (Hatefi *et al.*, 1961a). This contains bound cytochromes *b* and *c*₁ as well as bound non-heme iron. Cytochromes *b* and *c*₁ can be isolated from this complex but when the complex is so fragmented all activity is lost and this activity cannot be reconstituted merely by mixing cytochromes *b* and *c*₁ in their original proportions (Rieske, Wharton, and Zaugg, 1962). Furthermore, non-heme iron gives a characteristic EPR signal when present in the intact complex. Once the link between cytochrome *b* and cytochrome *c*₁ is severed this signal disappears and cannot be re-established by mixing the two cytochromes together (Rieske, Wharton, and Zaugg, 1962).

Each of the four complexes of the elementary particle contains a heavy metal (Fe or copper) that shows a characteristic electron spin resonance signal; the metal is not part of a tetrapyrrole nucleus (Beinert and Lee, 1961; Beinert and Sands, 1959; Doeg and Ziegler, 1962a; Ziegler and Doeg, 1962; Griffiths and Wharton, 1961b). Iron is the metal in three of the complexes while copper is the metal of the last complex—the reduced cytochrome *c* oxidase. To distinguish the iron that is responsible for the electron spin resonance signal from heme iron we shall refer to this iron as Fe_{non-heme}. In

each of the complexes the latter undergoes cyclical reduction and oxidation and this cycle is inhibited by specific reagents. Thenoyl-trifluoroacetone specifically inhibits the oxidation-reduction of the Fe_{non-heme} of succinic-Q reductase (Doeg and Ziegler, 1962b); Amytal the oxidation-reduction of the Fe_{non-heme} of DPNH-Q reductase (Hatefi *et al.*, 1961a); antimycin A the oxidation-reduction of the Fe_{non-heme} of the QH₂-c reductase (Rieske, Wharton, and Zaugg, 1962); and cyanide the oxidation-reduction of the copper of reduced cytochrome c oxidase (Griffiths and Wharton, 1961b). The available evidence points to the conclusion that these metals in each of the four complexes are bona fide members of the respiratory chain. The specific reductant or oxidant for each of the complexes also serves as the reductant or oxidant of the heavy metal atom. For example, the non-heme iron of succinic-Q reductase is reducible by succinate (but not by DPNH) and oxidizable by coenzyme Q. The copper of reduced cytochrome oxidase is reducible by reduced cytochrome c and oxidizable by molecular oxygen.

At present there is no obvious explanation of the purpose which these metal atoms serve in the four complexes. Copper is the atom preferred by nature for reaction with molecular oxygen and this could well be the role of copper in reduced cytochrome c oxidase. But the role of non-heme iron in the other three complexes is still unknown. Perhaps there is an intimate connection between the binding of phosphate and the heavy metal; at present this hypothesis has no supporting evidence.

There is more total iron in the electron transport chain than reducible iron (about 2:1) (Doeg and Ziegler, in press). Again the exact role of this additional iron is unknown at present. The iron and copper are tightly bound to protein and can be released only under special conditions. There is no equilibration of externally added radioactive inorganic iron with the bound non-heme Fe of the electron transport chain and the same holds true for copper.

When an over-all process is broken up into discrete segments there must be components that are shared by contiguous complexes, i.e., components that are at the end of the chain in the first complex and the beginning of the chain in the second. All other components belong to one or another complex and are firmly bound to the protein. The components that move from complex to complex are the only mobile elements in the electron transport system and these raise special problems. They must be sufficiently mobile to

bridge the gap between complexes; yet they cannot be held so loosely that they would be lost too readily from the particle.

The mobile coenzymes are DPN, coenzyme Q, and cytochrome c. DPN may be detached from the complex with the greatest ease; coenzyme Q is detached more difficultly and cytochrome c is intermediate in this respect. Coenzyme Q is utterly insoluble in water and this property ensures that it will not be extracted from the particle in an aqueous medium. Extraction of coenzyme Q requires the use of organic solvents (acetone, ether, etc.) (Lester and Fleischer, 1959). Cytochrome c may be isolated in the form of a water-soluble protein of molecular weight 13,000 with strongly basic properties (isoelectric point ca. pH 11) (Keilin and Hartree, 1938; Theorell and Åkesson, 1941). When cytochrome c is mixed with solubilized mitochondrial phospholipids (in the micelle state) it forms a complex that is more soluble in hydrocarbons than in water. There is a considerable body of evidence suggesting that cytochrome c exists in the mitochondrion in a lipid-soluble form (Ambe and Crane, 1959). The transition from the water-soluble to the lipid-soluble form can be exceedingly rapid and is reversible. Salts profoundly affect the equilibrium point. It is this reversible transition from one form to the other that explains why, under some conditions, e.g., in the presence of phosphate ions, cytochrome c can be partially extracted into aqueous media. Under usual conditions extraction of cytochrome c cannot be achieved except with the aid of bile acids or by the addition of phospholipase to the particle. The extraction of DPN proceeds so readily that extreme care must be taken to prevent it. Fresh mitochondrial suspensions in 0.5 M sucrose retain pyridine nucleotide, providing certain precautions are taken (Lester and Hatefi, 1958). With aging, or with any change in shape and dimensions of the mitochondrion, pyridine nucleotide readily diffuses out of the particle. Sonication of the mitochondrion, or exposure to bile acids, leads to complete loss of pyridine nucleotide. In point of fact, this nucleotide is retained only when the mitochondrion is maintained in the anaerobic state under special conditions, e.g., in the presence of Versene or bovine serum albumin, or in the aerobic state when the mitochondrion is actively carrying out oxidation leading to the synthesis of ATP.

Evidence is not lacking that DPN, cytochrome c, and even coenzyme Q are all weakly bound to the particle and that anything that interferes with this binding leads to a loss of function. By

virtue of this binding to protein the bound coenzyme exhibits certain properties that are not characteristic of the free nucleotide. It appears likely that the bound coenzyme undergoes a different kind of oxidation-reduction than does the detached coenzyme. This may make the difference between coupling and noncoupling.

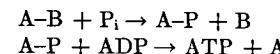
XI. Oxidative Phosphorylation

It is sobering to reflect that, despite the great insight which has been achieved into the molecular pattern of the mitochondrion, the coupling process which is the *raison d'être* of the biochemical machine still completely eludes us. The conventional machine either works or doesn't work. It may work badly but there is no change in principle when performance is less than perfect. The mitochondrion has the idiosyncrasy that electrons can flow through the chain in a way that leads to coupling or in a way that does not. When 2,4-dinitrophenol is added to a normal mitochondrion, electron flow is uncoupled from ATP synthesis (Cross *et al.*, 1949; Loomis and Lipmann, 1948). But the curious point is that this disengagement can proceed even without the addition of a special reagent. On standing at 0°, or after freezing, a mitochondrial suspension may lose its coupling capacity. So far as one can tell there is no qualitative change in the mitochondrion. The same components are present; the sequence of electron flow is the same; yet a major change in behavior has taken place. There are two possible causes of this disappearing act. Let us assume that, coincident with electron flow, a bond is established between components A and B and that this bond is then so exploited or manipulated that its energy can be utilized for synthesis of ATP. If the enzymatic machinery used to affect this conversion is damaged, then ATP synthesis cannot be achieved. Instead of the energy of the bond between A and B being conserved for synthetic purposes—a function which is the essence of the coupling process, the compound AB breaks down spontaneously to its component parts and the bond energy is dissipated as heat. In some instances this is probably the correct diagnosis of the fault in the machinery. For example, a particle can be prepared by sonication of mitochondria that can couple or not, depending upon whether a soluble component released during the sonication procedure is added back (Linnane and Titchener, 1960; Pullman *et al.*, 1958). If it is present, electron flow is coupled; if it is not, electron flow is uncoupled. In this particular instance we know that the soluble

component is concerned with the manipulation of the primary high energy bond that results in the formation of ATP. Thus, at least in this instance we can be sure that it is not the inability to form the bond that underlies uncoupling but rather the absence of one or more factors concerned in the transfer of the energy of this bond to the high energy bond of ATP.

It is important to recognize that the primary high energy bond must involve those electron transport components that undergo oxidation-reduction, whatever the mechanism of coupling. Thus, in the formulation of A-B as the primary high energy bond it is implicit that either A or B or both are oxidation-reduction components of the electron transport chain.

Perhaps this concept should be expanded a bit in order to facilitate the discussion of other possibilities. The conversion of the energy of A-B to that of ATP could proceed in any of several ways and since the argument that follows does not rest on any particular mechanism we shall consider only the following possibility:



If the transphosphorylating mechanism is damaged A-B does not interact with inorganic phosphate but decomposes hydrolytically to A and B.



Let us return to the particle obtained by sonication referred to above. The rate of electron flow is exactly the same whether the coupling factor is present or not. In terms of the A-B formulation the rate with which the energy of A-B can be transformed to that of ATP must be exactly the same as the rate with which A-B is spontaneously broken down to A and B since the rate of electron flow is the same whether the particle is coupled or not. When the coupling factor is present the second choice is out. In the presence of the coupling factor A-B cannot form A and B hydrolytically but only by interaction with P_i and ADP successively. Dinitrophenol rules out the rearrangement leading to ATP synthesis; the coupling factor rules out the hydrolytic breakdown. This problem of choice at the molecular level is what makes the problem of coupling so fascinating and so frustrating.

There is yet a second possibility as to the cause of uncoupling. Suppose that electron flow does not lead to a stable bond between

A and B. The bond could be formed but its lifetime is so short that enzymatic manipulation is interdicted. This is probably the real problem in preserving the coupling function. During electron flow, bonds are formed between let us say three pairs of components (A-B, C-D, E-F). Only when the mitochondrion is in perfect alignment so to speak are these bonds stable. If the particle undergoes some distortions, or modification, the alignment of A with B, C with D, or E with F is affected adversely and instead of a stable bond a metastable bond is formed that cannot be exploited by the transphosphorylating enzymes.

This is the uncertainty principle at the mitochondrial level. The intact mitochondrion is stable but too complex to study. When simpler particles are prepared that are more suitable for the analysis of mechanism then the instability of the primary high energy bond nullifies the advantages of simplicity. The difficulty is probably only a temporary one but it is one of the major roadblocks to progress in the understanding of the coupling process.

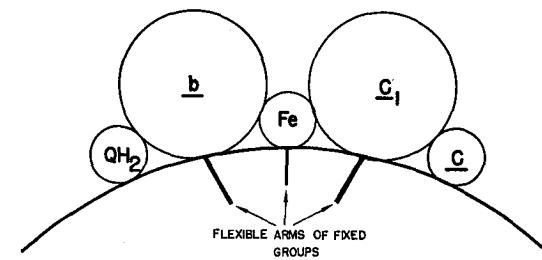
The heart of the coupling problem is the nature of A-B. The bond between A and B, formed by electron flow through the chain, is the consequence of energy conservation. That is the primary coupling process. The conversion of A-B to A and B in presence of P_i and ADP with concomitant synthesis of ATP is a secondary process unconnected with the primary coupling process. We have good reason to suspect that three of the four complexes of the elementary particle form the high energy bond. But what is unknown is which pair of oxidation-reduction components in each of the three complexes is involved and what kind of a bond is formed.

There are some 10 pairs of oxidation-reduction groups in the electron transport chain and only three phosphorylations. Since oxidative phosphorylation is a process of high efficiency it would be not only inefficient to trap the energy in only three of the ten oxidation-reductions but it is doubtful whether, at the most, more than one oxidation-reduction couple could supply sufficient energy for the synthesis of ATP (Clark, 1960; Krebs and Lowenstein, 1960). If the electron transport chain of each complex could behave as a single oxidation-reduction system the entire energy of the oxidation could be trapped and each complex would then be responsible for the esterification of one molecule of phosphate by ADP.

Let us consider any of the complexes of the electron transport chain from this standpoint. The QH_2 -cytochrome c reductase may

be looked upon as a complex of cytochromes b and c_1 with non-heme iron in some fashion being linked to both cytochromes and serving as an electronic bridge between them (Hatefi *et al.*, 1961a; Rieske, Wharton, and Zaugg, 1962). The protein moiety of cytochrome b contains a site to which reduced coenzyme Q is specifically attached while the protein moiety of cytochrome c_1 contains a site to which cytochrome c is specifically attached. Then in addition lipid, to the extent of some 28% by weight, is present in the complex. The most realistic way of dealing conceptually with the complex is to treat it

FIVE FUNCTIONAL GROUPS OF QH_2 -C REDUCTASE



PROTEIN OF QH_2 -C REDUCTASE

FIG. 7. Alignment of the 5 oxidation-reduction groups of QH_2 -cytochrome c reductase to form a continuous chain. The attachment of the fixed groups (*b*, Fe, c_1) to the protein via flexible side chains (such as lysine or other amino acids) allows for flexibility of rotation and fluctuation.

as the smallest common denominator or unit of electron transport. In a metalloflavoprotein the metal and the flavins to which it is linked are not considered as separate entities but as part of the same resonating complex (Mahler, 1956). The QH_2 -cytochrome c reductase (as well as the three other complexes) contains multiple oxidation-reduction groups that may be sufficiently close together to form a continuous chain in which electrons can be transferred smoothly from one component to the next. It is as if the components were hybridized to form a single oxidation-reduction system that catalyzes the oxidation of QH_2 by cytochrome c (cf. Fig. 7). Regardless of the mechanism invoked for electron transport the inescapable fact remains that the energy from several steps must be conserved in order to accomplish the synthesis of ATP.

Evidence is accumulating that electron flow is effected within the four complexes by the interaction of adjacent functional, oxidation-reduction groups. These groups which are linked to their respective proteins by flexible, molecular arms (e.g., a lysine group could serve in that capacity) are sufficiently close that they overlap during rotation or fluctuation. The principle underlying multicomponent complexes with electron transfer function is that of proteins fixed in position within the complex and of oscillating or revolving prosthetic groups.

The lipid coat that covers each of the complexes probably provides the medium in which electron flow through the oxidation-reduction components can take place. The stability of the A-B, C-D, and E-F bonds might be greatly weakened in the presence of water and the problem of preserving the coupling function could have as a major parameter the problem of preserving the integrity of the lipid-protein interfaces on which the exclusion of water depends.

XII. The Assembly of Mitochondria *in Vivo*

The intact mitochondrion contains at least 20 different proteins and nine separate complexes. The very complexity of this transducing system has encouraged speculation that mitochondria are laid down as a unit on a template and there is a school of thought that attributes self-duplicating properties to mitochondria. There is no convincing evidence available for such postulations. On the contrary, there is considerable evidence of a stepwise assembly process. The reconstruction of the electron transport chain by spontaneous combination of the four component complexes provides evidence that no directive influence other than complementarity of contouring surfaces is needed for this assembly (Hatefi *et al.*, 1961b). Structural protein combines spontaneously with the cytochromes and phospholipid (Criddle *et al.*, 1961). Here again the possibility of rapid and spontaneous interaction of mitochondrial components is realized; specificity of structure is the only qualification essential to ensure the right combinations.

There is reason to believe that all the parts of the mitochondrion are produced at the same time but not necessarily at the same site and that they reassemble spontaneously to form the mitochondrion just as two separated strands of nucleic acid can reform the original double-stranded unit. The studies of D. R. Sanadi (Fletcher and Sanadi, 1961) on the turnover of liver mitochondria have pointed

up two remarkable facts: (1) the half life of a mitochondrion is about 10 days; (2) the components making up the mitochondrion are synthesized at the same rate.

XIII. Pharmacology and Transducing Systems

The mode of action of some of the most potent and dramatic pharmacological agents (aspirin, digitalis, phloridzin, codeine, morphine, nicotine, belladonna, etc.) as well as of most of the hormones (ACTH, insulin, Pituitrin, adrenaline, etc.) is still an enigma despite three score years and more of intensive study by an army of investigators. More and more evidence is accumulating that the transducing machines of living systems—skeletal muscle, nerve, heart muscle, kidney tubules, elements of the cell membranes of cerebral cortex, etc., are the main targets of these powerful pharmacological and regulatory agents. The study of the mitochondrion has pinpointed the site of action and led to the elucidation of the mechanism of action of cyanide, Amytal, 2,4-dinitrophenol, and antimycin A, to mention the most outstanding agents. At the high dilutions in which these reagents act, their effects are exclusively on mitochondrial enzymes. The point to be made is that the frontiers of the science of pharmacology are concerned with the transducing systems and that, until the intricacies and details of the transducing systems are worked out, the fog surrounding the mode of action of the most powerful agents in the armory of the pharmacologists will remain. Some of the structures and components in any transducing system are unique to that system and certain pharmacological agents, by virtue of their affinity for or complementarity to these structures, are capable of modifying profoundly the behavior of the transducing system containing these structures. Once the reagent and the structure that it affects are paired off, and the chemical basis of their interaction worked out, then a science of pharmacology can be visualized in which reagents are designed for a particular structure in a transducing system. The work of the cell goes on predominantly in the transducing machines and these are the logical target points for reagents intended to control and regulate cellular activities.

XIV. Universality of the Molecular Principles of Mitochondrial Structure and Function

Mitochondria have been isolated from a wide variety of sources—animal, plant, and bacterial. The available evidence suggests

that, regardless of source, mitochondria contain essentially the same components and fulfill the same basic function. While occasionally there are significant differences in mitochondrial form these are not fundamental differences and much the same applies to some differences that have been observed in the electron transport process. The mitochondrion may be rod-like or spherical and even one and the same mitochondrion may exist in either of these forms depending upon conditions. Osmotic factors play an important role in determining whether a mitochondrion will exist in one form or another; thus, this mode of variation is not an expression of basic differences in structure. The number of cristae per mitochondrion varies enormously. In liver mitochondria there are relatively few and the spaces between cristae are extensive (Pallade, 1956; Sjostrand, 1959). In mitochondria of heart or flight muscle the cristae fill the interior almost completely and there is correspondingly little space between cristae (Chapman, 1954; Ziegler *et al.*, 1958). In a general way, it may be stated that mitochondria (e.g., from liver) that carry out auxiliary enzymatic functions, i.e., processes that are not connected with the primary transduction, have extensive intercristal spaces whereas mitochondria that concentrate on the transduction process, e.g., those of heart or flight muscle, lack such extensive intramitochondrial spaces. The available evidence suggests that most if not all of the auxiliary enzyme systems are localized in the intercristal spaces. Many of these auxiliary enzyme systems carry out metabolic processes that may require ATP, e.g., the synthesis of hippuric acid (Sarkar *et al.*, 1952) from glycine and benzoic acid; such a requirement may underlie the localization of some synthetic processes in the mitochondrion. The point at issue here is that some mitochondria carry extra enzymatic baggage not concerned in the transduction process and this extra baggage may be unique to mitochondria from one particular source. In any event, this mode of mitochondrial variation is not an expression of a basic difference. The most disconcerting aberrations in mitochondrial form and function have been observed in bacterial systems and for some considerable time these were taken to mean that bacteria do not have mitochondria. One great virtue of the bacterial systems is that they have compelled a more logical and rigorous definition of what constitutes a mitochondrion.

Some, if not most, bacteria do not have an internal body or structure which the electron microscopist would identify as a full-

blown mitochondrion (Kubai *et al.*, 1961). However, bacteria do contain double-membraned structures, usually attached to the cell membrane, that are indistinguishable in dimensions from those of the cristae of typical mitochondria. Indeed, when subcellular particles having typical mitochondrial function are isolated from bacteria their particles seem to correspond to the double-membraned structures seen in electron microscopy. The bacteria have taught us that what is basic in mitochondrial structure is the double-membraned arrangement of the cristae and what is expendable or at least of secondary importance, is a regular system of cristae enclosed by an external membrane. Some bacteriologists make a to-do about the fact that the respiratory particles are invaginations of the bacterial cell membrane rather than bodies in the interior. This again is rather a parochial difference and does not affect the heart of the mitochondrial problem. A much more serious dilemma is posed by the observation that some bacterial particles do not carry out the citric acid cycle and this cannot lightly be disposed of.

Basically the mitochondrion is a device for coupling electron flow to synthesis of ATP. Generally the ultimate sources of electrons are the substrates of the citric acid cycle. But this is not mandatory. In some animal mitochondria, α -glycerophosphate (Sacktor and Cochran, 1958) or β -hydroxybutyrate (Green *et al.*, 1937) is the source of electrons rather than citric cycle substrates. Regardless of whether this substitution is complete or partial the point has been made that the source of electrons is not invariant and, by the same token, not a fundamental consideration. The coupling of electron flow to synthesis of ATP is the constant and the source of the electrons is the variable. In bacteria a wide variety of oxidizable substances may serve as the source of electrons. But the same electron transport and coupling machinery intervenes regardless of the particular substance used as the oxidizable substrate.

The electron transport particles from *Azotobacter vinelandii* have been shown to contain exactly the same components (the two flavoproteins, the four cytochromes, coenzyme Q, non-heme iron, copper, lipid, etc.) as does the mitochondrion of beef heart muscle and in essentially the same proportions (Bruemmer *et al.*, 1957). Whenever analytical work of this kind has been carried out with mitochondria from various sources the results have led to much the same conclusion: the basic chemical properties and organizational

principles of the mitochondrion are invariant. Apparently, the mitochondrion was developed early in evolutionary history and, apart from some minor variations, the pattern has remained invariant thereafter. The mitochondrion had to be perfected before oxygen could be exploited as an oxidizing agent for energetic purposes and this perfection appears to have been completed before the unicellular organisms evolved to more complex forms.

The mitochondrion is part of the legacy of all aerobic cells from the unicellular organisms to the most complex multicellular forms of life. Underneath the bewildering variation and variability of living systems there are structures such as the mitochondrion which have resisted the erosion of the evolutionary process.

REFERENCES

- Ambe, K. S., and Crane, F. L. (1959). *Science* **129**, 98.
 Ambe, K. S., and Venkataraman, A. (1959). *Biochem. Biophys. Research Commun.* **1**, 133.
 Basford, R. E., and Green, D. E. (1959). *Biochim. et Biophys. Acta* **33**, 185.
 Beinert, H., and Lee, W. (1961). *Biochem. Biophys. Research Commun.* **5**, 40.
 Beinert, H., and Sands, R. H. (1959). *Biochem. Biophys. Research Commun.* **1**, 171.
 Belitzer, B. A., and Tsibikova, E. T. (1939). *Biokhimiya* **4**, 516.
 Bruemmer, J. H., Wilson, P. W., Glenn, J. L., and Crane, F. L. (1957). *J. Bacteriol.* **73**, 113.
 Chance, B., and Williams, G. R. (1956). *Advances in Enzymol.* **17**, 65.
 Chapman, J. B. (1954). *J. Morphol.* **95**, 237.
 Clark, W. M. (1960). "Oxidation-Reduction Potentials of Organic Systems," Williams & Wilkins, Baltimore, Maryland.
 Crane, F. L., Glenn, J. L., and Green, D. E. (1957). *Biochim. et Biophys. Acta* **22**, 475.
 Criddle, R., and Bock, R. (1959). *Biochem. Biophys. Research Commun.* **1**, 138.
 Criddle, R. S., Bock, R. M., Green, D. R., and Tisdale, H. D. (1961). *Biochem. Biophys. Research Commun.* **5**, 75.
 Cross, R. S., Taggart, J. V., Covo, G. A., and Green, D. E. (1949). *J. Biol. Chem.* **177**, 655.
 Doeg, K. A., and Ziegler, D. M. (1962a). *Biochim. et Biophys. Acta* (in press).
 Doeg, K. A., and Ziegler, D. M. (1962b). *Arch. Biochem. Biophys.* **97**, 37.
 Fernández-Morán, H., unpublished experiments.
 Fleischer, S., and Brierley, G. (1961a). *Biochem. Biophys. Research Commun.* **5**, 367.
 Fleischer, S., and Brierley, G. (1961b). *Biochim. et Biophys. Acta* **53**, 609.
 Fleischer, S., and Klouwen, H. (1961). *Biochem. Biophys. Research Commun.* **5**, 378.
 Fleischer, S., Brierley, G., Klouwen, H., and Slatterback, D. B. (1962). *J. Biol. Chem.*, in press.
 Fleischer, S., Klouwen, H., and Brierley, G. (1961). *J. Biol. Chem.* **236**, 2936.
 Fletcher, M. J., and Sanadi, D. R. (1961). *Biochim. et Biophys. Acta* **51**, 356.
 Fowler, L. R., and Hatefi, Y. (1961). *Biochem. Biophys. Research Commun.* **5**, 203.
 Fowler, L. R., and Richardson, S. H. (1962). In preparation.
 Gaffron, H., Brown, A. H., French, C. S., Livingston, R., Rabinowitch, E. I., Strehler, B. L., and Tolbert, W. E., eds. (1957). "Research in Photosynthesis." Interscience, New York.
 Goldberger, R., Smith, A. L., Tisdale, H. D., and Bomstein, R. (1961). *J. Biol. Chem.* **236**, 2788.
 Green, D. E. (1958). *Harvey Lecture Ser.* **52**, 177.
 Green, D. E., and Burkhard, R. K. (1961). *Arch. Biochem. Biophys.* **92**, 312.
 Green, D. E., and Fleischer, S. (1960a). In "Enzymes in Health and Disease" (D. M. Greenberg and H. A. Harper, eds.), pp. 51-72. C. C. Thomas, Springfield, Illinois.
 Green, D. E., and Fleischer, S. (1960b). In "Metabolic Pathways" (D. M. Greenberg, ed.), Vol. I, p. 41. Academic Press, New York.
 Green, D. E., and Fleischer, S. (1961-1962). Unpublished observations.
 Green, D. E., and Hatefi, Y. (1961). *Science* **133**, 13.
 Green, D. E., and Oda, T. (1961). *J. Biochem. (Japan)* **49**, 742.
 Green, D. E., Dewan, J. G., and Leloir, L. F. (1937). *Biochem. J.* **31**, 934.
 Green, D. E., Tisdale, H. D., Criddle, R. S., Chen, P. Y., and Bock, R. M. (1961a). *Biochem. Biophys. Research Commun.* **5**, 109.
 Green, D. E., Tisdale, H. D., Criddle, R. S., and Bock, R. M. (1961b). *Biochem. Biophys. Research Commun.* **5**, 81.
 Griffiths, D. E., and Wharton, D. C. (1961a). *J. Biol. Chem.* **236**, 1850.
 Griffiths, D. E., and Wharton, D. C. (1961b). *J. Biol. Chem.* **236**, 1857.
 Hartley, G. S. (1955). In "Solutions of Soap-Like Substances," Progress in the Chemistry of Fats and Other Lipids (R. T. Holman, W. O. Lundberg, and T. Malkin, eds.), p. 19. Pergamon Press, New York.
 Hatefi, Y. (1959). *Biochim. et Biophys. Acta* **34**, 184.
 Hatefi, Y., Haavik, A. G., and Griffiths, D. E. (1961a). *Biochem. Biophys. Research Commun.* **4**, 441.
 Hatefi, Y., Haavik, A. G., and Griffiths, D. E. (1961b). *Biochem. Biophys. Research Commun.* **4**, 447.
 Hatefi, Y., Haavik, A. G., and Jurtschuk, P. (1961c). *Biochim. et Biophys. Acta* **52**, 106.
 Jurtschuk, P., Sekuzu, I., and Green, D. E. (1961). *Biochem. Biophys. Research Commun.* **6**, 76.
 Kalckar, H. M. (1939). *Biochem. J.* **33**, 631.
 Keilin, D. (1933). *Ergeb. Enzymforsch.* **2**, 239.
 Keilin, D., and Hartree, E. F. (1938). *Proc. Roy. Soc. B125*, 171.
 Krebs, H. A., and Lowenstein, J. M. (1960). In "Metabolic Pathways" (D. M. Greenberg, ed.), Vol. I, p. 129. Academic Press, New York.
 Kubai, D., Ziegler, D. M., and Ris, H. (Nov. 1961). *Am. Soc. Cellular Biol.* p. 119.
 Lardy, H. A. (1956). *Proc. Intern. Congr. Biochem., 3rd Congr., Brussels, 1955* p. 287.
 Lehninger, A. L., and Kennedy, E. P. (1949). *J. Biol. Chem.* **179**, 957.

- Lester, R. L., and Fleischer, S. (1959). *Arch. Biochem. Biophys.* **80**, 470.
 Lester, R. L., and Fleischer, S. (1961). *Biochim. et Biophys. Acta* **47**, 358.
 Lester, R. L., and Hatefi, Y. (1958). *Biochim. et Biophys. Acta* **29**, 103.
 Linnane, A. W., and Titchener, E. B. (1960). *Biochim. et Biophys. Acta* **39**, 469.
 Linnane, A. W., and Ziegler, D. M. (1958). *Biochim. et Biophys. Acta* **29**, 630.
 Loomis, W. F., and Lipmann, F. (1948). *J. Biol. Chem.* **173**, 807.
 Mahler, H. R. (1956). *Advances in Enzymol.* **17**, 233.
 Oda, T., Blair, P., and Green, D. E. (1961-1962). Unpublished studies.
 Okunuki, K., Sekuzu, I., Yonetani, T., and Takemori, S. (1958). *J. Biochem. (Japan)* **45**, 847.
 Pallade, G. E. (1956). In "Enzymes: Units of Biological Structure and Function" (O. H. Gaebler, ed.), p. 185. Academic Press, New York.
 Pullman, M. E., Penefsky, H., and Racker, E. (1958). *Arch. Biochem. Biophys.* **76**, 227.
 Rhodes, W. C., and McElroy, W. D. (1958). *J. Biol. Chem.* **233**, 1528.
 Rieske, J., Tisdale, H. D., and Green, D. E. (1961-1962). Unpublished studies.
 Rieske, J., Wharton, D., and Zaugg, W. (1962). In preparation.
 Rouser, G. (1958). *Am. J. Clin. Nutrition* **6**, 681.
 Sacktor, B., and Cochran, D. G. (1958). *Arch. Biochem. Biophys.* **74**, 266.
 Sarkar, N. K., Beinert, H., Fuld, M., and Green, D. E. (1952). *Arch. Biochem.* **37**, 140.
 Schneider, W. C., and Potter, V. R. (1949). *J. Biol. Chem.* **177**, 893.
 Sekuzu, I., Jurtschuk, P., and Green, D. E. (1961). *Biochem. Biophys. Research Commununs.* **6**, 71.
 Sjostrand, F. S. (1959). *Rev. Modern Phys.* **31**, 301.
 Theorell, H., and Akesson, A. (1941). *J. Am. Chem. Soc.* **63**, 1804.
 Wald, G. (1956). In "Enzymes: Units of Biological Structure and Function" (O. H. Gaebler, ed.), Academic Press, New York.
 Weber, H. H. (1958). "The Motility of Muscle and Cells." Harvard Univ. Press, Cambridge, Massachusetts.
 Webster, G. (1962). *Biochem. Biophys. Research Commununs.* (in press).
 Widmer, C., and Crane, F. L. (1958). *Biochim. et Biophys. Acta* **27**, 203.
 Ziegler, D. M. (1961). In "Biological Structure and Function" (T. W. Goodwin and O. Lindberg, eds.), Vol. II, p. 253. Academic Press, New York.
 Ziegler, D. M., and Doeg, K. A. (1959). *Biochem. Biophys. Research Commununs.* **1**, 344.
 Ziegler, D. M., Linnane, A. W., Green, D. E., Dass, C. M. S., and Ris, H. (1958). *Biochim. et Biophys. Acta* **28**, 524.

Respiration-Linked Mechanochemical Changes in Mitochondria¹

ALBERT L. LEHNINGER

*Department of Physiological Chemistry,
The Johns Hopkins University School of Medicine,
Baltimore, Maryland*

In this paper I should like to outline some experimental work and ideas on those changes in permeability and conformation of the membranes of mitochondria which are involved in the uptake and extrusion of water. These "mechanochemical" or conformational changes in the membranes are closely linked to intermediate steps in the mechanism by which adenosine triphosphate (ATP) is formed during respiratory chain phosphorylation. This discussion will be presented against the background of a more general thesis, namely that the oxidoreduction energy of electron transport is used to drive three distinct activities of mitochondria: (a) the synthesis of ATP from adenosine diphosphate (ADP) and phosphate, (b) active transport of certain ions, and (c) the mechanochemical changes stated above.

The more traditional conception of energy transduction in the respiration chain is that the respiratory energy is conserved only as the phosphate bond energy of ATP, and that the latter is then used, in secondary reactions, to drive such endergonic activities as active transport, the mechanochemical changes, as well as biosynthetic reactions. However there is now increasing evidence from work in our laboratory and others that the specific key events in both the respiration-dependent specific ion binding and the respiration-dependent mechanochemical changes are brought about by those catalysts or structures in the respiratory assembly which also bring about the formation of ATP. Thus there are visualized three different modalities of respiratory energy transduction. It is believed that this concept of polymodal energy transduction, which we have stated before (Leh-

¹ Original work described in this paper was supported by grants from the National Institutes of Health, the National Science Foundation, The Nutrition Foundation, and the Whitehall Foundation.

ninger, 1955, 1956, 1959a,b, 1960a, 1961a), as have Price *et al.* (1956) and Mitchell (1961a,b), can provide new experimental and theoretical approaches to the problem of the mechanism of oxidative phosphorylation, which has very largely been approached in the past in one dimension only, namely with the traditional conception that the respiratory chain is a multienzyme system whose sole function is to conserve oxidoreduction energy as ATP.

That phosphorylating respiration is the driving force for active transport in aerobic cells is now widely acknowledged. The basic idea that active transport of certain ions is driven by asymmetric discharge of H^+ and OH^- ions from the respiratory chain into separate compartments was first postulated by Lundegardh (1940), and has since been refined by Conway (1953), Davies and Ogston (1950), and others [reviewed by Robertson (1960)]. Such ideas on the use of the respiratory chain as a "redox pump" have more recently been supplemented by speculations based on the proton exchanges accompanying the three phosphorylations in the chain. Davies and Ogston (1950), and particularly Mitchell (1961a,b), have pointed out that the respiration-coupled formation of ATP from ADP and phosphate, formally a dehydration, could lead to discharge of the H_2O as H^+ and OH^- ions into separate compartments. These ideas are now central and important principles in contemporary thought on the molecular mechanism of active transport. They are given some support by the recent findings of Skou (1957, 1960) and Jarnefelt (1961) that ATP hydrolysis in nerve membrane fractions is stimulated by K^+ and Na^+ , as well as the similar findings of Post *et al.* (1960) on the ATPase of erythrocyte membranes. These authors have postulated that these ATPase systems are located in cell membranes in such a fashion that the hydrolysis of ATP drives the separation of Na^+ and K^+ across the membrane. The phosphorylating digitonin fragments of the mitochondrial membranes (cf. Lehninger *et al.*, 1958) are capable of causing the binding of K^+ during respiration; Gamble (1957) has shown that K^+ binding is rather specific and does not require ADP or ATP, but simply the passage of electrons along the respiratory chain. However binding of K^+ is blocked by 2,4-dinitrophenol under these circumstances. It must be concluded that binding of K^+ is driven by intermediate reactions in energy coupling which occur prior to formation of ATP. Vasington (1962) has shown that Ca^{++} is also bound by these fragments during respiration, but not in the

presence of uncoupling agents. These findings, as well as those on mitochondrial swelling and contraction described below, thus show that ATP itself is not necessary to drive these reactions but that they may be driven by some preceding high-energy intermediate.

Now let us consider the directions of contemporary research on the mechanism of oxidative phosphorylation. It is now some thirty years since the basic pattern of the respiratory chain was recognized by Keilin and by Warburg, and some twenty years since phosphorylations coupled to the respiratory chain were first postulated to occur by Belitser and Tsibakova. In the intervening period, many important experimental findings have of course been made concerning details of electron transport and the sites and properties of the coupled oxidative phosphorylations, but even with the increasing tempo of recent research in this field, it cannot be said that we are yet really close to a description of the mechanism of oxidative phosphorylation. Virtually all the research on respiratory chain phosphorylation of this period has followed in its philosophy the expectation that this complex enzyme system will turn out in principle to resemble such well-understood multienzyme systems as the glycolytic cycle and the urea cycle, in that it would involve the sequential action of a series of enzymes, each acting on the product of the preceding member, so that finally a chemical flow-sheet or "metabolic map" for oxidative phosphorylation might be written out, in terms of specific enzymes and specific intermediates or prosthetic groups.

Even the discovery that the enzymes of the respiratory chain occur tightly bound in the membranes of the mitochondria, with their typical stratified lipid-protein ultrastructure, has not significantly shaken general confidence that the mechanism of electron transport and oxidative phosphorylation can ultimately be solved by the approaches and postulates of classic enzymology. It has of course placed the problem on a more difficult plane, because of the technical difficulties of "solubilization" of individual enzymes and of the reconstruction of the system on specific sites of a solid-phase assembly of enzyme molecules. Techniques for both, while still quite empirical, are in fact improving. Several recent investigations have indeed achieved some significant successes in specific reconstitution of the respiratory chain. These include the rebinding of soluble purified succinic dehydrogenase to the cytochrome chain in deficient membrane fragments, described by Keilin and King (1960), the recombination of particles containing separate segments of the respir-

atory chain to reconstitute a complete but non-phosphorylating DPN-linked respiratory chain operating through coenzyme Q in Green's laboratory (cf. Fowler and Hatefi, 1961), the recoupling of phosphorylation to electron transport by a soluble ATPase in Racker's laboratory (Pullman *et al.*, 1960; Penefsky *et al.*, 1960), and the reconstitution of dinitrophenol sensitivity of the soluble, purified ATP-ADP exchange enzyme recently studied in our laboratory (Wadkins and Lehninger, 1960). Thus there is in fact some encouragement that classic enzyme isolation and reconstruction methods, suitably modified to encompass "insoluble" systems, may yet lead to a classic "flow-sheet" description of respiratory chain phosphorylation. However, progress has really been rather slow in the light of the highly developed and sophisticated enzymology of today, and it seems possible that this is owing to some failure to recognize that the respiratory chain also has the osmotic and mechanochemical functions described above, which might also be exploited in the experimental attack on oxidative phosphorylation.

These comments are not intended to imply a retreat into the nihilistic position that this problem is hopelessly difficult to solve or that the classic "enzyme-system" approach has nothing to offer. Actually it is my expectation that many, if not most of the individual steps of electron transport and of ATP formation will be approachable by classic enzymology and will turn out to be enzyme systems in the classic "flow-sheet" sense, but where the "intermediates" are non-diffusing, bound prosthetic or active groups on the enzymes. However at the three key coupling points in the respiratory chain I expect there will be found to exist molecular structures which transduce oxidoreduction energy not only into the ATP precursor, but also into osmotic and mechanochemical events, for which it will not be possible to write a simple enzymatic reaction sequence without specifying some structural, directional, or compartmenting factors. Thus both classic enzymology and the principles of molecular arrangement in ordered structures will have to be employed to solve this problem at the molecular level. But even this view may not be radical enough. Mitchell has recently proposed, in an important and thoughtful article (Mitchell, 1961), that respiratory chain phosphorylation proceeds with no intermediates at all and is in principle the consequence of the separation of H^+ and OH^- into two compartments by a reversible ATPase, situated anisotropically in the membrane dividing these compartments, as the ATPase is

dehydrating ADP and phosphate to form ATP. The driving force for the ion separation and thus the ATP formation is provided by "sinks" for the H^+ and OH^- ions. These "sinks" are created by the OH^+ and H^- , respectively, which are viewed as being discharged into these compartments by an anisotropically located electron-transferring system. Thus in Mitchell's view the search for "intermediates" by enzymologists may be redundant; the entity to be sought is a "compartmented," or anisotropic ATPase, which has also been postulated by Skou (1957, 1960) and more recently by Järnefelt (1961) to account for ATP-driven Na^+ and K^+ movements in nerve.

Intrinsically it is not at all necessary to have an "insoluble" or "particulate" system to convert oxidoreduction energy enzymatically into so-called phosphate bond energy of ATP. The well-known soluble glyceraldehyde 3-phosphate dehydrogenase system is an efficient energy-coupling system, as is the internal oxidoreduction catalyzed by enolase. These two reactions are in fact the energetic mainsprings of most anaerobic cells. In principle, a series of similar soluble enzyme systems, of the classic "flow-sheet" type, could also perform efficient energy conversion into ATP on passing electrons from an electronegative substrate to molecular oxygen in aerobic cells. But the course of biochemical evolution evidently did not occur this way, and in all aerobic cells the enzymes of the energy-coupling respiratory chains are fixed in membranes showing structural anisotropy of lipid and protein: in higher cells in the mitochondria and in aerobic bacteria in the protoplast membranes. The evolutionary location of the respiratory chains in such membranes could well have been selected biologically because of the necessity to use the energy developed by the respiratory chain to drive the directional flow of electrolytes necessary to preserve constancy of the intracellular milieu, and to drive the conformational changes which presumably control the entrance and exit of metabolites and water.

Now let us examine more closely the mechanochemical aspects of mitochondrial energy coupling which we have been studying. Many investigators, beginning with Raaflaub (1953), have observed that isolated mitochondria undergo swelling with water uptake in the presence of specific "swelling" agents of natural occurrence, which include Ca^{++} , phosphate, thyroxine, phloridzin, oxidized and reduced glutathione (Neubert and Lehninger, 1962), the disulfide hormones (Lehninger and Neubert, 1961), and higher fatty acids,

when these are added in amounts making no significant change in the tonicity of the medium (cf. reviews by Lehninger, 1960b, 1962). This swelling, which is easily followed optically or by direct gravimetric measurement of volume, is linked to the process of respiration, since water uptake is blocked by respiratory inhibitors such as cyanide and antimycin A, or by anaerobiosis. This type of swelling is also inhibited by the uncoupling agent 2, 4-dinitrophenol and this fact provides further important evidence for an enzymatic relationship between water uptake and respiration. However, fully coupled respiration in the presence of phosphate acceptor does not lead to swelling; the swelling is associated more specifically with respiration in the absence of phosphate acceptor, i.e., the "State 4" of Chance and Williams (1956). The swelling proceeds with an increase in permeability toward test solutes such as sucrose, which normally does not penetrate rapidly (Lehninger *et al.*, 1959).

Contraction of the swollen mitochondria, with extrusion of water, can be brought about either by restoring fully coupled respiration and phosphorylation by addition of ADP, or, in the absence of respiration, by simple addition of ATP + Mg⁺⁺. For the ATP-linked contraction, respiration is unnecessary; neither cyanide or dinitrophenol block contraction driven by ATP. On the other hand, three characteristic inhibitors of oxidative phosphorylation, namely azide, oligomycin, and sucrose, completely block ATP-induced contraction (Lehninger, 1959c,d, 1962). Thus it may be concluded that water uptake by the mitochondria is dependent on electron transport, but does not require the entire, complete process of oxidative phosphorylation, whereas the water extrusion is dependent on a more terminal stage of the energy-coupling system responsible for forming ATP, and can occur in the absence of respiration.

Mitochondrial swelling and contraction have been studied *in vitro* under two rather different sets of conditions. In one situation, so-called "small-amplitude" swelling-contraction cycles have been studied, mainly in the work of Packer (1960) and Holton (1957), in which the optically measured amplitude of the cycle is at the most 1 or 2% of the original mitochondrial volume; the respiration may remain tightly coupled throughout several such low-amplitude swelling-contraction cycles. Under these conditions the swelling-contraction cycle has properties suggesting that it is a simple reversible phenomenon dependent on the concentration of some high energy intermediate of oxidative phosphorylation. But this simple picture

is probably not correct, in view of the properties of the so-called "large-amplitude" swelling-contraction cycles (cf. review by Lehninger, 1962). If mitochondrial swelling proceeds long enough (10–20 minutes) the mitochondria may swell to double or triple their original volume, with easily and directly measured uptake of water, which is extruded again on addition of ATP. These "large-amplitude" cycles are reversible, since ATP can bring about extrusion of the water which was taken up, but during such cycles the normally tight coupling between respiration and phosphate acceptor is lost and the mitochondria suffer some irreversible damage to the mechanism of oxidative phosphorylation. This type of "large-amplitude" mitochondrial swelling and contraction thus represents a rampant, uncontrolled swelling process isolated from its normal counter-opposing restraints in the cell. But it has the great advantage of "loosening" the normally tightly coupled relationships between respiration and phosphorylation and the osmotic and mechanical events, so that the critical points connecting them may be identified.

Such "large-amplitude" studies, as mentioned above, have revealed that swelling requires respiration and that contraction may proceed independently of electron flow, but in this case requires the azide- and oligomycin-sensitive step of oxidative phosphorylation to activate it. The independence of contraction from electron flow is also clearly shown by the fact that mitochondria can be left in the fully swollen state for several hours and over this time all ability for organized respiration and phosphorylation is lost, but such mitochondria still contract on addition of ATP. Thus swelling and contraction are not necessarily the simple reverse of each other, a conclusion which has an important bearing on the nature of the physical changes occurring in the mitochondrial structure. In any case one would not expect the swelling-contraction cycle to be simply a reversible increase and decrease of permeability toward some solute of the medium. There is some evidence that the water uptake is the consequence of an increase in permeability toward solutes of the medium, entry of which would be accompanied by uptake of water. However a *decrease* of permeability of swollen mitochondria by addition of ATP could hardly be expected to cause extrusion of water, but could only keep it locked within, along with the solute. There is increasing evidence that the water extrusion is the result of a contractile or conformation change in disposition of the mitochondrial membranes (cf. Lehninger, 1962).

A second important conclusion has followed from study of the properties of the "large-amplitude" cycles. Intact mitochondria, as mentioned before, are capable of causing active ion movements between the intramitochondrial and extramitochondrial compartments, coupled to respiration. It might be thought that water uptake and extrusion in mitochondria, such as that described above, is in reality a secondary consequence of the movements of ions into and out of mitochondria. However this possibility has now been excluded and a clear differentiation between ion movement and mitochondrial contraction has been achieved, although normally both are linked to the fully coupled chain. When mitochondria are aged in the swollen state they lose not only the ability to respire and phosphorylate but also they lose all activity in the specific binding of K^+ from the medium. However such mitochondria still contract on addition of ATP in an oligomycin-sensitive reaction. Furthermore, ATP-induced contraction is relatively independent of the ionic environment (Lehninger, 1961b).

One of the most promising entries into the problem of the enzymology of the ATP-induced contraction is the finding that it involves a specific mitochondrial protein, which we have called "Contraction-factor," or "C-factor." As mentioned before, mitochondrial swelling requires not only a state of active respiration, but the presence of a specific swelling agent. ATP can reverse swelling caused by any of these swelling agents, with the exception of swelling induced by reduced glutathione. This agent causes a more severe swelling than most others (Lehninger and Schneider, 1959; Lehninger, 1959d). The failure of $ATP + Mg^{++}$ to contract glutathione-swollen mitochondria has been traced down to the fact that glutathione causes detachment or leakage of a specific, heat-labile, soluble protein (C-factor) from the mitochondria into the suspending medium (Lehninger and Gotterer, 1960). This protein is required in contraction, and following its detachment from the mitochondrion it becomes so dilute in the suspending medium it is no longer effective in supporting the ATP-induced contraction. However when excess C-factor is added back to the medium, then glutathione-swollen mitochondria contract readily in the presence of ATP.

We have recovered this C-factor protein either from suspending media in which mitochondria had been exposed to glutathione, or from sonic extracts of mitochondria. One of the most striking of its properties is the fact that when it is added in increasing concentra-

tions to glutathione-swollen mitochondria in the presence of ATP, it produces graded increases in the final equilibrium level of contraction which is attained, rather than increases in the rate of contraction. Thus the swollen mitochondria can literally be titrated with C-factor in the presence of excess ATP until nearly all the excess water is extruded. This property clearly suggests a stoichiometric requirement of C-factor by the membrane in the presence of ATP, rather than a catalytic participation (Lehninger and Gotterer, 1960). Another interesting property is that ATP and C-factor are apparently bound to the membrane in some rather specific interdependent manner, since the contractile effect occurs only after one specific sequence of addition. If C-factor is added to glutathione-swollen mitochondria *after* the ATP, then no contraction results. On the other hand, if C-factor is added first, followed by ATP, then full contraction is observed (Neubert and Lehninger, 1962). Mg^{++} ions are necessary for the C-factor activity.

We have also found, using the GSH-swollen mitochondria as a bioassay for C-factor activity, that this factor may be extracted in soluble form from all types of mitochondria tested, i.e., from brain, liver, kidney, muscle, and heart of different species and from the Ehrlich ascites tumor of mice (Rose *et al.*, 1962). It can also be detected in extracts of phosphorylating mitochondrial membrane fragments such as prepared with digitonin or by sonic methods. Thus it is always found in preparations capable of oxidative phosphorylation. Curiously, and this may be very significant, C-factor activity has been extracted also from mature erythrocytes and from protoplasts of *Micrococcus lysodeikticus* (Rose *et al.*, 1962). It will be recalled that the erythrocyte as well as certain bacterial protoplasts also undergo metabolism-linked or ATP-linked swelling-contraction cycles. C-factor may thus be a common enzymatic denominator among these membranous structures (Lehninger, 1961a; Rose *et al.*, 1962).

Still another significant finding is the fact that C-factor activity is found in the soluble fraction of the cytoplasm of a number of tissues, but not in the microsomes or nuclei. This and other findings suggest that there may be a dynamic equilibrium of C-factor between the mitochondria and the bathing cytoplasm, which is dictated by the level of glutathione or by the ratio of oxidized and reduced glutathione (Neubert and Lehninger, 1962). Thus soluble cytoplasmic factors can powerfully affect the shape and properties

of mitochondria in the cell and may help make understandable the fact that mitochondria in intact, unfixed cells, as in the fibroblast for example, may undergo rather enormous changes in volume and shape, presumably reversible. These *in vivo* transformations of mitochondria are inhibited by cyanide and dinitrophenol, as well as other respiratory inhibitors, as shown by phase-contrast films (Frederic, 1958).

If there is in fact a functional relationship between the respiratory chain, the ATP-forming enzymes, and the mechanochemical changes occurring in mitochondrial swelling and contraction, then any specific enzymatic activity of C-factor protein, in addition to its contractile activity, would be of the greatest significance. We have found that purified C-factor does not catalyze the ATP-ADP or the ATP- P_i^{32} exchange reactions nor does it have ATPase activity when freshly prepared. However Drs. Neubert and Gregg have shown that C-factor preparations can greatly increase the P:O ratio of mitochondria or membrane fragments under special conditions. C-factor preparations also often show M-factor activity; the latter factor is identified by its ability to confer dinitrophenol sensitivity on the ATP-ADP exchange reactions in phosphorylating membrane fragments (Wadkins and Lehninger, 1960). Thus it is clear that C-factor not only has activity in the contractile response of the membrane to ATP, but also may be an enzymatic component involved in formation of ATP.

Another experimental approach to the biochemistry of the mechanochemical changes in the membranes has been yielded by recent light scattering studies of membrane fragments prepared by sonic treatment. Suspensions of such fragments, which retain capacity for oxidative phosphorylation, undergo large changes in light absorption when supplemented with succinate and ATP. The less easily sedimented fragments also are observed to undergo changes in angular dissymmetry of light scatter (cf. Gotterer *et al.*, 1961), particularly at the ratio 135°/45°, when ATP is added. Both types of effect are prevented when oligomycin is present. These experiments not only confirm the dependence of mechanochemical changes on enzymes of oxidative phosphorylation, but also demonstrate that mechanochemical changes can take place in the well-washed membrane fragments, in the absence of soluble components of the mitochondrial matrix.

From this experimental evidence, we have postulated that some

of the enzyme molecules comprising the "respiratory assemblies" which are embedded in the protein monolayer of the mitochondrial membranes, are what may be termed "mechanoenzymes," capable of conformational changes depending on respiratory state (Lehninger, 1959c,d, 1960b, 1961a,b, 1962). More specifically it is suggested the oxidation-reduction state of the electron carriers, or the occurrence of the coupling enzymes in either a phosphorylated or dephosphorylated forms, depending on the presence of ATP, can bring about conformational changes either in individual molecules of the assembly, as in their tertiary structure, or in the entire assembly by alterations of the binding between adjacent molecules, for example, through electrostatic or conformational changes. Since the enzymes of the respiratory assemblies constitute from 25 to 50% of the protein of the mitochondrial membranes, it can be understood that such conformational changes in individual respiratory assemblies can bring about changes in permeability of the membrane or in the manner of its folding to form cristae, or possibly in its flat dimensions, to account for the respiration-driven water uptake and extrusion (cf. Lehninger, 1959c,d, 1960a,b, 1961a,b, 1962).

The terms "mechanochemistry" and "mechanoenzymes" which I have used in this connection have been borrowed of course from Engelhardt, who first used them to describe the relationship between the ATPase of myosin and the "mechanical" changes occurring in this contractile system (Engelhardt and Ljubimowa, 1939; Engelhardt *et al.*, 1941). But these terms were not borrowed simply because of some superficial resemblances between the ATP-linked mechanical changes in mitochondria and in the myofibril. Actually there are many profound similarities between these two systems which deserve close attention. First, both myosin and mitochondrial membranes possess ATPase activity in a latent form. Hydrolysis of ATP is not a normal event in oxidative phosphorylation, and perhaps not in the contractile event in the actomyosin system. In both cases ATP hydrolysis is caused by an aberrant or side reaction in which a water molecule replaces the normal reactant; such ATP hydrolysis may be evoked by structural changes in the mitochondrial membrane, or by lack of proper orientation of actin and myosin such as occurs in an actomyosin solution. Both ATPase activities are stimulated by 2,4-dinitrophenol (cf. Greville and Reich, 1956); this finding is rather significant because dinitrophenol is not known to attack any other enzymatic reactions. Both mitochondrial

membranes and myosin catalyze O^{18} exchanges between H_2O and phosphate (cf. Chan *et al.*, 1960; Levy *et al.*, 1960). Still another similarity is the finding that both types of system are similarly and biphasically dependent on sulfhydryl groups. Silver ions cause stimulation of ATPase activity of mitochondria and of myosin at low concentrations, but inhibition at high concentrations. Titration of sulfhydryl groups of myosin (Kielley and Bradley, 1956) and mitochondrial membrane fragments (Cooper, 1960; Kielley, 1962) has shown occurrence of one class of sulfhydryl group which normally inhibits ATPase activity and a second class which is necessary for ATPase activity after the first type has been titrated.

All these similarities are made most compelling by the fact that both systems participate in "mechanoenzymatic" activities, in which orientation and placement of the component molecules or submolecules relative to each other may be related to binding or splitting of ATP. These similarities suggest at the least a common evolutionary precursor of the two systems, and also furnish a suggestive background for assessing the possible significance of the C-factor activity shown by actin. Clearly study of oxidative phosphorylation can benefit not only from study of the mechanochemical activity of mitochondria, but also from contemplation of the "mechanochemistry" of the actomyosin system, in particular its subunits first discovered by Albert Szent-Györgyi and his school: F-actin, G-actin, and the L- and H-meromyosins.

The conception of "mechanoenzymes" or mechanoproteins" is neither novel or radical, nor does it involve any particularly new principles of protein structure. It is now clear that many protein and enzyme molecules may undergo changes in tertiary structure or conformation, depending on binding of substrates or on their oxidation-reduction state. Such "mechanochemical" changes are seen in hemoglobin and in cytochrome c, as a function of oxygenation or oxidation-reduction state, respectively, for example, and mechanochemical or conformational changes are in fact implicit in current conceptions of the interaction of enzymes with substrates [cf. the "induced fit" theory of Koshland (1958)].

It is hoped that this essay has developed at least the bare bones of the relationship between chemical and mechanical coupling in oxidative phosphorylation and the idea that both may be reflections of the function and properties of the same set of molecules concerned in respiratory energy transduction. The third leg of this proposition,

that osmotic events are coupled to the respiratory chain, will not be developed further here. The stimulating article of Mitchell (1961b) referred to above postulates that chemical coupling and osmotic coupling of respiratory energy are in fact one and the same process, for which he uses the term "chemiosmotic" coupling. Mitchell has also recognized the need for accounting for the mechanochemical changes and he suggests that they can be accounted for by electrical and mechanical stresses across the membrane while ion separation is taking place. Mitchell's denial of the occurrence of "intermediates" in oxidative phosphorylation can perhaps be challenged as being unwarranted and even unnecessary. Possibly the word "intermediate" should be more accurately defined. However there is certainly ample basis for his view (1961b) "if the processes that we call metabolism and transport represent events in a sequence not only can metabolism be the cause of transport, but transport can be the cause of metabolism . . . transport and metabolism, as usually understood by biochemists, may be conceived advantageously as different aspects of one and the same process of vectorial metabolism."

The evidence now accumulating thus sustains, at least in outline, the conception developed here, namely that the enzyme molecules comprising the respiratory assembly, which are embedded in the mitochondrial membrane, not only bring about coupled formation of ATP, but also constitute the elements for carrying out osmotic and mechanical work. Perhaps recognition of these three levels of energy transduction will provide more rational approaches to reconstruction of the respiratory chain from its components. But just as the components of the actomyosin system can be put together in several different ways to yield various types of physical or chemical response in the presence of ATP, depending on their orientation relative to each other, so can we expect to achieve various types of reconstruction of function in reassembling the catalysts of the respiratory chain. Some of these may show electron transfer and a coupled phosphorylation of ADP, just as an artificial actomyosin thread can contract with ATP under certain conditions. But such a thread is patently a far cry from the structural organization of the myofibril and it is therefore important that experimental reconstructions of the respiratory chain and oxidative phosphorylation be scrutinized carefully, to see that they provide also the structural elements necessary for osmotic and mechanochemical coupling.

REFERENCES

- Chan, P. C., Lehninger, A. L., and Enns, T. E. (1960). *J. Biol. Chem.* **235**, 1790.
 Chance, B., and Williams, G. R. (1956). *Advances in Enzymol.* **17**, 65.
 Conway, E. J. (1953). *Intern. Rev. Cytol.* **2**, 419.
 Cooper, C. (1960). *J. Biol. Chem.* **235**, 1815.
 Davies, R. E., and Ogston, A. G. (1950). *Biochem. J.* **46**, 324.
 Engelhardt, W. A., and Ljubimowa, M. N. (1939). *Nature* **144**, 668.
 Engelhardt, W. A., Ljubimowa, M. N., and Meitina, R. A. (1941). *Compt. rend. acad. sci. U.R.S.S.* **30**, 644.
 Fowler, L. R., and Hatefi, T. (1961). *Biochem. Biophys. Research Commun.* **5**, 203.
 Frederic, J. (1958). *Arch. biol. (Liège)* **69**, 167.
 Gamble, J. L., Jr. (1957). *J. Biol. Chem.* **228**, 955.
 Gotterer, G. E., Thompson, T. E., and Lehninger, A. L. (1961). *J. Biophys. Biochem. Cytol.* **10**, 15.
 Greville, G. D., and Reich, E. (1956). *Biochim. et Biophys. Acta* **20**, 440.
 Holton, F. A. (1957). *Biochem. J.* **66**, 37P.
 Järnefelt, J. (1961). *Biochim. et Biophys. Acta* **48**, 104.
 Keilin, D., and King, T. E. (1960). *Proc. Roy. Soc. B* **152**, 163.
 Kielley, W. W. (1961). *5th Intern. Congr. Biochem., Symposium V, Moscow*, 1961, Pergamon Press, New York. In press.
 Kielley, W. W., and Bradley, L. B. (1956). *J. Biol. Chem.* **218**, 263.
 Koshland, D. E., Jr. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 98.
 Lehninger, A. L. (1955). *Harvey Lectures, Ser.* **49**, 176.
 Lehninger, A. L. (1956). In "Enzymes: Units of Biological Structure and Function" (O. H. Gaebler, ed.), p. 217. Academic Press, New York.
 Lehninger, A. L. (1959a). In "A Symposium on Molecular Biology," Chicago, 1957-1958 (R. E. Zirkle, ed.), p. 122. Univ. of Chicago Press, Chicago, Illinois.
 Lehninger, A. L. (1959b). *Rev. Modern Phys.* **31**, 136.
 Lehninger, A. L. (1959c). *J. Biol. Chem.* **234**, 2187.
 Lehninger, A. L. (1959d). *J. Biol. Chem.* **234**, 2465.
 Lehninger, A. L. (1960a). In Symposium Issue. *Federation Proc.* **19**, 952.
 Lehninger, A. L. (1960b). *Ann. N.Y. Acad. Sci.* **86**, 484.
 Lehninger, A. L. (1961a). In "Symposium on Biological Structure and Functions, Stockholm, 1960" (T. W. Goodwin, ed.), Academic Press, New York. Vol. II, p. 31.
 Lehninger, A. L. (1961b). *Biochim. et Biophys. Acta* **48**, 324.
 Lehninger, A. L. (1962). *Physiol. Rev.* In press.
 Lehninger, A. L., and Gotterer, G. S. (1960). *J. Biol. Chem.* **235**, 8P.
 Lehninger, A. L., and Neubert, D. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 1929 (1961).
 Lehninger, A. L., and Schneider, M. (1959). *J. Biophys. Biochem. Cytol.* **5**, 109.
 Lehninger, A. L., Wadkins, C. L., Cooper, C., Devlin, T. M., and Gamble, J. L., Jr. (1958). *Science* **128**, 450.
 Lehninger, A. L., Ray, B. L., and Schneider, M. (1959). *J. Biophys. Biochem. Cytol.* **5**, 97.

MECHANOCHEMICAL CHANGES IN MITOCHONDRIA 435

- Levy, H. M., Sharon, N., Lindemann, E., and Koshland, D. E., Jr. (1960). *J. Biol. Chem.* **235**, 2628.
 Lundegardh, H. (1940). *Lantbruks-Högskol. Ann.* **8**, 233.
 Mitchell, P. (1961a). In "Symposium on Biological Structure and Function, Stockholm 1960" (T. W. Goodwin, ed.), Academic Press, New York. Vol. II, p. 581.
 Mitchell, P. (1961b). *Nature* **191**, 144.
 Neubert, D., and Lehninger, A. L. (1962). *J. Biol. Chem.* **237**, 952 (1962).
 Packer, L. (1960). *J. Biol. Chem.* **235**, 242.
 Penefsky, H. S., Pullman, M. E., Datta, A., and Racker, E. (1960). *J. Biol. Chem.* **235**, 3330.
 Post, R. L., Merritt, C. R., Kinsolving, C. R., and Albright, C. D. (1960). *J. Biol. Chem.* **235**, 1796.
 Price, C. A., Fonnesu, A., and Davies, R. E. (1956). *Biochem. J.* **64**, 754.
 Pullman, M. E., Penefsky, H. S., Datta, A., and Racker, E. (1960). *J. Biol. Chem.* **235**, 3322.
 Raaflaub, J. (1953). *Helv. Physiol. et Pharmacol. Acta* **11**, 142, 157.
 Robertson, R. N. (1960). *Biol. Revs.* **35**, 231.
 Rose, T. H., Neubert, D., and Lehninger, A. L. (1962). *J. Biol. Chem.* In press.
 Skou, J. C. (1957). *Biochim. et Biophys. Acta* **23**, 394.
 Skou, J. C. (1960). *Biochim. et Biophys. Acta* **42**, 6.
 Vasington, F. V. (1961). *J. Biol. Chem.* In press.
 Wadkins, C. L., and Lehninger, A. L. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1576, 1582.

Psychophysics Considered at the Molecular and Submolecular Levels¹

FRANCIS O. SCHMITT

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts

I. Introduction	437
II. Electrophysiological Properties of Neurons and Neuronal Nets	439
III. Limitations of Electrophysiological Techniques for Psychophysical Investigations	440
IV. Suggestions for Psychophysical Research	441
A. Investigation of Presently Known but Poorly Understood Systems	442
B. The Search for Processes Yet Undiscovered	447
References	454

I. Introduction

Man is distinguished by his ability to perceive, store, recall, and utilize information through processes of learning; to manifest conscious, cognitive behavior; and, by transducing his thoughts to spoken and written symbols, to communicate his ideas—however imperfectly—to his fellow man. These faculties, in the last few centuries, have culminated in the development of science as we know it; this year we celebrate the quadricentennial of the birth of Francis Bacon, considered by many the founder of the modern scientific method. The enormous impact of the scientific revolution—perhaps better called “evolution”—particularly in pure and applied physics is apparent in all aspects of modern man’s life. The biomedical sciences have checked the ravages of disease and are now attacking aging itself. With the resulting substantial

¹ These studies were aided by a research grant (B-24) from the National Institute of Neurological Diseases and Blindness of the National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare; by a contract (NR 101-100) between the Office of Naval Research, U.S. Department of the Navy and the Massachusetts Institute of Technology; and by grants from the Trustees under the Wills of Charles A. and Marjorie King, and from the Lou and Genie Marron Foundation.

extension of life expectancy and with increased production and distribution of food, a population explosion has been set off which already poses grave ethical problems. Perhaps man will find solutions to these problems and, at the same time, take a giant step forward in his evolutionary development by gaining fundamental knowledge of the nature of his own being—regarded by many contemporary thinkers as the greatest challenge in the whole of science.

What might be the nature of such a quantum step in evolutionary development that could vastly improve man's adaptation not only to his physical and biotic environment but also to himself and to his civilization, a civilization which at its present stage of development seems poorly adapted to man's needs in the age of the atom? Perhaps an important advance might be achieved if man turned his analytical and creative abilities, exemplified in his development of modern theoretical physics, to the formulation of a scientific frame of reference even more powerful from the human viewpoint, namely that of theoretical biology, still virtually non-existent. This would demand an attack at all levels of complexity: by the holistic methods of descriptive, taxonomic, evolutionary, behavioral sciences; by the analytical, physiological, biochemical, and biophysical sciences; and probably by the invention of new mathematical methods of dealing with problems of a type never before encountered in physics.

Investigation of the physical and chemical mechanisms of the human mind may well provide the central target of such an attack, because substantial advances in this area may open up unexplored dimensions of human intercommunication which might prove a turning point in the human adventure of this planet—and perhaps elsewhere. It would still remain to be seen whether such advances would be utilized constructively or destructively from the human point of view.

The present essay² is a modest contribution to an analysis of the problems and possibilities of psychophysics³ considered at the molecular and submolecular levels.

² Portions of this paper were presented in an address on the occasion of the dedication of the Stanley Cobb Research Laboratories of Neuropsychiatry of the Massachusetts General Hospital, Boston, Massachusetts, October 14, 1960.

³ The term "psychophysics" is here used according to the definition given

II. Electrophysiological Properties of Neurons and Neuronal Nets

As a means of transmitting the coordinating digital informational signals required as organisms increased in size in early stages of evolution, neurons and neuron nets were evolved. With its rich synaptic integration both at the cell-body region and at the terminals and with its intervening axonal communication channel, the neuron is admirably adapted to the coordination needs even of the highest organism. Excitation and propagation of membrane-limited bio-electric changes ("impulses") over these neuronal nets are the elementary processes dealt with in neurophysiological investigation of central nervous system function.

Biophysical analysis of the mechanism of this self-propagating alteration of membrane properties during impulse transmission has culminated in what may be characterized as the inorganic electrochemistry of nerve conduction. These studies deal with the flow of particular inorganic ions, chiefly sodium and potassium, that seem to be the chief charge carriers in the extracellular and intracellular phases and through the membrane during excitation. Knowledge of the structure and properties of membranes in general is growing rapidly (Sjöstrand, 1959; Fernández-Morán, 1959, 1961a,b, 1962; Hodge, 1959; Robertson, 1961b; Ponder, 1961), though the structural, chemical, and enzymatic details concerned with the fast, reversible membrane change during nerve excitation remain almost completely unknown.

When the impulse traveling down the axon reaches junctional tissue, transmission of excitation across the intercellular gap (order of 10^2 Å wide) is thought to be mediated by highly potent hormonal transmitter molecules, such as acetylcholine and noradrenaline, liberated in quantal amounts into the junctional fluid. Little is known about the site of synthesis of the hormones within the neuron and even less about the mechanism of their discharge by the impulse and of the regulation of the amount available at endings in relation to frequency of firing (see Castillo and Katz, 1956; Katz, 1958).

in Webster's New International Dictionary, 2d ed., Unabridged, 1959: "The scientific study of the relations between mental and physical processes," rather than that commonly adopted by experimental psychologists, i.e., a study of "responses to physical stimuli and especially the perception of physical magnitudes."

Electrophysiology, i.e., the study of the generation, transmission, and transduction of bio-electric signals (10^{-6} to 10^{-1} volts, 10^{-4} to 10^{-1} seconds duration, 10^{-1} to 10^2 meters per second propagation velocity) in the neurons of the central and peripheral nervous system, as recorded from macro- or micro-electrodes, has developed into a highly sophisticated science indispensable for understanding the integrative and other "household" tasks, including the neuronal processing of information external and internal to the organism (Bionics Symposium, 1960; Rosenblith, 1961), and the function of localized regions of the central nervous system in normal and abnormal states (Magoun, 1960). Application of the techniques of computer science and information theory ("Mechanisation of Thought Processes," 1959) extends still farther the role of electrophysiology in neurology, neuropathology, psychology, psychiatry, and other branches of biomedical sciences. However, to make substantial progress in psychophysics it will be necessary also to penetrate below the level of the neuron, to deal with the properties of macromolecules and of complex macromolecular systems.

III. Limitations of Electrophysiological Techniques for Psychophysical Investigations

Until relatively recently, the view was widely entertained that neuronal firing in circulating or reverberating loops constitutes the basis of long-term memory traces. Complex neuronal pathways having once been traversed continue to fire, giving rise to the circulating or reverberating loops. Network firing is undoubtedly an essential factor in alertness, awareness, and short-term memory processes. But the reverberating circuit theory begs the question of the molecular mechanism of long-term memory recording, storage, and recall, for if a neuronal circuit becomes easier to fire because of previous excitations, it does so because of processes and structures at the molecular level.

The reverberating neuronal circuit theory for long-term memory phenomena may be ruled out because if animals are chilled to severe hypothermia or subjected to deep anesthesia or to convulsive electro-shock, all electrical activity recordable from the brain stops during the treatment and for variable periods thereafter. Long-term memory is neither destroyed nor reduced after recovery. According to Morrell (1962), "Even with micro-electrode methods the cellular events attendant upon memory persistently escaped detection." Clearly the permanently stored engram or memory

trace must be sought below the level of neurons and their nets as functional units, i.e., at the level of molecules or macromolecules.

Only a fraction of the brain is involved in receiving sensory information from outside, processing this information, and sending out impulses over efferent fibers to the effectors (muscles, glands, etc.) in the form of "orders" based on "decisions" reached by the central processing. Actually the number of axons carrying impulses to and from the brain is small compared with the number of cells (neurons and glia) in the brain and is very small compared with the number of synaptic connections between neurons in the brain. Much brain activity apparently eludes detection by conventional electrophysiological methods; thinking, remembering, learning, and other cognitive processes may include important components that are electrically silent to these methods. Indeed, it is possible that the potentials that are recorded from the central nervous system "may result from different types of activity, so that even their similar appearance may not indicate physiological or pharmacological relationships," according to Grundfest (1961), who suggests that "a searching re-examination must follow regarding the significance of recorded electrocortical data, which can mean a great deal in one context and be almost meaningless in another."

What is the nature of the needed re-examination? Where lies the greatest promise for advance in psychophysical research? Progress in the investigation of the memory of the race, i.e., genetic mechanisms, became rapid, indeed autocatalytic, when geneticists looked below the level of the organism or the cell to the level of macromolecules (nucleic acids and nucleoproteins), the lowest level of structure possessing the diversity required for an hereditary code yet susceptible of rigorous investigation by biochemical and biophysical methods. Similar considerations apply to the study of immunological mechanisms, the chemical "memory" of the organism. The need for a similar trend in psychophysics has been evident for some time, but little evidence is yet available. In the following section are offered illustrative examples of some avenues of approach which have suggested themselves from the writer's experience; different and highly profitable ones will doubtless occur to others.

IV. Suggestions for Psychophysical Research

If a broad interdisciplinary investigation of mental processes were undertaken on the scale and with the tempo that the subject deserves, the metabolism of ideas would doubtless be very rapid, the

half-life of hypotheses relatively short. Such a program would include not only a search for undiscovered mechanisms, structures, and substances but also an intensified re-examination of those, long recognized, whose functions remain unknown. Illustrative examples of each type may be cited.

A. INVESTIGATION OF PRESENTLY KNOWN BUT POORLY UNDERSTOOD SYSTEMS

1. *Role of Satellite Cells*

Neurons are invariably associated with satellite cells. In the central nervous system these satellite or neuroglial ("glial") cells, which are about ten times more numerous (ca. 10^{11} cells) than neurons, assume several forms subserving different functions. For example, one variety is thought to furnish a specialized means of assuring rapid diffusion of solutes between neurons and blood. Another variety coats axons of the central nervous system with many wrappings of its surface membrane to provide these axons with myelin sheaths. Myelination of the fibers in peripheral nerves (as distinguished from the central nervous system) involves a similar enfolding of a different type of satellite cell called "Schwann cells" (see Causey, 1960).

It is usually assumed that, apart from their role in development, e.g., in myelinogenesis, the satellite cells have a trophic function; they are the "soup kitchen" for the neurons. The types of necessary metabolites that might thus be provided by the satellite cells were specified in only vague terms.

In peripheral nerves, where the farthest reaches of the axon may be remote from the center of neuronal life, the cell body, it was natural to picture the Schwann cells as supplying necessary metabolites—possibly energy-giving compounds such as adenosine triphosphate. Actually little is known about the biochemical or physiological role of Schwann cells. If there is a high interchange of solutes between them and the axoplasm of the neuron, this molecular traffic must pass across the thin (ca. 100–200 Å) aqueous channels between the axon surface membrane and the limiting membrane of the Schwann cell. Since many of these metabolites are organic electrolytes, they may influence the rapid influx and efflux of inorganic ions thought to have electromotor function. It would be well to bear in mind that the satellite cells, in forming thin aqueous

junctional channels, may also provide substructural continua for biophysical phenomena presently unspecifiable but conceivably of significance in psychophysical processes.

In the central nervous system the need for an extraneuronal, cellular metabolic trophic supply line, apart from the specialized type of blood-neuron pipeline, is less obvious. From the ultrastructural and biochemical evidence, the neuron cell body is a highly active biosynthetic type of cell resembling in this respect exocrine secretory cells. Among the substances synthesized by neurons—and by glial cells—are nucleotides and ribonucleic acids (RNA). Hydén and his group have developed elegant micro methods by which the metabolic turnover of these substances, as well as proteins and other metabolites, may be determined both in neurons and in glial cells (Hydén, 1959, 1960; Hydén and Lange, 1961; Edström *et al.*, 1961; Egyházi and Hydén, 1961; Cummins and Hydén, 1961). The results indicate an intimate functional relation between neurons and glial cells which immediately surround them. Glia appear to supply neurons with energy-rich compounds particularly during increased brain activity. Nucleotides and RNA base changes showed a reciprocal relationship between neurons and glia suggesting that glia may provide bases needed for the neuron's high steady-state RNA synthesis. Such processes are of central interest in theories that regard RNA or protein (which requires specific RNA for its synthesis) as the macromolecular basis of memory codes.

When cultivated *in vitro*, glia have interesting electric and mechanical properties which may have oscillatory character (Glees, 1955; Windle, 1958; Hild, 1962). When stimulated electrically, glia appear to be capable of contraction and of electrical response (Tasaki and Chang, 1958; Chang and Hild, 1959). The physiological role of these processes is not clear.

Galambos (1961, 1962) suggests that glia may play not a secondary, accessory, trophic role in brain processes but that, possibly organized in networks with "gliaptic" junctions (analogous to synaptic junctions in neuronal nets), the glia may be repositories of the informational memory codes and may actually program the neurons when this code is read out.

Clearly the role of neuro-satellite cells in mental as well as in somatoregulatory processes deserves detailed and vigorous investigation.

2. Role of High Steady-State Metabolic Activity of Neurons

The gaseous and chemical metabolism of the neuron is among the highest of tissue cells. The ultrastructure of the neuron, with its greatly developed endoplasmic reticulum richly provided with ribosomes (comprising the Nissl substance of classic neurohistology), is characteristic of a busy biosynthetic cellular factory. This mobilization of energy and rapid biosynthesis of RNA and protein are hardly required for the production of the electric action waves by which the neurons transmit impulses; these are energy-conservative processes requiring but a few per cent of the total energy mobilized.

The need for the production of neuroplasm is obvious in developmental phases when, following the differentiation of the neuroblast into the neuron, the axon is spun out from the cell body; thus is achieved the evolutionary need for communication and integrative coordination over large distances. In the differentiation of neurons, it is essential that cell division be prevented.⁴ Chaos would result if, after the immeasurably complex interweaving of neurons in embryogenesis had produced a normal brain, neuronal cell bodies then divided to any substantial degree. Possibly these two factors—rapid synthesis of neuroplasm and suppression of mitosis—are related in both an evolutionary and a physiological sense. The differentiated cell would be capable of maintaining a steady-state production of neuroplasm, as has been demonstrated by Weiss (1962) and collaborators in the constant movement of axoplasm peripherally down the axon, and could regenerate a new axon following injury.

In cellular mechanisms as vital, for both the life of the individual organism and for evolutionary adaptation, as those of the nervous system it is probable that cellular processes may simultaneously serve more than one function. Thus the high protein and RNA biosynthetic activity of neurons may be intimately concerned not only with the generation or regeneration of the axon but also with the production of the macromolecular substance of the engram, the

⁴ Certain microorganisms deprived of vitamin B₁₂ (or deoxyribosides) suffer a similar inhibition of cell division through breakdown of DNA synthesis and similarly assume long filamentous forms (Beck *et al.*, 1962). The role of vitamin B₁₂, long known to be involved in certain neuropathies, and the roles of other factors vital to DNA synthesis in the differentiation of neuroblasts into neurons are under investigation in this laboratory.

memory tape which records, stores, and plays back experiential information originally piped in over sensory channels.

3. The Properties and Functional Role of Neurofilaments

With the light microscope fine fibrous strands have been observed in the cell bodies and axons of some kinds of fresh, unfixed neurons (Weiss and Wang, 1936) and in substantially all neurons which have been subjected to appropriate histological treatment (Hild, 1959). This ubiquitous occurrence of neurofibrils strongly suggests that these structures play a significant physiological role. But, after a century of study, nothing is yet known concerning the function of this fibrous protein. In electron micrographs of thin sections the unit of neurofibril structure, the neurofilament, is a smooth contoured structure 80–100 Å in diameter. The neurofibrils of histology represent bundles of neurofilaments; they pursue a tangential perinuclear course in the cell body and pass down the axon without demonstrable discontinuity through nodal regions and possibly the entire way to the terminal region of the axon.

The fibrous protein, isolated from the axoplasm of giant fibers of the large squid *Dosidicus gigas*, has been purified and characterized physicochemically. It had been tacitly supposed for years that the filaments are constructed of a parallel alignment of highly asymmetric protein macromolecules. However, Davison and Taylor (1960) proposed the novel suggestion that the structural unit may be a relatively small (M.W. order of 10⁴) globular molecule; linear aggregation of such molecules into strands and helical wrapping of the strands about a common axis would generate a structure similar to the protein component of tobacco mosaic virus. The proposed neurofilament structure is illustrated diagrammatically in Fig. 1. Like tobacco mosaic virus the neurofilament may possess a hollow core ca. 20–30 Å in diameter. This is not a necessary requirement of the helical strand structure; however, holes of this order of size have subsequently been observed in high resolution electron micrographs of neurofilaments in transverse sections by several investigators (Gray, 1959; Fernández-Morán, 1961a; Palay, 1962). Amino acid analysis shows it to be an acidic protein. No nucleic acid or lipid is present.

This novel ultrastructure, particularly the aqueous core, of the neurofilaments and their great length, perhaps extending to the axon terminals, invite speculations concerning the function of neuro-

filaments (Schmitt and Davison, 1961). Whether the protein is involved in any way in engram production or readout is unknown. The fact that the protein is sensitive to proteases provides a means of testing the validity of functional hypotheses, at least to the extent that the experimental neuron is of sufficient diameter to permit injection of protease preparations.

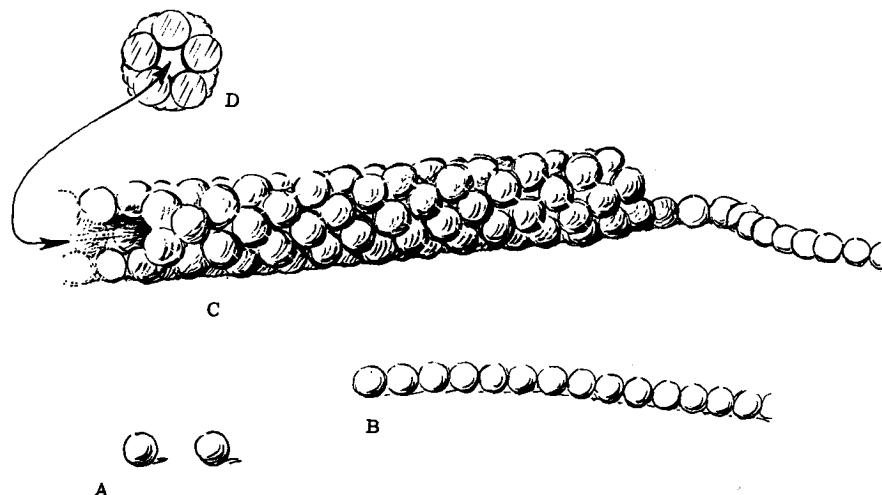


FIG. 1. Diagrammatic representation of the suggested molecular organization of the neurofilament. Magnification factor ca. 3×10^6 . A: Globular subunit, $S_{20,w} = 2-4$ Svedberg units. B: Linear aggregation of subunits into strands. C: Neurofilament showing how helical packing of subunit strands might produce a filament of uniform width (80–100 Å) with a hollow core 20–30 Å in diameter. D: End view indicating subunits and hollow central core.

This diagram is not intended to depict the number of strands comprising the neurofilament since this number is unknown.

4. Origin and Function of the Oscillatory Electrical Activity of the Brain

One of the most characteristic properties of the brain is the production during waking states of oscillatory electrical potential waves, the most prominent of which is the ca. 10-cycle α -rhythm recordable from scalp electrodes, as the electroencephalogram. Despite years of investigation described in a voluminous literature there exists little consensus among the most competent authorities concerning either the site and mode of generation of the oscillations

or their physiological function. They serve a technically useful function in the diagnosis of brain malfunction. Intimately related to processes essential for consciousness, the oscillatory discharges disappear during sleep except during dreaming with which process they also seem to be associated. It seems probable that the phenomenon may play an important physical role in mental processes, though the conjectures thus far offered from the current electrophysiological viewpoint have not been highly productive.

B. THE SEARCH FOR PROCESSES YET UNDISCOVERED

Though a rich harvest may be expected from the investigation of structures and processes already known in the central nervous system, the most exciting challenge, especially to physicists, chemists, mathematicians, and engineers interested in psychophysics is the search for concepts and phenomena not presently part of the fabric of the life sciences—or perhaps of physics—which may play a vital role in processes of consciousness, learning, memory, cognitive behavior, and perhaps aspects of man's being not yet considered amenable to study by science. Some of these, already dimly perceived or slowly taking shape, can be roughly characterized, as in the illustrative examples here presented.

1. Macromolecular Diversity and Specificity as Physical Basis for Long-Term Memory Engram

After it was realized that electrophysiological techniques were unlikely to reveal the nature of the long-term memory trace (engram), consideration was given to characteristics other than excitation and propagation of action-potential impulses but still at the rather gross neuronal level: plastic structural alterations, redistribution, disconnections, and reconnections of the terminal differentiations of axons and dendrites were among these factors. However, the last decade has established that the necessary properties, particularly the enormous diversity required, are provided only by large macromolecules polymerized from a set of individual monomer species: proteins, considered by Katz and Halstead (1950), have some 20 amino-acid monomer types; nucleic acids (RNA), considered by Hydén (1959, 1960), Morrell (1961a,b, 1962), Gerard (1960, 1961), John (1960), Corning and John (1961), and others, provide 4 mononucleotide types. If the polymer under consideration contained but a thousand residues (a relatively small polymer),

20^{1000} combinations would be possible for proteins and 4^{1000} for RNA. Though such numbers have little real significance, they illustrate the fact that giant polymers could readily account for the 10^{15} to 10^{20} bits of information thought to be processed in the course of a 70-year human life (Von Neumann, 1958).

In such hypotheses of physicochemical writing and reading of experience, three processes are distinguishable: (1) the fixation of experience (to use a term employed by Gerard, 1960, 1961); (2) the delocalization and long-term storage of this information; and (3) the recall or readout of the information in conscious, cognitive behavior. These processes, which were the subject of detailed discussion in a recent seminar lecture series (F. O. Schmitt, ed., 1962), will be briefly considered.

a. *Fixation of Experience: The Transduction of Sensory Inputs into Stable Macromolecular Form.* For purposes of simplification we may assume that this transduction is from an electrical signal, e.g., a characteristic flux of ions or other charge carriers as some function of time, to a modification of the sequence of monomer species in a large linear macromolecule. Chemical transmitter hormones may conceivably be also involved, though their role in synaptic transmission in the brain is not certain, and it is difficult to visualize a pattern of hormone concentration as one can a modulated frequency ion flux pattern resulting from the temporal and spatial summation of sensory action currents. Even if a chemical transmitter were involved, it would probably not be in the vital transduction process itself, but would merely cause the electrical firing of postsynaptic membranes; the resulting electrical disturbance would still be the transducing modality.

For present purposes it seems desirable to keep the discussion general, avoiding even such fundamental specification as to whether the transducing, electrical parameter is restricted to neuronal nets or may act transneuronally through ion flux, electrical field, potential, or other physical parameter.

It need not be specified whether the transduction occurs in neuron cell bodies or perhaps in glial somata (see Galambos, 1961, 1962) from which the engrams are transferred into the neurons for programming purposes.

For clues concerning the nature of the transduction it is useful to consider current concepts of specificity determination both in genetics, the long-term memory of the race, and in immunology

which, in the antigen-antibody relationship, deals with the long-term chemical memory of the body and provides defense against disease. The transduction of the coded information of the chromosomal DNA for the chemical operation of protein (enzyme) synthesis is thought to occur by an *instructional* process whereby a particular length of DNA is precisely though complementarily copied by an equivalent length of ribonucleic acid (RNA) which, after detachment from the DNA, carries the instructional message to the cytoplasmic sites of synthesis in the ribosomal nucleoprotein where the code is read out as a specific sequence of amino-acid residues in the synthesized protein (Brenner *et al.*, 1961; Jacob and Monod, 1961). In the immunological case the mechanism may be *selectional* whereby the protein (γ -globulin) coding macromolecule is not instructed by the antigen to assume a specific configuration unique for that antigen but rather the antigen selects, from the innumerable possible variations produced by the cell, one configuration which functions thereafter as a specific antibody for that antigen (Lederberg, 1960; Gitlin, 1962).

Conceivably if a substantial fraction of the current (ion flux) produced by an action wave in sensory neurons and their immediate extraneuronal environment passed into the neuronal or glial cells where transduction occurs, the electrical effect may be exerted upon a biosynthetically active steady-state system either: (1) in an instructional manner, through a *qualitative* alteration by changing the *sequence* of the constituent residues (nucleotides in the case of RNA or amino acids in the case of proteins) incorporated into the macromolecule from a pool of precursors at the moment of transduction; or (2) by a selectional method, through a *quantitative* alteration by changing the *length* of the macromolecule synthesized at the moment of transduction without changing the residue sequence in the synthesized fragment.

Hydén's (1959, 1960) suggested mechanism is essentially the instructional type: the ion flux generated by the incoming action waves over sensory neurons affects the ionic equilibrium in the cytoplasm and hence changes nucleotide bases incorporated at certain sites. Particular bases are thus exchanged for other bases in the surrounding pool; the RNA specification thereby altered, in turn, alters the protein biosynthesized under the influence of RNA. The protein or its complement is deposited in the postsynaptic structure where it causes the neuron to fire on arrival of a subse-

quent modulated frequency input identical with that which originally triggered the formation of the specific RNA and protein macromolecules. Any cell having in its synaptic membrane a macromolecular engram which is the code for a specific bit of sensory information would respond in readout. This scheme avoids strict localization within brain cells; individual neurons may participate as subunits of many neuronal circuits. No details of the mechanism by which such a transduction is effected were suggested by Hydén or, so far as we know, by other authors who have invoked molecular configuration as engrams for stable long-term memory traces.

It may be profitable to consider also mechanisms of the selectional type, one of the various possibilities of which may be described for illustrative purposes. Self-replication is not unique to DNA or virus RNA. It has recently been suggested (Ochoa, 1961; Reddi, 1961) that certain types of RNA, after formation upon the DNA template, may, at an unspecified locus in the cell but presumably not in the ribosomes, be capable of self-replication ("stencil RNA"). Conceivably the transducing electrical modality, at the moment it encountered the steady-state stencil RNA replicating system, might eject partially synthesized RNA fragments from the stencil RNA template. Any fragment larger than a critical liminal size might be the engram thus derived from the normal sequence of nucleotide residues by a temporal selection process.

b. *Storage of Engram. Replication with Accompanying Delocalization.* If the memory trace is a specific configurational and constitutional variant of a particular macromolecular species, the fact that such traces may persist for many years would seem to require either that the engram is extraordinarily stable to metabolic attack—a highly unlikely possibility—or that it has the capability of self-replication, hence of indefinite survival. Self-replicability would possibly illuminate another seeming mystery, namely that the engram does not remain localized to the neuronal net in which it was first laid down but becomes widely distributed in the brain (Lashley, 1950). Self-duplication is a property of DNA and viral RNA and possibly also intracellular RNA of the "stencil" type. While proteins seemingly lack this ability, they are undoubtedly involved in cellular processes of information capture, storage, and recall and may play a vital role in engram processing. The search would thus seem to be narrowed to an RNA or to an RNA-protein type of macromolecule, consistent with the evidence thus far

adduced regarding RNA synthesis and turnover as related to neuronal function (Hydén, 1959, 1960) and to the facts brought out in the elegant and highly promising experiments of Morrell (1961a,b, 1962). In these experiments a chronic epileptogenic lesion produced in a small area of the surface of one cerebral hemisphere causes the appearance in a few days of epileptiform neuronal firing in the homotopic region of the opposite hemisphere ("mirror focus"). This secondary discharge, at first dependent on the firing in the primary lesion, may become independent; neurons untouched directly have learned a new type of behavior. Histochemical studies showed that alterations in RNA accompany this learned change.

The fact that ribonuclease may under certain conditions remove or destroy learned or conditioned memories (Corning and John, 1961) is consistent with the RNA hypothesis but is in itself not definitive because of the danger of contamination of the enzyme preparation with traces of other enzymes, such as proteases, and because the system under investigation is not a simple one involving the reaction of a pure substrate with a pure enzyme but is a highly complex compartmentalized system of cells or entire organisms.

c. *Readout (Recall).* The mechanism by which macromolecular representation of memory engrams, perhaps widely distributed in the cells of the brain, may be read out in conscious (or subconscious) recall is the least understood of the hypothesized psychophysical processes. It is supposed by some (Hydén, 1959, 1960) that the presence of the engrams or their complements in synaptic membranes facilitates the firing of these neurons when again confronted with the same stimulus as originally caused the engrams to be formed. How this activation or transduction from some physico-chemical property of specific macromolecules, in an organized system of ultrastructural components, to excitation of neuronal membranes occurs remains a mystery. The vastness of this area of ignorance need not be unduly disturbing for, despite the enormous amount of research and the epochal achievements of molecular genetics in the last two decades, almost nothing is known about readout or recall in the genetic process. This enigma looms more impressive if the regulatory, modulating processes of growth and development are included in the ultimate readout of the preformed morphogenetic code. However, this factual and theoretical void detracts little from the DNA macromolecular hypothesis of genetic determination. Development of useful hypotheses of macromolecular coding

of mental processes may similarly have to await much preliminary research, but these difficulties need not delay the initiation of a concerted attack on the seemingly simpler processes of recording and storage of the engram.

2. Role of Fast Reactions

As electron microscopy and other techniques of ultrastructure research have been developed and fruitfully applied in the study of cell and tissue function, it has become apparent that the most fundamental cellular processes depend for their operation on an exquisite organization at the molecular and macromolecular levels. Enzymes are grouped in specific arrays to facilitate the processing of substrate molecules in coupled reactions, such as in the Krebs cycle or the electron transport chain in mitochondria. There apparently exist idioms of solid-state structures representing the basis for the vital biochemical reactions of energy interconversion, mobilization, and coupling (Lehninger, 1959; Green, 1960). High resolution electron microscopy now reveals these structural idioms (Fernández-Morán, 1961b, 1962); further improvement of techniques may lead to new vital molecular mechanisms.

High resolution in space is now possible: 10 Å resolution in the electron microscopy of biological materials is no longer an unusual accomplishment, and 5 Å resolution or better is probably not far off. The increase in our knowledge which this revelation will unfold will probably be far more spectacular than that which followed the first biological applications of electron microscopy. By contrast our capabilities for resolution in time are still relatively low. Intimate life processes are fast reactions. To make substantial progress especially in the psychophysical field, methods of studying individual fast reactions and their coupling mechanisms must be vastly improved. It seems probable that aspects of solid state physics, physical chemistry, and quantum chemistry will become essential ingredients in these new advances [see the recent symposia, "International Colloquium of Fast Reactions in Solutions," 1960, and Bioenergetics (Augenstein, 1960)]. Among the types of fast fundamental transfer processes important in aqueous biomolecular systems are: (1) transfer of elementary charged particles, both of electrons (producing semiconductivity in molecular aggregates) and of protons (through systems of organized hydrogen bonds, as in hydrate layers about macromolecules); and (2) energy transfer

including excitation transfer (important in photoreceptors), charge transfer interaction (donor-acceptor interaction), and "electron transport" (as in energy mobilization in mitochondria). These were dealt with in a recent lecture seminar series (F. O. Schmitt, ed., 1960, see especially Kasha, 1960).

Because such reactions are likely to be directly concerned in the various transductions that make possible cognitive behavior—probably in the fixation of sensory information as stable macromolecular configurations and in the readout of this information in memory recall—this subject seems appropriate for presentation in a volume dedicated to Albert Szent-Györgyi. More than anyone else he has proclaimed the importance of submolecular processes underlying fundamental biochemical and biophysical phenomena (Szent-Györgyi, 1957, 1960).

In a recent series of lectures (F. O. Schmitt, ed., 1962) on the role of macromolecular specificity in biological memory, Kasha (1962) examined the various types of transient molecular phenomena that might serve as a basis for models for the coding and recall of long-term memory traces. These included molecular electronic excitation, intramolecular energy transfer, excitation interaction in photoexcited macromolecular and lamellar systems, reversible photochemical processes, photoreversible photochemical processes and reversible electroluminescence. The last named, as applied perhaps in the far infrared or microwave frequencies, seemed a promising area for investigation. The importance of biologically structured microfields and stochastic memory models was considered by O. H. Schmitt (1962).

Coupled with the search for solid-state quantum-chemical and "bionic" (see "Bionics Symposium," 1960) models, should be the simultaneous investigations of the microcomponenetry of neurons and satellite cells by high resolution electron microscopy and by other biophysical and biochemical methods. If one could investigate (with millimicro-electrodes) the electrical properties of the macromolecular substructure of brain cells (and other cells!), one might discover high fields across intercellular interfaces (such as are known to occur across the axon surface membrane) and possibly highly specific transient phenomena of electrostatic or electromagnetic nature associated with such fields.

Half a century passed after biological evidence for animal electricity was obtained by Galvani before Du Bois-Reymond, a

biophysically inclined physiologist who happened to be an expert with galvanometers, provided experimental evidence concerning the physical nature of the phenomenon. It is to be hoped that unnecessary time lags will be avoided in contemporary psychophysical research through close cooperative interaction by scientists from various fields who share a common deep interest in the study of the mind.

REFERENCES

- Augenstine, L. G., ed. (1960). "Bioenergetics." *Radiation Research, Suppl.* **2**, 685 pp. Academic Press, New York.
- Beck, W. S., Hook, S., and Barnett, B. H. (1962). *Biochim. et Biophys. Acta*. In press.
- "Bionics Symposium (Living Prototypes—The Key to New Technology)" (1960). Wright Air Development Division Technical Report 60-600, December, 499 pp. U.S. Air Force, Wright-Patterson Air Force Base, Ohio.
- Brenner, S., Jacob, F., and Meselson, M. (1961). *Nature* **190**, 576-581.
- Castillo, J. del, and Katz, B. (1956). In "Progress in Biophysics" (J. A. V. Butler, ed.), Vol. 6, pp. 121-170. Pergamon Press, New York.
- Causey, G. (1960). "The Cell of Schwann," 120 pp. E. & S. Livingstone, London.
- Chang, J. J., and Hild, W. (1959). *J. Cellular Comp. Physiol.* **53**, 139-144.
- Corning, W. C., and John, E. R. (1961). *Science* **134**, 1363-1365.
- Cummins, J., and Hydén, H. (1961). Presented at the V International Congress of Biochemistry, Moscow, August 10-16.
- Davison, P. F., and Taylor, E. W. (1960). *J. Gen. Physiol.* **43**, 801-823.
- Edström, J.-E., Eichner, D., and Schor, N. (1961). In "Regional Neurochemistry" (S. S. Kety and J. Elkes, eds.), pp. 274-278. Pergamon Press, New York.
- Egyházi, E., and Hydén, H. (1961). *J. Biophys. Biochem. Cytol.* **10**, 403-410.
- Fernández-Morán, H. (1959). In "Biophysical Science—A Study Program" (J. L. Oncley *et al.*, eds.), pp. 319-330. Wiley, New York.
- Fernández-Morán, H. (1961a). In "Macromolecular Complexes" (M. V. Edds, Jr., ed.), pp. 113-158. Ronald Press, New York.
- Fernández-Morán, H. (1961b). *J. Assoc. Research Nervous Mental Diseases, Suppl.* In press.
- Fernández-Morán, H. (1962). In "Macromolecular Specificity and Biological Memory," Proceedings of Lecture Seminar Series, Spring Term 1961, in the Department of Biology, M.I.T. (F. O. Schmitt, ed.). M.I.T. Press, Cambridge, Massachusetts.
- Galambos, R. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 129-136.
- Galambos, R. (1962). In "Macromolecular Specificity and Biological Memory," Proceedings of Lecture Seminar Series, Spring Term 1961, in the Department of Biology, M.I.T. (F. O. Schmitt, ed.). M.I.T. Press, Cambridge, Massachusetts.
- Gerard, R. W. (1960). In "Handbook of Physiology," Section 1: Neurophysiology (H. W. Magoun, Section ed.), Vol. III, pp. 1919-1965. American Physiology Society, Washington, D.C.
- Gerard, R. W. (1961). In "Brain Mechanisms and Learning" (J. F. Delafresnaye, A. Fressard, R. W. Gerard, and J. Konorski, eds.), pp. 21-32. Blackwell, Oxford.
- Gitlin, D. (1962). In "Macromolecular Specificity and Biological Memory," Proceedings of Lecture Seminar Series, Spring Term 1961, in the Department of Biology, M.I.T. (F. O. Schmitt, ed.). M.I.T. Press, Cambridge, Massachusetts.
- Glees, P. (1955). "Neuroglia, Morphology and Function," 110 pp. C. C Thomas, Springfield, Illinois.
- Gray, E. G. (1959). *J. Anat.* **93**, 420-433.
- Green, D. E. (1960). *Radiation Research, Suppl.* **2**, 504-527.
- Grundfest, H. (1961). *Ann. N.Y. Acad. Sci.* **92**, 877-889.
- Hild, W. (1959). In "Handbuch der mikroskopischen Anatomie des Menschen" (W. v. Möllendorf and W. Bargmann, eds.), Vol. IV/4, pp. 1-184. Springer, Berlin.
- Hild, W. (1962). In "Macromolecular Specificity and Biological Memory," Proceedings of Lecture Seminar Series, Spring Term 1961, in the Department of Biology, M.I.T. (F. O. Schmitt, ed.). M.I.T. Press, Cambridge, Massachusetts.
- Hodge, A. J. (1959). In "Biophysical Science—A Study Program" (J. L. Oncley *et al.*, eds.), pp. 331-341. Wiley, New York.
- Hydén, H. (1959). In "Biochemistry of the Central Nervous System," Proceedings of the 4th International Congress of Biochemistry, Vol. III, pp. 64-89. Pergamon Press, London.
- Hydén, H. (1960). In "The Cell" (J. Brachet and A. E. Mirsky, eds.). Vol. IV, Part 1, pp. 215-323. Academic Press, New York.
- Hydén, H., and Lange, P. (1961). In "Regional Neurochemistry" (S. S. Kety and J. Elkes, eds.), pp. 190-199. Pergamon Press, New York.
- "International Colloquium of Fast Reactions in Solutions" (1960). *Z. Elektrochem.* **64**, No. 1.
- Jacob, F., and Monod, J. (1961). *J. Mol. Biol.* **3**, 318-356.
- John, E. R. (1960). In "Toward a Definition of Mind" (J. Scher, ed.). Free Press, Glencoe, Illinois.
- Kasha, M. (1960). In "Fast Fundamental Transfer Processes in Aqueous Biomolecular Systems," Proceedings of Lecture Seminar Series, Spring Term 1960, in the Department of Biology, M.I.T. (F. O. Schmitt, ed.). Cambridge, Massachusetts.
- Kasha, M. (1962). In "Macromolecular Specificity and Biological Memory," Proceedings of Lecture Seminar Series, Spring Term 1961, in the Department of Biology, M.I.T. (F. O. Schmitt, ed.). M.I.T. Press, Cambridge, Massachusetts.
- Katz, B. (1958). *Bull. Johns Hopkins Hosp.* **102**, 275-312.
- Katz, J. J., and Halstead, W. C. (1950). *Comp. Psychol. Monogr.* **20**, 1.
- Lashley, K. (1950). In "Physiological Mechanisms in Animal Behavior," Symposia of the Society for Experimental Biology, No. IV, pp. 454-482. Cambridge University Press, London and New York.
- Lederberg, J. (1960). *Science* **131**, 269-276.
- Lehninger, A. L. (1959). In "Biophysical Science—A Study Program" (J. L. Oncley *et al.*, eds.), pp. 136-146. Wiley, New York.

- Magoun, H. W., ed. (1960). "Handbook of Physiology," Section 1: Neurophysiology, Vol. III. American Physiological Society, Washington, D.C.
- "Mechanisation of Thought Processes" (1959). National Physical Laboratory Symposium No. 10, Vols. I and II, 980 pp. Her Majesty's Stationery Office, London.
- Morrell, F. (1961a). *Physiol. Revs.* **41**, 443-494.
- Morrell, F. (1961b). In "Brain Mechanisms and Learning" (J. F. Delafresnaye, A. Fessard, R. W. Gerard, and J. Konorski, eds.), pp. 375-389. Blackwell, Oxford.
- Morrell, F. (1962). In "Macromolecular Specificity and Biological Memory," Proceedings of Lecture Seminar Series, Spring Term 1961, in the Department of Biology, M.I.T. (F. O. Schmitt, ed.). M.I.T. Press, Cambridge, Massachusetts.
- Ochoa, S. (1961). Paper presented at Symposium on a Multidisciplinary Research Program in a Mental Hospital. McLean Hospital, Belmont, Massachusetts, May 15.
- Palay, S. L. (1962). In "Macromolecular Specificity and Biological Memory," Proceedings of Lecture Seminar Series, Spring Term 1961, in the Department of Biology, M.I.T. (F. O. Schmitt, ed.). M.I.T. Press, Cambridge, Massachusetts.
- Ponder, E. (1961). In "The Cell" (J. Brachet and A. E. Mirsky, eds.), Vol. II, pp. 1-84. Academic Press, New York.
- Reddi, K. K. (1961). *Science* **133**, 1367.
- Robertson, J. D. (1960a). In "Molecular Biology" (D. Nachmansohn, ed.), pp. 87-151. Academic Press, New York.
- Robertson, J. D. (1960b). In "Progress in Biophysics" (J. A. V. Butler and B. Katz, eds.), Vol. 10, pp. 343-418. Pergamon Press, New York.
- Robertson, J. D. (1961a). In "Regional Neurochemistry" (S. S. Kety and J. Elkes, ed.), pp. 497-534. Pergamon Press, New York.
- Robertson, J. D. (1961b). In "Electron Microscopy in Anatomy" (J. D. Boyd, F. R. Johnson, and J. D. Lever, eds.), pp. 74-99. Edward Arnold, London.
- Rosenblith, W. A., ed. (1961). "Sensory Communication," 844 pp. The M.I.T. Press, Cambridge, Massachusetts, and Wiley, New York.
- Schmitt, F. O., ed. (1960). "Fast Fundamental Transfer Processes in Aqueous Biomolecular Systems," Proceedings of Lecture Seminar Series, Spring Term 1960, in the Department of Biology, M.I.T., Cambridge, Massachusetts.
- Schmitt, F. O., and Davison, P. F. (1961). *Actualités Neurophysiologiques* (A.-M. Monnier, ed.), pp. 355-369. Masson, Paris.
- Schmitt, F. O., ed. (1962). "Macromolecular Specificity and Biological Memory," Proceedings of Lecture Seminar Series, Spring Term 1961, in the Department of Biology, M.I.T. M.I.T. Press, Cambridge, Massachusetts.
- Schmitt, O. H. (1962). In "Macromolecular Specificity and Biological Memory," Proceedings of Lecture Seminar Series, Spring Term 1961, in the Department of Biology, M.I.T. M.I.T. Press, Cambridge, Massachusetts.
- Sjöstrand, F. S. (1959). In "Biophysical Science—A Study Program" (J. L. Oncley *et al.*, eds.), pp. 310-318. Wiley, New York.
- Szent-Györgyi, A. (1957). "Bioenergetics," 143 pp. Academic Press, New York.

- Szent-Györgyi, A. (1960). "Introduction to a Submolecular Biology," 135 pp. Academic Press, New York.
- Tasaki, I., and Chang, J. J. (1958). *Science* **128**, 1209-1210.
- Von Neumann, J. (1958). "The Computer and the Brain," 85 pp. Yale Univ. Press, New Haven, Connecticut.
- Weiss, P. (1961). In "Regional Neurochemistry" (S. S. Kety and J. Elkes, eds.), pp. 220-242. Pergamon Press, New York.
- Weiss, P., and Wang, H. (1936). *Anat. Record* **67**, 105-117.
- Windle, W. F., ed. (1958). "Biology of Neuroglia," 340 pp. C. C Thomas, Springfield, Illinois.

The Complex Copper of Nature¹

EARL FRIEDEN

*Department of Chemistry,
Florida State University, Tallahassee, Florida*

I. Introduction	461
II. Biological Importance of Copper	462
A. The Copper Enzymes and Proteins	462
B. The Biological Role of Ceruloplasmin	464
C. <i>In Vivo</i> Significance of Copper	467
III. How Copper Exists in Nature	468
A. The Copper Enzymes	472
1. Attachment of the Metal Ion to the Protein	472
2. Specificity of the Metal Ions	472
3. Kinetic Role of the Metal Ion	473
B. The Copper Proteins	476
C. General Properties of the Copper Proteins and Enzymes	476
1. Molecular Weight and Copper Content	476
2. Optical Characteristics	477
3. Oxidation State of Copper Ion in Copper Proteins	478
4. Inferences from the Effect of Ligands on Copper Enzymes	479
5. The Interaction of Copper Ions with Proteins	483
6. Active Site of the Copper Proteins	486
IV. Copper Ion in Model Catalytic Systems	489
References	495

I. Introduction

It is appropriate that a volume dedicated to the recognition of Albert Szent-Györgyi's research contributions include a chapter devoted to the complex role of copper in nature. In 1928 Szent-

¹ Supported in part by grant, G-14015, from the National Science Foundation and by a contract with the Division of Biology and Medicine, United States Atomic Energy Commission. For the unpublished work originating from the author's laboratory, the contributions of Charles Walter, Yashwant Karkhanis, Leonard B. Spiegel, and Mrs. Anne Wahlborg are gratefully acknowledged.

Györgyi reported the isolation of vitamin C, ascorbic acid.² In 1930, Szent-Györgyi discovered an enzyme in cabbage leaves which catalyzed the oxidation of ascorbic acid to dehydroascorbic acid. This enzyme, ascorbic acid oxidase, was later shown to contain copper as its prosthetic group. In the early literature a debate arose about the enzymatic nature of this reaction. The effectiveness of cupric ion itself at pH's 5–9, in catalyzing the oxidation of ascorbate as well as other substrates, represents one of the foremost examples of a simple enzyme model. The so-called "autocatalysis" of ascorbate is so commonly encountered that only many years of careful research by Dawson and his co-workers (1945, 1950) finally established the unequivocal existence of a plant ascorbic acid oxidase—a homogeneous copper protein with catalytic properties distinct from those of copper itself.

The work of Szent-Györgyi involves biologically active copper at another point. Much of Szent-Györgyi's early work was concerned with oxidative metabolism and cycles. It is now thought likely that the principle terminal oxidases of plants and animals contain copper ion as an essential functional component. After prolonged controversy it has been agreed that cytochrome oxidase contains copper ion, but the precise role of copper ion remains unsettled (Griffiths and Wharton, 1961). A possible exception among the oxidases are the flavoproteins, but their importance in terminal oxidation is dubious. Thus copper enzymes are intimately related to the biological oxidations that have captured so much of the attention of Szent-Györgyi and his associates.

II. Biological Importance of Copper

A. THE COPPER ENZYMES AND PROTEINS

The most salient function of copper in either plants or animals is its apparent essential role in the principal terminal oxidases, particularly cytochrome oxidase, ascorbic acid oxidase, and tyrosinase.

² Szent-Györgyi enjoyed the exercise of his privilege of naming the new unknown substance which later proved to be vitamin C. Recognizing that the new compound was a monosaccharide whose structure he did not know, he submitted the name "ignose" to the British journal, *Nature*. When pressed by the editors for a less facetious name, he offered as a revision, "God-knons." Later he gave it a transitional name of "hexuronic acid" with a subsequent renaming to ascorbic acid to identify its physiological function.

Indeed the remarkable sensitivity of either plants or animals to cyanide ion poisoning probably arises from inhibition due to the effect of cyanide ion on the copper ion moiety of the enzyme. However, for cytochrome oxidase, the heme of the enzyme may also be involved. The intimate connection between copper and cytochrome oxidase is now well established in both *in vivo* and *in vitro* experiments. Decreases up to eightfold in cytochrome oxidase activity in copper-deficient animals have been reported by Gubler and associates. It has not been established whether this is due to a direct or indirect effect on the synthesis of this important enzyme. Baker and Nelson found that over 85% of the oxygen uptake of respiring potato slices is catalyzed by tyrosinase. In barley embryos, as the plant matures, the burden of terminal oxidation shifts from cytochrome oxidase to ascorbic acid oxidase, according to James (1953).

It is at the terminal oxidase stage that copper enzymes may be intimately involved in the susceptibility of tissues to radiation damage. According to a theory of J. Schubert of Argonne National Laboratory, radiation produces organic peroxides which in turn oxidize Cu(I)-proteins to Cu(II)-proteins such as cytochrome oxidase. The Cu(II)-proteins become "fixed" in this oxidation state and can no longer react with molecular oxygen thus impairing respiratory metabolism. The evidence for these postulates is limited at present. Schubert reports that stabilization of Cu(I), direct peroxide destruction, and reduction of tissue oxygen can protect against all types of ionizing radiations even when given after exposure. Conversely, Cu(II) stabilizers produce increased sensitivity to radiation.

Certain important specialized functions involve copper proteins in lower forms. Hemocyanin serves as an extracellular oxygen carrying protein in certain animal groups, particularly crustacea and mollusks. In this respect these copper proteins are replacing iron-containing proteins. In animals, too, the formation of the skin pigments, such as melanin, involve tyrosinase action. The skin pigments not only serve to protect the animal against excessive light absorption, but also play a role in protective color adaptation. In insects melanins become so dense that sclerotization results, contributing an added mechanical protection for the interior of the insect against dehydration. But perhaps more important for certain plants and lower animals is the possible antibiotic action of the quinone oxidation products produced when the skin is broken. On injury, tyro-

sinase activity is initiated resulting in rapid darkening of the immediate injured area. These chromogenic oxidation products aid in preventing bacteria and parasites from growing in the exposed region. Pigmentation in melanoma tumors has been correlated with the presence of tyrosinase enzymes. No tyrosinase is found in unpigmented tumors related to melanoma. Tyrosinase has been found in the germs of red and black chicken feathers, but the enzyme was missing from the feather germs of white leghorns. The presence of tyrosinase has also been associated with coloring in guinea pigs.

The work of Kun and Fanshier (1959) established the role of copper in transsulfuration in a manner related to transoxygenation. They have isolated a copper enzyme, β -mercaptopyruvate transsulfurase, from rat liver which transfers the sulfur from the substrate to an appropriate acceptor, such as sulfite ion. This represents an important extension of the metabolic role of the copper enzymes.

Recently Katoh and Takamiya have isolated a new copper protein, plastocyanin, from the chloroplasts of many plants which may participate in photosynthesis. This copper protein accounts for about one half of all the copper in the leaves of certain plants and is more abundant than cytochrome c. The chlorophyll to plastocyanin ratio is estimated as less than 300:1, whereas the chlorophyll to cytochrome c ratio is about 400:1. Since plastocyanin appears to be absent from the purple bacteria, Katoh *et al.* (1961) suggest that it may act in the oxygen-evolving mechanism which is lacking in the photosynthetic system of the purple bacteria. Plastocyanin can also stimulate and participate in photoreduction, serving as a natural Hill oxidant. Thus another possible role is its participation as an oxidation-reduction link in electron transport in plant metabolism.

B. THE BIOLOGICAL ROLE OF CERULOPLASMIN

The function of ceruloplasmin is only beginning to be appreciated (Laurell, 1960). Though it has weak catalytic properties, these need not be related necessarily to its most significant biological activity. Early hopes to show a correlation between serum ceruloplasmin and schizophrenia have not been realized. However, it has been suggested that an increase in ceruloplasmin may occur as a "biochemical protective" measure in removing certain biologically active aromatic amines through its action as an enzyme.

Since ceruloplasmin accounts for over 90% of the total serum

copper ion, it was first assumed that this serum protein was intimately connected with the transfer of copper. However, ceruloplasmin does not seem to have a role in the transport of copper ion from the gut to the liver, since copper ion does not appear to become associated with ceruloplasmin until after transport to the liver. Ceruloplasmin could be a convenient storage form of copper ion, providing a natural copper ion-buffering system. Present evidence suggests that, for the human, Cu(II) is excreted through the feces rather than the urine. The biochemical requirements for copper appear to be relatively constant. Because of the high toxicity of copper, a rapid biochemical regulatory mechanism may be required to dispose of ingested copper within several hours after it has entered the blood stream. Ceruloplasmin may provide a mechanism in which a toxic level of free copper ion in the liver or the serum can be avoided. In terms of the stability of many oxidizable biological constituents, such as ascorbate, polyphenols, and psychogenic amines, free Cu(II) is more catalytically active at pH 7.4 than is the copper ion of ceruloplasmin. Ceruloplasmin therefore may be regarded as an essentially innocuous form of copper ion. It is also possible that ceruloplasmin serves to maintain copper ion in an effectively soluble form at nonacidic pH's.

A remarkable number of potential functions of ceruloplasmin have been suggested by M. Shimizu and co-workers as a result of preliminary clinical studies in Japan. It is postulated that ceruloplasmin is a controlling factor in the hematopoietic system, accelerating absorption and storage of iron resulting in a marked increase of iron in the liver. Ceruloplasmin increases the number of young cells of the erythropoietic and the granulopoietic systems, eventually leading to a hyperplastic condition of the bone marrow. In anemia, the ratio of ceruloplasmin copper to total serum copper is reduced but a normal ratio can be restored by the administration of ceruloplasmin in small quantities. Ceruloplasmin also increases liver respiration and liver catalase activity of rabbits. It is claimed also that improvement in certain anemias can be produced by ceruloplasmin. An increase in ceruloplasmin occurs simultaneously with anemia in rabbits caused by bleeding, phenylhydrazine hemolysis, and hemolysis due to fibrinolytic enzymes. Reduced ceruloplasmin is said to promote increased cellular division of already existent erythrocytic precursors. The usual oxidized ceruloplasmin accelerates the growth of the young cells of the granulopoietic sys-

tem. Ceruloplasmin in human serum is elevated in infections, pregnancy, anemia, and malignant tumors. It is depressed in Wilson's disease and chronic nephritis. It is emphasized that these findings and proposals of Shimizu *et al.* remain to be confirmed.

Recently F. C. Brown and J. B. White observed that fortified heart muscle particles slowly reduced ceruloplasmin under anaerobic conditions. Under aerobic conditions, oxidized ceruloplasmin inhibited the electron transport system. Though there is no evidence for ceruloplasmin in these tissues, a possible role of ceruloplasmin in regulating oxidation is indicated.

Ceruloplasmin may also be a long sought mammalian ascorbic acid oxidase. The evidence for the presence of such an enzyme in plants, fungi, and bacteria is unequivocal, but prior to the report of Holmberg and Laurell (see Laurell, 1960), the existence of a mammalian enzyme of this type was not proved. Recently Scheinberg and Sternlieb (1960) have contended that the ascorbic acid oxidase activity of ceruloplasmin was due to Cu(II) impurities in the ceruloplasmin preparations. Using appropriate precautions to exclude Cu(II), Walter and Frieden have re-examined this question and have noted important quantitative and qualitative differences between the catalysis of this oxidation by ceruloplasmin and Cu(II) at pH 5.2. These data have convinced us that ceruloplasmin possesses authentic ascorbic acid oxidase activity. Our findings are summarized below.

1. The kinetic parameters of the Cu(II) and ceruloplasmin reactions at equivalent copper concentrations differ by almost several orders of magnitude. The ceruloplasmin-catalyzed oxidations of ascorbate show typical Michaelis-Menten kinetics with a K_m of $6.6 \times 10^{-6} M$ and a velocity constant of 3.2 per atom of copper. The Cu(II) reaction is usually measured in the first-order range when ascorbate is less than $10^{-4} M$ with a calculated K_m of $7.0 \times 10^{-8} M$ and a velocity constant of 150 per Cu(II). Illustrative data are shown in Fig. 1.

2. Ceruloplasmin copper ion is reversibly reduced and decolorized by ascorbate.

3. Neocuproine does not affect either the ascorbic acid oxidase or *p*-phenylenediamine oxidase activity of ceruloplasmin. Neocuproine completely inhibits the Cu(II)-catalyzed oxidation of ascorbic acid at pH 6.8.

4. EDTA inhibits all of the ascorbic acid oxidase activity and

about 60% of the *p*-phenylenediamine oxidase activity of ceruloplasmin. However, all the original oxidase activity can be restored toward both these substrates when the EDTA-ceruloplasmin complex is dialyzed. These experiments indicate that these catalytic

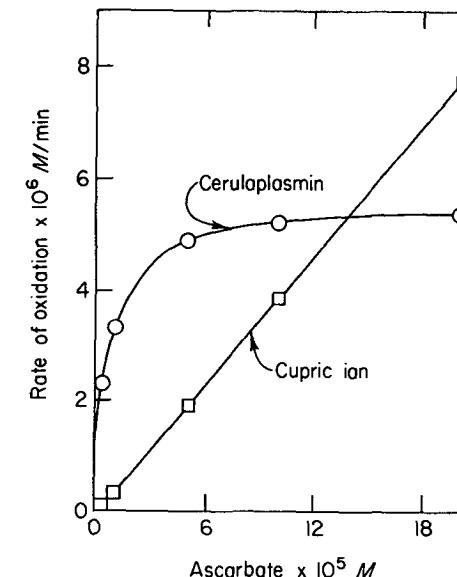


FIG. 1. Kinetic difference between ceruloplasmin and Cu(II)-catalyzed oxidation of ascorbate at pH 5.2 and 0.20 M acetate. The preparations used were equivalent in Cu(II) concentrations at $1.76 \times 10^{-6} M$. Data from C. Walter and E. Frieden, in press.

activities are associated with a nondialyzable portion of ceruloplasmin preparations.

C. In Vivo SIGNIFICANCE OF COPPER

The significance of copper in other metabolic functions is only beginning to be appreciated. Classically, as for most trace substances, deficiency and toxicity symptoms have served as the guide to deductions concerning the possible role of copper. Several surveys of this topic have appeared including the interesting recent review by Scheinberg and Sternlieb (1960) emphasizing copper metabolism in the higher mammal. Therefore, only a brief synopsis of this information will be presented.

Copper deficiency in animals is probably rarer than copper toxicity. Yet, as might be expected, copper deficiency has been a more useful tool in yielding clues as to the role of copper. In general, the essentiality of copper has been clearly shown for a variety of animals. The principal metabolic disturbances arising as a result of copper deficiency are as follows:

1. Reduction in the absorption and utilization of iron leading to hypochromic and microcytic anemia and a deficiency of heme enzymes including cytochrome oxidase.
2. A defect in phospholipid biosynthesis resulting in demyelination and "swayback" disease in lambs.
3. Impairment of osteoblastic activity leading in some animals to skeletal changes similar to those reported in scurvy.
4. Abnormalities in keratin and pigment formation resulting in wool defects in lambs and achromotrichia in rats.

Copper toxicity leads to a complex syndrome in the experimental animal including anemia, liver disease, hereditary hepatolenticular degeneration (Wilson's disease) in humans, and more typical poisoning symptoms. Its rarity in humans is due to a careful control of absorption and excretion of excess copper ion. It has also been suggested that a contribution to the aging process may arise from the accumulation in tissues of certain heavy metal ions, particularly copper and iron. Cu(II) will inhibit many enzyme systems at low concentrations including several copper enzymes.

Lipner has found evidence for the endocrine control of the copper content of certain rat tissues. The total copper level of heart, kidney, and liver of Wistar rats was inversely related to the thyroid state of the animal. To account for these observations, it was suggested that in the hyperthyroid state there is reduction of inhibitory copper, resulting in an enhancement of oxidative metabolism, perhaps via the cytochrome oxidase system. Suzuki has reported earlier that thyroxine activates cytochrome oxidase preparations. Earlier, we explained the activation of ascorbic acid oxidase by thyroxine and analogs by removal of traces of inhibitory Cu(II) or through changes involving sulphydryl groups (Frieden and Maggiolo, 1957).

III. How Copper Exists in Nature

Copper ion in plants and animals appears to exist in the form of a series of copper proteins. Little free copper ion is found even

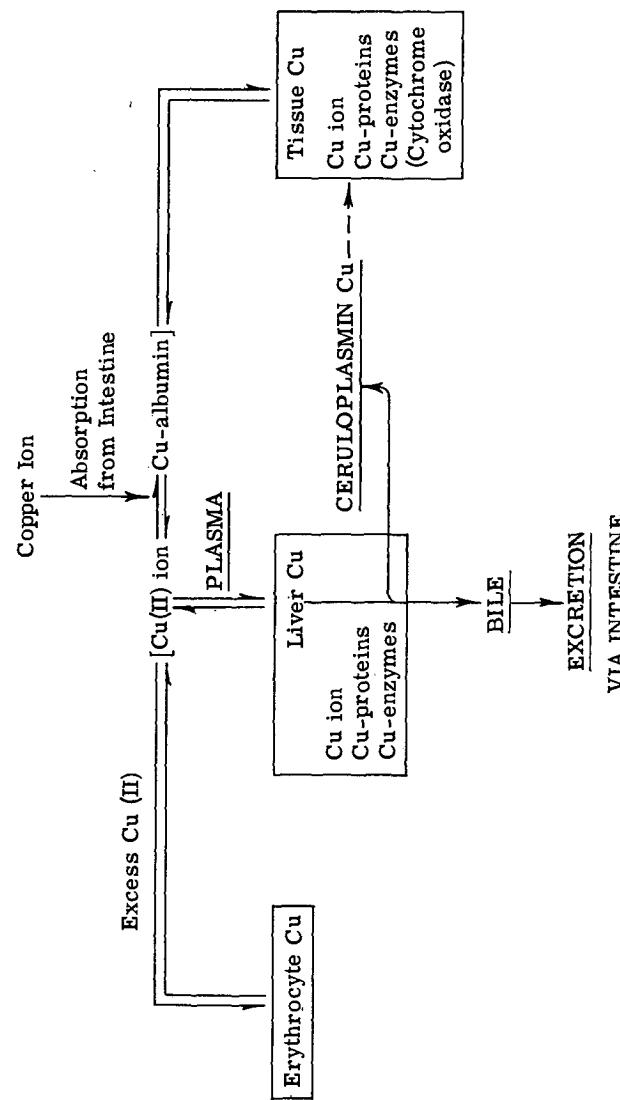


Fig. 2. Possible routes of copper metabolism in mammals. The dotted lines indicate pathways for which little or no evidence is available.

TABLE I
SOME PROPERTIES OF THE COPPER ENZYMES

Enzyme	Best source	% Cu	10^3 gm	Est. mol. weight	Cu/mole	Enzymatic activity
1. Tyrosinase (plant)	Mushrooms	0.20	100	3-4	Phenol and polyphenol-oxidation	
2. Tyrosinase (mammalian)	Melanoma	—	—	—	Tyrosine, dopa → melanin	
3. Insect phenoloxidase	Blowfly	—	500	—	Oxidation of dopamines	
4. Laccase	Lacquer tree	0.22	120	4	Oxidation of aromatic amines, phenols	
5. Ascorbic acid oxidase	Squash	0.25	150	6	Ascorbate → dehydroascorbate	
6. Cytochrome oxidase	Heart, liver	0.09	70	1	Reduced cytochrome c oxidation	
7. Uricase	Liver	0.007-0.067	120	1	Uricate → allantoin	
8. β -Mercaptopyruvate transsulfurase	Liver	0.17	35	1	β -Mercaptopyruvate → pyruvate	
9. <i>Pseudomonas</i> blue protein	<i>Pseudomonas</i>	0.35	17	1	Electron transport intermediate	
10. Plastocyanin	<i>Chlorella</i> , green leaves	—	—	—	Photosynthesis	
11. Ceruloplasmin	Serum	0.34	150	8	p-Phenylenediamine oxidation, ascorbate oxidation	

in the circulatory fluids such as blood. In the adult male human, for example, an injection of radioactive Cu⁶⁴ first appears in the serum, bound to the serum albumin, rapidly enters the liver and then subsequently emerges in the form of a copper protein, ceruloplasmin. Undoubtedly a finite small amount exists as copper ion, probably Cu(II). Saltman and co-workers have found that the uptake of copper ion by rat liver slices can be explained by assuming the presence of specific copper ion binding sites. A summary of the metabolism of copper ion in the human is pictured in more detail in Fig. 2.

There are many facets of copper metabolism which are not known which may reveal new types of naturally occurring copper compounds. For example, we do not know the form of copper in the bile nor how copper ion is absorbed in the gut. Low molecular weight copper complexes are probably involved at both these sites. Certainly the existence of other copper compounds cannot be excluded. Cu(II) interacts strongly with biological constituents other than proteins, e.g., nucleic acids and their components, certain amino acids, particularly thyroxine, cysteine, histidine, and certain carbohydrates such as ascorbic acid (see Frieden and Alles, 1958). Recently Wacker and Vallee have isolated from liver Cu(II) in combination with ribonucleic acid. Otherwise, only copper proteins have been reported and these in relatively low concentrations. The copper content of human tissues varies from a high of 5 mg/100 gm dry liver to 0.7 mg/100 gm dry bone. In the body fluids the copper content varies from 100 µg/100 ml blood to 10 µg/100 ml spinal fluid. Thus it is clearly possible that other forms of copper ion may exist in biological systems.

What is a copper protein? It is a conjugated protein in which copper ion is firmly bound to the protein as a prosthetic group. The copper ion cannot be removed by exhaustive dialysis at neutral pH's. The copper ion remains associated with the protein during the isolation process and is not contributed by contaminating copper ion in the reagents or water used during extraction and purification. Tables I and II contain a summary of the principal isolated copper proteins to date. It is striking that this list of copper proteins has more than doubled since this subject was reviewed by Dawson *et al.* (1950) and others in the "Symposium on Copper Metabolism" in 1950, the first of the well-known McCollum-Pratt Symposia.

A. THE COPPER ENZYMES

For purposes of this discussion, the copper proteins have been segregated into those copper proteins which have been demonstrated to have catalytic activity and those which, thus far, do not show any enzymatic activity. Thus a copper enzyme is a copper protein for which indisputable catalytic activity has been demonstrated. The copper enzymes comprise an appreciable segment of a group of metalloenzymes whose significance and numbers are constantly growing. This group of enzymes may be sharply distinguished from the metal ion-activated enzymes on three principal points:

1. Attachment of the Metal Ion to the Protein

The principal distinguishing characteristic between metalloenzymes and metal ion-activated enzymes is the firmness of the attachment of the metal ion to the protein. For the metalloenzymes, there is usually no detectable dissociation of the metal ion, the dissociation constant of the metal ion-protein complex being certainly less than $10^{-10} M$. The metal ion is nondialyzable. During purification, the ratio of metal ion to the protein steadily increases, finally reaching a constant value corresponding to at least one atom of metal ion per minimal molecular weight of the protein. In contrast, the dissociation of the metal ion of the metal ion-activated enzymes is easily detected and measurable. During purification the metal ion tends to be lost, resulting in a decrease in the metal ion to protein ratio. The metal ion also may be lost during prolonged dialysis.

2. Specificity of the Metal Ions

The specificity with respect to the metal ions for the metalloenzymes is ordinarily very high. The metal ions usually involved are Zn, Cu, Fe, and Mo, with suggested possibilities for Mn and Co. Usually, related metal ions will not substitute for each other for the activity of the metalloenzymes. The metal ion-activated enzymes do not show high specificity, and similar ions readily substitute for one another. Virtually all the metal ions of the periodic table have been shown to activate some enzyme system, including such exotic metal ions as Nd, La, or Sm ions. However, fluidity in our thinking about these criteria must be maintained. Recently Coleman and Vallee prepared apocarboxypeptidase and found that the Zn ion of this

naturally occurring metalloenzyme could be replaced by Co, Ni, and to a lesser extent Mn ions, for the restoration of peptidase activity. For the esterase activity of this enzyme, Hg, Cd, and Pb ions could also substitute. Of the eight transition metal ions tested, only Cu ion did not restore the activity of either catalytic activity of apocarboxypeptidase. This observation could reflect the inability of Cu(II) to form appropriate catalytic species but it could also indicate an alternative reaction in which Cu(II) oxidized certain key groups in apocarboxypeptidase resulting in an inactive form of the apoenzyme. Such interactions of proteins with Cu(II) have been reported by our laboratory as well as several others (Frieden, 1958).

3. Kinetic Role of the Metal Ion

A crucial test for the role of the metal ion in either group of enzymes is the partial or complete dependence of enzymatic activity on the presence of the metal ion. For the metalloenzymes, there is no activity in the complete absence of metal ion. Upon the addition of the metal ion, other parameters being constant, there is a direct proportionality between added metal ion and enzyme activity. The preparation of the apoenzyme of the metalloenzyme may prove to be difficult because of the firmness of the bonds between the metal ion and the protein. Metal ion-activated enzymes are easy to prepare free of their metal ions, but frequently show some, albeit low, activity. When the appropriate metal ion is added, modified Michaelis-Menten kinetics are obtained.

All the copper enzymes listed in Table I are metalloenzymes. To date there are no verified examples of a copper ion-dependent or -activated enzyme. This may be due to the especially high affinity of copper ion for proteins. Of all the metal ions found in biological systems, Cu(II) probably binds proteins most firmly. For a protein to be guaranteed the use of copper ion for its catalytic machinery, copper ion had to be incorporated firmly into its helices. This assures the retention of copper ion by the copper enzymes despite the presence of competing proteins. One possible exception is the nitrite reductase of *Neurospora*, crude preparations of which may increase in activity when Cu(II) is added. Many enzymes lose activity when Cu(II) is added, perhaps because of the oxidation of sulphydryl groups.

In his review in 1950 on the role of metal ions in enzyme systems,

Lehnninger (1950) postulated three main possible functions for metal ions in enzyme systems, which are not necessarily mutually exclusive. The high affinity for copper ion by proteins limits the possible role of copper ion in the catalytic function of the copper enzymes. Copper ion could not be expected to exert its enzyme control by antagonizing the activating effect of some other metal ion on a particular enzyme system. But, as with ceruloplasmin, copper ion may be useful as a binding group, bringing the enzyme and substrate together in proper juxtaposition. Finally, it is logical to assume that copper ion must be a part of the actual catalytic center of the enzyme, particularly in view of the demonstrated changes in its oxidation state during catalytic activity.

The listing of the copper enzymes as in Table I involves certain arbitrary definitions as to what constitutes a different type of enzyme or enzyme action. For example, although only three types are listed, there probably are more than three types of phenoloxidases, e.g., *Neurospora* tyrosinase, another insect tyrosinase, etc. It is not yet certain that even the comparatively well-characterized mushroom enzyme is not a combination of two enzymes with monophenolase and polyphenoloxidase activity, respectively. The specificity of the enzymes known as laccases, is relatively ill defined. Singer and Kearney (1954) define the laccases as essentially phenoloxidases, differing in specificity from tyrosinase in that they have no activity on monophenols, but oxidize *p*-phenylenediamine and hydroquinone as well as *o*-dihydricphenols. While ceruloplasmin does not fit this latter definition precisely, it resembles a laccase more than a tyrosinase. Ceruloplasmin also catalyzes the oxidation of ascorbic acid, so complicating its classification that it has been listed separately.

Omitted from this list are two enzymes earlier reported to contain copper as an integral part of the protein. Both butyryl coenzyme A dehydrogenase and δ -aminolevulinic acid dehydrase have been shown not to require copper for their enzymatic activity. Presumably the earlier reports of their copper content were due to contamination of the protein from the source or during isolation. Additional confirmation is also needed to verify that uricase is a copper protein. Mahler (1961) *et al.* have reported that neither dialysis against cyanide ion nor exposure to acidic pH's removes the relatively small amount of copper ion from uricase. These two procedures almost invariably abstract copper from all authentic copper proteins, except cytochrome oxidase.

TABLE II
SOME PROPERTIES OF THE COPPER PROTEINS^a

Protein	Best source	% Cu	Est. mol. weight 10^3 gm	Cu/mole	Possible function
11. Ceruloplasmin	Serum	0.34	150	8	Many possible functions; see text
12. Hemocyanin	Lobster plasma Snail plasma	0.16 0.19	780 6,700	20 200	Oxygen transport
13. Cerebrocuprein	Brain	0.30	35	2	Possibly copper storage, etc.
14. Erythrocuprein ^b	Human red blood cells	0.35	34	2	Possibly copper storage, etc.
15. Hemoeprein ^b	Beef red blood cells	0.35	30	2	Possibly copper storage, etc.
16. Hepatocuprein	Horse liver, ox liver	0.34	35	2	Possibly copper storage, etc.
17. Milk copper protein	Milk	0.19	—	—	Possibly copper storage, etc.
18. <i>Rhus vernicifera</i> blue protein	Japanese lacquers	0.33	25	1	Unknown

^a There have been additional reports of less well-characterized copper proteins from various sources, including Brewer's yeast (0.12% Cu), vaccinia virus (0.05% Cu).

^b A clear distinction between these two proteins is yet to be confirmed.

B. THE COPPER PROTEINS

For the list of copper proteins in Table II, it has not been possible to describe accurately the biological role of these proteins. Although it is not required that a protein have a specific function, this is frequently the case. Likewise, it is not required that the biological activity of a copper protein be associated with its copper content, but this is a logical hypothesis upon which to design future experiments. Thus we suspect a copper protein in a given tissue as being involved in copper uptake, transport, or release, or serving as a copper buffer or in the catalytic machinery of that particular tissue. Ceruloplasmin may have a role in copper transfer, but for most of the other proteins it can only presently be suggested that they are involved in copper ion storage. Any globular protein in nature is suspect for catalysis until all possible substrates and optimum conditions have been exhausted. The optimum conditions may be particularly important for the copper proteins. For example, to show the relatively weak catalytic activity of ceruloplasmin, careful selection of the proper ionic strength and pH must be exercised. It is suggested that when a more complete study of other copper proteins is made, catalytic activity may appear when optimum environmental conditions are found. It might be speculated that it would be physiologically convenient to have available a series of oxidative enzymes such as the copper proteins which would appropriately serve as enzymes only when certain pH's or salt concentrations are reached. The alkaline and acid phosphatases already illustrate this point as far as pH is concerned.

C. GENERAL PROPERTIES OF THE COPPER PROTEINS AND ENZYMES

1. Molecular Weight and Copper Content

Except for the hemocyanins, all the copper proteins have a molecular weight of 1.5×10^5 or less. Several of the copper proteins share a common protein to copper subunit of about 17,000 gm of protein per copper atom. This is reminiscent of heme proteins such as hemoglobin which have a similar protein to heme ratio. By analogy, the copper proteins might also be expected to be composed of dissociable subunits corresponding to the α , β , or other hemoglobin chains, but subunits of a much larger size have been reported

only for the high molecular weight hemocyanins. The copper enzymes cannot be classified readily on this basis as their protein to copper ratios vary from 17,000 to over 100,000. The hemocyanins probably represent a special case for the copper proteins, but reflect the general trends in the molecular evolution of the oxygen-carrying proteins. Where the oxygen-carrying proteins have remained extracellular, aggregation into giant molecules has occurred, presumably to diminish their contribution to the osmotic pressure of the circulatory fluids. Only the two extremes of hemocyanin molecular size are given in Table II.

2. Optical Characteristics

Most of the copper proteins and the copper enzymes are a beautiful blue or blue-green color, with absorption bands in 600–665 m μ region. Uricase and ox-liver hepatocuprein are reported to be colorless. Tyrosinase had long been associated with extensive pigmentation, but when different purification procedures were developed, it also proved to be virtually colorless. This absence of color is consistent with the suggestion of Kertesz of a Cu(I) state in tyrosinase. Kertesz based his contention on the appearance of the Cu(I)-cuproine complex formed after removal of the copper ion from tyrosinase in glacial acetic acid. As Frieden (1958) has emphasized, this test can be misleading, since there are several side-chain groups in proteins which rapidly reduce Cu(II) to Cu(I). Robert reported that these side reactions are more pronounced under acid conditions. Thus it is believed that the unequivocal demonstration of the Cu(I) in tyrosinase must await other methods and modifications of chemical methods such as those of Felsenfeld. Particularly suitable will be those techniques which do not involve the removal of the copper ion from the protein such as electron spin resonance measurements as already employed by Nakamura (1958) and Malmström *et al.* (in 1958) for laccase.

It has been realized for some time that the visible spectrum of copper ion in copper proteins was considerably enhanced. Simple cupric ion compounds rarely exceed molar extinction coefficients $E_{max} = 200$, but for hemocyanin, ceruloplasmin, and ascorbic acid oxidase, $E_{max} = 750, 1200$, and 767 , respectively. These bands disappear on removal of the copper ion with cyanide or in some cases on reduction of the copper ion to cuprous ion. Orgel (1958) suggests that these spectral enhancements are due to the presence of low

lying charge-transfer bands. Since the structures contributing to this involve oxygen binding, either as a charge-transfer band, or as a band of cupric ion modified by a large charge-transfer component, it provides a good diagnostic test for oxygen carriers. Hemocyanin has an absorption band at 345 m μ which, like the 600 m μ band, can be removed by the removal of the copper ion with cyanide.

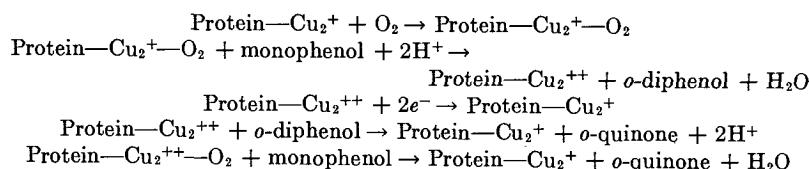
Thus far most of the interest has centered on the visible absorption due to copper. The ultraviolet absorption spectrum, representing primarily the protein absorption, may also furnish valuable clues to subtleties in the structure of the copper proteins. According to Kertesz, homogeneous tyrosinase has an unusually high absorption at 280 m μ , $E_{1\text{cm}}^{1\%} = 27$. This could indicate an unusually higher proportion of tyrosine or tryptophan or perhaps a residual highly chromogenic impurity. Ceruloplasmin also has a somewhat elevated $E_{1\text{cm}}^{1\%}$ of 15.5. But typical protein values of about 9.7 and 11.5 have been reported for cerebrocuprein and uricase.

3. Oxidation State of Copper Ion in Copper Proteins

To serve as the prosthetic group for terminal oxidative catalysis, copper ion in copper proteins must be able to accept electrons from an appropriate substrate and presumably transfer these electrons directly to oxygen. This is believed to entail a continuous shuttle from the Cu(II) to the Cu(I) state or vice versa. In the blue copper enzymes it is assumed that once the source of electrons is exhausted and in the presence of sufficient oxygen, the enzyme-copper ion remains in the Cu(II) state. Only in the presence of substrate can it exist in the Cu(I) state. This has been confirmed in the case of the copper ion of laccase by Nakamura (1958). From magnetic susceptibility measurements, Nakamura concluded that in the oxidized or native form of plant laccase, Cu(II) was exclusively present. When laccase was reduced by ascorbic acid, the copper of the enzyme was quantitatively reduced to Cu(I). Malmström *et al.* have also reported evidence for a similar shift in the oxidation state of fungal laccase. It is also suggested from the observations of Joselow and Dawson who found that the copper ion of ascorbic acid oxidase exchanged with Cu⁶⁴ only when the enzyme was operating and presumably the enzyme-copper ion could be in the reduced state. The problems involved in the detection of Cu(I) in proteins (Frieden, 1958) were mentioned in the previous section. Once again the desir-

ability of the use of physical methods on the intact protein to determine the state of copper ion in copper proteins is emphasized.

The interesting but complicated enzyme system tyrosinase may represent a special case. Mason (1956) has proposed a mechanism which accounts for a Cu(I) resting state of the enzyme and its recognized phenolase and diphenolase activities:



The possible role of complexes analogous to cupryl ion, CuO⁺ and percupryl ion, CuO₂⁺ in tyrosinase action has been suggested by L. L. Ingraham in his recent book, "Biochemical Mechanisms" (Wiley, New York, 1962).

On the basis of Cu⁶⁴(II) exchange experiments with both resting and functioning tyrosinase, Dressler and Dawson (1960) postulated that the oxidation of a monophenol may not proceed via an *o*-dihydric phenol. The Cu⁶⁴(II) exchange data support the suggestion of two distinct activity sites or enzymes in tyrosinase, i.e., catecholase and cresolase activity centers. The copper ion at the cresolase activity sites is nonexchangeable in contrast to its exchangeability at the catecholase activity locus. Mason (1956) has pointed out the functional parallel between the heme and the copper metalloproteins. Each series contains an oxygen carrier dependent upon reduced metal (hemoglobin and hemocyanin), an electron carrier (cytochrome and cytochrome oxidase), and an oxidizing enzyme (peroxidase and tyrosinase).

4. Inferences from the Effect of Ligands on Copper Enzymes

Some inferences of value can be drawn from the various effects of certain ligands on particular copper enzymes, summarized in Table III. Cyanide ion is probably the most effective of all copper ion affecting reagents, probably because its smaller size permits maximum access to the copper ion of the copper proteins. It also appears to compare favorably with other ligands in the relative intensity of their interaction with copper ion. Thus, cyanide ion inhibits all known copper enzymes. Frequently cyanide ion rapidly

TABLE III
EFFECT OF METAL ION LIGANDS ON SOME COPPER ENZYMES

	CN	CO	EDTA	DIECA	1,10-Phenanthroline	Neocuproine	8-Hydroxyquinaline	Thyroxine Cu(II)	CN- removes Cu from protein
Ascorbic acid oxidase ^a	I	±	A	I	5	None	I	A	I 6-7
	4.0				1	5.3	4	None	None $\frac{1}{2}$ of total
Ceruloplasmin ^a	I	None	I	I	None	None	I	None	
	5.5		5.3		4				
Cytochrome oxidase	I	I	None	±	None	None	A		No
	5.3				I				
Tyrosinase ^a	I	I	I	I	4.5	4.7	I	None	Yes
	4.9		4.5						
Uricase	I	I	I	I			I		No
	5.3								
Laccase	I	None	None	±	±	I	I		
β -Mercaptopyruvate transulfurase									
Cu(II)*—Ascorbate oxidation	I	I	I	A	I	I	I	I	

* Key. Numbers below symbols = $-\log(I)$ necessary for at least 50% inhibition. I = Inhibition; A = Activation. ± No more than 35% inhibition at 10^{-8} to $10^{-4}M$.

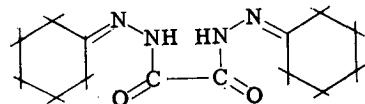
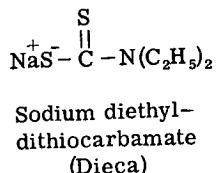
^a Data from published and unpublished work in this laboratory.

removes copper ion from the copper proteins by complexing or reducing copper, producing in some cases reactivatable apoenzymes. An exception is cytochrome oxidase which cyanide ion inhibits and is assumed to be responsible for the lethal effect of cyanide in humans. It is not certain that cyanide ion directly affects the copper ion of this particular copper enzyme, because it is the only copper enzyme which has one heme group for every copper. It is, of course, well known that the activity of virtually every heme enzyme is destroyed by cyanide ion.

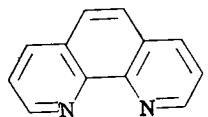
Let us examine other ligands frequently used to study copper enzymes. Their chemical structures are depicted in Fig. 3. It is, of course, misleading to describe these compounds as "copper ion reagents." It is true that in most comparisons of transition metals, these ligands chelate copper ion better than Ni(II), Fe(II) or (III), Co(II), Zn(II), Mn(II). But the difference is quantitative rather than qualitative. Specificity of interaction in a given biological system may be achieved. But it must be proved for each case and generalizations from the particular case must be approached with caution. For example, 1,10-phenanthroline chelates so efficiently with Fe, Zn, and Cu ions in many proteins that no claim for the specificity of this reagent would appear to be legitimate. Sodium diethyldithiocarbamate (DIECA) has been used by James (1953) to assess the importance of copper enzymes in terminal plant respiration. His approach was to use cyanide ion to detect the respiration due to both iron and copper terminal oxidation and DIECA for the copper enzymes alone. Any residual respiration was believed to be due to flavoproteins. While these are interesting and suggestive data, the use of DIECA is not as clean-cut as implied by James (1953). The reagent is subject to considerable instability and its specificity for copper enzymes is questionable.

Several extremely useful metal ion ligands have become available during recent years—so recently, in fact, that they have not been used as widely as deserved. A particularly interesting reagent is neocuproine which is reasonably water soluble. Neocuproine is probably more specific for copper ion than its simpler analog, 1,10-phenanthroline (Fig. 3). However, it is not correct to describe neocuproine as an entirely specific Cu(I) reagent. Rather neocuproine is preferential for Cu(I), since we have shown that it also chelates Cu(II). The effect of these larger copper ion ligands is not as predictable as with cyanide. Neocuproine immediately inhibits

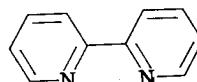
tyrosinase noncompetitively, but it does not inhibit the *p*-phenylene-tyrosinase or ascorbic acid oxidase activity of ceruloplasmin or cytochrome oxidase. In contrast, EDTA instantaneously inhibits the ascorbic acid oxidase activity of ceruloplasmin and uricase, requires



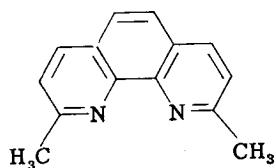
bis-Cyclohexanone-dioxaldihydrazone
(Cuprizone)



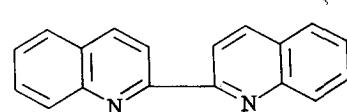
1,10-Phenanthroline



2,2'-Bipyridine



2,9-Dimethyl-1,10-phenanthroline
(Neocuproine)



2,2'-Biquinoline
(Cuproine)

FIG. 3. Structure of frequently used ligands in the study of copper enzymes and other metalloenzymes.

several hours to inhibit tyrosinase, but does not affect cytochrome oxidase activity.

The variation in the response of the copper enzymes to different copper ion ligands is probably due to four factors:

a. *The affinity of the ligand for protein copper.* All the ligands usually employed have an intense tendency to chelate or complex free copper ions. While there is some variation in their relative interactions, these variations are not believed to be the determining factor in producing the differing effects of various ligands. However, it is possible that subtle differences may be exaggerated when copper ion is partially bonded to proteins.

b. *Number and type of bonds with the copper ion.* The residual or displaceable bonds available to the copper ion ligand limit the type of complex that can be formed. We may assume that the firmness of the copper bonds in most copper enzymes indicates at least four to six bonds between copper ion and the protein unless two adjacent sulphydryl groups are involved. This may even be reflected in the rate of the reaction between the ligand and the copper enzyme. For example, EDTA reacts instantaneously with ceruloplasmin, but only slowly with tyrosinase.

c. *Steric accessibility of the copper site.* The neighboring geometry of the copper site may be such as to restrict the approach or entry of some of the larger ligands. Cyanide ion is probably the smallest of the ligands and is the only reagent known to inhibit all the copper enzymes, and cyanide ion and carbon monoxide are the only compounds that seem to be able to penetrate to the copper or heme group of cytochrome oxidase. Steric considerations may explain why ceruloplasmin is affected by 1,10-phenanthroline but not by its 2,9-dimethyl derivative (neocuproine) (see p. 495).

d. *Oxidation state of the copper ion.* Since the ligands used vary in their effectiveness in complexing Cu(I) or Cu(II), it is obvious that the oxidation state of copper could influence their effect on any copper enzyme. The susceptibility of tyrosinase to neocuproine may be due to the claimed existence of this enzyme in the Cu(I) state (Kertesz, 1957).

5. The Interaction of Copper Ions with Proteins

It is unfortunate that at this writing we cannot accurately describe the structure of the copper site in any copper protein as we can for iron in the iron-heme proteins. Certain general comments are in order and the inferences which have been made as to the nature of the copper site for specific copper proteins also will be considered.

All copper proteins have in common the fact that copper ion is so

firmly bound that it will not be lost by dialysis at neutral pH's. As the pH recedes below 5 for most copper proteins, copper ion begins to dissociate, but only exploratory experiments on this point have been made. Thus at pH's 5-8, no dissociation of copper ion can be detected by dialysis procedures. This firm bond between copper ion and proteins is an important fact which distinguishes copper proteins from many other metallo-proteins. It furnishes the copper enzymes with a built-in label for the identification of the catalytically active site or, at least, the copper locus in the enzyme molecule. Work is in progress in our laboratories in which we are attempting to isolate and identify the catalytically active sites of several copper enzymes such as tyrosinase and ceruloplasmin. Several years ago, preliminary studies on tyrosinase revealed a marked resistance of tyrosinase to proteolysis by chymotrypsin and trypsin when the tyrosinase to protease ration was 50-100 to one. This may reflect inactivation of these proteases by tyrosinase. Only the more vigorous proteolytic action of Carlsberg subtilisin resulted in 70% loss of tyrosinase action after 20 hours. EDTA, $10^{-2}M$, partially protected the activity of the enzyme from subtilisin. The logic here is to block selectively proteolysis in the neighborhood of the copper ion and so to preserve the peptides in the neighborhood of the copper. Caution must be observed in interpreting data from this type of experiment. Mason and co-workers studied the action of proteases on hemocyanin and raised the possibility of the migration of the copper ion during the course of proteolysis. The protection of the copper site with copper ion complexers such as EDTA, 1,10-phenanthroline, might obviate this difficulty.

The specific binding of copper in the copper proteins must be distinguished from the strong but dissociable bonding between cupric ion and various other proteins such as serum albumin. The binding of Cu(II) by proteins and amino acids has been long recognized, but only the recent conclusions of Klotz (1953) will be presented here. Cu(II) is one of the transition metals which bind very strongly with proteins and may involve several functional groups of the protein including sulphydryl, imidazole, indole, free amino, free carboxylate ions, and peptide nitrogens. The strength of interaction of Cu(II) with certain free amino acids may have been somewhat exaggerated, but the intensity of the complexation with cysteine, histidine, peptides, and proteins is very high. With some proteins, it is evident that groups other than sulphydryl are involved.

At high pH's the spectral characteristics of copper-protein complexes suggest that C—N bonds, possibly peptide nitrogen, are involved. Klotz has noted a parallelism between the binding constant and the isoelectric point of several proteins, as shown in Table IV. Presumably the lower the isoelectric pH compared to pH 6.5 the greater the negative charge on the protein. Thus electrostatic forces would favor Cu(II) protein interaction as:



For serum albumin Klotz *et al.* obtained evidence for the involvement of carboxyl groups, since the methylated protein showed reduced binding whereas acetylated albumin showed no loss of ability

TABLE IV
CORRELATION OF THE AFFINITIES OF CUPRIC ION FOR PROTEINS
AND THEIR ISOELECTRIC pH^a

Protein	F°, kcal/mole	Isoelectric pH
α-Casein	-7.30	4.0
β-Casein	-7.16	4.5
Serum albumin	-6.49	4.7
β-Lactoglobulin	-5.81	5.2
Lysozyme	-4.39	11.0

^a For purposes of comparison a unit weight of 100,000 has been taken for each protein. Measurements made at 0°C, pH 6.5, 0.2 M acetate solution.

to interact with cupric ion. From a study of the spectra of copper-protein complexes at different pH's, Klotz concluded that at higher pH's, Cu . . . NH₂ linkages become possible just as they do in simple cupric ion-peptide models. Tetracoordinated cupric ion shows a maximum at about 600 m μ , even if two of the ligands are carboxyl groups. Since most of the naturally occurring and synthetic copper proteins also absorb in the same region, analogous bonding involving two of the bonds to amine groups in the protein may be postulated.

The extent of binding of copper ion by any ligand is highly dependent on the oxidation state of copper. Cupric ion complexes are frequently more stable than the corresponding cuprous ion complex. Also in the transition of cupric to cuprous, a large change in geometry occurs. Cu(II) complexes are usually square planar when tetra-

coordinated or octahedral where all six coordinates are filled. Cu(I) complexes are tetrahedral when tetracoordinated with virtually no tendency to hexacoordinate. In the transition from Cu(II) to Cu(I) this change in geometry must be satisfied. Understandably the release of copper ion and its exchange with copper ion in solution will be facilitated. Thus Joselow and Dawson found that the copper ion of the enzyme, ascorbic acid oxidase, exchanged with Cu⁶⁴ of the solution only when the enzyme was active and probably going to the Cu(I) state.

Finally, an ever present difficulty in assessing the interaction of Cu(II) and proteins is the multiplicity of reactions that may be encountered. Earlier we referred to the reduction of Cu(II) to Cu(I) by certain amino acids and numerous proteins in the presence of neocuproine. If H₂O₂ is also present, Phelps, Putnam, *et al.*, found that Cu(II) catalyzes several reactions including modification of the heterocyclic and aromatic amino acids, oxidation of the sulfur-containing amino acids and finally a cleavage of the molecule into several large fragments. Thus the possibility that Cu(II) is producing significant alterations in the structure of a particular protein cannot be overlooked.

6. Active Site of the Copper Proteins

It has been stressed that we do not know the pertinent facts about the location of copper ion in the copper proteins. Only tentative conclusions can be drawn from indirect data. For example, the pH optimum of the copper enzymes usually falls within the pH range 6-7, presumably the histidine dissociation range. The fact that copper ion is released below pH 5 is interpreted as indicating the importance of certain side carboxylate groups in the binding of copper ion. For β -mercaptoproprylate transsulfurase, Kun and Fanshier (1959) concluded that both SH groups and protein-bound copper participate in enzymatic catalysis. They propose a probable mechanism of transsulfuration which assumes the proximity of Cu ion and two neighboring sulfhydryl groups. The possible importance of sulfhydryl groups in the action of ascorbic acid oxidase was also suggested by Frieden and Maggiolo (1957). But Dawson reports that not all preparations of this enzyme show sensitivity to the organic mercurial reagents such as *p*-chloromercuribenzoate. Since ascorbic acid oxidase is highly subject to activation by low concentrations of many compounds including certain sulfur-containing

compounds, apparent inhibition would be observed if the organic mercurial reacted with an activating substance. Some years ago we showed the interaction of phenylmercuric chloride with thyroxine and several of its analogs. It is of interest that both ascorbic acid oxidase and β -mercaptoproprylate transsulfurase are inhibited by 10⁻⁴ M or less Cu(II) (Frieden and Maggiolo, 1957; Kun and Fanshier, 1959). Presented in more detail below are summaries of what is currently known about the copper ion sites of ceruloplasmin and hemocyanin, which have been most thoroughly studied in this regard, primarily because these two proteins are the most readily available in a pure or crystalline state.

a. *The copper site of ceruloplasmin.* The binding of copper in ceruloplasmin has been recently discussed by G. Curzon (1960). The distinction between the linkages between Cu(II) serum albumin and ceruloplasmin are clearly evident. The Cu ion of Cu(II)-albumin can be dialyzed free, though with difficulty, removed by at least one adsorbent, and shown to form a yellow complex with the reagent, DIECA. Ceruloplasmin Cu is stable under the above conditions, but it can be removed after a combined treatment with ascorbate and DIECA and after heat denaturation. The latter fact emphasizes the importance of conformation, especially geometrical relationships arising from intra- and interhelical bonding of copper ion. From acid-base titrations and electrophoretic analysis of ceruloplasmin and its apoenzyme, Scheinberg and Sternlieb (1960) suggested that the Cu(II) was bound to two negatively charged carboxyl groups in the native protein. From his and Curzon's suggestions, the copper site of ceruloplasmin may be tentatively pictured as in Fig. 4 (see p. 488). There is also a remote possibility that copper ion may be linked to glucose, mannose, or xylose, found in ceruloplasmin by Osaki (1961), although it is known that Cu(II) complexes more effectively with most amino acids than with these monosaccharides. Osaki (1961) also reported that phenylalanine was at both termini of ceruloplasmin.

A unique feature of ceruloplasmin is that there appears to be two types of copper atoms present. Both chemical and kinetic data suggest that about one half of the copper atoms are distinct from the others. Scheinberg and Morell found apoceruloplasmin would accept only four to five of its original eight copper atoms. They also found that only about four of the eight copper ions of ceruloplasmin were exchangeable with Cu⁶⁴ in the presence of ascorbate. On pro-

longed digestion of ceruloplasmin with chymotrypsin, Curzon found that only about one half of the copper ion became dialyzable. In our laboratory, we have extended the earlier suggestion of Humoller *et al.* that ceruloplasmin has at least two different catalytically active sites, both undoubtedly involving copper atoms. As depicted in Fig. 5, site I serves as an active site for both *p*-phenylenediamine and ascorbate. This site can also be shown to be inhibited by CN⁻,

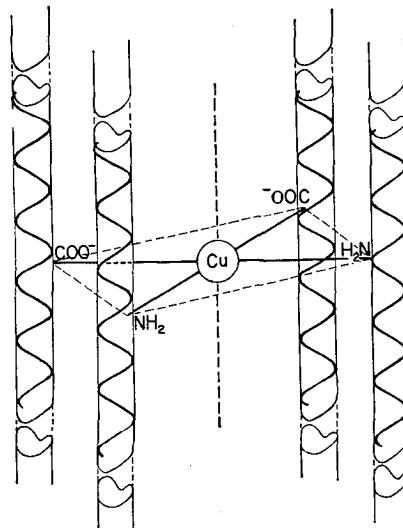


FIG. 4. A possible structure for a copper site in ceruloplasmin. The copper ion is bonded by at least four groups in either separate or adjoining peptide chains. Two sites of the possible six sites are purposely left available for catalytic function. (Modified after Curzon.)

EDTA, and 1,10-phenanthroline. Site II is exclusively for *p*-phenylenediamine catalysis and is only affected by CN⁻. While site I is obviously more accessible than site II, its availability does not extend to other ligands such as neocuproine and thyroxine which do not affect any of the catalytic activities of ceruloplasmin.

b. The copper site of hemocyanin. Although it has been known since 1847 that copper is present in hemocyanin, the structure of its copper site is still obscure despite numerous efforts to answer this question. The implication of Cu-S-protein linkages based on indirect evidence has been challenged. Felsenfeld, using unpurified *Limulus* serum, removed the copper ion from hemocyanin and estimated the stability constant of the Cu⁺⁺-apohemocyanin complex

as 10¹⁷ to 10¹⁹, a value in a range similar to the stability constant of cuprous cysteinate 1.5×10^{19} . Klotz *et al.* observed a close relationship between the spectra of hemocyanin and bovine serum albumin-copper complex, and the presence of Cu(I) in nonoxygenated hemocyanin. Accordingly, he postulated that each copper was bound to a single thiol group. But Thomson, Hines, and Mason determined the number of sulfhydryl groups in hemocyanin before and after removal of copper by amperometric titration and showed that for each four atoms of copper removed, only one sulfhydryl group was exposed. Lontie reported that the copper of apohemocyanin could be quantitatively restored even after blocking its thiol

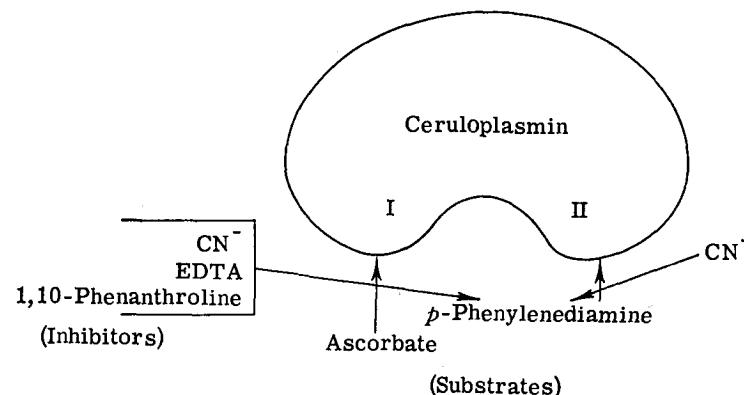


FIG. 5. The two sites of ceruloplasmin showing substrate and inhibitor specificities. Neocuproine and thyroxine do not inhibit the oxidation of ascorbate at site I or the oxidation of *p*-phenylenediamine at either site.

groups with typical thiol reagents. Lontie's suggestion that imidazole groups are involved in the copper binding is, however, based on evidence as indirect as the earlier evidence for thiol groups. Perhaps the only convincing proof will come from the isolation and proof of structure of the copper ion-containing fragment under conditions which exclude the migration of the copper ion.

IV. Copper Ion in Model Catalytic Systems

No metal ion or other nonprotein catalyst surpasses copper salts in their versatility as catalysts for an impressive variety of reactions. Only iron salts, platinum, and perhaps a few other metals compare with copper ions in their catalytic prowess. The recognition of the virtually unique catalytic features of copper ions and copper

TABLE V
REACTIONS CATALYZED BY Cu(II) IONS OR Cu(II) COMPLEXES

Type of reaction	Reactant(s)	Product(s)
Peroxidatic action	Catechol + H ₂ O ₂ 3-Indole acetic acid + H ₂ O ₂ Human γ -globulin + H ₂ O ₂ Tyrosine + H ₂ O ₂ + ascorbic acid	o-Benzoquinone (Unknown) 2 Protein fragments
Catalase action	H ₂ O ₂	H ₂ O + $\frac{1}{2}$ O ₂
Transamination	α -Amino acid + pyridoxal α -Amino acid + α' -keto acid (in pyridine solution)	α -Keto acid α' -Amino acid + α -keto acids
Decarboxylation	Oxaloacetic acid Dimethyloxaloacetic acid	Pyruvic acid + CO ₂ α -Ketoisovaleric acid + CO ₂
Hydrolysis	α -Amino acid esters R-Amides Bisthiophenalethylenediamine Diisopropylfluorophosphate Methyl isopropylphosphorofluoridate	α -Amino acid + alcohol R-Acid + NH ₃ Monoethylenediamine Cu(II) + thiophenaldehyde HF + Diiodopropylphosphoric acid HF + Methylisopropylphosphoric acid
Oxidation	H ₂ (cuprous salts in pyridine or quinoline) Cysteine, glutathione Uric acid Acetoin Reducing sugars Isonicotinic acid hydrazide Primary arylamines 2,6-Disubstituted phenols	H ₂ O Cystine, oxidized glutathione Dehydrouric acid Biacetyl Sugar acids Isonicotinic acid, isonicotinamide diisonicotinyl hydrazine Azo compounds Polymers

TABLE V (Continued)

Type of reaction	Reactant(s)	Product(s)
Oxidation	Ascorbic acid <i>o</i> - or <i>p</i> -Phenylenediamine Phenol + morpholine Tyrosine + trace dopa 3,4-Dihydroxyphenylalanine (dopa) Adrenaline Catechol, hydroquinone Homocatechol Pyrogallol	Dehydroascorbic acid, hydrogen peroxide <i>o</i> - or <i>p</i> -Benzoquinone + NH ₃ Dimorpholine + <i>o</i> -benzoquinone Melanin Dopa 3,4-quinone, melanin Adrenochrome <i>o</i> -Benzoquinone, <i>p</i> -benzoquinone 4-Methyl- <i>o</i> -benzoquinone Purpurogallin

complexes occupies hundreds of pages in the chemical and the biochemical literature. The breadth and complexity of this field is so vast that the challenge of attempting to present a comprehensive review on this subject has yet to be met. The concluding pages of this paper are intended only to stimulate the interest of the reader in the mysteries of copper ion catalyses.

TABLE VI
EFFECT OF LIGANDS ON THE Cu(II) CATALYZED OXIDATION OF ASCORBATE

Designation	Name	Probable ligand to Cu(II) ratio	Inhibition of catalysis ^a
EDTA	Ethylenediamine tetraacetic acid	1	50% at $5 \times 10^{-7} M$ 100% at $7 \times 10^{-7} M$
Thyroxine	3-[4-(4-Hydroxy-3,5-diiodophenoxy)-3,5-diiodophenyl] alanine	2	50% at $2 \times 10^{-6} M$ 100% at $1 \times 10^{-6} M$
Neocuproine	2,9-Dimethyl-1,10-phenanthroline	2	50% at $1 \times 10^{-6} M$ 100% at $1.5 \times 10^{-6} M$
Cuprizone	Biscyclohexanone-dioxaldihydrazone	2	50% at $1 \times 10^{-5} M$ 100% at $5 \times 10^{-5} M$
Phenanthroline	1,10-Phenanthroline	2	50% at $1.5 \times 10^{-6} M$ 0% at $1 \times 10^{-6} M$ Higher rate above $1 \times 10^{-6} M$
2,2'-Bipyridine	2,2'-Bipyridine	2	30% at $1 \times 10^{-6} M$ 0% at $1 \times 10^{-6} M$ Higher rate above $1 \times 10^{-6} M$

^a Reaction conditions: 30°C, Cu(II) = $6.0 \times 10^{-7} M$, Ascorbate = 0.020 M, PO₄ buffer (pH 7.0) = 0.010 M.

The adaptability of copper ion as a catalyst for a wide variety of reactions is indicated in Table V. That copper ion catalyzes the oxidation of a large number of compounds including phenols, sulfhydryl compounds, aromatic amines, and other compounds is well known. Less well publicized are the catalytic participations of copper salts in transamination, hydrolysis, decarboxylation, deamination, and peroxidation. These additional catalytic activities might serve

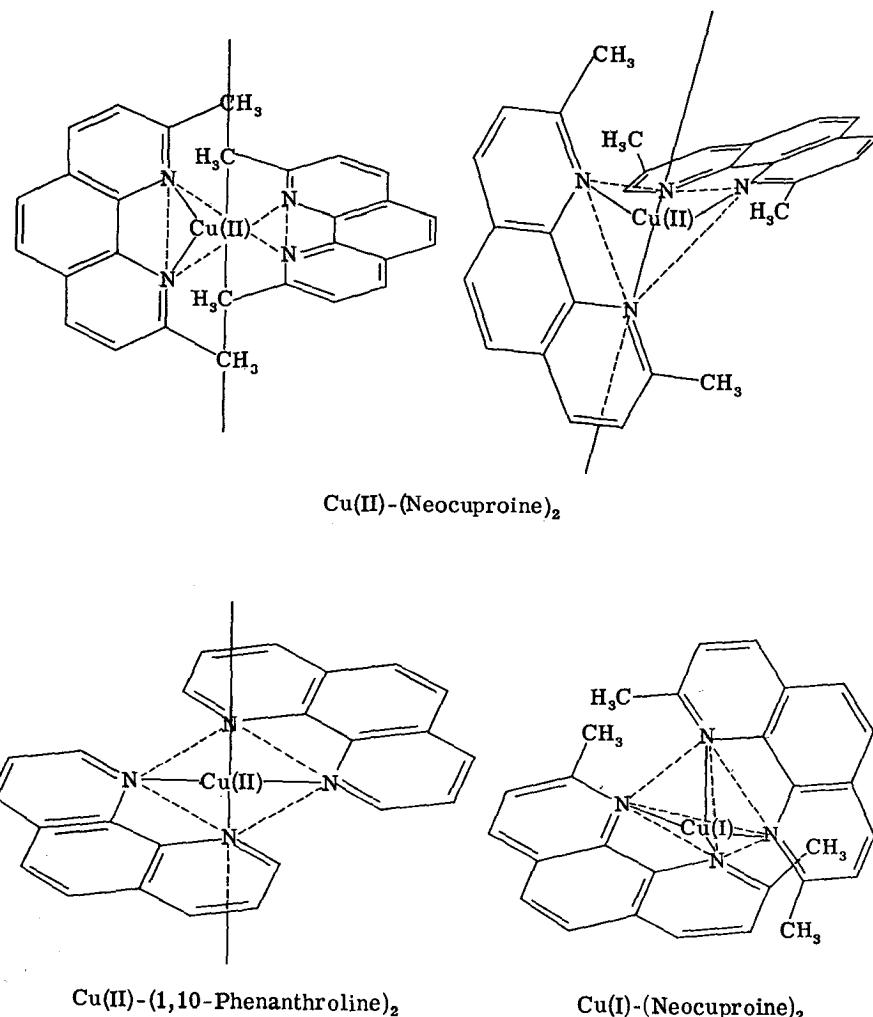
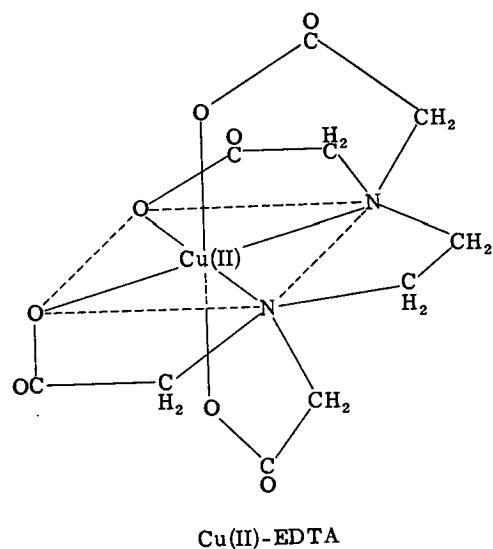


FIG. 6. Postulated structures of phenanthroline complexes of Cu(I) and Cu(II). The two upper structures are different views of the same chelate.

as a guide for a directed exploration for new and unique copper enzymes. More often than not, copper ion alone will catalyze reactions identical to those catalyzed by copper enzymes, though usually less effectively. Thus the protein apoenzymes appear to aid in attaching and directing the substrate to the copper site and perhaps



Cu(II)-EDTA

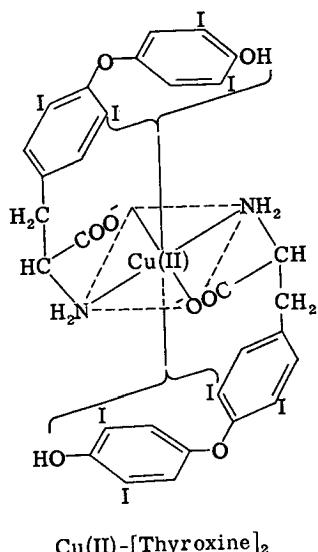
Cu(II)-[Thyroxine]₂

Fig. 7. Postulated structures of the Cu(II) chelates with EDTA and thyroxine.

also in improving the rate constant, but they rarely appear to confer unique catalytic properties on copper ion.

During recent years we have studied the effect of numerous ligands in one of the most useful copper ion model systems, the Cu(II)-ascorbate-oxygen system. Most of these ligands inhibit ascorbate oxidation by chelating Cu(II) and/or Cu(I). Several compounds form complexes which show some catalytic activity. Activity of copper chelates has also been reported by Butt and Halloway and also by Khan and Martell. Data with representative ligands are summarized in Table VI. A simple explanation for the above phenomena may be available. The probable structures of a number of important copper ion chelates is shown in Figs. 6 and 7. Two of these chelates, Cu(II)-1,10-phenanthroline and Cu(II)-2,2'-bipyridine, show catalytic activity in ascorbate oxidation. All the other chelates inhibit this reaction by complexing copper ion. The resulting chelates are catalytically inactive. The two active chelates involve complexing with four of the six possible sites of Cu(II) as portrayed in Fig. 6. It is also possible that 1,10-phenanthroline and 2,2'-bipyridine might themselves undergo reversible oxidation and reduction. The inactive chelates involve all six sites directly or involve structures which may sterically hinder access to the remaining two sites. An alternative explanation is that certain of these ligands, e.g., neocuproine, bind Cu(I) so effectively that there is no chance for the reoxidation of Cu(I) by oxygen. This hypothesis explaining the activities of Cu(II) in such a model system might be useful in predicting structures of proteins at the copper ion locus. It would be expected that copper ion in the oxidative enzymes would have at least two free or displaceable sites which must be accessible to the substrate and to the electron acceptor, usually oxygen.

REFERENCES*

This paper was not intended to be an exhaustive review of the literature and therefore reflects the interests and bias of the author. For purpose of

* Note added in proof: Since the completion of this paper, several significant reviews on closely related topics have appeared including:

1. Westerfeld, W. W. et al. (1961). In "Biological Aspects of Metal Binding" (Symposium on Metal Chelation), *Federation Proc.* **20**, 1-263.
2. Tuchida, A. et al. (1961). *Symposium on Copper Proteins and Copper Enzymes Biophysics* **1**, 42-53 (in Japanese).
3. Adelstein, S. J., and Vallee, B. L. (1961). *New Eng. J. Med.* **265**, 892-897, 941-946.

brevity, ten review references are given, and fourteen recent references which will lead to the earlier papers are cited.

Reviews on Copper Metabolism and Copper Proteins

- Dawson, C. R., and Mallette, M. F. (1945). *Advances in Protein Chem.* **2**, 179-248.
- Dawson, C. R., Mallette, M. F., and others (1950). In "A Symposium on Copper Metabolism" (W. D. McElroy and B. Glass, eds.). Johns Hopkins Press, Baltimore, Maryland.
- Klotz, I. M. (1953). In "The Proteins" (H. Neurath and K. Bailey, eds.), Vol. I, Part B, pp. 727-806. Academic Press, New York; Klotz, I. M., and Klotz, T. A. (1955). *Science* **121**, 477.
- Laurell, C. B. (1960). In "The Plasma Proteins" (F. W. Putnam, ed.), Vol. I, pp. 360-369. Academic Press, New York.
- Lehnninger, A. L. (1950). *Physiol. Revs.* **30**, 393.
- Mahler, H. R. (1961). In "Mineral Metabolism" (C. L. Comar and F. Bronner, eds.), Vol. I, Part B, pp. 743-879. Academic Press, New York.
- Mason, H. S. (1956). *Nature* **177**, 79.
- Orgel, L. E. (1958). In "Metals and Enzyme Activity" (E. M. Crook, ed.), pp. 8-20. Cambridge Univ. Press, London and New York.
- Scheinberg, I. H., and Sternlieb, I. (1960). *Pharmacol. Revs.* **12**, 355.
- Singer, T. P., and Kearney, E. B. (1954). In "The Proteins" (H. Neurath and K. Bailey, eds.), Vol. II, Part A, pp. 135-160. Academic Press, New York.

Recent References

- Curzon, G. (1960). *Biochem. J.* **77**, 66.
- Dressler, H., and Dawson, C. R. (1960). *Biochim. et Biophys. Acta* **45**, 508, 515.
- Frieden, E. (1958). *Biochim. et Biophys. Acta* **27**, 414.
- Frieden, E., and Alles, J. (1958). *J. Biol. Chem.* **230**, 797.
- Frieden, E., and Maggiolo, I. W. (1957). *Biochim. et Biophys. Acta* **24**, 42.
- Griffiths, D. E., and Wharton, D. C. (1961). *J. Biol. Chem.* **236**, 1850, 1857.
- James, W. O. (1953). *Biol. Revs.* **28**, 245-60.
- Katoh, S., Suga, I., Shiratori, I., and Takamiya, A. (1961). *Arch. Biochem. Biophys.* **94**, 136.
- Kertész, D. (1957). *Nature* **180**, 506.
- Kun, E., and Fanshier, D. W. (1959). *Biochim. et Biophys. Acta* **32**, 338.
- Nakamura, T. (1958). *Biochim. et Biophys. Acta* **30**, 44, 538, 640.
- Osaki, S. (1961). *J. Biochem.* **50**, 29.
- Szent-Györgyi, A. (1928). *Biochem. J.* **22**, 1387.
- Szent-Györgyi, A. (1930). *Science* **72**, 125.

Cancer and Necrosis Induced Selectively by Hydrocarbons

CHARLES HUGGINS

*The Ben May Laboratory for Cancer Research,
University of Chicago, Chicago, Illinois*

I. Introduction	497
II. Methodology	499
III. Molecular Structure of the Special Mammary Carcinogens	501
IV. Influence of Heredity on Chemical Carcinogenesis in Rat	505
V. Selective Cell Death Caused by 7,12-DMBA	506
VI. Discussion	508
VII. Conclusions	510
Acknowledgment	511
References	511

"Die Wissenschaft ist alles rein menschlich."
H. FISCHER TO OTTO WARBURG

I. Introduction

This paper is dedicated to Szent-Györgyi. It is concerned with special mammary carcinogens—compounds which induce cancer of the breast in animals with great frequency and rapidity after a solitary feeding (Huggins *et al.*, 1961a). This is one of the most remarkable of natural phenomena.

The biological responses to aromatic hydrocarbons are characterized by a high degree of selectivity. A single feeding of a massive but tolerable amount of some of the polynuclear aromatic hydrocarbons to rats yields catastrophic changes. These pathological lesions are (1) cancer and (2) necrosis. While certain cells become malignant and massive areas of cellular death are found in other organs, the generality of the cells of the body are uninjured and emerge quite uninjured from their encounter with the hyperconjugated hydrocarbons.

The phenomenon under discussion is highly useful in the experimental induction of mammary cancer since, under appropriate conditions, cancer of the breast arises in every animal within a few weeks. A single feeding of effective carcinogens is more efficient than

the repeated feeding technique which has been described earlier (Huggins *et al.*, 1959a).

Cancer of the breast is one of the most noble problems of medicine. Among Western women, mammary cancer has the highest rate of incidence of any malignancy. In the Orient the frequency of cancer of the breast is much lower.¹ Women who eat with sticks are less vulnerable to this disease than occidental females are but the food conveyors *qua* chopsticks are not to be implicated as the protective agents.

There are two requirements for an acceptable solution of the cancer problem: (1) A rather sharp image of the essential nature of the cancer cell. (2) Prevention and cure of cancer in man. Cancer of the breast has proven to be vexatious and rather difficult since only a partial solution has been achieved. But some progress was made when it was learned that a proportion of the mammary cancers are dependent on hormones of the ovaries and adrenal glands (Huggins and Bergenstal, 1952). Even extensive mammary cancer which is hormone dependent undergoes a profound regression which can be very prolonged when the supporting hormones are withdrawn, and this concept has had some clinical usefulness throughout the world.

Normal and hormone-dependent cancer cells differ essentially in their response to hormone withdrawal. Normal cells dependent upon hormones undergo atrophy, merely becoming dormant and shrinking in size when their host is deprived of the supporting hormones. Hormone-dependent cancer cells differ profoundly from these normal cells of origin; the cancer cells die when the supporting steroid hormones are abstracted. Clearly the steroids are not merely accelerators of growth of the cancer cells—the very life of the cancer cells depends on their presence.

Mammary cancer research has been unsatisfactory because of a lack of technology. Techniques rule science—theory remains speculative and not substantive without an experimental test of its validity. Until the development of the method to be described, mammary cancer research suffered from two defects which made it inefficient and largely unmanageable: (1) The method of induction in animals was very slow; always many months or as long as 2 years were required to evoke mammary cancer in rodents. The results of

¹ In 1956–1957, the incidence of mammary cancer in white women in the United States was 27.08/100,000; the comparable incidence in Oriental women in Japan was 3.26 (Segi, 1960).

experiments in this field were meager but under these sluggish conditions it is commendable that anything was learned. (2) The incidence of mammary cancer was always less than 100% so that statistical treatment of the results, with its inherent difficulties, was necessary. The technology has been solved by the application of hydrocarbons to rats. Maisin and Coolen (1936) painted the skin of mice with 3-methylcholanthrene and observed that, in addition to skin cancer, breast cancer arose in the painted mice in an appreciable number (18%) after about 1 year. This was the first indication of the predilection of hydrocarbons for mammary gland vis-à-vis cancer. The present techniques take advantage of this selective action.

II. Methodology

The most efficient substance that we have found to induce mammary cancer in rodents is 7,12-DMBA.² When 7,12-DMBA (20 mg), dissolved in sesame oil (1 ml), was administered by stomach tube, once only to intact female rats of the Sprague-Dawley strain, age 50–65 days, cancer developed in every rat in a few weeks (Huggins *et al.*, 1961). The age and sex of the rat were critical in this experiment. The strain of the rats was important.

After a single feeding of aromatic hydrocarbons, cancers of three kinds developed commonly; other cancers arose less frequently (Table I).

1. Mammary cancer arose in every animal under the stated conditions.

2. Cancer of sebaceous glands near the ear was observed in rats whose life had been preserved for 6+ months.

3. Sarcoma arose at the site of injection of hormones (Huggins *et al.*, 1959b). This tumor developed following the repeated injection of steroids dissolved in sesame oil. Sarcoma also formed where gonadotropin in saline had been injected repeatedly (Huggins *et al.*, 1959b). The effect of the solutions devoid of hormones has not been investigated.

Notably, the carcinogen was placed in the stomach and traversed the intestine and some of it passed through liver (evident from strong hepatic fluorescence) but cancer was not observed in these regions. It arose selectively and in distant sites.

² 7,12-DMBA, 7,12-dimethylbenz[a]anthracene. NQO, 4-nitroquinoline-N-oxide.

The hydrocarbons need not pass through the entire alimentary tract to be absorbed in effective amounts. The injection of a solution of 7,12-DMBA, 20 mg, into the lumen of the cecum (Huggins *et al.*, 1961b) gave rise to cancer in 6 of 12 rats. Indeed, the hydrocarbons are not required to traverse any part of the intestine to evoke distant cancer. A single intravenous injection of an emulsion of 7,12-DMBA, 1 mg, induced mammary cancer in every rat. Moreover the implantation of a compressed pellet of 7,12-DMBA in the spleen was followed by mammary cancer in a proportion of the animals, and mere traces of the hydrocarbon were absorbed from the intrasplenic pellet in the process (Huggins *et al.*, 1961a). Whereas notably smaller quantities are required to induce cancer by parenteral administration than when hydrocarbons are administered by mouth, the single feeding of 7,12-DMBA is a convenient method of extreme simplicity, and it gave rise invariably to mammary cancer under the stated conditions in our experience.³

The cancers evoked by a single feeding or a single intravenous injection of 7,12-DMBA arise in two discontinuous series. Mammary cancer was palpable in 3 weeks to 3 months (Table I) whereas other cancers were detected after 3 months, most commonly about 6 months after the exposure.

Not all of the many millions of mammary cells became malignant after feeding effective amounts of the carcinogens. A stoichiometric relationship (Huggins *et al.*, 1961a) exists between the dose of the effective carcinogens and the rate of incidence of tumors together with their number. A minimal effective dose gives rise to 1 cancer only. The maximal number of mammary cancers observed in our laboratory after a single feeding is 21.

The mammary cell has precise requirements to become malignant after a single exposure to the carcinogen (Huggins *et al.*, 1959b). The administration of hydrocarbons to hypophysectomized rats gave rise to no cancer because the cells were atrophic. Similarly, when the mammary glands were hyperplastic and undergoing very

³ It will be recalled that a single exposure of Sprague-Dawley female rats, age 40 days, to a sublethal dose of γ -radiation (400 r) is also followed by a high incidence of mammary tumors which are evoked selectively. Cancer does not arise at random in many tissues but preferentially in the mammary gland (Shellabarger, C. J., Cronkite, E. P., Bond, V. P., and Lippincott, S. W., The occurrence of mammary tumors in the rat after sublethal whole-body irradiation, *Radiation Research*, 6, 501-512, 1957).

active growth mammary cancer did not arise. The endocrine status of young adult female rats, as reflected in mammary gland structure, is exquisitely suited for development of cancer of the breast.

TABLE I
NEOPLASMS INDUCED IN RATS BY FEEDING POLYNUCLEAR HYDROCARBONS

Neoplasm	Compound fed ^a	No. of cases	Cancer detected, days		
			Range	Median	Mean
Mammary cancer	7,12-DMBA	90	28-92	41	42.8 \pm 11
Carcinoma of ear duct glands	7,12-DMBA	21	101-171	140	136.6 \pm 22
Sarcoma at site of injections of hormones	3-Methylcholanthrene	39	88-289	175	180.6 \pm 46
Leukemia	3-Methylcholanthrene	6	105-171	127	139
Renal carcinoma	7,12-DMBA	2	168-196		
Bladder carcinoma	3-Methylcholanthrene	1	339		

NOTE: The compounds dissolved in sesame oil were administered by stomach tube.

\pm , Standard deviation.

^a 7,12-DMBA, 20 mg, single feeding. 3-Methylcholanthrene, 10 mg daily, fed 6 days each week, age 50-100 days.

In our colony of rats, estradiol-17 β is inactive or at most a very feeble inducer of mammary cancer, while tumors of the pituitary arise readily. Eleven females (the ovaries and uterus had been removed) were injected with estradiol-17 β , 50 μ g daily, from age 65 days for 1 year. Mammary cancer was detected in 1 rat only and at age 358 days, whereas 5 of the 11 animals developed pituitary tumors of large size.

III. Molecular Structure of the Special Mammary Carcinogens

By definition, the special mammary carcinogens induce cancer of the breast and, less frequently other neoplasms at selective sites, after a single feeding to Sprague-Dawley female rats, age 50-65 days. Ten compounds with this property have been recognized (Fig. 1). With the exception of Compound XII, all others of these 10 substances had been found by earlier workers to produce skin

cancer or sarcoma in mice (Hartwell, 1951; Nakahara *et al.*, 1957; Shubik and Hartwell, 1957; Takayama, 1960) following appropriate local application of the carcinogens. Their properties as special mammary carcinogens are novel. In our colony, mammary cancer has

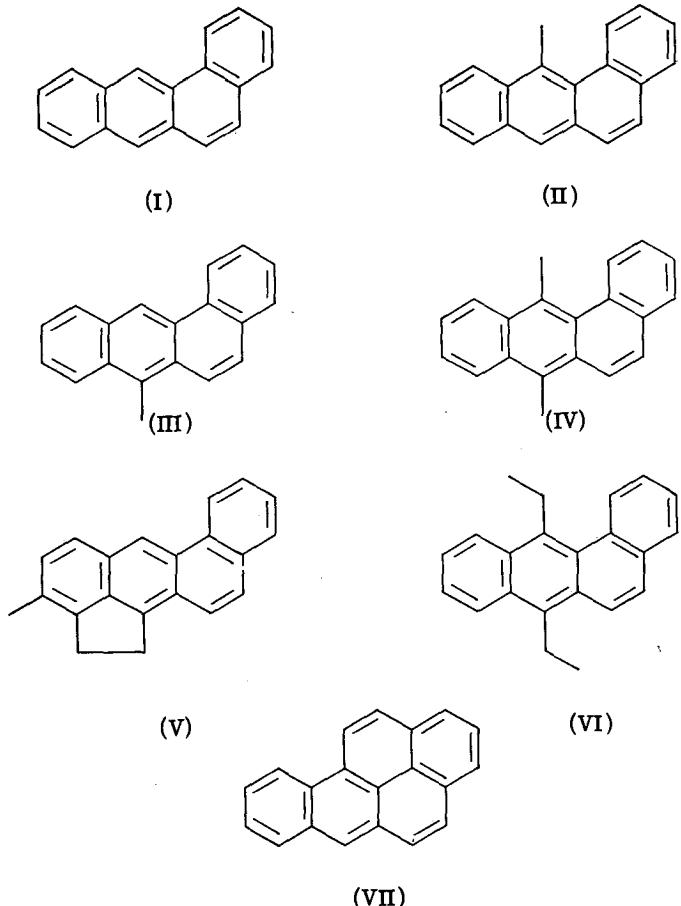


FIG. 1.

been observed in 3 animals among ca. 200 young untreated female Sprague-Dawley rats.

Benz[a]anthracene (I) was not carcinogenic but cancer-producing activity ensued with the addition of a methyl group in the *meso* positions; 12-methyl-(II) and, better, 7-methylbenz[a]anthracene

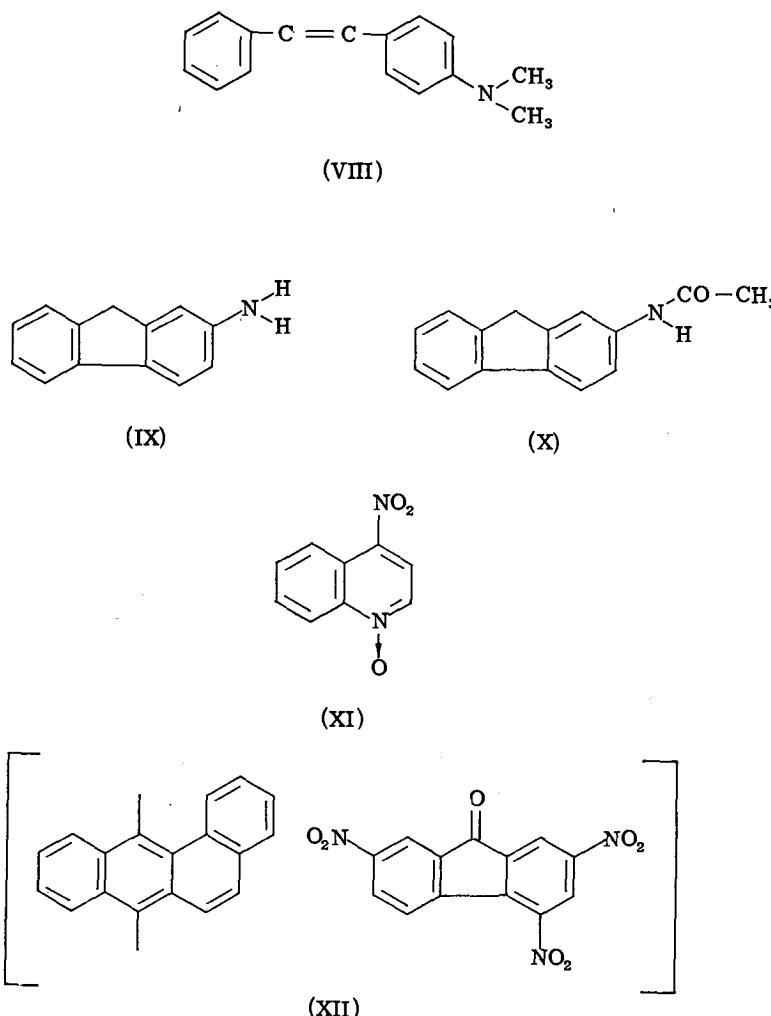


FIG. 1. (cont'd)

(III) were somewhat carcinogenic. 7,12-DMBA (IV) was the most effective special carcinogen encountered in our experiments. Whereas methyl groups at select positions of benz[a]anthracene enhance its cancer-producing activity, ethyl groups do not share this function; 7,12-diethylbenz[a]anthracene (VI) did not induce mammary cancer (Table II).

The addition of a ring to benz[*a*]anthracene (I) to form benzo[*a*]pyrene (VII) greatly increased carcinogenicity of (I). Benzo[*a*]pyrene (VII) was similar quantitatively to 3-methylcholanthrene (V) in this regard.

The phenanthrene structure is not a necessary molecular requirement for mammary carcinogens. An ethylene bridge between aromatic rings is adequate; 4-dimethylaminostilbene (VIII) synthesized by Haddow (Haddow *et al.*, 1948) was a moderately effective carcinogen.

TABLE II
INDUCTION OF MAMMARY CANCER BY A SINGLE FEEDING OF
POLYNUCLEAR HYDROCARBONS

No.	Compound	Dose, mg	No. of rats	Mam- mary cancer	%
I	Benz[<i>a</i>]anthracene	200	18	0	0
II	12-Methylbenz[<i>a</i>]anthracene	100	12	2	16.7
III	7-Methylbenz[<i>a</i>]anthracene	100	13	4	30.8
IV	7,12-Dimethylbenz[<i>a</i>]anthracene	15	19	100	100
V	3-Methylcholanthrene	100	46	43	93.4
VI	7,12-Diethylbenz[<i>a</i>]anthracene	100	20	0	0
VII	Benzo[<i>a</i>]pyrene	100	9	8	88.9
VIII	4-Dimethylaminostilbene	7.5	18	5	27.8
IX	2-Aminofluorene	100	10	2	20
X	2-Acetylaminofluorene	100	28	5	17.9
XI	4-Nitroquinoline- <i>N</i> -oxide	20	10	2	20
XII	Complex (7,12-DMBA · TNF) ^a	40	8	3	37.5

NOTE: The compounds dissolved in sesame oil were administered by stomach tube to female Sprague-Dawley rats, age 50 days.

^aA 1:1 complex of 7,12-DMBA with 2,4,7-trinitro-9-fluorenone was administered as a suspension in sesame oil.

Two fluorene derivatives caused cancer after a single feeding. An amino group (IX) or an acetyl amino group (X) at position 2 in fluorene conferred carcinogenicity on fluorene which was inactive as a carcinogen.

Mammary cancer was induced by a single feeding of 4-nitroquinoline-*N*-oxide (XI) to rats. This compound appears to be a good electron acceptor. Our colleague, Dr. H. G. Williams-Ashman (1961) studied complex formation of NQO (final concentration

0.013 M) with aromatic compounds (0.06 M) in methylene chloride. NQO formed deeply colored solutions at -196°C with a variety of electron donors including Compounds IV, V, and IX. The complexes are presumably charge transfer complexes in the sense defined by Mulliken.

Our colleague, Dr. Ronald G. Harvey, prepared a 1:1 complex of 7,12-DMBA and 2,4,7-trinitro-9-fluorenone from hot ethanol. This complex (XII) was emulsified in sesame oil and fed once to rats; a single feeding of the complex, 40 mg, induced cancer in 3 of 8 rats (Table II). It would appear that the insoluble complex became dissociated in the intestine and that products were absorbed which incited mammary cancer.

IV. Influence of Heredity on Chemical Carcinogenesis in Rat

Genetic influences in the development of mammary tumors were investigated in females of 2 strains of rats, each strain having a different responsiveness to polynuclear aromatic hydrocarbons in the development of cancer. The strains were heterozygous rather than inbred by brother-sister mating. All of the rats were fed 7,12-DMBA, 20 mg, by stomach tube at age 50 days.

TABLE III
BENIGN AND MALIGNANT MAMMARY TUMORS IN SPRAGUE-DAWLEY AND
LONG-EVANS FEMALE RATS AND THEIR RECIPROCAL HYBRIDS
AFTER FEEDING 7,12-DMBA^a

	Total No.	Benign fibroadenoma	Mammary cancer
SD ♀ × SD ♂	90	0	90
LE ♀ × LE ♂	22	18	4
SD ♀ × LE ♂(F ₁)	24	3	21
LE ♀ × SD ♂(F ₁)	27	6	21

NOTE: 7,12-DMBA, 20 mg, dissolved in sesame oil was administered by stomach tube at age 50 days.

^aSydnor and Huggins (1962).

Sprague-Dawley (SD) rats are albino and are highly susceptible to development of malignant neoplasms of the mammary gland under the stated conditions. Ninety SD females were fed 7,12-DMBA; all developed mammary carcinoma while no rat developed benign tumors (fibroadenoma) of the breast (Table III) in 100 days.

Long-Evans (LE) are piebald and have a low susceptibility to development of mammary carcinoma after feeding aromatic hydrocarbons. Twenty-two rats were fed a single dose of 7,12-DMBA and all developed mammary tumors. Whereas 4 of these rats developed mammary cancer, 18 animals developed fibroadenoma of the breast (Sydnor and Huggins, 1962).

SD and LE strains were interbred and females of the reciprocal crosses were given a *single feeding* of 7,12-DMBA, 20 mg, at age 50 days. The differences in the strains in evocation of mammary cancer disappeared by interbreeding.

In F₁ daughters (SD ♀ × LE ♂) 21 of 24 of the rats developed mammary cancer, whereas benign tumors of the breast appeared in the remaining 3 rats (Table III). In F₁ daughters (LE ♀ × SD ♂) 21 of 27 of the rats developed mammary carcinoma while fibroadenoma arose in the other 6 rats (Table III).

From this experiment it is evident that the susceptibility to malignant mammary neoplasms is transmitted as a dominant trait. Moreover the transmission of this dominant trait occurs as efficiently from father as from mother.

Susceptibility (or resistance) in these strains is concerned exclusively with pituitary function. Puberty occurs early in SD rats and much later in the LE strain. A single shot of 7,12-DMBA at 50 days suffices to induce mammary cancer in every SD rat. When 7,12-DMBA was given intravenously to LE rats on 4 occasions at 3 week intervals after 50 days when rhythmic pituitary function had been established (periodic estrus), resistance disappeared and every LE female developed cancer of the breast.

V. Selective Cell Death Caused by 7,12-DMBA

It is known that certain aromatic hydrocarbons cause systemic effects aside from cancerization. Picard and Laduron (1934) injected emulsions of benzo[a]pyrene into mice and observed atrophy of spleen and bone marrow; these workers concluded that reticuloendothelial cells are specially damaged by the hydrocarbon.

A single feeding of 7,12-DMBA, 20 mg, to rats age 50 days, was followed by leucopenia evident on day 2 (Fig. 2). The circulating leucocytes reached their low point (ca. 50% of normal values) on day 7 but recovery began shortly thereafter and normal levels were reached after 3 weeks. Leucopenia was accompanied by death of many of the cells in spleen and thymus. There was also degeneration

and death of myelocytic cells in bone marrow with bleeding in this tissue.

A remarkable phenomenon was observed in female rats fed a single dose of 7,12-DMBA, 30 mg; an intravenous injection of 5 mg had equal effectiveness. Selective necrosis of the adrenal cortex occurred in every rat (Huggins and Morii, 1961). The changes were observed only in zona fasciculata and zona reticularis; zona glomerulosa and medulla remained uninvolved. Death of cells in the

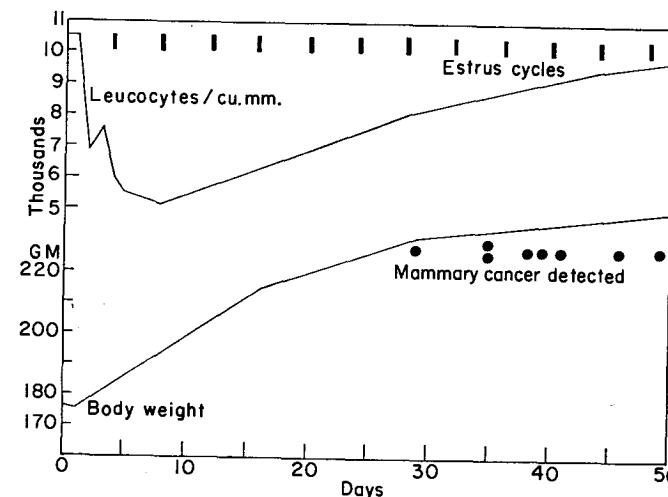


FIG. 2. 7,12-DMBA, 20 mg, was fed to eight rats at day 0.

involved cortical zones was detected within 24 hours. Apoplexy of the adrenals was visible on days 2-7, and most pronounced on day 3. Later regeneration of the adrenal cortex took place, originating in the zona glomerulosa, and the necrotic zones became calcified. Adrenal necrosis occurs in the absence of the pituitary: the hypophysis was removed and 7,12-DMBA was injected intravenously immediately thereafter and adrenal apoplexy was observed on day 3.

The adrenal zones vulnerable to 7,12-DMBA are those in which hydrocortisone is produced; the resemblance of structure between hydrocortisone and 7,12-DMBA is considerable.

It is significant that rats maintained a regular 4-day estrus cycle (Fig. 2) demonstrated in the vaginal smear, despite administration of 7,12-DMBA in amounts sufficient to cause selective cell death in

adrenal cortex and hemopoietic apparatus. Clearly, the extensive and selective damage in the bone marrow and adrenal cortex did not occur in other areas of the adrenal gland or in the pituitary and the ovary.

VI. Discussion

The techniques which have been described represent useful methods for inducing special cancers with speed and extreme simplicity. The conditions necessary for induction of the tumors are highly restricted but are easily satisfied.

One of the strains of rats (SD) which was studied was highly susceptible to the induction of mammary cancer by 7,12-DMBA given at age 50 days whereas another strain (LE) was much more resistant. Susceptibility is transmitted as a dominant trait and as effectively by the father as by the mother. Susceptibility (or resistance) is determined exclusively by pituitary function exerted through structures (ovary, mammary gland) dependent upon it.

The toxic effects of mammary carcinogens are radiomimetic. While both irradiation and aromatic hydrocarbons damage the bone marrow, γ -radiation damages the gonads while hydrocarbons spare the gonads. Again, 7,12-DMBA causes selective adrenal apoplexy whereas irradiation exerts no such effect.

Remarkably irradiation and polynuclear aromatic hydrocarbons are similar in their oncogenic activity: *a single shot of γ -radiation or of specific polynuclear hydrocarbons results in the same effect—mammary tumors are evoked selectively and in high incidence under the stated conditions.*

Selectivity in the site of injury is one of the most striking findings in the experiments which have been described in this paper. Many of the tissues remained essentially uninjured by the hydrocarbons which were simultaneously producing grave lesions in other organs. Necrosis and cell death occurred in bone marrow and in two of the four layers of the adrenal gland after a rather large dose of 7,12-DMBA; in these same animals the ovary and the pituitary were uninjured. Moreover only a rather small number of mammary cancers were induced in the breast by a single administration of the hydrocarbons. Not all of the many millions of mammary cells in a susceptible rat exposed to a single onslaught of the aromatic compounds became malignant—far from it. The largest number of cancers observed after a solitary assault was 21 and the number of

breast tumors observed bore a direct stoichiometric relation to the dose of the hydrocarbon (Huggins *et al.*, 1961a). It is obvious that vulnerability of cells of the body is unequal. In the breast, damage depends in some way on its special biochemical status at the time of the attack (Huggins *et al.*, 1959b).

The damage inflicted on the susceptible adrenal zones by 7,12-DMBA is extensive so that the cell dies but the damage suffered by cells in other organs destined to become cancer is less severe—a dead cell cannot turn into a cancer. Thus, no cancers arose in the damaged adrenals whereas in the vulnerable mammary gland no necrosis was observed.

The damage inflicted on the vulnerable cell which results in cancer cannot be a random spray-like effect. True, a mutation results which expresses itself as a malignant cell, but in the mutated mammary cell, now become malignant, lactic dehydrogenase (Rees and Huggins, 1960) is extremely active. The nucleic acid complex which governs the cellular synthesis of lactic dehydrogenase cannot be destroyed or appreciably weakened by the hydrocarbon. Other genes can be strongly repressed but not that which induces lactic dehydrogenase. Preservation of the cell's ability to synthesize considerable quantities of lactic dehydrogenase is a law of cancer.

All of the special mammary carcinogens are flat molecules with conjugated double bond systems and possess substituent groups of a special sort or, what is equivalent, an additional ring at a salient position in the molecule. The parent molecular species (benz[a]-anthracene, fluorene, stilbene) are devoid of carcinogenic activity; *potency of hydrocarbons in inciting cancer depends on the contribution of the substituents or their equivalent to the ring structure.* Substituents with effective potency are methyl (two methyl groups at special molecular sites are far more effective than a single $-\text{CH}_3$ group at any site), amino, and acetyl amino groups. Ethyl groups are ineffective substituents. The addition of an extra ring (*meso* attached) to benz[a]anthracene to form benzo[a]pyrene or cholanthrene creates compounds nearly as effective as 7,12-DMBA. The common property of $-\text{CH}_3$, $-\text{NH}_2$, $-\text{NHCOCH}_3$ and the salient rings in their ability to donate electrons into the conjugated π -electron system of the rings.

Szent-Györgyi (Szent-Györgyi *et al.*, 1960) discovered that carcinogenicity of aromatic hydrocarbons is connected with their ability to form charge-transfer complexes with local acceptors donating an

electron. Because benzo[a]pyrene is devoid of polar functions such as —OH or —NH₂ groups, the only possible bonding of the intact molecule to biological systems is that of a charge-transfer complex formation (Yang *et al.*, 1961). But charge transfer per se is not enough to cause cancer. Steric factors are involved. In this wise, two compounds, 9,10-dimethylanthracene and 7,12-dimethylbenz[a]anthracene form strong charge-transfer complexes with 2,4,7-tri-nitro-9-fluorenone; of these, only 7,12-DMBA is carcinogenic.

The steric factor is of high importance in the induction of cancer by the special mammary carcinogens. Our colleague, Dr. Yang, has pointed out that there is a direct increase in carcinogenicity as aromatic hydrocarbons become sterically similar to steroids. "For a polynuclear aromatic hydrocarbon to be carcinogenic, it must bear steric resemblance to an active steroid. Among polynuclear aromatic hydrocarbons of similar electronic properties, the closer the steric resemblance to a steroid the higher is the carcinogenicity" (Yang *et al.*, 1961). Science is greatly indebted to A. and B. Pullman (1955) for their magnificent studies of electronic status of hydrocarbons vis-à-vis carcinogenicity. The steric factor supplements the Pullman hypothesis.

Whereas polynuclear aromatic hydrocarbons induce cancer through precise electronic and steric properties which permit selective damage of cells, it is evident that there is no single chemical inducer of cancer; viruses are included in this consideration. It is known that a hit with high energy irradiation (e.g., γ -rays) can accomplish malignant transformation selectively. Any method which will cause a special and restricted damage, as yet incompletely identified, of chromosomal nucleic acid will result in cancer.

Whereas many facets of the induction of cancer by a single dose of polynuclear aromatic hydrocarbons seem to be rather well understood, one great question remains unsolved. Why does cancer of the breast develop invariably while cancer of the uterus never is evoked?

VII. Conclusions

Remarkably, a single sublethal dose of γ -radiation or a single shot of specific polynuclear hydrocarbons evokes an identical oncogenic effect; both agents result in the preferential and selective induction of mammary cancer.

In the induction of cancer by polynuclear aromatic hydrocarbons the following factors are involved:

Solubility, absorption, transport, steric fit, charge transfer, deoxyribonucleic acid, selective special damage. Too great damage to the cell results in its death.

When these conditions are satisfied, the special mutation which characterizes cancer results and it is heritable in somatic cells.

ACKNOWLEDGMENT

We wish to thank Professors H. G. Williams-Ashman, R. G. Harvey, N. C. Yang, and Szent-Györgyi himself for many valuable discussions.

Professor A. Haddow generously furnished 4-dimethylaminostilbene; Professor W. Nakahara kindly gave us 4-nitroquinoline-N-oxide.

This work was aided by grants from the American Cancer Society and from the Jane Coffin Childs Memorial Fund for Medical Research.

REFERENCES

- Haddow, A., Harris, R. C. J., Kon, G. A. R., and Roe, E. M. F. (1948). *Phil. Trans. Roy. Soc. London, Ser. A* **241**, 147.
- Hartwell, J. L. (1951). "Survey of Compounds Which Have Been Tested for Carcinogenic Activity," 2nd ed. National Cancer Institute, Bethesda, Maryland.
- Huggins, C., and Bergenstal, D. M. (1952). *Cancer Research* **12**, 134.
- Huggins, C., and Morii, S. (1961). *J. Exptl. Med.* **114**, 741.
- Huggins, C., Briziarelli, G., and Sutton, H., Jr. (1959a). *J. Exptl. Med.* **109**, 25.
- Huggins, C., Grand, L. C., and Brillantes, F. P. (1959b). *Proc. Natl. Acad. Sci. U.S.* **45**, 1294.
- Huggins, C., Grand, L. C., and Brillantes, F. P. (1961a). *Nature* **189**, 204.
- Huggins, C., Morii, S., and Grand, L. C. (1961b). *Ann. Surg.* **154**, No. 6. Supplement, 315.
- Maisin, J., and Coolen, M.-L. (1936). *Compt. rend. soc. biol.* **123**, 159.
- Nakahara, W., Fukuoka, F., and Sugimura, T. (1957). *Gann* **48**, 129.
- Picard, E., and Laduron, H. (1934). *Compt. rend. soc. biol.* **115**, 1739.
- Pullman, A., and Pullman, B. (1955). *Advances in Cancer Research* **3**, 117.
- Rees, E. D., and Huggins, C. (1960). *Cancer Research* **20**, 963.
- Segi, M. (1960). "Cancer Mortality for Selected Sites in 24 Countries (1950-1957)." Tōhoku University, Sendai.
- Shubik, P., and Hartwell, J. L. (1957). "Survey of Compounds Which Have Been Tested for Carcinogenic Activity," Supplement 1. National Cancer Institute, Bethesda, Maryland.
- Sydnor, K. L., and Huggins, C. (1962). In press.
- Szent-Györgyi, A., Isenberg, I., and Baird, S. L., Jr. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1444.
- Takayama, S. (1960). *Gann* **51**, 139.
- Williams-Ashman, H. G. (1961). Personal communication.
- Yang, N. C., Castro, A. J., Lewis, M., and Wong, T.-W. (1961). *Science* **134**, 386.

Some Perspectives of Antispace Research in Biology

HERMAN M. KALCKAR¹

*McCollum-Pratt Institute and the Department of Biology,
The Johns Hopkins University, Baltimore, Maryland*

I. Introduction	513
II. Pressure Studies on the Architecture of Proteins	515
III. Volume Changes and Denaturation	516
IV. Pressure Studies on Peptide Synthesis and Peptide Degradation	517
V. Studies on Volume Changes in Connection with the Splitting of Peptide Bonds	517
VI. Future Perspectives	518
References	521

I. Introduction

The biological aspects of space research are challenged by those of antispace research—research on the effect of high pressure on physicochemical and biochemical reactions. It is therefore not an unworthy topic for a birthday article dedicated to Albert Szent-Györgyi.

Homer tells us that in "Hades," the underworld of Greek mythology, we shall meet the great heroes of the past. In Poseidon's Hades—the abyssal depths of the sea—also called the hadal or hadopelagic zone, we know that the temperature is very monotonous, always around 4°C. More characteristic for the submarine abyss, however, is the fact that the hydrostatic pressure is very high, ranging from several hundred atmospheres to above one thousand atmospheres. Yet real life exists under these formidable pressures. Several expeditions starting with Alexander Agassiz in 1888 have studied fauna in the depths. The Danish Galatea expedition and the Russian Zenkevich expedition have recently explored the hadal trenches of 30,000 feet in the Philippines and in the Kurile-Kamchatka area. Even at these depths they found a variety of living creatures such as holothurians, echinoid worms, amphipod

¹ Present address: Biochemical Research Laboratory, Harvard Medical School and Massachusetts General Hospital, Boston, Massachusetts.

crustaceans, and sea anemones (Bruun, 1957). In less extreme but still so-called hadal depths (ca. 15,000 feet) eels with fairly well-developed eyes have been encountered (Bruun, 1957). Simpler organisms such as the interesting barophilic bacteria described by Zobell (1952) may well give us the first clue to the biochemical mechanisms underlying biological adaptation to high hydrostatic pressures.

The beginning of laboratory work on the effect of hydrostatic pressure on living organisms commences with Regnard in 1891. This pioneer physiologist first sensed the biological perspectives which this physical vector presents. Regnard found, for instance, that fermentation of sugar by yeast will tolerate 400 atmospheres but will stop at 600 atmospheres (1891). The counterpart is Zobell's work on barophilic bacteria (1952) in which the tolerance of various reactions toward ordinary atmospheric pressure was studied.

Experimental studies on biological reactions as related to pressure have been pursued along two main lines—pressure studies and volumetric studies. A classic treatise of this topic is to be found in the monograph by Johnson, Eyring, and Polissar (1954). The purpose of the present essay is merely to put the spotlight once more on this partially overlooked field. It is felt to be of special interest to try to point out the potential value of pressure studies in modern molecular biology.

It should be made clear at this point that the pressure employed in the experiments to be discussed is always of the hydrostatic type. Each cell or protein molecule is completely surrounded by a liquid medium and this medium serves to transmit the pressure equally in all directions. Any local compression between impinging solid surfaces is apt to give distortions and injuries even at relatively low pressures. Another important feature is the absence of any gas phase in the system. Pressure applied on the medium of a supernatant atmosphere drives extra quantities of the atmospheric gases into solution. This can give rise to many physicochemical effects whether the gases are reactive or inert. However, application of pure hydrostatic pressure should not exert any effect on reaction equilibria or reaction rates unless the reaction under study involves a change in volume. Such a change in volume can be due either to a direct alteration in volume of the solute or to an indirect effect on the solvent such as electrostriction of water molecules through formation of charged groups.

Pressure chambers were designed as early as 1890 by Regnard (1891). Around 1914 Bridgman developed chambers for the use of very high pressures. Pressure chambers with windows for observations and measurements of possible biological responses toward pressure were developed by Marsland and Brown (1936). Recently, Disteche developed a glass electrode for measuring alterations in ionization constants under high hydrostatic pressures and thereupon measured the ionization constants of a number of simple electrolytes in a pressure chamber as well as in the depth of the ocean (1959).

In order to effect a change in a biochemical reaction (alteration of a rate constant or of an equilibrium constant) by means of hydrostatic pressure it is a prerequisite that the reaction in question be accompanied by volume changes. If the volume change of the reaction is positive, the reaction rate tends to decrease by an increase in pressure whereas if a volume constriction occurs, an increase in pressure would tend to increase the rate of reaction. Likewise many equilibria may be effected in the same way. This is only another aspect of the picture formulated by Eyring and Stearn (1939) in their theory on absolute rate.

It should be pointed out that volume changes may be due to many different types of rearrangements. For instance, if we talk about polymers we must not only take into consideration the possible volume changes of the polymer itself. The solvent has to be taken into account as well. This will be evident during the discussion on proteins. Techniques for following changes in volume brought about by chemical reactions were initiated by Rona and Fischgold (1933) and Sreenivasaya *et al.* (1934) and were further developed by Weber (1930) and by Cohn and co-workers (1934) in their classic studies which paved the way for the work of Linderstrøm-Lang (Linderstrøm-Lang and Lanz, 1938; Linderstrøm-Lang and Jacobsen, 1941). Although Linderstrøm-Lang pointedly emphasized that his methods dealt with relative changes and that Weber's work should be referred to for absolute data, his elegant and simple microvolumetric methods provided a decisive stimulus to the development of this field.

II. Pressure Studies on the Architecture of Proteins

The first study on the effect of hydrostatic pressure on protein was performed by Bridgman almost fifty years ago (1914). It was

found that a solution of ovalbumin exposed to 5000 to 6000 atmospheres underwent "pressure coagulation." Later Bridgman and Conant (1929) described the same effect on hemoglobin exposed to 9000 atmospheres. This denaturation and flocculation effect was later studied in more detail by Macheboeuf and Basset (1934), Lauffer and Dow (1941), Johnson *et al.* (1954), Laidler (1958), and most recently by Suzuki and his group (1958, 1960). Most of these studies are solubility studies performed after termination of the pressure phase. However, by means of chambers with windows it was possible to study the state of specific proteins under pressure. Johnson and his group made ingenious quantitative observations on bioluminescent microorganisms under pressure and Suzuki and co-workers recorded changes in the absorption spectra of hemoglobin under pressure.

Work on luminescent microorganisms showed that high hydrostatic pressure inhibits bioluminescence. However, at temperatures around 30° which markedly suppress bioluminescence at one atmosphere, application of high hydrostatic pressure exerts a protective action against the thermo-inactivation (Johnson *et al.*, 1954). This corresponds to observations on a variety of proteins in which it was found that high hydrostatic pressure delays the rate of irreversible denaturation at higher temperatures which are close to the coagulation point (Suzuki, 1958; Lauffer and Dow, 1941). These results are in harmony with Linderstrøm-Lang's observations on volume changes in connection with denaturation (Linderstrøm-Lang and Jacobsen, 1941).

III. Volume Changes and Denaturation

Although experiments on the denaturation of β -lactoglobulin first indicated a lack of volume change by denaturation (Linderstrøm-Lang and Jacobsen, 1941), later experiments (Linderstrøm-Lang, 1949) indicated that this presumably was due to a partial renaturation of β -lactoglobulin. In the presence of 38% urea a considerable volume change took place. The denaturation gave rise to a volume constriction of 6 μ liter per gram which corresponds approximately to 250 ml per mole protein. The establishment of a volume constriction with urea denaturation of β -lactoglobulin also throws light on the anomalously large volume constriction which takes place upon tryptic digestion of this protein (see Section IV). This in turn encouraged Linderstrøm-Lang to propose the theory

that a disintegration of the secondary structure seems to be a prerequisite for the splitting of the peptide bonds of a number of proteins.

IV. Pressure Studies on Peptide Synthesis and Peptide Degradation

Several workers have been intrigued by the possibility of driving protein synthesis by a combination of the catalytical action of proteolytic enzymes, a high concentration of amino acids (or low-molecular weight peptides), and an exposure of the system to high hydrostatic pressures. Bresler and his co-workers at the Polymer Institute in Leningrad have carried out such types of studies (1958). They reported that glycine in high concentration in the presence of pancreatic juice gave rise to the formation of a tripeptide of glycine if exposed to pressures of 6000 atmospheres. Later Bresler extended these pressure studies to crystalline serum albumin. The serum albumin was hydrolyzed for 18 hours with trypsin until the average molecular weight had dropped below 1000 (corresponding to peptides averaging 6 to 7 amino acids). Exposure of this digest to a pressure of 6000 atmospheres at 38° for 18 hours brought about the formation of high-molecular weight albumins much like the original material. Besides the characteristic physicochemical properties of albumin the pressure apparently also restored the antigenic characteristics. These studies so far represent the only successful syntheses of polypeptide chains carried out by means of application of pressure.

Myogen, an enzyme which shows activity as an aldolase against fructose diphosphate, has also been studied along these lines by Bresler (1958). Exposure to high pressure was found to inactivate the aldolase activity except if glucose was present in large amounts in which case some activity was preserved.

V. Studies on Volume Changes in Connection with the Splitting of Peptide Bonds

Although it might have been expected from Bresler's observations (1958) that processes like peptide splitting would yield volume expansion, Weber demonstrated as early as 1930 that peptide splitting in aqueous solutions actually gives rise to a volume constriction. This is caused by the appearance of additional charges (e.g., one dipeptide with one positive charge is split into two amino acids giving two dipoles) which surround themselves with com-

pressed water molecules. As emphasized by Linderstrøm-Lang and Jacobsen (1941) the charges of opposite signs are so close together in the free amino acid that their resulting influence upon surrounding water molecules is somewhat smaller than that of an ion pair. Peptides larger than 3 or 4 units exert an effect close to that of an ion pair. They estimated that the enzymatic breaking of a terminal peptide bond would give a volume constriction of about 13 ml per mole split whereas a splitting in the middle of a longer peptide chain should give a constriction 30 to 40% higher. Linderstrøm-Lang and Jacobsen found, however, that in the enzymatic splitting of β -lactoglobulin a constriction as high as 50 ml per mole (i.e., almost 400% of the value found for the release of one amino acid) was found. They suspected that in this case a change in secondary structure was involved too (see Section II).

VI. Future Perspectives

One of the fields of biology in which pressure studies are likely to make themselves felt is in the study of macromolecules. Pressure chambers with windows fitted to spectrophotometers, spectropolarimeters, and fluorometers will become available if they are not so already.

Some intriguing studies carried out by McElroy and de la Haba (1949) point toward the relevance of studies on the effect of pressure gradients on the structure of *deoxyribonucleic* acid (DNA) and *polynucleotides*. It was found that high hydrostatic pressure decreases the mutation rate of *Neurospora*. The application of about 600 atmospheres pressure to a conidial suspension containing nitrogen mustard depressed the number of morphological mutants by a factor estimated to be 45 to 50%. These observations are obviously difficult to interpret at present in the context of modern molecular biology. Yet they may in time suggest certain simple experiments on DNA or polynucleotide biosynthesis. It should be mentioned here that the equilibrium between single-stranded and double-stranded DNA may well be significantly influenced by hydrostatic pressure. It is known that the buoyant density of single-stranded DNA is higher than that of double-stranded DNA. Corresponding volume change accompanying the transition from double to single strands may possibly ensue (presumably volume constriction in the direction stated). This in turn would make the equilibrium susceptible to pressure changes. Depending on the molecular volume changes,

it is even possible that the pressures brought upon DNA in the cesium chloride gradient of the ultracentrifuge might affect the distribution of the two molecular species with a significant increase in abundance of the single-stranded species. Some of the pressure effects generated in the liquid of the cell of the ultracentrifuge are now under study (Hearst, Ift, and Vinograd, 1961).

The complicated apparatus of *protein synthesis* centers around the submicroscopic particles, the ribosomes. One or more steps in this process may be significantly affected by high pressure. This applies likewise to the apparatus of regulation (repression and de-repression). Mutants with a temperature-sensitive repressor system have been described recently (Horiuchi *et al.*, 1961).

Johnson's (1958) early pressure studies on bioluminescence foreshadowed studies on purified enzymes (cf. Johnson *et al.* 1954). McElroy and his co-workers (1949) have made fundamental contributions toward our knowledge of the properties and function of crystalline firefly luciferase (McElroy and Seliger, 1961). Pressure studies on this system as well as on the purified bacterial luciferase may widen our knowledge of these systems further. There are a number of other enzyme systems which might lend themselves toward useful pressure studies in a window chamber. Chance's hydrogen peroxide catalase complexes (1950) which give rise to characteristic spectral changes are of considerable interest. The steady state levels of the enzyme substrate complexes might be markedly affected by increasing the pressure several hundredfold. Another enzyme which might lend itself particularly well to optical observations during exposure to high pressure is the crystalline glutamate dehydrogenase system. This system which catalyzes the reduction of diphosphopyridine nucleotide (DPN) by glutamate was first found by Frieden (1959) to undergo reversible transitions in molecular weight under the influence of various nucleotides. Thus adenosine diphosphate (ADP) stabilizes the active large molecule whereas reduced DPN (DPNH) facilitates the splitting into smaller inactive protein units. Interest in this system was further stimulated by the recent work of Tomkins and co-workers (1961) who showed that a number of biologically active steroids are able to bring about a splitting much like that effected by DPNH yet independent of this nucleotide. In this case ADP also counteracted splitting. Of additional interest is the novel finding that the active low-molecular weight glutamate dehydrogenase has acquired a

greatly increased catalytical action toward the dehydrogenation of alanine. The ability of the hydrophobic ring structure of steroids to change the specificity of an enzyme is clearly of great interest and one wonders how this reversible system would be influenced by high pressure. The nature of this splitting is not as yet understood.

Of great interest in this context are the observations by Fincham on *Neurospora* mutants with a temperature-sensitive glutamate dehydrogenase (1957, 1959). It so happens that certain *Neurospora* mutants have a defect in their glutamate dehydrogenase which renders the enzyme highly sensitive to temperature. However, unlike most of the other temperature-sensitive enzymes in which the enzyme is active below 25° to 30°C and inactive at 37° this type of mutant behaves in exactly the opposite manner. The enzyme is inactive at temperatures around 20° to 23° and only becomes active in the temperature range between 35° and 50°. It seems evident from what has been said earlier that hydrostatic pressure added as an additional vector might have an effect on either type of aberrant temperature-sensitive enzyme. For instance, the type of temperature-sensitive enzyme which is rendered inactive above 25° to 30° might tend to become more stabilized at the transition temperature if subjected to the effect of high hydrostatic pressure. In the second type it would be particularly intriguing to see whether an increase in pressure below the temperature activation point would "break the ice."

The study of changes in *secondary and tertiary structure* as related to pressure seems especially promising in the case of ribonuclease. The use of optical rotation as a criterion for helix formation would be particularly appropriate because it would make it feasible to make quantitative observation of changes under pressure. This applies too to studies on synthetic polypeptides, the proline helix, actomyosin, and hemoglobin and related heme proteins.

Recently Racker (1961) and his colleagues discovered a most interesting protein which carries adenosinetriphosphatase (ATPase) activity and which is required for mitochondrial oxidative phosphorylation. They found that this protein is highly sensitive to cooling which brings about inactivation, seemingly irreversibly. It might be most appropriate to subject this system to changes in pressure as well as temperature. Perhaps the system is related to that of myosin contraction and relaxation. Lehninger has found, for instance, that ATP elicits contraction in mitochondria (1960).

The protein opsin forms chromogenic complexes with vitamin A aldehyde. According to Wald and his co-workers (1957) rhodopsin, upon exposure to a light quantum is converted to lumirhodopsin in which the shape of vitamin aldehyde has been converted to the straight all-trans-form. The all-trans-chromoprotein lumirhodopsin yields a second all-trans-chromoprotein metarhodopsin in which the reactive groups in opsin (two sulphydryl groups and one proton-binding group) are exposed. Vertebrate metarhodopsin is unstable and hydrolyzes to opsin and all-trans-retinene. The formation of metarhodopsin which is suspected to be responsible for triggering the visual excitation process may very well be associated with volume changes. Hence the steady state level of this intermediary may be affected by pressure.

Wald and co-workers (1957) found an interesting correlation between the location of the absorption maximum of rhodopsins of marine fishes and their depths of habitat. Thus the lancet fish and other species living at depths greater than 400 meters (ca. 200 fathoms) and therefore exposed to hydrostatic pressures of at least 50 atmospheres or more contain a yellow visual pigment which has been called chrysopsin. However, since the vitamin aldehyde part is the usual neo-b (11-cis)-retinene whereas the opsin is different, Wald prefers to call these pigments simply deep-sea rhodopsins. This adaptation may not be related to hydrostatic pressure but rather to the fact that only blue light reaches the great depths and therefore a yellow-orange pigment would be more sensitive to this end of the spectrum. Yet the difference in the structure of opsin may still be advantageous in securing a rapid triggering of visual excitation under the high hydrostatic pressures at the depths of the "Seven Seas."

ACKNOWLEDGMENTS

It is my pleasure to thank Drs. Chance, Edsall, Harrington, and Seliger, and Mr. R. Newton for many interesting and useful discussions on this topic.

REFERENCES

- Bresler, S. E. (1958). *Advances in Protein Chem.* **13**, 478.
- Bridgman, P. W. (1914). *Proc. Am. Acad.* **49**, 11.
- Bridgman, P. W., and Conant, J. B. (1929). *Proc. Natl. Acad. Sci. U.S.* **15**, 680.
- Bruun, A. F. (1957). *Geol. Soc. Am., Mem.* **67**, 641-672.
- Chance, B. (1950). *Advances in Enzymol.* **12**, 153.
- Cohn, E. J., McMeekin, T. L., Edsall, J. T., and Blanchard, M. H. (1934). *J. Am. Chem. Soc.* **56**, 784.

- Disteche, A. (1959). *Rev. Sci. Instr.* **30**, 474.
- Eyring, H., and Stearn, A. E. (1939). *Chem. Revs.* **24**, 253.
- Fincham, J. R. S. (1957). *Biochem. J.* **65**, 721.
- Fincham, J. R. S. (1959). *Ann. Rev. Biochem.* **28**, 343.
- Frieden, C. (1959). *J. Biol. Chem.* **234**, 809 and 815.
- Hearst, J. E., Ift, J. B., and Vinograd, J. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 999.
- Horiuchi, T., Horiuchi, S., and Novick, A. (1961). *J. Mol. Biol.* **3**, 703 (1961).
- Jacobsen, C. F. (1941). *Compt. rend. trav. lab. Carlsberg, Sér. chim.* **24**, 281.
- Johnson, F. H., Eyring, H., and Polissar, M. (1954). "The Kinetic Basis of Molecular Biology," p. 297. Wiley, New York.
- Laidler, K. J. (1958). "Chemical Kinetics of Enzyme Action," p. 210. Oxford Univ. Press (Clarendon Press), London and New York.
- Lauffer, M. B., and Dow, R. B. (1941). *J. Biol. Chem.* **140**, 509.
- Lehninger, A. L. (1960). *Federation Proc.* **19**, 952.
- Linderstrøm-Lang, K. (1949). *Cold Spring Harbor Symposia Quant. Biol.* **14**, 117.
- Linderstrøm-Lang, K., and Jacobsen, C. F. (1941). *Compt. rend. trav. lab. Carlsberg, Sér. chim.* **24**, 1.
- Linderstrøm-Lang, K., and Lanz, H. (1938). *Compt. rend. trav. lab. Carlsberg, Sér. chim.* **21**, 315.
- McElroy, W. D., and de la Haba, G. (1949). *Science* **110**, 640.
- McElroy, W. D., and Seliger, H. (1961). In "Symposium on Light and Life" (W. D. McElroy and B. Glass, eds.), p. 219. Johns Hopkins Press, Baltimore, Maryland.
- Macheboeuf, M. A., and Basset, J. (1934). *Ergeb. Enzforsch.* **3**, 303.
- Marsland, D., and Brown, D. E. S. (1936). *J. Cellular Comp. Physiol.* **8**, 167.
- Racker, E. (1961). Symposium on Regulation of Metabolism, 45th Ann. Meeting, Fed. Am. Soc. Exptl. Biol.
- Regnard, P. (1891). "Recherches experimentales sur les conditions physiques de la vie dans les eaux," 501 pp. Masson, Paris.
- Rona, P., and Fischgold, H. (1933). *Biochem. Z.* **261**, 66 and 366.
- Sreenivasaya, M., Sastry, B. N., and Sreerangachar, H. B. (1934). *Biochem. J.* **28**, 351.
- Suzuki, K. (1958). *Rev. Phys. Chem. Japan* **28**, 24.
- Suzuki, K., and Kitamura, K. (1960). *Rev. Phys. Chem. Japan* **29**, 81.
- Tomkins, G. M., Yielding, K. L., and Curran, J. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 270.
- Wald, G., Brown, P. K., and Brown, P. S. (1957). *Nature* **180**, 969.
- Weber, H. H. (1930). *Biochem. Z.* **218**, 1.
- Zobell, C. E. (1952). *Science* **115**, 507.

Water¹

IRVING M. KLOTZ

*Department of Chemistry,
Northwestern University,
Evanston, Illinois*

I. Introduction	523
II. Pure Substance	523
A. Some Physical Properties	523
B. Molecular Structure and Behavior	526
III. Hydration	530
A. Ionic Solutes	530
B. Apolar Solutes	536
C. Proteins	539
D. Tissues	544
IV. Some Speculations on Energy and Charge Transport	545
References	549

I. Introduction

The title of this paper, "Water," presumably allows me to range over a wide area of knowledge, which is not my intention. Rather, I wish to emphasize a molecular, or even an atomic, viewpoint and to limit myself to a few subjects that might be of physiological significance. It is my intention to start with a description of some of the molecular properties of water and of its solutions of ionic and apolar molecules. With these facts and some of the concepts they have generated, I shall proceed to an interpretation in terms of hydration of some aspects of the behavior of protein macromolecules. Finally, I shall speculate about the possible significance of these ideas in regard to the physiological properties of tissues.

II. Pure Substance

A. SOME PHYSICAL PROPERTIES

Let us start with a consideration of a few selected properties of the pure substance.

¹ Based on a lecture given at the Marine Biological Laboratory, Woods Hole, Massachusetts, July 1961.

There are, of course, three states of water, but one of these, the gaseous state, is of little significance for our purposes. I shall make a table, therefore, comparing some of the characteristics of only the liquid and ice states. As the first item (Table I) I have listed the concentration of hydrogen ions due to the self-dissociation of water itself. As everyone is aware, the value for the liquid state is approximately $1 \times 10^{-7} M$. Many may not be aware that a value has also been measured (Eigen and De Maeyer, 1958) in the ice state and is about $1 \times 10^{-10} M$ (see Table I).

TABLE I
A FEW SELECTED PROPERTIES OF WATER

Property	Liquid	Ice
Conc. of H^+ ion: C_{H^+}	$1 \times 10^{-7} M$	$\sim 1 \times 10^{-10} M$
Electrical conductance	$1 \times 10^{-8} \text{ ohm}^{-1} \text{ cm}^{-1}$	0.3×10^{-8}
Mobility of H^+	$0.0036 \text{ cm}^2 \text{ volt}^{-1} \text{ sec}^{-1}$	~ 0.3
Rate constant $H^+ + OH^- \rightarrow H_2O$	$10^{11} \text{ liter mole}^{-1} \text{ sec}^{-1}$	$10^{13} - 10^{14}$
Mobility of Li^+	0.0004	$< 10^{-8}$
Diffusion coefficient	$2.4 \times 10^{-5} \text{ cm}^2 \text{ sec}^{-1}$	8×10^{-11}

Turning to the second item in Table I, electrical conductance, we find that in contrast to the H^+ ion concentrations which are about a thousandfold different between the liquid and ice state, the specific conductivities (at 0°C) are more nearly the same, 1×10^{-8} in the liquid, about 0.3×10^{-8} in the solid state (Eigen and De Maeyer, 1958). It is very surprising to find that the hydrogen ion concentration can differ by a thousandfold factor and yet the conductance differ very little. The significance of this apparent discrepancy in behavior can be realized from a consideration of the molecular basis of electrolytic conductance. In essence, what we are comparing in conductance is the ability to carry current, in other words, to transport charges across a boundary in a particular unit of time. The transport of charges across the boundary, furthermore, depends on two factors, the number of charges or ions which are available (per unit volume) in the solution, and the speed (mobility) with which they move. Even if a few ions are present, if they move very fast, the conductance will be great because the number of charges which pass a specific plane in the solution will be great. The mobility of hydrogen ion in liquid water, 0.0036, is

a reasonable value in solution for a small ion as we shall see later. Since the conductance in the ice state is only slightly different from that in the liquid state, and yet the number of hydrogen ions in the ice state is very much smaller, it follows that the mobility of H^+ in the ice state must be a very much larger number than the mobility of hydrogen ion in the liquid state. Thus we find (Table I) for the mobility of hydrogen ion in ice a value of about 0.3 (Eigen and De Maeyer, 1958), almost 100 times greater than the 0.0036 for the hydrogen ion in the liquid state. We arrive at the important conclusion from conductivity measurements that the mobility of hydrogen ion in ice is an exceedingly large number.

Values of mobility, in essence, refer to the speed with which an ion will move at a voltage drop of 1 volt per centimeter, a rather small value. In practice, in cellular physiology, voltage drops would be much greater than 1 volt per centimeter and it follows, of course, that the speed of motion of hydrogen ions in ice at such high voltages could be very much greater too.

Related to the very unusual mobility of hydrogen ion in ice is the very large difference in the rate constants observed for the combination of an H^+ ion and an OH^- ion in water in the liquid versus the ice state. In the liquid state, as Table I shows, the value of the rate constant (Eigen and De Maeyer, 1958) is about 1×10^{11} . In effect, this means that the rate of the reaction is roughly the reciprocal of this value and hence that H^+ ion and OH^- ion in water can combine in about 10^{-11} seconds. This is essentially the rate which would occur if the H^+ ion and the OH^- ion approach each other at the speed with which diffusion toward each other can occur. Turning to the ice state we find that the rate of combination of H^+ ion and OH^- ion is even larger, by a factor of 100–1000 (Eigen and De Maeyer, 1958), than the value for the same reaction in the liquid state. Thus again we see there must be something very unusual about the mobility of H^+ or OH^- ions in the ice state, for they come together even faster than they would in a diffusion-controlled reaction in the liquid state.

Let us turn, now, to another property of ions in water which will be useful in connection with our discussion of the unusual mobility of hydrogen ions. Listed next in Table I is the mobility of lithium ion, the value of which is 0.0004. This value is about 10 times lower than that for hydrogen ion in the liquid state but still not an unusually small number. In contrast, in the ice state the

mobility of lithium is an exceedingly small number, much less than 10^{-8} . Furthermore, comparing hydrogen ion with lithium ion, we see that for the latter ion there is an enormous drop in the mobility in the solid ice state, whereas in the case of the former, hydrogen ion, there is not only a slow-down in ice but an actual increase in the mobility.

Finally, in Table I, the diffusion coefficients of a water molecule in liquid water (Wang *et al.*, 1953) and in ice (Gränicher, 1958) are listed to show that the value in ice, again, is slowed down quite substantially compared to that in liquid water, by almost a factor of a million.

B. MOLECULAR STRUCTURE AND BEHAVIOR

How do we interpret some of these unusual properties of water?

Let us look at Fig. 1 which is a schematic diagram, summarizing the essential information on the structure of ice and liquid water.

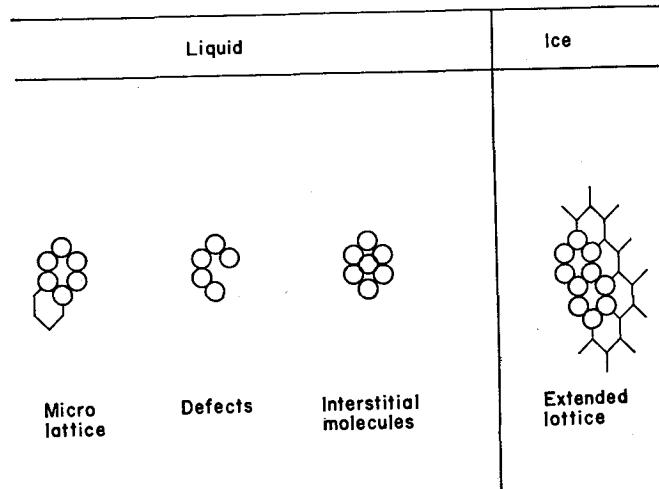


FIG. 1. Schematic diagram of structure of water in liquid and in solid states.

As is indicated very schematically in merely two dimensions, the basic arrangement of water molecules in ice is in a hexagonal array (Bernal and Fowler, 1933; Lonsdale, 1958) with an extended lattice, penetrating or extending far out into the ice crystal. Each molecule is bonded to its neighbors by hydrogen bonds. The lattice has a

large number of holes in it (one within each hexagon), as can be seen from the schematic structure (Fig. 1), and is thus a very open structure. In fact, if it were completely close-packed, the density of ice would be more nearly 1.6 (Gränicher, 1958), rather than slightly less than 1 as is actually found.

Let us turn to the liquid structure. There are three important distinctions to make in contrasting the liquid with the ice structure. The actual molecular packing is still uncertain and may be hexagonal (Morgan and Warren, 1938) or pentagonal (Pauling, 1959). For purposes of schematic illustration (Fig. 1), we will assume that the crystal structure is a hexagonal one, as in the case of ice. One important distinction we immediately must emphasize then, is that in the liquid the crystal is a microcrystal or a microlattice, with the hexagons extending only, perhaps, 10 Å into the body of the liquid instead of essentially continuously throughout the entire solid, as they do in the ice state. A second distinction in the liquid state is that there are defects in this crystalline lattice, as is indicated in Fig. 1. Thirdly, in the liquid, we find interstitial molecules, that is, a few water molecules that have managed to get into the open space in the hexagons. It is these interstitial molecules that account for the somewhat higher density of liquid water around 4°, than that of water slightly below that temperature. In the liquid state, too, water molecules, are connected by means of hydrogen bonds. In the liquid state, however, the hydrogen bonds break and form much more rapidly than in the case of ice. Such behavior is evident, for example, from the diffusion coefficient. The diffusion coefficient is effectively a measure of the distance that a water molecule will go due to thermal motion alone, under certain concentration conditions. And as is evident (Table I), the value for the molecule of water in liquid water is almost a million times greater than in ice, or vice versa, the value in ice is a million times slower, and so we conclude that the time for breaking of bonds in the liquid state is much shorter, the process much faster, than it is in the solid state.

Recognizing that water is a hydrogen-bonded structure, let us examine the mechanism of the mobility of hydrogen ion and of other charges in water in the liquid or in the solid state (see Fig. 2).

The top line (Fig. 2) shows the mechanism for the transport of protons. As is indicated, if a proton happened to approach a water molecule at the left of the chain, it could attach itself to that water molecule. If it did so, and formed a covalent bond, then one of the

hydrogens which previously had been covalently bonded to that oxygen, and hydrogen-bonded to a neighboring oxygen, would now flip over toward the right, to the oxygen with which it had been previously hydrogen-bonded, and would form a new covalent bond. In turn, the middle oxygen or the middle water, would transfer one of its hydrogens over to the right as a hydrogen came from the left. In this way, as we see on the right of the arrow of the top entry of Fig. 2, we would, in essence, produce a free proton at the right side of the chain of water molecules. This chain-bucket mechanism allows a hydrogen ion which has appeared on the left side of the chain to essentially produce a hydrogen ion on the right side of the chain. Thus, the very high mobility of hydrogen ions in water is not due

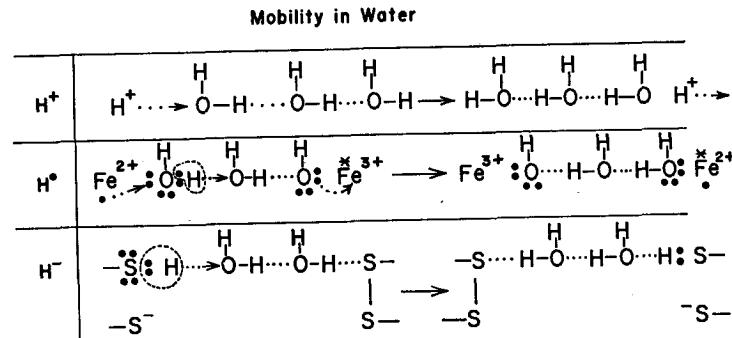


FIG. 2. Mobility of H^+ , H^- , and H^- through chains of water molecules.

to the mobility or transport of the specific hydrogen ion which was introduced at the left, but rather of a new one which appeared at the end of a hydrogen-bonded water brigade, at the right-hand side of the chain (Fig. 2).

A similar mechanism has been suggested for the transport of an electron, or of two electrons, in water. In the case of a single electron attached to a hydrogen we have a hydrogen free radical; and it has been suggested (Dodson, 1952) that in oxidation-reduction reactions which often occur between metals of different (positive) charge, actually it may be a hydrogen radical which is carrying the electrons. For example, if we have, in solution, some Fe^{2+} ion and mix it with some radioactive $*Fe^{3+}$ ion, we may find after a short while that the radioactive $*Fe^{3+}$ ion disappears and radioactive $*Fe^{2+}$ ion has appeared. Effectively, this means that an electron must have been transferred from the original ferrous to the original radioactive ferric.

But it is very difficult to conceive that two ions of such enormously high charge could approach each other closely enough so that an electron could tunnel its way from the one ion to the other. On the other hand, we recognize that, in solution, these ions are hydrated and draw, as is indicated in the middle entry of Fig. 2, a number of water molecules, hydrogen-bonded to each other, with one on the left also coordinated with the ferrous ion. Furthermore we assume that the ferric ion is very likely to be coordinated with an OH group because of its very strong acidic properties. Then we can see, as is indicated in Fig. 2, that an electron from the Fe^{2+} ion could go into the oxygen valence shell, if the water molecule with that oxygen would give up a hydrogen plus an electron, or hydrogen-free radical, to the neighboring water molecule on the right. The neighboring water molecule, in turn, can relay the effect, that is, it can receive the hydrogen plus an electron, if, in turn, it will give up one of its hydrogens plus an electron, or one of its H^- free radicals. This transfer can continue along the chain of water molecules until an H^- free radical moves over to the OH^- which is attached to the ferric ion. It can move over to the OH^- attached to the ferric ion if one of the electrons on the oxygen attached to the ferric ion is now allowed to jump over to the $*Fe^{3+}$ to convert it to $*Fe^{2+}$. In this way, we get the transition indicated in the middle entry of Fig. 2 with the final product, shown on the right-hand side of the arrow, consisting of the original $*Fe^{3+}$ converted to $*Fe^{2+}$ ion, and the original Fe^{2+} ion now being Fe^{3+} ion, the electron having been transferred by means of an H^- free radical through a bridge of water molecules.

A similar mechanism is shown in the last entry in Fig. 2 for the transport of an H^- with two electrons, a hydride ion (Klotz *et al.*, 1958). In this case, the transport is presented as occurring from a reducing agent consisting of an SH group to an oxidizing agent consisting of an SS group. Again (on the left side of the arrow in the bottom row of Fig. 2), water molecules are represented as forming a bridge between the SH group and the SS. If a hydride ion from the SH, that is, an H atom with 2 electrons, moves from the SH with the 2 electrons to the adjacent water molecule, then the S group remaining on the left-hand side is now deficient in 2 electrons and can form an SS bridge with a neighboring S⁻ or SH group. The H⁻ in turn can move to the water but only if the oxygen can give up 2 of its electrons, as otherwise it would still be saturated. The oxygen could give up 2 of its electrons if one of its hydrogens would move

to a neighboring water molecule along with 2 electrons from that oxygen, in other words, if a hydride ion will move again to a further adjacent H_2O molecule. Finally, if the water molecule adjacent to the SS group will give up a hydride ion to the SS group, then we would form, as is indicated on the right-hand side of the arrow in Fig. 2, an S^- (or SH) since 2 electrons would have been carried over to the SS group. Again, then, we see the basis for a long-range oxidation-reduction, involving, in this case, 2 electrons, the essential mechanism being that a hydride ion carry the 2 electrons through a bridge of water molecules.

On the basis of these brigade mechanisms it is easy to understand why hydrogen ion shows unusual mobilities in the ice state as compared to liquid water. In the mobility of H^+ ion, as we visualize the mechanism, an H^+ is transported from one water molecule to another. It follows, therefore, that the rate of over-all transport will depend on whether the hydrogen bridges are extensive and whether hydrogen ion, furthermore, can jump across the bridge from one oxygen to the other. In liquid water, the hydrogen bridges are not too extensive, but, as we indicated, are formed in any moment only for a short distance and are continually changing. And so, even though the mobility of hydrogen ion is greater than that of lithium because there are some bridges, it is not enormously greater than that of a normal ion in water. On the other hand, in the ice state, the hydrogen bridges are formed over enormous distances and the limiting factor in the rate is merely the limiting rate at which hydrogen ion can tunnel its way from one water molecule to another. Because these bridges in ice are formed over such an enormous distance, the mobility of hydrogen ion is an enormously large value, approaching that of semiconductors.

Likewise, we can see why the rate constant or speed of the combination of hydrogen ion with hydroxide ion is so enormously large in the case of ice. Once again, the H^+ ion, and indirectly the OH^- ion, can move through the water lattice with enormous speed because of the brigade mechanism.

III. Hydration

A. IONIC SOLUTES

Let us turn, now, from the pure substance, water, to solutions in water, or, more specifically, to questions of hydration.

I should like to start by considering the hydration of ionic solutes. Let me list first a few of the properties of ions in solution. Assembled in Table II are a few typical hydration energies (Conway, 1952; Eigen and De Maeyer, 1958). As can be seen, for hydrogen ion the hydration energy is a very large number, 276,000 cal/mole. That of some of the other univalent cations, lithium, sodium, potassium, is not quite so large, and is different for each one of these cations, the trend showing that the hydration energy is dependent not only on the charge but also on the nature of the ion. Likewise, to indicate the very strong charge dependence Table II lists values for divalent calcium and for trivalent aluminum ions. The exceedingly large energies of hydration show, of course, that there must be an enormously strong interaction between each of these ions and water molecules.

TABLE II
HYDRATION ENERGIES OF IONS

Ion	$-\Delta H$ (kcal mole $^{-1}$)
H^+	276
Li^+	131
Na^+	116
K^+	92
Ca^{++}	410
Al^{+++}	1149

Let us turn, next, to a few values of the heat capacities of ionic compounds in solution (Harned and Owen, 1958). For this purpose I should like to look, first, at a graphical presentation of the results, as is shown in Fig. 3. If we plot the heat capacity as a function of the moles of solute, we might indicate at the outset, by the broken horizontal line, the value which one would observe for the pure solvent itself. It is a striking fact, that if you then add a salt to make a solution, the heat capacity of that solution is always less than that of the pure solvent itself (Fig. 3). Numerically, some examples are shown in Table III. With sodium chloride, for example, the heat capacity of the solid is a positive number, meaning, of course, one must put in some heat to raise the temperature. Yet the effective value in solution, \bar{C}_p , so-called, is a negative number, corresponding to the negative slope in Fig. 3, and indicating that when you add

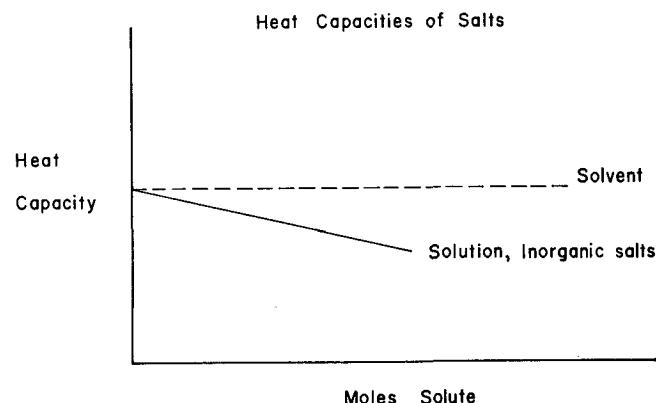


FIG. 3. Decrease in heat capacity upon formation of aqueous solution of inorganic salt.

sodium chloride to the solution it takes less heat to raise the temperature of that solution 1° than it took to raise the temperature of the same amount of pure solvent before the sodium chloride was put in.

Equally interesting properties of salts can be found in many books on electrolytic compounds (Conway, 1952; Harned and Owen,

TABLE III
HEAT CAPACITIES OF IONIC COMPOUNDS

Substance	C_p for solid	\bar{C}_{p_2} in aqueous solution
LiCl	$12 \text{ cal mole}^{-1} \text{ deg}^{-1}$	-15.63
NaCl	11.9	-23.8

1958). Another property has been indicated in Table IV, the volumes of some ionic compounds. As the table shows, for salts such as lithium chloride and sodium chloride, the volume of 1 mole of solute is of the order of 20 cc. The effective volume of the solute when it is added to an aqueous solution is less. In other words, the sum of the volumes of the pure solvent plus the pure solid solute actually turns out to be a greater number than the observed volume of the mixture in solution; there has been a decrease in total volume on dissolving. It is perhaps interesting, in connection with this table, to point out, furthermore, that for a few substances, of which sodium

hydroxide is an example, the volume of a solution composed of a mole of sodium hydroxide dissolved in a liter of water is actually less than the volume of a liter of the pure water itself. That is the significance of the figure of -6.7 in the table under \bar{V}_2 for this substance in solution. In connection with ionic compounds the table also shows that a dipolar ion such as glycine is more or less similar to a completely dissociated ionic compound such as sodium chloride, in that the effective volume in aqueous solution is less than that of the pure crystalline compound.

TABLE IV
VOLUMES OF IONIC COMPOUNDS

Substance	V for solid	\bar{V}_2 in aqueous solution
LiCl	$20.5 \text{ cc mole}^{-1}$	17.00
NaCl	27.0	16.40
NaOH	18.8	-6.70
Glycine	46.7	43.199

These properties, plus a variety of other phenomena of ionic aqueous solutions, have led to a rather generally accepted model for ionic hydration in water. This model (Frank and Wen, 1957) is shown in Fig. 4. In essence, we visualize around the central ion, indicated in the figure as a positive ion, four to six very strongly and rigidly held water molecules, irrotationally bound, as they are sometimes called. These strongly held water molecules are meant to be a representation of the very large energies of hydration. If we move to the outer edge of the diagram we would have normal liquid water, hexagonal crystals, fluctuating crystals, with a tendency to bring other water molecules into the same hexagonal microcrystalline array. In between these two regions is an intermediate region which, as we have indicated, would be disorganized, because of two competing attractive effects. The centro-symmetric field of the ion is attempting to align these intermediate water molecules in the specific array shown around the cation at the center of the diagram. On the other hand, a normal liquid is attempting to align these water molecules in a hexagonal array. The net result of these competing effects is, however, that the water in the intervening region is disorganized rather than organized.

On the basis of this model one can account for the essential properties of ionic aqueous solutions. Let me mention just the heat capacities listed in Table III. There are a number of factors which change heat capacities in solution. I want to emphasize only one. In essence, the heat capacities are a measure of the amount of heat necessary to raise the temperature of the solution 1° . In pure water this heat

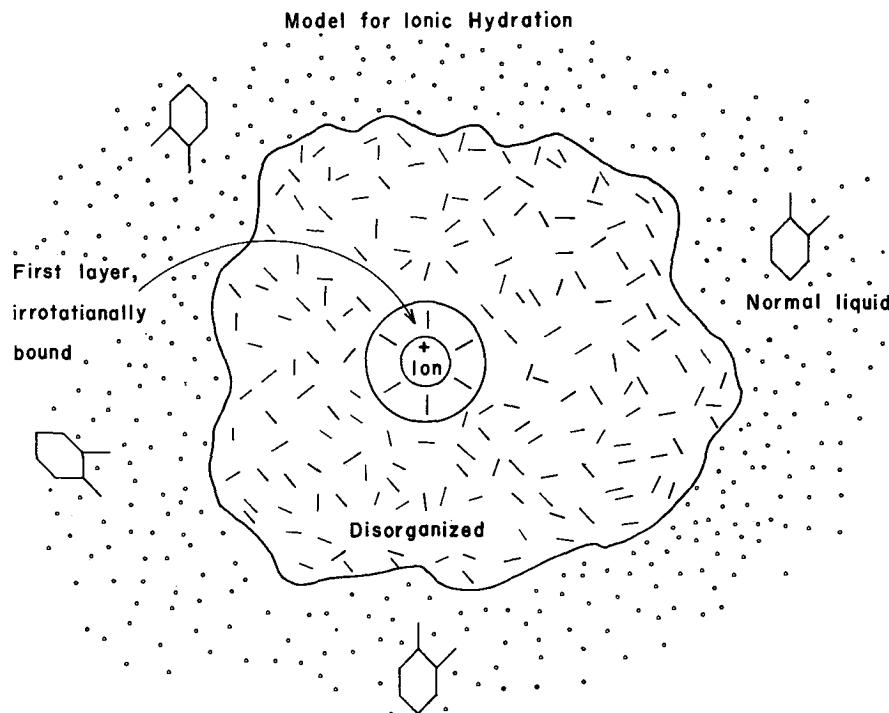


FIG. 4. Model of structure of water in neighborhood of inorganic ions.

is used largely to melt some microcrystals. If an ion has already disorganized some of the water in the solution, then it is reasonable to expect that it should take less heat to raise the temperature of the water 1° since the ion, itself, has accomplished some of this disorganization by its own efforts. In this way, then, in terms of the disorganized middle region, we can account for various heat capacity effects and particularly for the changes of heat capacity with the size of the ions (Frank and Wen, 1957).

I should like to insert, at this point, a little peripheral informa-

tion which we will not particularly need subsequently, but which might be of interest in connection with the discussion of the hydration of ions. Even though some of the ions hold water in the first layer very rigidly, as is indicated by the very strong hydration energies, we must keep in mind that the lifetimes of these bound water molecules are not necessarily very long in terms of our common experience, although at the molecular level they are very long, indeed. Some typical lifetimes are shown in Table V. For orientation purposes, let us look, first, at the time of a molecular vibration, 10^{-13} sec. Water in a liquid cluster has a lifetime of about 10^{-10} sec (Frank, 1958), one thousand times longer than the period of a molecular vibration, nevertheless, not a very long time in human

TABLE V
LIFETIMES OF BOUND WATER

(Molecular vibration)	10^{-13} sec
Water in liquid cluster	10^{-10}
First layer hydration	
Mn^{2+}	10^{-7}
Cu^{2+}	10^{-6}
Ni^{2+}	10^{-4}
Al^{3+}	$>10^{-4}$

experience. When water is bound in the first layer to an ion, then it is held in that state for a longer time before it exchanges with the outside water. Values for monovalent ions do not seem to be available but I have been able to find some for some divalent ions and trivalent ions (Connick and Poulson, 1959; Jackson *et al.*, 1960) and these will give the proper orders of magnitude. As is shown in Table V, for manganese ion the lifetime of a bound water molecule is about 10^{-7} sec. A divalent ion may have a specific effect; as is shown in the table, copper has a somewhat stronger hold on the water, holding it for about 10^{-6} sec, nickel holding it for 10^{-4} sec. Charge, as you would expect from the increasing hydration energy, is an important factor, and, as is shown in the table, aluminum holds water for longer times, somewhat more than of the order of 10^{-4} sec. Now 10^{-4} sec may seem like a very short time but, on the molecular scale, it is about one billion times longer than the period of vibration or rotation of a normal molecule.

B. APOLAR SOLUTES

Let us turn, now, to an examination of some of the properties of apolar solutes as they reflect on the hydration of these apolar molecules.

I should like to consider schematically some of the heat capacity results, which are shown in Fig. 5, that have been obtained with a largely apolar molecule, tetrabutylammoniumbromide. As Fig. 5 shows, if we plot heat capacity of the solvent alone, we could draw

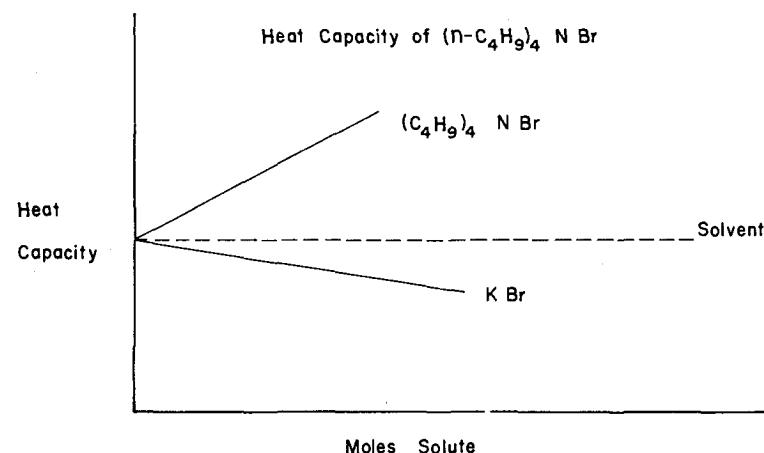


FIG. 5. Effect of apolar substituents in $(n\text{-C}_4\text{H}_9)_4\text{NBr}$ upon heat capacity in aqueous solution.

a horizontal (broken) line. Had a salt such as KBr been added to this solvent, the heat capacity of the solution would have dropped. On the other hand, if an organic apolar group is added to the solvent the heat capacity rises very substantially. Part of the rise can be attributed to the contributions of motions of the apolar butyl group, but these contributions can be estimated by standard thermodynamic methods, and would account, at best, for only about a half of the increase. We may note, then, that if we interpret the effect of KBr in causing the heat capacity to drop as due to a disorganization of water, then it follows that the rise in heat capacity produced by the apolar butyl groups must be attributed to an *organizing ability* of these groups on water molecules.

There is a good deal of independent evidence that these apolar molecules can increase the structural organization of water. In particular, these tetrabutyl compounds form crystallizable hydrates which melt very substantially above the melting point of ordinary ice, despite the fact that the formula of these hydrates (generally, 1 tetrabutylammonium molecule and 32 water) shows, on a molar basis, that the compounds contain over 90% water. The melting point, for example, of the tetrabutylammoniumhydroxide hydrate is about 30° , of the fluoride about 25° . The crystal structure of these hydrates is being worked out, but actually more precise information is known about some other hydrates, some simpler hydrates, and so I should like to turn to them next.

TABLE VI
HEATS OF HYDRATION OF APOLAR HYDRATES

Apolar molecule	$-\Delta H^\circ$, kcal mole $^{-1}$
Ar	16.6
Kr	13.9-16.5
Cl ₂	16.0-17.7
H ₂ S	16.5
PH ₃	15.0
SO ₂	16.6-19.0
CH ₄	14.5-17
C ₂ H ₂	15
C ₂ H ₄	15
C ₂ H ₆	15
CH ₃ Cl	15-18.1
CH ₃ SH	16.6

In Table VI, I have listed some of these hydrates, together with values for their heats of hydration. As can be seen in the table, molecules of very different chemical properties form crystal hydrates; argon and krypton, for example, inorganic gases with essentially no chemical properties; chlorine, a very reactive inorganic gas; methane, CH₄, an organic material. Among the organic materials we can compare methane with C₂H₂, acetylene, or with C₂H₆, ethane; or just to emphasize that one need not be restricted only to pure hydrocarbon compounds, but can also have functional groups, we might look at the last two entries in the table, in particular, methylmercaptan. There are a variety of other molecules which also form apolar

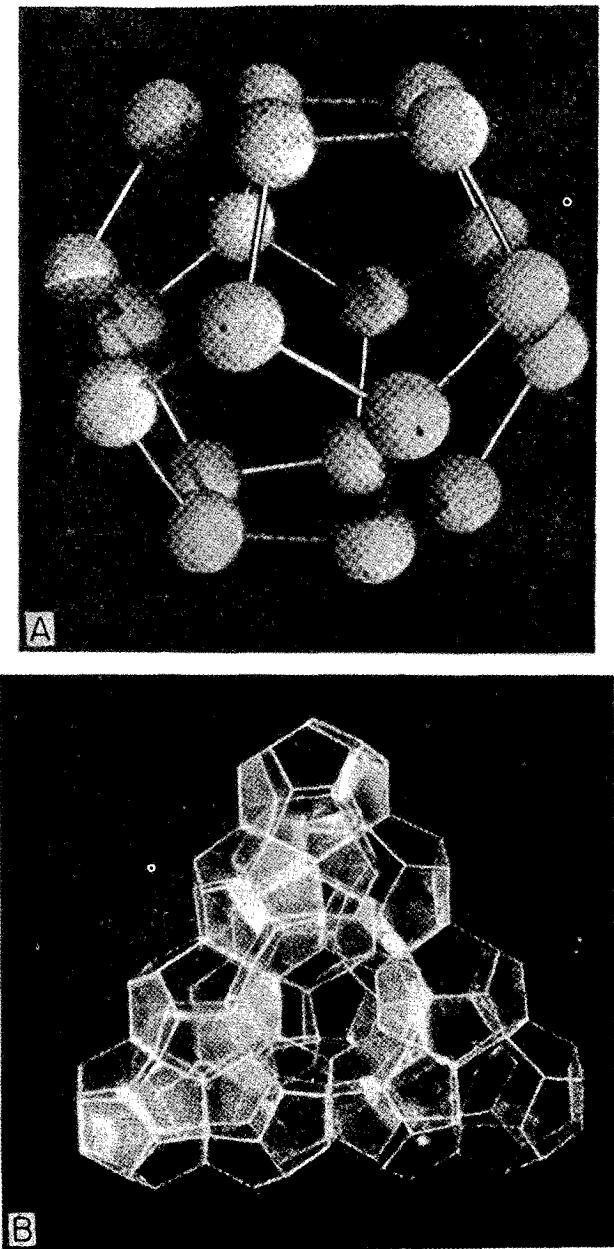


FIG. 6. (A) Model of pentagonal dodecahedron formed in hydrates of apolar molecules. Each ball represents a water molecule. (B) Multiple polyhedra forming a large lattice in hydrates of apolar molecules.

hydrates and I shall mention some of them, subsequently. For the moment, then, we can see that molecules of very different chemical properties form these hydrates.

What is perhaps even more striking is that the values of the hydration energies for these materials are almost uniformly identical. The entries in the second column of Table VI show values which fall within 1 or 2 kcal of 16 kcal, with an experimental error of 1–2 kcal. Thus, we can conclude that the stability of these hydrates arises not from any specific interaction of the apolar molecule with the water, but rather from the stability of the water itself in the presence of these trapped molecules. The crystal structures of these materials have actually been worked out (Stackelberg and Müller, 1954) and I have indicated schematically, in Fig. 6, the essential results. These crystal hydrates have water arranged not in the hexagonal array of normal ice but in a pentagonal array. As is indicated in Fig. 6A, five water molecules combine to form a planar pentagon and these pentagons can be combined as faces of very large polyhedra, often dodecahedra (Fig. 6A) but also higher polyhedra (Fig. 6B). These polyhedra form cages of very open structures. These crystal hydrates would be very unstable if all of the cages were completely empty. In practice, when some, not necessarily all, of these holes are filled by one of the molecules listed in Table VI, then we obtain a stable hydrate, in fact, a hydrate so stable that in many cases it will not melt until the temperature is some 30° or more above the melting point of ordinary ice.

Thus, we see that apolar molecules tend to organize water around them into a pentagonal semicrystalline arrangement. With this background information, let us turn our attention, now, to hydration in proteins as macromolecules.

C. PROTEINS

Let me start by showing in Table VII a comparison of some hydrate formers with some corresponding groups which occur as side chains in proteins. As is indicated in this table, CH_4 forms a crystal hydrate; correspondingly, we have a CH_3 side chain in the alanine residue of proteins. Likewise, on the left-hand side, we see propane molecules corresponding to the side chain of valine, isobutane corresponding to the leucine side chains. Again, not to leave the impression that we can only have organic hydrocarbon residues, we note that one can also have a $\text{CH}_3\text{-SH}$ hydrate former or a $\text{CH}_3\text{-S-CH}_3$

hydrate former corresponding to the cysteine and methionine side chains of proteins, respectively.

Recognizing that such side chains do exist in proteins, and that they exist in juxtaposition to each other, let us indicate, schematically, in Fig. 7, what one might expect to find if each of these side chains can induce a stabilized arrangement of water molecules into a crystalline array. As is shown in Fig. 7, we indicate a protein helix with some intramolecular hydrogen bonds and with

TABLE VII
COMPARISON OF HYDRATE FORMERS WITH AMINO ACID RESIDUES

Some molecules forming crystal hydrates	Some amino acid side chains
CH ₄	—CH ₃ (Ala)
CH ₃	CH ₃
CH ₂	—CH
CH ₃	CH ₃ (Val)
CH ₃	CH ₃
CH ₃ —CH	—CH ₂ —CH
CH ₃	CH ₃ (Leu)
CH ₃ —SH	—CH ₂ —SH (Cys)
CH ₃ —S—CH ₃	—CH ₂ —CH ₂ —S—CH ₃ (Met)
	—CH ₂ —  (Phe)

side chains protruding to the right. If each of these side chains happens to be an apolar material, then one would expect, on the basis of our preceding remarks, to find some organization of water. In the protein one would expect to find, furthermore, a cooperative effect between adjacent side chains so that the organization of water would extend out, somewhat, from the surface of the protein, and, in any event, reinforce the crystalline nature of the hydration layers. From our viewpoint (Klotz, 1958, 1960), then, one would expect that some of the stabilization that is inherent in proteins, comes from the stabilization which is found in apolar hydrates, merely due to the fact that the water is crystallized. We would attribute apolar bond stabilization in these macromolecules to the formation of

crystalline-type hydrates by the protein side chains, *hydrotactoids*, similar to the apolar hydrates which are found in simple molecules.

On this basis one can understand the normal effects of temperature in denaturing proteins. We would attribute the effect of temperature to the disorganization or melting of the *hydrotactoids*. Likewise, we would attribute the denaturing effect of urea to its ability to disorganize the hydration lattice. Whether or not urea can break hydrogen bonds within the helix, it would still be able

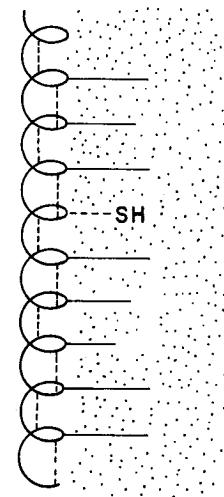


FIG. 7. Schematic diagram of *hydrotactoid* formation around apolar groups of protein macromolecule.

to form hydrogen bonds with water. In that way, it would disrupt the hydration lattice and thereby decrease the stabilizing effect originating from the presence of apolar groups.

On the basis of this picture, one can understand a variety of masking effects which are observed in proteins. As is indicated in Fig. 7, let us suppose that one of the side chains happens to be an SH group in the midst of some apolar groups, and let us keep in mind that it, itself, is essentially an apolar group. All of these could form a hydrate or a crystalline array of water molecules around that region of the protein. What do we mean by masking? In essence, that this SH group in this example cannot react very readily with a normal mercaptan-titrating agent, for example, silver. We can see

that if the water happens to be in a crystalline array it would be harder for the silver ion to penetrate the water to get to the SH group than it would be for the silver to get to an SH group on a simple mercaptan not attached to a protein molecule. Likewise, we can understand why the presence of urea would abolish the masking. If urea would break down the ice-like lattice, then the silver ion can now move in. As we indicated earlier, using lithium ion as an example of a monovalent ion, the mobility of lithium ion is much, much smaller in ice than it is in liquid water. Likewise, then, it seems reasonable to assume that the mobility of a silver ion would be much, much less in ice than it would be in liquid water, and that the basis of the masking, then, is the inability of the silver to penetrate the hydrotactoid lattice.

Similar masking phenomena are observed with artificially introduced side chains. For example, the acid-base character of an $-N(CH_3)_2$ on an apolar molecule attached to proteins is much different from that of the small molecule itself (Klotz, 1960), shifts in pK 's being in a direction typical of masking. In this case one can hardly invoke an interpretation, often used in explaining masking of naturally occurring groups, that the protein is folded in such a way as to sterically block access to the group of interest. For, with the artificially introduced group, the side chain was attached long after biosynthesis was complete.

There are a number of other features of the behavior of proteins and of their properties which indicate that water is very intimately involved in maintaining the structure of proteins (Klotz, 1958). But I want to discuss briefly just one additional set of experiments, that of Harrington and von Hippel (1961) on the collagen-gelatin transformation. As I have indicated schematically in Fig. 8, starting at the bottom, we have good reason to believe that collagen is composed of three strands, arranged in a specific helical array—not the α -helix but the polyproline helix, which is a very long stretched-out helix. These three strands give collagen a very rigid net structure. But the strands can be separated and collagen converted to gelatin, as is indicated in Fig. 8, by heat—actually by a temperature of only about 45° . The three strands are then separated and each one becomes, essentially, a random coil. This transition can be reversed I through a series of steps which I will not discuss here; I merely want to examine one stage, marked *b* on the diagram. At stage *b* there is good reason to believe that the polyproline helix has

been regenerated, but if we keep the concentration of gelatin low enough, the helices do not come together to form the collagen macromolecule. As I mentioned, this type of helix is a stretched-out one, one in which there cannot be hydrogen bonds within the macromolecule itself. In other words, there are no intramolecular hydrogen bonds stabilizing this helix, in part because of the actual geometric

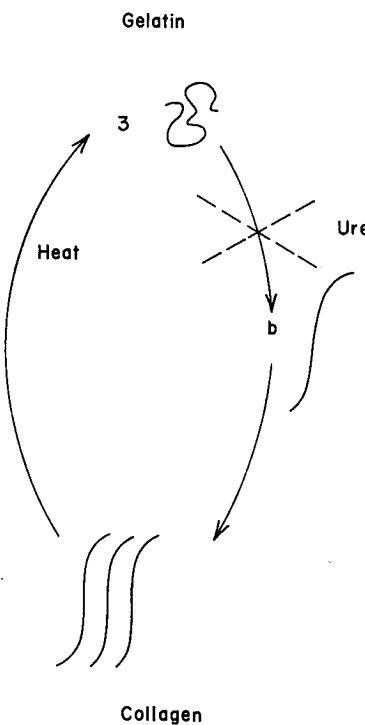


FIG. 8. Schematic representation of the collagen \leftrightarrow gelatin transformation.

configuration of the helix and also in the case of collagen and gelatin because the molecule contains a very large proportion of proline residues which do not have NH groups to donate to hydrogen bonds. Despite the fact that there are no hydrogen bonds within the macromolecule, within the stiff polyproline-type helix, Harrington and von Hippel (1961) observed that the formation of stage *b* could be inhibited, in fact, could be prevented by the presence of urea. Urea cannot be breaking any $NH \cdots O=C$ hydrogen bonds within the single

b strand because there are none in the first place. Urea, therefore, must be doing something else and it seems fairly reasonable to attribute its effect to a disruption of the contribution of water to the stability of the helix. Harrington and von Hippel (1961) have visualized a water molecule bridging two C=O groups within the macromolecule, adjacent carbonyl groups within the helix. I would be inclined to suspect that some of the apolar groups are contributing some stabilization by the formation of ice-like lattices. In either event, it is quite clear that water is involved in maintaining the particular fold which is characteristic of the collagen structure, in other words, water is involved in the stabilization of protein configuration.

D. TISSUES

Let us turn to a consideration of some aspects of the hydration of tissues. In this case, not too much information is actually available. Considering collagen as a tissue, as in rat tail tendon, for example, we can cite some observations which have recently been described by Berendsen (1960). The basis of Berendsen's work is some measurements of nuclear magnetic resonance. For our purposes we need not go into the theory of this particular technique, but merely to recognize the fundamental principle. A hydrogen nucleus has a magnetic moment. Consequently, molecules having hydrogen will interact with a magnetic field. And because they react with a magnetic field they will absorb electromagnetic radiation of a particular frequency, in the usual type of apparatus, at around 40 Mc. An important point for our purpose is that the nature of the signal is very much dependent on the rotatability of the hydrogen. If the hydrogen happens to be attached to a very mobile molecule then the signal is a very sharp absorption. On the other hand, if the hydrogen is on a molecule which is partly or completely immobilized, then the signal is very greatly broadened, and in some cases may even show a particular fine structure. In the case of collagen, Berendsen found that, if the sample could be oriented, one obtained a very specific fine structure. The details he could ascribe, essentially, to an organization of water molecules in chains around the collagen macromolecule in a pentagonal arrangement. The signals are unequivocally due to water molecules under restrictions in rotation and while they may not be absolutely in pentagonal arrays, they seem to have a definite lattice structure. X-ray investigations

(Esipova *et al.*, 1958) of rat tail tendon also show water molecules in ordered structure around collagen chains.

As regards other tissues, some evidence is beginning to accumulate that immobilized water exists in these, too. Particularly interesting are some of the recent observations over a period of a few years by Odeblad (1959, 1960), again using nuclear magnetic resonance techniques, with vaginal contents and eye tissues. There are clear-cut indications in these and other tissues, that the water is immobilized in a specific oriented array although the nature of the array is not known.

IV. Some Speculations on Energy and Charge Transport

How does this information contribute to an understanding of behavior at a physiological level?

As one example, let us take a very interesting recent paper by Pauling (1961), presenting a new idea for a general theory of anaesthesia. In essence, what Pauling has done is to make a list of some general anaesthetics, e.g., chloroform, N₂O, CO₂, ethylene, cyclopropane, nitrogen, argon, and xenon, and point out that these are substances that have no strong covalent chemistry, that do not form hydrogen bonds, and, in fact, looking particularly at argon and xenon, that have essentially no chemistry at all. If they are all general anaesthetics, how do they work? The only type of interaction that argon or xenon has that we are aware of is, as I indicated earlier, the ability to form crystal hydrates. In essence then, Pauling says, the mechanism of action of these materials as general anaesthetic agents must be that they increase the rigidity of the hydration which is normally found in protein side chains, by cooperating with these protein side chains to extend the region of crystal formation. In other words, he attributes the effect of these reagents to the formation of hydrate microcrystals in the encephalic fluid. He assumes, further, that because of the trapping of various side chains and ions in these crystals, the impedance of the network is increased and this is the basis of the narcotic or anaesthetic effect. On the basis of what I have indicated earlier in this paper in regard to changes in hydrogen ion transport in ice, one might offer as an alternative description of the effect of these narcotic agents, the creation of local short circuits rather than increased impedance; but this is a further step in theory upon which no real information is available. The essential part is the attributing of the effect of the generalized anaesthetic

agents to the formation of hydrate microcrystals. It is of interest, as Pauling notes, that on the basis of this kind of conception, one would predict that merely lowering the temperature of the brain would also create narcosis, and this seems to be true. Likewise, the theory would also predict that increasing pressure would abolish the anaesthetic effect, since, in general, hydrate crystals are of larger volume than the separate components, and this, too, seems to be a known physiological effect.

I have used Pauling's theory of general anaesthesia merely as one illustration of the kind of picture that hydration in the form of *hydrotactoids* leads to in the interpretation of physiological behavior. Similar ideas have been used (Klotz and Luborsky, 1959) to account for other combinations of substrates with biological macromolecules. I should also like to sketch, in a general way, the kind of picture we need to develop to incorporate some of our observations and concepts with regard to water into our visualization of cellular and subcellular phenomena. We should start with a view of the cell including large regions containing predominantly water. Furthermore, we assume that in the matrix of apolar groups attached to macromolecules the hydration water will form microcrystalline lattices or chains whose extent and direction will depend on the spatial disposition of the apolar groups. Then there can be three types of conduction or transport processes that might be at the molecular basis of some fundamental biochemical and physiological phenomena.

The first involves conduction by H^+ ion. As was emphasized at the beginning of this paper, an H^+ ion can be transported by a water brigade mechanism (Fig. 2, top row), and therefore current can be transported very fast and along specific directions in an ice-like hydration lattice. At potential differences that are observed in tissues, the rate at which an H^+ ion could move would be of the order of many meters per second.

Proton conduction may also play a role in a more subtle way, in the processes catalyzed by enzyme macromolecules. A large host of enzymatic transformations are of a hydrolytic nature, and the catalysis may involve a generalized acid or base mechanism (Bender, 1960). There are many basic or acidic side-chains in proteins, and these are usually assumed to be in the immediate neighborhood of the "active site" (X) at which the substrate molecule becomes bound. As is illustrated in Fig. 9, the basic (B) and acidic ($B-H$) groups could be substantially removed from the neighborhood of the

active site X and still perform their function, uptake or provision of a proton, through the mediation of the chain of water molecules. This picture offers a basis for interpreting the effects on enzymatic activity of a wide number of reagents that modify particular side-chains on the macromolecule, not all of which are likely to be in the vicinity of the site of attachment of the substrate. Modification of specific basic or acidic sides would, of course, directly remove their

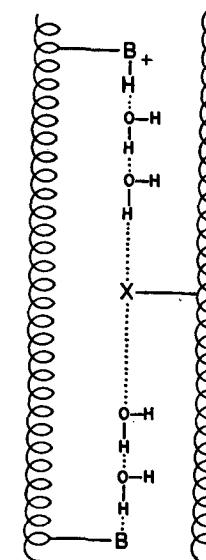


FIG. 9. Schematic diagram of long-range basic or acidic catalysis of reaction occurring at site X of an enzyme. B represents a basic residue, $B-H^+$ an acidic residue. It is implicitly assumed that extent and direction of hydration lattice is determined by apolar side chains which are not shown on the diagram.

potential contribution to the catalysis, but so would changes in the apolar side-chains for these are instrumental in maintaining the water bridge between B or BH and X. Likewise, the more general inhibiting effects of substances such as urea on enzymatic activity may be viewed as a consequence of the disruption of the organization of the hydration lattice.

A second class of transport phenomena through water may involve H^- free radicals or H^- hydride ions (Fig. 2, middle and bottom rows), which provide, in essence, a long-range oxidation-reduction mechanism (Klotz *et al.*, 1958). The direction in which the oxidation-

reduction could occur would be determined again by the particular arrangement of apolar groups, for these establish the specific pathway provided by the hydration "ice" lattice for the transport of electrons associated with H^{\cdot} or H^- . Oxidation-reduction could occur between widely separated groups. Figure 2 (middle row) reminds us of the widespread occurrence of iron pigments in oxidative enzymes. One should also keep in mind that the copper oxidases could operate in a similar fashion. In both cases the metal could be deeply embedded within the macromolecule, without direct access to the substrate, and yet electrons could be transported by H^{\cdot} or H^- through the organized water. This mechanism of "energy-transport" could also operate between different substrates or between coenzymes and substrates, or between different enzymes in multiple-enzyme systems, particularly those in a fixed matrix such as in mitochondria or chloroplasts. Likewise, transport could occur in rigid media in which thermal molecular motions are locked in. Sulfhydryl and disulfide groups (Fig. 2, bottom row) could very well participate in this mechanism of electron transport without being directly involved in the locale of binding of the substrate.

This second transport mechanism also provides a framework for the interpretation of many light-activated phenomena in protein behavior which seem to involve energy transfer over long distances. For example, in the myoglobin Fe-CO complex, the action spectrum for the release of CO indicates that electromagnetic energy absorbed by amino acid residues in the protein can affect the spatially removed heme group (Bücher and Kaspers, 1947). The absorbed radiation may produce a free radical (in molecules other than myoglobin by splitting an S-S bond symmetrically), by perhaps releasing an H^{\cdot} radical from tyrosine or aromatic rings. Thereafter, one can readily visualize the oxidation of the heme iron taking place through the transfer of an electron, by means of H^{\cdot} transport through an H_2O attached to the Fe, to a distant free radical, which in turn forms a stable bond with an H^{\cdot} free radical. It is significant in this connection to note that S-S linkages, as well as aromatic amino acid residues, absorb ultraviolet light in the region of $250 \text{ m}\mu$ (Barltrop *et al.*, 1954) and that such light can dissociate disulfides into —S $^{\cdot}$ free radicals (Russell and Tobolsky, 1954).

Finally, one can visualize a third mechanism by means of which hydrotactoids might influence the movement of charged particles and, thereby, associated physiological responses. As I indicated in

Table I, the mobility of a monovalent cation, such as lithium, is very much dependent on the state of water. In liquid water Li^+ moves rapidly, in ice very slowly. Although data are not available for Na^+ , there is every reason to presume that it would behave similarly to Li^+ . One might ask, therefore, whether the striking changes in permeability of the nerve membrane to Na^+ following a stimulus may not be due to the conversion of hydration water from an ice-like state to a liquid one. If water in the membrane is ice-like in the resting state of the nerve, its permeability to Na^+ would be very poor. When the hydration water is liquefied, the mobility of Na^+ should increase by a factor of 10^5 - 10^6 . After the passing of the stimulus, the initial state could be re-established within a period of microseconds, as the lifetimes of Table V indicate.

The comments of this section have been largely conjectural in nature, and I do not want to pursue what is largely guesswork to an unreasonable extent. I just wanted to indicate, particularly in connection with the discussion of hydration of tissue, that we have a model here which may offer a very interesting avenue for the interpretation of a variety of phenomena at the physiological level.

To summarize, the main thesis of this paper has been that water possesses some very interesting and unique properties at the molecular level. These properties are affected by solutes. And if we keep in mind that biological tissue contains 80 to 90% water, it seems most likely that physiological behavior may reflect molecular features, not only of the solutes, but, equally importantly, of solvent molecules, too.

REFERENCES

- Barltrop, J. A., Hayes, P. M., and Calvin, M. (1954). *J. Am. Chem. Soc.* **76**, 4348.
- Bender, M. (1960). *Chem. Revs.* **60**, 53.
- Berendsen, H. J. C. (1960). *Biol. Bull.* **119**, 287.
- Bernal, J. D., and Fowler, R. H. (1933). *J. Chem. Phys.* **1**, 515.
- Bücher, T., and Kaspers, J. (1947). *Biochim. et Biophys. Acta* **1**, 21.
- Connick, R. E., and Poulson, R. E. (1959). *J. Chem. Phys.* **30**, 759.
- Conway, B. E. (1952). "Electrochemical Data," p. 132. Elsevier, Houston, Texas.
- Dodson, R. W. (1952). *J. Phys. Chem.* **56**, 852.
- Eigen, M., and De Maeyer, L. (1958). *Proc. Roy. Soc. A* **247**, 505.
- Esipova, N. G., Andreeva, N. S., and Gatovskaya, T. V. (1958). *Biophysics (U.S.S.R.)* **3**, 505.
- Frank, H. S. (1958). *Proc. Roy. Soc. A* **247**, 481.
- Frank, H. S., and Wer, W.-Y. (1957). *Discussions Faraday Soc.* **24**, 133.

- Gränicher, H. (1958). *Proc. Roy. Soc. A* **247**, 453.
- Harned, H. S., and Owen, B. B. (1958). "The Physical Chemistry of Electrolytic Solutions," 3rd ed. Reinhold, New York.
- Harrington, W. F., and von Hippel, P. H. (1961). *Arch. Biochem. Biophys.* **92**, 100.
- Jackson, J. A., Lemons, J. F., and Taube, H. (1960). *J. Chem. Phys.* **32**, 553.
- Klotz, I. M. (1958). *Science* **128**, 815.
- Klotz, I. M. (1960). *Brookhaven Symposia in Biol.* **13**, 25.
- Klotz, I. M., and Luborsky, S. W. (1959). *J. Am. Chem. Soc.* **81**, 5119.
- Klotz, I. M., Ayers, J., Ho, J. Y. C., Horowitz, M. G., and Heiney, R. E. (1958). *J. Am. Chem. Soc.* **80**, 2132.
- Lonsdale, K. (1958). *Proc. Roy. Soc. A* **247**, 424.
- Morgan, J., and Warren, B. E. (1938). *J. Chem. Phys.* **6**, 666.
- Odeblad, E. (1959). *Ann. N.Y. Acad. Sci.* **83**, 189.
- Odeblad, E. (1960). *Nature* **188**, 579.
- Pauling, L. (1959). In "Hydrogen Bonding" (D. Hadzi, ed.), p. 1. Pergamon Press, New York.
- Pauling, L. (1961). *Science* **134**, 15.
- Russell, K. E., and Tobolsky, A. V. (1954). *J. Am. Chem. Soc.* **76**, 395.
- Stackelberg, M. V., and Müller, H. R. (1954). *Z. Elektrochem.* **58**, 25.
- Wang, J. H., Robinson, C. V., and Edelman, I. S. (1953). *J. Am. Chem. Soc.* **75**, 466.

Quantum Biochemistry

From Quantum Chemistry to Quantum Biochemistry¹

ALBERTE PULLMAN AND BERNARD PULLMAN²

*Institut de Biologie Physico-chimique, Université de Paris,
Paris*

"I have no doubts that the coming century will witness a profound revolution, extension of biology, the establishment of a quantum-mechanical biochemistry" (Szent-Györgyi, 1957)

I. The Prospects of Quantum Biochemistry	553
II. The Method of Quantum Biochemistry	556
A. The Choice	556
B. The Possibilities	561
III. The Problems	564
A. Electronic Delocalization in Biochemistry	564
B. Nucleic Acids	566
C. Proteins	568
D. Energy-Rich Phosphates	573
E. Coenzymes	575
F. Electron-Donor and Electron-Acceptor Properties of Biochemicals	576
G. Chemical Carcinogenesis	578
IV. Conclusions	580
References	580

I. The Prospects of Quantum Biochemistry

From a historical point of view, quantum chemistry is a baby science. From the human point of view it is, however, a grown-up, well-established individual. Its contributions toward the understanding of the nature of the chemical bond, the electronic structure of molecules, the dependence of the molecular properties upon

¹ This work was sponsored by the U.S. Public Health Service, Grant CY 3073 (National Cancer Institute), and the Division of Biology and Medicine of the U.S. Atomic Energy Commission.

² This paper was prepared while the authors were writing professors at the Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida.

the characteristics of this structure, have been enormous and decisive.

In recent years the field of application of quantum chemistry has suddenly been widely broadened, and we are observing, presently, a deep penetration of the ideas and methods of this science into the field of biochemistry. In fact, quantum biochemistry is quickly establishing itself as an important branch of what appears to be, nowadays, one of the essential fields of scientific research: molecular biology.

Quantum biochemistry! The mere combination of these two words relating, one to most complex abstract mathematics, and the other to most complex phenomena of life, seems daring and, to some people, even too daring. Of course, the existence of objectors is a normal phenomenon accompanying the birth of any new branch of science and, in fact, some of the objections presently raised against the application of the ideas of quantum mechanics to biology are exactly the same as those which were raised, initially, against the application of these ideas to chemistry. The following quotation, a quarter of a century old, has still its full meaning: "To be satisfied, one must adopt the mental attitude and procedure of an optimist rather than a pessimist. The latter demands a rigorous postulational theory, and calculations devoid of any questionable approximation or of empirical appeals to known facts. The optimist, on the other hand, is satisfied with approximate solutions of the wave equation. . . . He appeals freely to experiment to determine constants, the direct calculation of which would be too difficult. The pessimist, on the other hand, is eternally worried because the omitted terms in the approximation are usually rather large, so that any pretense of rigor should be lacking. The optimist replies that the approximate calculations do, nevertheless, give one an excellent 'steer' and a very good idea of 'how things go,' permitting the systematization and understanding of what could otherwise be a maze of experimental data codified by purely empirical valence rules" (Van Vleck and Sherman, 1935). The successful development of quantum chemistry shows that the optimists were right. In science, optimists are probably always right and the progress, finally, always surpasses all expectations.

To be an optimist does not mean, of course, to underestimate the difficulties which are connected with an enterprise such as the establishment of a quantum biochemistry, especially at its very

beginning. The difficulties are, in the first place, technical, the present methods of quantum chemistry being too simple and the subject too complex. They are also human, as the study of quantum biochemistry involves the knowledge of many different and difficult fields of science which it is not easy to find combined in one person. Answers to these problems may, perhaps, be found in the following lines:

"There is the spreading of quantum chemistry into biology. . . . The indications here seem to be quite clear: that more and more use will be made of wave-mechanical ideas in this exciting field. . . . The work itself, however, will almost inevitably be crude, and of the type that it is now fashionable to call semiempirical. Group I exponents will throw up their hands in horror at such attempts to estimate the electrostatic forces in an energy-rich phosphate bond; even Group II members will mistrust the complete neglect of many terms and integrals which are known to be large. But in the establishment of correlations and primitive patterns of understanding it does not do to be too fussy. A rough track through the jungle precedes the construction of metalled highway. And there is much experience possessed by professional biologists which could be linked with the deeper levels of interpretation associated with the quantum theory, to the enrichment of both. But let no one here make his claims too easily or too definitely. Biological systems are much more perverse than are laboratory chemical systems. In this field the prizes are immense—no less than the understanding and control of life itself. The future here may be far off. But, as was first said in rather different circumstances, 'Ce n'est que le premier pas qui coûte'; and there are very few sensible people who would wish to deny that in this ultimate human enterprise there is to be no contribution from quantum chemistry" (Coulson, 1960).

"The advancing front of science now lies in the borderlands between the older sciences, and the investigator who would be successful today must explore many fields, learn many skills, and dare to apply to a challenging problem in one area the insight he has gained in studying many. That man is lost who would 'stick to his last,' specializing more and more on less and less. The great problems of life and light will yield only to those whose knowledge of light suffuses their knowledge of life, whose knowledge of life quickens their knowledge of light" (Glass, 1961).

"The unknown offers an insecure foothold. What admits no

doubt in my mind is that the Creator must have known a great deal of wave mechanics and solid state physics, and must have applied them. Certainly, He did not limit Himself to the molecular level when shaping life just to make it simpler for the biochemist. . . . At present, the number of those who master both sciences, biology and quantum mechanics, is very small. Maybe it will never be very great owing to the limited nature of human life and brain. Both sciences claim a whole mind and lifetime, so, at least for the present, developments depend on some sort of hybridization" (Szent-Györgyi, 1960).

A different type of an answer can also be given: "As for myself, I like only basic problems, and could characterize my own research by telling you that, when I settled in Woods Hole and took up fishing, I always used an enormous hook. I was convinced that I could catch nothing anyway, and I thought it much more exciting not to catch a big fish than not to catch a small one" (Szent-Györgyi, 1961). Those who have seen Szent-Györgyi at work on fishing in Woods Hole, know how many big fishes he actually has caught.

The application of quantum mechanics to biochemistry raises, at the very beginning, two fundamental questions: (1) which method of quantum chemistry is the best suited for biochemistry, and (2) what are the basic problems of biochemistry which *a priori* are the best suited for undergoing a quantum mechanical investigation?

II. The Method of Quantum Biochemistry

A. THE CHOICE

Quantum chemistry provides us with two fundamental methods for the study of the electronic structure of molecules: *the valence bond method*, whose simplified qualitative version is referred to frequently as *the resonance theory*, and *the molecular orbital method*. Both represent approximate procedures for obtaining approximate solutions of the Schrödinger equation relative to molecules. This equation is the basic equation of the quantum theory whose resolution provides the electronic energy levels and the distribution of the electronic cloud in chemical systems. Approximate procedures are needed because we are unable, at present, to solve rigorously the Schrödinger equation for any atomic or molecular system beyond the very simplest ones.

Both methods have met with outstanding success in the field of

QUANTUM CHEMISTRY TO QUANTUM BIOCHEMISTRY 557

organic chemistry. In particular, the resonance theory has been of great value for the development of our ideas about the electronic structure of molecules, and for the penetration of these ideas among the organic chemists, because of its availability in a non-mathematical form and its constant use of concepts with which chemists are well acquainted. The purely qualitative formulation frequently adopted for this theory does not diminish the intrinsic value of its contribution, although it represents a serious handicap and sometimes leads to flagrant mistakes.

At first sight, it would consequently appear tempting to consider the valence bond method as the most promising tool for the application of quantum mechanical procedures to biochemistry. However, a more careful consideration of the basic principles of this method, even without any detailed investigation of its mathematical procedure, quickly shows that it is indeed far too complex to be of any real use in the study of the structure of biochemicals. Consider, for instance, the case of conjugated (resonating) molecules, which, as is well known, are the best "substrates" for this method and for whose study classic chemistry seems particularly inadequate. In this type of molecule, of which benzene is a classic example, we are interested essentially in the behavior and properties of the so-called mobile or π -electrons which are the electrons of the double bonds in the usual chemical formulas. The reason is that these electrons are the most important ones as far as a great number of essential chemical, physicochemical, and biochemical properties of the molecule are concerned. The electrons of the conventional single bonds (the so-called σ -electrons), form the rigid, basic skeleton in which the π -electrons move. Now, it is also well known that the usual chemical formulas are inadequate for describing the real distribution of the π -electrons; the basic principle of the valence bond method is that, to obtain an approximately satisfactory account of this situation we must consider, simultaneously, a number of structures associated with such molecules, structures which differ only in the apparent distribution of these mobile electrons, e.g., the well-known two Kekulé and three Dewar structures for benzene. None of these formulas has any physical meaning when considered alone; only their mental superposition may give us the proper picture of the actual molecular structure.

Quantum mechanics describes this situation by saying that the true wave function of the system (that is the mathematical function

which gives the true distribution of the electrons of the system) is a linear combination of the wave functions which can be associated with the different possible classic structures of the compound (that is the mathematical functions which translate into the quantum mechanical language the distribution of the mobile electrons associated with each structure). Then, the valence bond method uses the general procedures of quantum mechanics for the calculation of the composite, true wave function,³ with the aid of the wave functions associated with the fragmentary representations contained in the chemical formulas. This is done by resolving a determinant whose order is equal to the number of structures in the "canonical set" that can be written for the compound: namely, five in the case of benzene which is thus easily manageable. *Unfortunately this order increases tremendously with the increase in molecular dimensions:* the canonical set already comprises 42 structures in naphthalene and 429 in anthracene!

Thus the treatment of the naphthalene problem already involves the solution of a determinant of the 42nd order, and the treatment of anthracene a determinant of the 429th order. Even when the calculations are simplified because of the existence of high symmetries, the work still represents a formidable task. Now, even the simplest biochemicals will frequently be of the size of these hydrocarbons and yet generally devoid of symmetry. Obviously the actual treatment of such compounds by the valence bond method is precluded.

In fact, the situation is even worse, because there are a number of *ionic* structures which should also be included in the calculation. Unfortunately, the mere number of the ionic structures (170 for benzene!) alone makes their inclusion impossible. Moreover, there are important mathematical difficulties involved in their inclusion. This is a great obstacle to the utilization of the valence bond method for the investigation of the electronic structure of biochemicals. Thus, while it seems probable that, in spite of their great number, the ionic formulas do not contribute greatly to the structure of the *ground state* of conjugated hydrocarbons, and may consequently be dismissed in calculations concerning these molecules, it is obvious that the situation is different in the heterocyclic or substituted derivatives of hydrocarbons and, broadly speaking, in any heterocyclic

³ This "true" wave function may of course be, in fact, more or less "true," according to the different possible approximations used in the calculations. Except for some very rare cases, it is never really "true."

compounds. In such molecules, the contribution of the ionic structures to the total wave function of the system may be very large and their significance for the properties of the system predominant. Now, if we put aside some carotenoids, which are purely hydrocarbon conjugated systems (and, even they contain methyl substituents, which have either to be neglected or treated as any other substituents), all of the fundamental conjugated biochemicals are heterocompounds containing a number of different kinds of atoms, and, consequently, ionic structures are essential. Thus, if only for these quoted mathematical complexities (and there is still a number of complementary ones), the valence bond method appears quite impracticable for a quantitative investigation of the electronic structure of biochemicals. It can also be shown, and probably many of the readers of this paper have had the personal experience, that the purely qualitative version of the method, the resonance theory, is much too crude and too ambiguous to be of any practical use in this field.

Consequently, we turn to the second fundamental method of quantum chemistry, the molecular orbital method, and ask ourselves whether it can be of more use in the field of biochemistry. At first sight there seems to be an important handicap for the biochemist: the method is essentially, necessarily, a mathematical method, in which nothing can be obtained without calculations. It does not even make use of the classic structural formulas which we commonly write. Discouraging prospects for a method pretending to deal successfully with molecular structure!

But in fact, as we know, this claim has been borne out in many branches of organic chemistry and it does not take long to realize that it may do just the same in biochemistry. There are some essential features of the method which point to its possible utility in biochemistry: (1) its quantitative applicability to huge molecules without any particular difficulty; (2) the simplicity of its mathematical apparatus which enables any biochemist with a general mathematical background to grasp easily its principles, its procedures, and the significance of its results, and after a short training, *to use it himself for the investigation of many problems.*

What is the basic principle of the method? Let us again restrict ourselves to conjugated molecules. Like the valence bond theory, the method of molecular orbitals looks for a proper description of the state of the π -electron pool of such compounds. But while the

first method tries to obtain such a description by superposing a number of representations of imaginary structures, the molecular orbital method proposes to arrive at the same result by studying each π -electron separately in the field of the remaining electrons and nuclei. Once the appropriate description of each π -electron of the molecule (its molecular wave function, or *molecular orbital*) is obtained, the description of the total molecular system may be achieved by an appropriate "summing up"⁴ of the descriptions related to each electron of the system.

This mode of approach is responsible for the wide applicability of the method and its relative simplicity. The effective calculation again consists of the solution of appropriate determinants, but the order of the determinants will be generally much smaller than in the valence bond method. In the molecular orbital method, the order of the matrix to be solved for a conjugated system is equal to the number of the π -electrons of the system (or in case of compounds containing atoms with "lone pairs" of electrons, to the number of atoms carrying the π -electrons). Thus it is 6 in benzene, 10 in naphthalene, and (only) 14 in anthracene. These are easily manageable matrices compared to the enormous determinants of the valence bond method; the more so, as the determinant elements themselves are much simpler in the molecular orbital method than in the valence bond one. This is easily understood if one remembers that, in the molecular orbital method, these elements correspond to one-electron wave functions while in the valence bond method they are related to wave functions representing chemical formulas, which are consequently polyelectronic. The low order of the determinants appearing in the molecular orbital method makes it possible also to use the same approximation for a series of widely different compounds.

Thus, from the purely theoretical point of view, the method of molecular orbitals obviously presents great advantages in favor of its utilization in biochemistry. On the other hand, the method exists, as all approximate methods do, in various stages of refinement. There are the so-called semiempirical LCAO (linear combination of atomic orbitals) approximation, the self-consistent field-molecular orbital approximation, the approximation of configuration mixing, etc. Although based on the same general principle, these various procedures are widely different from each other in

⁴The term "summing up" is put in quotation marks because it may cover a procedure more complex than a simple sum.

their mathematical development and precision. Consequently, they also differ widely in the labor they require. The use of the higher approximations implies complex mathematical procedures which can be dealt with only by qualified specialists of the quantum mechanical methods. On the contrary, the simple LCAO approximation may be utilized by chemists who have only a general mathematical training. Fortunately, it also appears that a great majority of fundamental biochemical structures and problems may be quite satisfactorily dealt with by this simple semi-empirical LCAO form of the method. Refinements are of course always welcome and always useful, but, in many problems, the essential results can be obtained with the use of the simple approximation. It is only in some particular cases that a more refined approximation obviously has to be used from the very beginning. We may thus say that *the molecular orbital method, particularly in its semiempirical LCAO approximation, at present appears to be the only available method, and at the same time the proper method, for the preliminary investigation of the electronic aspects of biochemistry*. It appears as the quantum mechanical method of biochemistry.

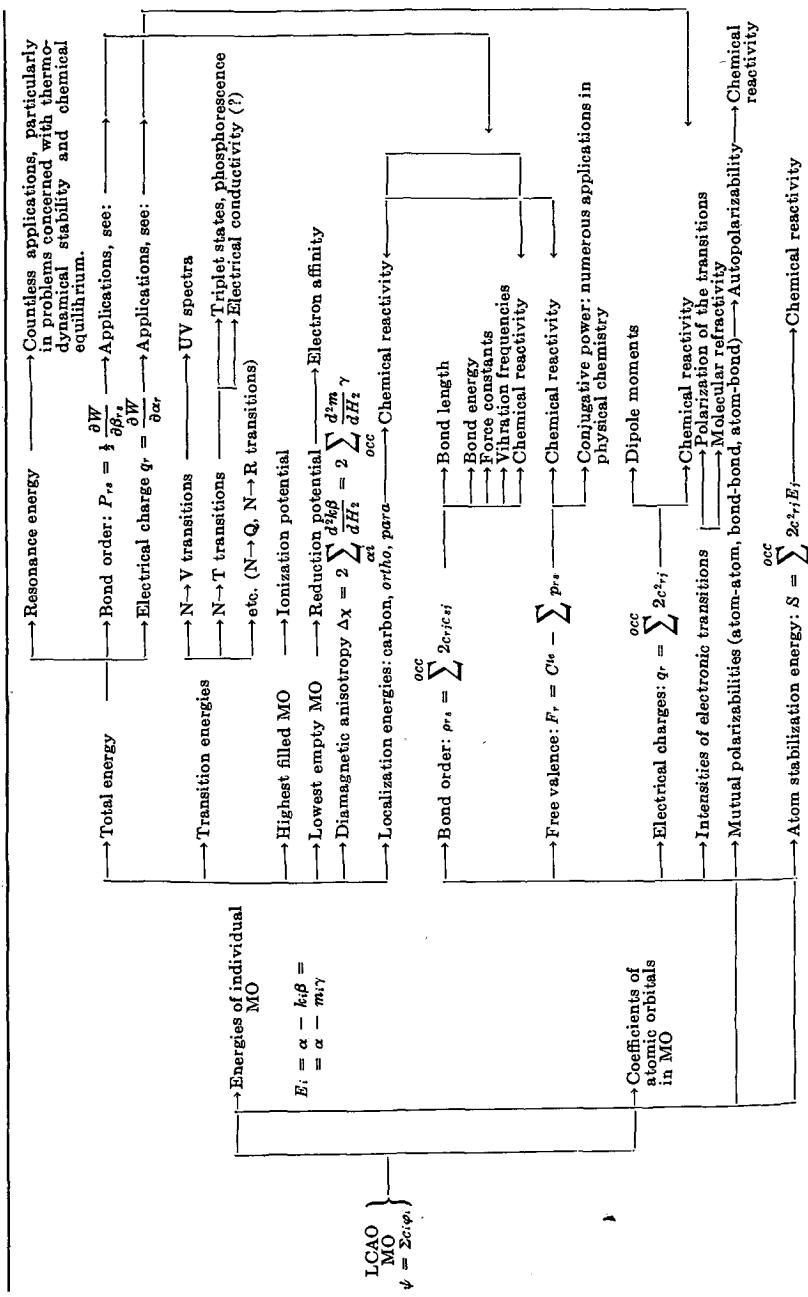
B. THE POSSIBILITIES

Once the choice of the method is settled, the next question is concerned with the possibilities of the method. What kind of information does it provide and what use can be made of the results of the calculations? A detailed answer to this question would be very long, or would involve the description of the whole field of quantum chemistry (see, e.g., B. Pullman and A. Pullman, 1952; Coulson, 1952; Kautzmann, 1957). Suffice it to say here that the method provides the values of essential energy and electronic indices characteristic of molecules or of their constituent atoms or groups of atoms. It enables thus, also, the location of the sites at which any particular index has its most significant value. These indices determine the essential chemical, physicochemical, and biochemical properties of molecules, and thus enable a deep understanding and a proper interpretation of their structure, behavior, and function.

The different possibilities of the method and the correlation between the calculated indices and the observable characteristics are indicated in more detail, although still very schematically, in Table I.

A few comments may be useful in connection with this table.

TABLE I
ELECTRONIC INDICES OF THE MOLECULAR ORBITAL METHOD AND THEIR APPLICATIONS



The two basic quantities obtained by calculations are the energies of the individual molecular orbitals (MO) and the coefficients of the atomic orbitals comprising the different molecular orbitals.

The energies of the individual molecular orbitals enable us to obtain the total energy of the mobile electrons and this quantity leads directly to the well-known *resonance energy* which indicates the gain in stability due to the delocalization of the π -electrons. The applications of resonance energy are countless, particularly in problems concerned with chemical equilibrium (e.g., tautomerism, oxidation-reduction potentials of reversible systems, acid and base strength). Other important structural indices of conjugated molecules like *bond order* or *electrical charge* are also directly related to the total energy. The energies of the individual molecular orbitals enable us also to establish the values of the different electronic transitions obtained by excitation of an electron from a bonding to an antibonding orbital.

The energies of the *individual* molecular orbitals are also of interest, particularly those of the highest filled and the lowest empty orbitals, which are related respectively to the ionization potential and the reduction potential of hydrocarbons.

We have also indicated, as deriving from the energies of the molecular orbitals, the so-called *localization energies*. These quantities are very important in the theory of chemical reactions and measure the essential varying part of the activation energy for substitutions and additions occurring on highly conjugated systems. In fact, they are not obtained directly from the same equations as those which lead to the energies of the molecular orbitals in the molecules, but from related secular determinants from which certain rows and columns have been struck out.

The quantities which are related to the coefficients of the atomic orbitals composing the molecular orbitals: bond orders, free valences, electrical charges, etc., are becoming nowadays classic indices of electronic structures with numerous applications. Another property which depends on these coefficients is the intensity of electronic transitions; the dipole moment of the transition, which is a fundamental quantity concerned with the intensity, is related to the coefficients of the atomic orbitals in the molecular orbitals involved in the transition.

Finally, there is a certain number of indices such as the mutual polarizabilities which depend both on the energies and coefficients.

III. The Problems

A. ELECTRONIC DELOCALIZATION IN BIOCHEMISTRY

The next fundamental question concerns the type of biochemical problem which *a priori* appear to be most suitable for investigation by the molecular orbital method.

The group of problems which immediately appears as the most promising one, in this respect, is connected with the significance and the role of electronic delocalization in biochemistry. As already stated, the molecular orbital method, especially in its simple LCAO approximation, is particularly suited for the study of the electronic structure of conjugated molecules and, thus, for the interpretation of the manifestations of electronic delocalization, which is the most outstanding feature of these compounds. In no way is the method really limited to this class of molecule and it may, of course, also be applied to saturated, non-resonating compounds. However, if significant results have to be obtained for this last type of molecule, more refined approximations have to be used, which are more difficult to handle.

Now, if we look at biochemistry from the point of view that we are concerned with here, one fundamental aspect of this science appears, which does not seem to have been noticed or stated before: *all the essential biochemicals which are related to, or perform, the fundamental function of the living matter are constituted of completely, or, at least partially, conjugated systems.* For instance, the three fundamental structural and functional units of the cell are undoubtedly the nucleic acids, the proteins, and the energy-rich phosphates. Now, the most significant constituents of the nucleic acids are the purine and pyrimidine bases, which are conjugated heterocycles. In the biologically important energy-rich phosphates, the mobile electrons of the phosphoryl group always interact electronically either with those of another similar phosphoryl group or with those of a resonating organic radical. No such interaction is present in the energy-poor phosphates. As to the proteins, although they appear, at first sight, to be essentially non-resonating entities containing only isolated conjugated fragments (each peptide link is such a fragment) (Pauling, 1958; A. Pullman and B. Pullman, 1959b), there are a number of indications that their over-all supramolecular structure

involves (as foreseen by Szent-Györgyi, 1941) some degree of a general electronic delocalization.

Nucleic acids, proteins, and energy-rich phosphates, although the most important representatives, are not the only conjugated constituents of the cells.

Pteridines, porphyrins, quinones, carotenoids, etc., are other important structural components of biochemicals belonging to the same type of conjugated compounds.

Another most striking observation in the same field concerns enzymes. There are hundreds of enzymes and they are essentially proteins. However, if one excepts the hydrolytic ones, most of the enzymes exert their catalytic activity in conjunction with a coenzyme. There is only a very limited number of essential coenzymes, about a dozen, and *practically all of these coenzymes are conjugated compounds.* Such is the case, in particular, with the oxidation-reduction coenzymes, DPN, TPN, FAD, FMN, the cytochromes, and the quinones. And it is also the case with the coenzymes involved in group transfer reactions: folic acid coenzymes, pyridoxal phosphate, thiamine pyrophosphate, etc.

As a matter of fact, among the fundamental organic constituents of the living cell, carbohydrates, fats, and steroids seem to be the only non-conjugated types of molecules. Among these, carbohydrates and fats represent merely the fuel for the driving of the machinery without being functional constituents of it. The steroids generally possess a π -electron section and, anyway, seem to be in some way involved in electron transfer phenomena (Talalay and Williams-Ashman, 1958; Hurlock and Talalay, 1958, 1959).

Finally, if it is added that a large number of drugs susceptible of acting strongly on the living cell are also conjugated systems, it becomes obvious that *the basic manifestations of life are intimately connected with the existence of highly conjugated compounds.* As these are relatively complicated structures and, as "nature does not indulge in luxuries" (Szent-Györgyi, 1960), it is obvious that some essential features in these types of compounds make them particularly suitable for being the vehicle of life. This feature can hardly be anything else but *electronic delocalization* which confers upon these molecules both complementary elements of stability (which may, for instance, be of fundamental importance in determining their resistance to radiation damage (B. Pullman and A. Pullman, 1960a).

and reaction possibilities not encountered in other types of molecules (pronounced resonance stabilization of activated complexes, long-range transmission of electronic perturbations, possibilities of energy transfer, etc.). The essential fluidity of life agrees with the fluidity of the electronic cloud in conjugated molecules.

To what degree and in what respect do the molecular orbital calculations help in understanding the role of the electronic delocalization in determining the structure and function of biochemicals? A partial answer to this question may best be obtained by looking at the results which have already been obtained for some groups of these molecules.

B. NUCLEIC ACIDS

As stated earlier, the determination of the electronic indices in conjugated molecules, enables us to understand their chemical and physicochemical properties. Such a determination has been carried out in detail for the purine and pyrimidine bases of the nucleic acids (B. Pullman and A. Pullman, 1959a, 1961). The principal results, summarized in Table II, illustrate the possibilities and the scope of such investigations.

A more pictorial representation of these general results is given in Figs. 1 and 2. They contain the fundamental base pairs as present in the nucleic acids (S = sugar, P = phosphate). Figure 1 indicates the sites at which the essential theoretical indices of electronic structure possess their most significant values. The indices are underlined by a heavy line when placed at the site of their most significant value and by a broken line at their next most important site. In Fig. 2, the same thing has been done for the different chemical or physicochemical properties determined by these indices. In some cases, indices or properties have been indicated on Figs. 1 and 2 without being underlined. This signifies that these indices or properties are important characteristics of the corresponding bases, but that it is not possible to ascertain and compare their relative values in the different units.

The representation enables a very quick localization of the main reactive centers of nucleic acids toward given types of reagents. In this respect, one of the striking features of Fig. 2 is that none of the quoted reactions appear heavily underlined for the adenine ring. (It is only the transformations which are concerned with its extracyclic NH_2 group which appear as occurring with particular ease

TABLE II
FUNDAMENTAL ELECTRONIC INDICES AND OUTSTANDING CHEMICAL AND PHYSICOCHEMICAL PROPERTIES OF THE NUCLEIC ACIDS

The electronic index	Its manifestation
<i>Resonance energy.</i> The greatest resonance energy per π -electron is that of the adenine ring. The next one is that of guanine. Resonance stabilization through hydrogen bonds is greater for the guanine-cytosine pair than for the adenine-thymine pair.	Resistance of purines and in particular of adenine to ionizing and ultraviolet radiations (B. Pullman and A. Pullman, 1960a). Higher temperature of denaturation of nucleic acids rich in guanine-cytosine pair, than of those rich in adenine-thymine pair (Marmur and Doty, 1959).
<i>Highest filled molecular orbital.</i> The highest filled molecular orbital is the highest in guanine followed by adenine.	The electron donor properties of purines, visible in their tendency to form charge-transfer complexes with different electron acceptors (B. Pullman and A. Pullman, 1958, 1960b).
<i>Lowest empty molecular orbital.</i> The lowest empty molecular orbital is the lowest in cytosine, followed by adenine.	The electron acceptor properties of these two bases, visible in their tendency to trap electrons (Gordy, 1959; B. Pullman, 1961).
<i>Basicity of ring nitrogens.</i> In the nucleic acids the most basic nitrogen is predicted to be N_7 of guanine and N_1 of adenine (Jardetzky and Jardetzky, 1960; Brookes and Lawley 1960). N-Oxidation occurs at N_1 of adenine (Stevens and Brown, 1958). Alkylation of nucleic acids by alkylating agents occurs essentially at N_7 of guanine (Lawley, 1957; Reiner and Zamenhof, 1957; Bantz and Freese, 1960). The principal site of chelation by metals is between N_7 and the O of guanine (Frieden and Alles, 1958).	Protonation and alkylation by alkylating agents occur at N_7 of guanine and N_1 of adenine (Jardetzky and Jardetzky, 1960; Brookes and Lawley 1960). N-Oxidation occurs at N_1 of adenine (Stevens and Brown, 1958). Alkylation of nucleic acids by alkylating agents occurs essentially at N_7 of guanine (Lawley, 1957; Reiner and Zamenhof, 1957; Bantz and Freese, 1960). The principal site of chelation by metals is between N_7 and the O of guanine (Frieden and Alles, 1958).
<i>Charge of the primary amino nitrogens.</i> Because of the delocalization of their lone pairs the nitrogens of the NH_2 groups are formally positive. The most charged among them (the less positive) is that of adenine, followed by that of cytosine.	Direct alkyl substitution at the amino group and the reaction with formaldehyde occur essentially on the NH_2 groups of adenine, followed by that of cytosine (Whitehead and Traverso, 1960; Fraenkel-Conrat, 1954; Staehelin, 1958).
<i>Charge of the carbon atoms carrying the amino groups.</i> These carbon atoms are formally positive. The most positive among them is that of guanine, followed by that of cytosine.	Deamination of DNA by NO_2H occurs essentially on guanine, followed by cytosine (Schuster, 1960).

TABLE II (Continued)

The electronic index	Its manifestation
<i>Charge of the oxygen atoms.</i> The most negative oxygen is that of cytosine.	The oxygen of cytosine may be the primary site of protonation of undenatured DNA (Dove <i>et al.</i> , 1959).
<i>Carbon-carbon bond orders.</i> The greatest C-C bond order is that of the 5-6 bond of thymine, followed by the 5-6 bond of cytosine.	Hydroperoxidation under the effect of ionizing radiations and hydration upon photolysis occur essentially at the C-C bond of thymine (B. Pullman and A. Pullman, 1960a). Permanganate oxidation occurs essentially at the C ₅ -C ₆ bond of cytosine, followed by that of thymine (Bayley and Jones, 1959).
<i>Indices of unsubstituted C atoms.</i> C ₈ of guanine has the greatest free valence, C ₅ of cytosine the greatest formal negative charge.	Both carbons are sites of electrophilic substitutions (Albert, 1957).
<i>Bipositivity of the C-N glycosidic linkage.</i> The formal positive charge of the sugar-carrying nitrogen and the bipositivity of the glycosidic linkage are greater for the purines than for the pyrimidines.	Preferential reactivity of the glycosidic linkages of the purines in enzymatic or acidic hydrolysis (B. Pullman and A. Pullman, 1959) and toward mercaptoacetic acid and diphenylamine (Jones and Letham, 1958; Peterson and Burton, 1959).

at that molecule.) The relatively great stability of adenine toward external agents of any type is thus immediately visible after a glance at Fig. 2. This impression is still substantiated by a look at Fig. 1 in which the only characteristic of the adenine ring which deserves the heavy underlining is its relatively great resonance energy. The particular role played by adenine in biochemistry may, perhaps, be linked with its particularly pronounced thermodynamic and kinetic stability. It may be useful to add that the results on the electronic structures of the purine and pyrimidine bases have been helpful, among others, for the establishment of correlation between structure and antitumour activity of purine antimetabolites (B. Pullman and A. Pullman, 1960b).

C. PROTEINS

In the case of the nucleic acids, it is essentially their *submolecular* structure which has been submitted to a molecular orbital investi-

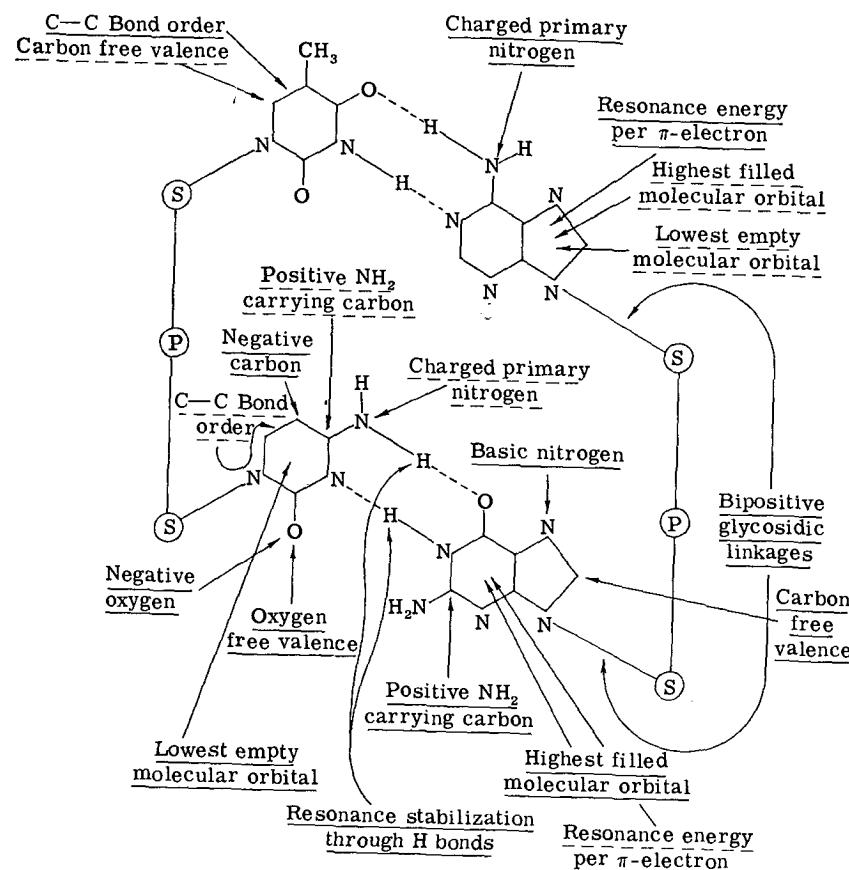
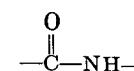


FIG. 1. Outstanding electronic indices of the nitrogenous bases of nucleic acids.

gation. In the case of the proteins the quantum mechanical investigations have been essentially concerned with some aspects of their supramolecular structure.

As is well known, the peptide backbone of proteins is formed of a periodical repetition of small resonating units



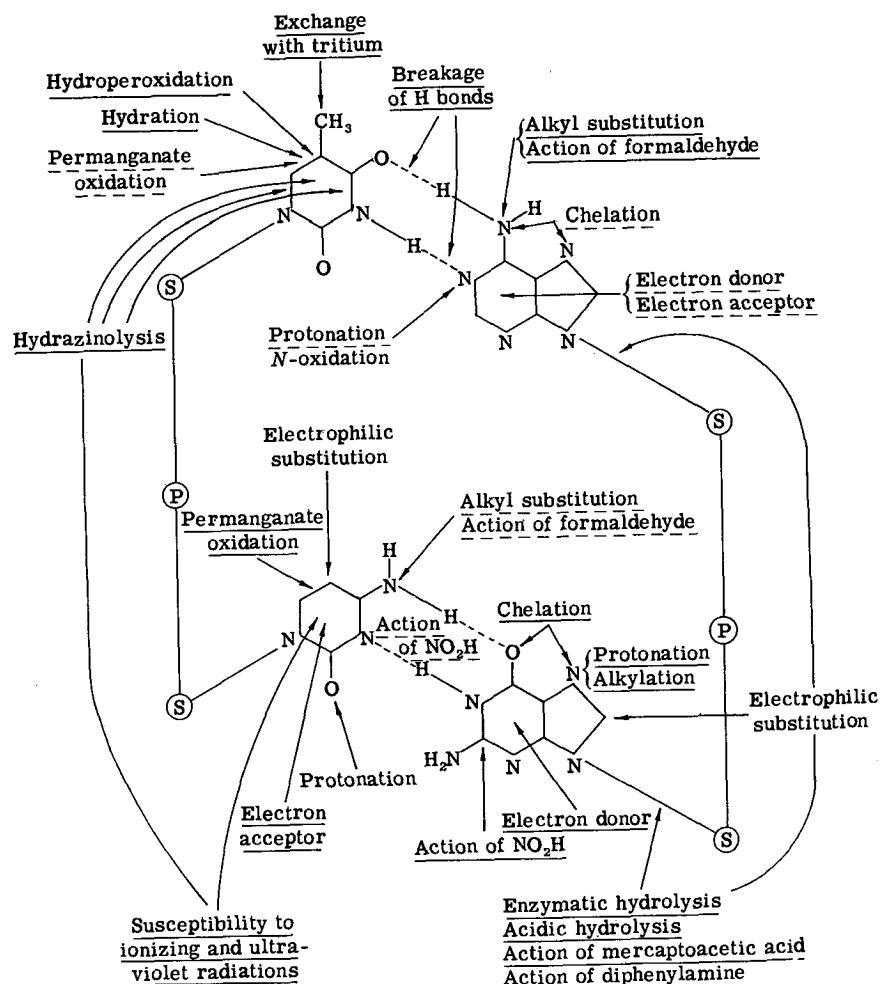


FIG. 2. Principal chemical and physicochemical properties of the nitrogenous bases of the nucleic acids.

which are, in the backbone, separated from each other by saturated carbons. Consequently no electronic delocalization can occur along such backbones. The peptide bonds are, however, united to each other by secondary cross linkages, namely hydrogen bonds. Inasmuch as these bonds may participate in electronic delocalization and transmit conjugation effects, the existence of such a net-

work represents the possibility of an extended electronic delocalization involving the whole molecular framework of the protein. The question is, of course, to what extent, if at all, conjugation effects may be transmitted across the hydrogen bonds. If a nonnegligible transmission effectively occurs, the peptide bonds may not be considered anymore as isolated four π -electron systems; they would fuse, more or less, following the extent of the transmission, into a giant, conjugated structure. Even in the hypothesis of a weak interaction across the hydrogen bridge, quite new effects may be anticipated from such a cooperative interplay of an enormous, practically infinite, number of resonating units. It is this particular aspect of a possible over-all conjugation, mediated by the hydrogen bonds between the peptide linkages, that has been considered.

This problem of a *general electronic delocalization extending over large surfaces of proteins* was explicitly stated in 1941 by Szent-Györgyi, who put forward the hypothesis that energy transfer in biological systems could be carried out by a mechanism analogous to the conductivity phenomena in crystals and, in particular, that the regular arrangement of the peptide linkages in the proteins could result in the existence of *energy bands*, similar to those present in semiconductors. This daring hypothesis inspired a great number of experimental researches on the semiconductive properties of proteins (Eley, 1959; Cardew and Eley, 1959; Eley and Spirey, 1960) and has also given rise to a great many discussions and even controversies.

From the theoretical point of view, Evans and Gergely (1949) have been the first to try the verifications of Szent-Györgyi's hypothesis, by carrying out calculations on the energy levels of proteins. The authors adopted a simple molecular model by assuming that the peptide units, of four π -electrons each, were able to interact with each other through their hydrogen bonds. Such an interaction would result in the formation of infinite conjugated layers (Fig. 3) in which electronic delocalization could give rise to energy bands.

Evans and Gergely have carried out LCAO molecular orbital calculations of energy levels corresponding to such an infinite layer. Recently, these calculations have been greatly refined in our laboratory, with the use of the self-consistent field molecular orbital method (Suard *et al.*, 1961). The refined results point to the probable existence of four energy bands in proteins, three of which are entirely

filled up with electrons and one of which is entirely empty. The calculated transition energy from the highest filled to the lowest empty band exceeds 5 ev. The conclusion to be drawn for this result is thus, that, inasmuch as polypeptides can be accommodated by the proposed model, pure proteins should be rather quite good insulators. The observed semiconductivity, which, following the quoted work of Eley and collaborators, involves an energy gap of 2-3 ev, is probably extrinsic, originating from extra energy levels, due to the presence of defects or impurities.

There is, however, one aspect of the results obtained by the self-consistent field method which may have a much greater significance for the catalytic role of proteins in enzymatic reactions. The calculations show that the ionization potential of a protein should be

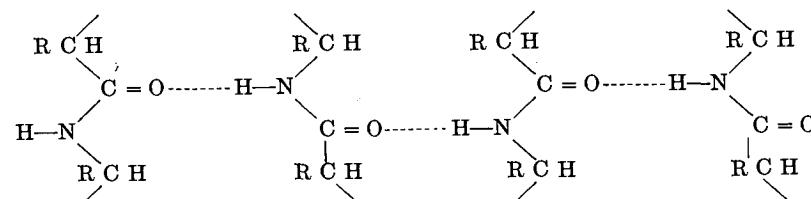


FIG. 3. Conjugation, via hydrogen bonds, in proteins.

about 0.9 ev smaller and its electron affinity about 0.6 ev greater than in the monopeptide unit. Such variations, although obviously not enormous, may nevertheless represent an advantage which is not negligible in the phenomena of catalysis due to the enzymatic proteins. It is recognized that a factor of essential importance in the catalysis of chemical reactions by semiconductors is the possibility of an electron transfer, in one direction or the other, between the catalyst and the substrate, such a transfer being able to lead to the formation of charge-transfer bonds or complexes (Leach, 1954; Garner, 1957). The facility of such charge transfer is thus of primary importance for the bonds between the catalysts and the substrates. Under these conditions the decrease of the ionization potential and the increase of the electron affinity, i.e., the increase of the electron-donor and the electron-acceptor properties of a polypeptide chain, with respect to the same characteristics of an isolated peptide unit, signify a net increase in the catalytic properties of these substances.

D. ENERGY-RICH PHOSPHATES

Why are energy-rich phosphates rich in energy? It means, why is the free energy of hydrolysis of certain selected phosphates a few kilocalories greater than that of the more usual ones?

Experiment and theory agree that the energy richness of the phosphates originates from a number of contributions, the most important ones being:

(1) *Opposing resonance* representing the difference between the resonance energy of the energy-rich phosphate and the sum of the resonance energies of its constituent fragments. This includes the *primary* opposing resonance resulting from the mere fusion of the fragments and the *complementary* opposing resonance due to deeper structural changes produced by this fusion.

(2) *Electrostatic interaction energy.* The energy-rich phosphates are characterized by a rather unusual distribution of electrical charges, which consists of a main backbone of at least three adjacent atoms carrying a net positive charge surrounded by a cloud of negatively charged atoms. As an illustration, we reproduce here the distribution of net charges in ATP (Fig. 4).

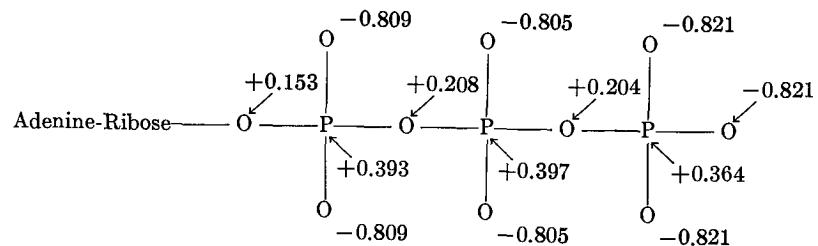


FIG. 4. The distribution of π -electrons in the pyrophosphate chain of ATP.

This distribution leads sometimes to strong electrostatic repulsion.

(3) *The enol-keto tautomerism* of the product of hydrolysis.

(4) *The free energy of ionization* of the product of hydrolysis.

The quantitative evaluation of these contributions to the energy of the different types of energy-rich phosphates are summed up in Table III (B. Pullman and A. Pullman, 1960c).

The essential contributions in each case are underlined. They are different in each case. The general agreement with experiment is satisfactory.

TABLE III
“ENERGY RICHNESS” OF PHOSPHATES (CONTRIBUTIONS TO THEIR FREE ENERGY OF HYDROLYSIS) IN KILOCALORIES PER MOLE

Compound	Experimental	Evaluated contributions						Free energy of ionization	Total
		Opposing resonance		Electrostatic ^b		Keto-enol	Tautomerism		
		Fundamental ^a	Primary	Complementary	repulsion	tautomerism			
ATP	7-8	3	2	0.6	2			7.6	
ADP	7-8	3	2	0.6	1.4			7	
Carboxyl phosphate	10-12	3	1.6	3	-0.7			3.2	10.1
Phosphoenopyruvate	11.5-12.5	3	0.2	0.8	-0.5	9			12.5
Guanidino phosphate	9-10	3	0.4	0.8	-0.7	-		?	?

^a Free energy of hydrolysis of the energy-poor phosphates taken as equal to about 3 kcal/mole.

^b The + sign means repulsion and the - sign means attraction.

E. COENZYMES

This is one field of biochemistry in which the application of the molecular orbital method has led to most fruitful results. The essential coenzymes are, as stated, conjugated organic molecules and the quantum mechanical investigation of their mode of functioning has brought into evidence the importance of electronic delocalization in determining their ability to function as coenzymes in the specific types of reaction in which they are involved. The theoretical studies included the oxidation-reduction coenzymes, pyridine nucleotides, and flavins (B. Pullman and A. Pullman, 1959b) as well as cytochromes (Pullman *et al.*, 1960), folic acid coenzymes (Perault and Pullman, 1960), pyridoxal phosphate (Perault *et al.*, 1960), thiamine pyrophosphate (Pullman and Spanjaard, 1960), retinenes (A. Pullman and B. Pullman, 1961), etc. The results are much too abundant to be even summarized here so that the interested reader must refer to the original papers. We would simply like to emphasize the general result of this research which shows that it is the presence of mobile electrons which, in one way or another, is responsible for the ability of these compounds to function as coenzymes. It has been shown, for instance, that in the respiratory coenzymes, the oxidation-reduction is accompanied by an instantaneous redistribution of the energies of the molecular orbitals and, in particular, of those of the lowest empty and the highest filled orbitals, in such a way that in each case a particularly low-lying empty orbital is associated with the oxidized form and a particularly high-lying filled orbital is associated with the reduced form. The oxidized form will then have a natural tendency to accept electrons and the reduced form to give them up, thus making these compounds particularly suitable to act as electron carriers. Such a redistribution of molecular orbitals can only occur in conjugated molecules, and, in fact, only in some very particular conjugated heterocycles.

Similarly, the studies on the group-transfer coenzymes have indicated the importance of the conjugation phenomena for the proper functioning of each of them, and have even led to the elucidation of some common features in their mechanism of action. Thus it has been shown that the essential driving force for the reactions catalyzed by the quoted coenzymes is, practically in every case, the transformation of the primary product of the interaction

of the coenzymes with a substrate into an intermediate product whose essential characteristics are, on the one hand, great energetical stability due to a high resonance energy, and, on the other hand, pronounced chemical reactivity due to the existence of large local accumulations of net, positive or negative, π -electron charges, as well as free valences. Consequently, these intermediates will display a great tendency to be formed, but also a great tendency to continue to react; in other words, they will be associated with low energies of activation for their formation but, equally, with low energy of activation for their subsequent transformations. These factors must play an essential role in the mechanism of enzymatic reactions and their discovery enables us to understand why most of the coenzymes are conjugated systems. It is only with such systems that these phenomena of simultaneous great stabilization and activation of the transitional compounds may be obtained.

It may be added that the studies on the folic acid coenzymes have recently been extended to the investigation of the mechanism of action of folic acid antimetabolites, in particular, of the antimetabolites active in cancer chemotherapy (Perault and Pullman, 1961).

F. ELECTRON-DONOR AND ELECTRON-ACCEPTOR PROPERTIES OF BIOCHEMICALS

Some other aspects of the possible use of the molecular orbital method in biochemistry must be stressed. One of these aspects concerns situations in which the lack of quantitative experimental data increases significantly the importance of the theoretical contribution. Such a situation is encountered, for instance, in the case of the fundamental problem of the electron-donor and electron-acceptor properties of biochemicals. All branches of biochemistry are involved in this problem which, in fact, is one of the very central problems of this science. From the viewpoint of physics, these properties may be related respectively to the ionization potential and the electron affinity of the compounds. The difficulty lies in the fact that, for practical reasons, none of these characteristics can be easily measured in such complex molecules as the biochemicals. Consequently, practically no quantitative experimental data exist concerning the electron-donor or electron-acceptor properties of these molecules. In view of the importance of such data this is a paradoxical situation.

Now, the molecular orbital method can compensate for this absence of experimental data very easily. As a matter of fact, its standard calculations give immediate quantitative knowledge about the relative electron-donor or electron-acceptor properties of the

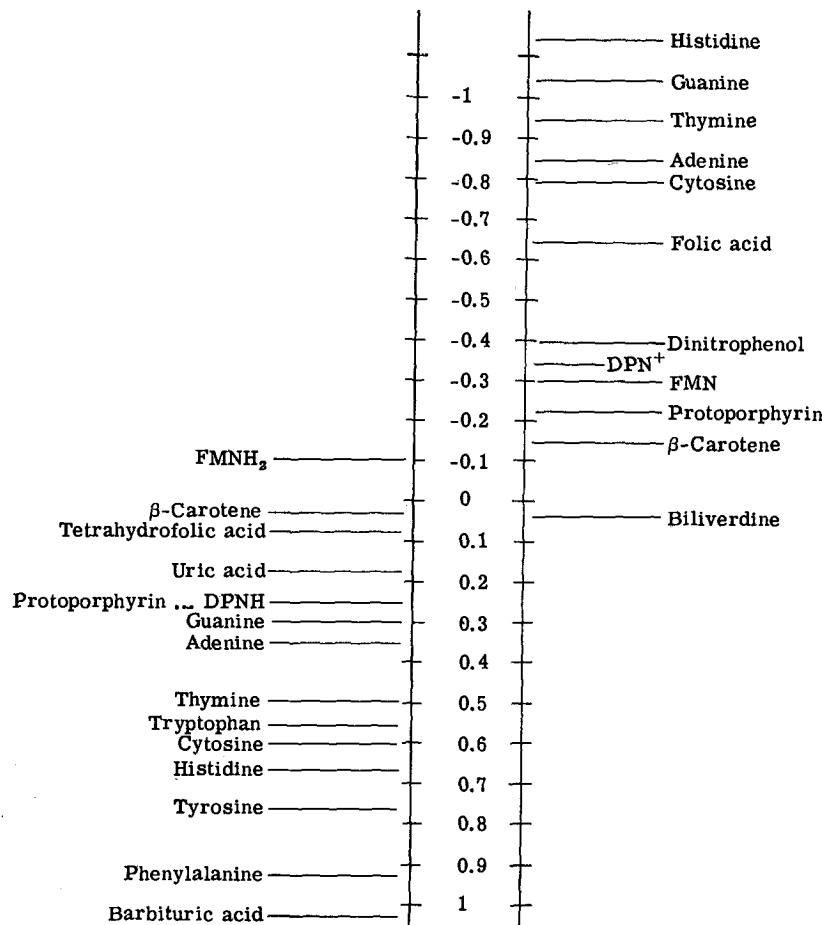


FIG. 5. Scale of electron donors and acceptors.

biochemicals. These two quantities are related respectively to the energy coefficient of the highest filled and the lowest empty molecular orbital. Without entering here into any details about the calculations (see B. Pullman and A. Pullman, 1958 and 1960b) we may

simply say that they have permitted the construction of a scale, in which about 300 biochemicals have been classified with respect to their relative electron-donor or electron-acceptor properties. A simplified version of this scale is given in Fig. 5 for some selected compounds. The electron-donor capacity of molecules is measured on the left side of the scale. The higher a compound is located on this side of the scale, the better electron donor it is. The electron-acceptor capacity is measured on the right side of the scale. The lower a compound is located on this side of the scale, the better electron-acceptor it is. These data have found numerous applications in a number of biochemical problems (B. Pullman and A. Pullman, 1958, 1959b, 1960b; Szent-Györgyi, 1960), in which they replace the missing experimental information.

G. CHEMICAL CARCINOGENESIS

Finally, we can hardly close this rapid review without quoting the work carried out on the correlations between the electronic structure and the carcinogenic activity of molecules. Besides its proper value, the work illustrates the possibilities of utilizing the molecular orbital method for the understanding of the mode of

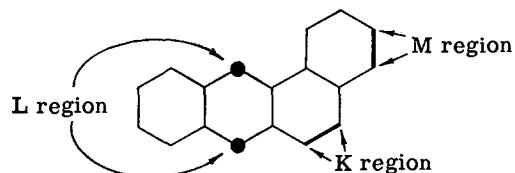


FIG. 6. Important regions in carcinogenic molecules.

action of drugs, whether harmful or chemotherapeutic. The subject is again too broad to be even summarized here (see A. Pullman and B. Pullman, 1955a,b). The theory has been able to correlate the existence of carcinogenic activity in, e.g., aromatic hydrocarbons with the characteristics of different specific regions in the molecules, known as the K, L, and M regions (Fig. 6). This subsequently led to a series of deductions and predictions about the nature of the interaction leading to carcinogenesis. Rather striking is the fact that practically all of the principal theoretical predictions seem to be verified by experiment. This can be seen from Table IV comparing some of the experimental and theoretical conclusions.

TABLE IV
EXPERIMENTAL AND THEORY IN CARCINOGENESIS

Experiment	Theory
1. "Binding of chemical carcinogens to tissue proteins is obligatory for the initiation of cancer" (Heidelberger, 1959).	... L'idée fondamentale ... est ... la nécessité d'une fixation électrique de la molécule sur un élément cellulaire défini ... Le problème est, selon toute probabilité, un problème de réactivité chimique (A. Pullman and B. Pullman, 1955b).
2. "In order for binding to occur, the intact ring system is required" (Heidelberger, 1959). "There is an apparent necessity for a completely aromatic system for binding" (Oliverio and Heidelberger, 1958).	Le problème de l'activité cancérogène met en jeu, selon toute probabilité, la molécule initiale et entière (A. Pullman and B. Pullman, 1955b).
3. "Binding takes place largely through the K region. The isolation of PDA (2-phenyl-phenanthrene-3'-2-di-carboxylic acid) ... was the first experimental verification of the Pullman's concept of the interaction of a carcinogen with cellular constituents at the K region" (Oliverio and Heidelberger, 1958).	The idea of the K region (a reactive bond) as the active center of carcinogens, through which the bonding to a cellular receiver occurs, is the fundamental idea of the theory, postulated by one of us as early as in 1945 (A. Pullman, 1945).
4. "The binding through the K region involves an addition reaction and is subject to steric hindrance. When one K region is blocked there is a possibility of binding to the second K region" (Oliverio and Heidelberger, 1958).	These are complete confirmations of the constant predictions of the theory, e.g., "The mechanism of action of carcinogenic molecules involves ... a chemical reaction between the carcinogen and cellular receiver. ... The reaction probably consists in the formation of an addition product or complex. ... The cellular receiver is probably of an electrophilic nature ..." (A. Pullman and B. Pullman, 1955a).
5. "The work reported here demonstrates that PDA is bound to the proteins through its carboxyl groups in amide linkage" (Bhargava and Heidelberger, 1956).	"Our scheme is in accordance with the theoretical speculations of the Pullman's with respect to a quinonoid bond between the hydrocarbon and the tissue" (Bhargava and Heidelberger, 1956).

It seems highly probable that the understanding of the mechanism of drug action will be among the principal results of the development of quantum biochemistry.

IV. Conclusions

We have tried to outline in this review some of the features of quantum biochemistry which are among its outstanding characteristics at the present moment. This includes the selection of the quantum mechanical method which appears the best suited for biochemistry and of the problems which, presently, appear as the most accessible to investigation. The method is the molecular orbital method and the problems seem to be those which are connected to electronic delocalization in biochemicals. We have indicated in a number of examples the type of problems which have been treated already and the nature of the results obtained.

It must be well understood that these indications and estimations are based on present-day possibilities and, are the authors' personal experience. Others may find it preferable to progress along different lines. Whatever it be, a day must obviously come when biochemistry will be entirely *expressible* in a quantum mechanical language. What seems certain, then, is that the elucidation of the structure of biochemicals, at the electronic level, will necessarily lead to a deeper understanding of their function and thus, finally, to the understanding of the mechanism of life itself.

REFERENCES

- Albert, A. (1957). In "The Chemistry and Biology of Purines," A Ciba Foundation Symposium, p. 97. Churchill, London.
- Bantz, E., and Freese, E. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1585.
- Bayley, C. R., and Jones, R. S. (1959). *Trans. Faraday Soc.* **55**, 492.
- Bhargava, P. M., and Heidelberger, C. (1956). *J. Am. Chem. Soc.* **78**, 3671.
- Brookes, P., and Lawley, P. D. (1960). *J. Chem. Soc.* p. 539.
- Cardew, H. H., and Eley, D. D. (1959). *Discussions Faraday Soc.* **27**, 115.
- Coulson, C. A. (1952). "Valence." Oxford Univ. Press, London and New York.
- Coulson, C. A. (1960). *Rev. Modern Phys.* **32**, 3.
- Dove, W. F., Wallace, F. A., and Davidson, N. (1959). *Biochim. Biophys. Research Commun.* **1**, 312.
- Eley, D. D. (1959). *Research* **12**, 293.
- Eley, D. D., and Spirey, D. I. (1960). *Trans. Faraday Soc.* **49**, 79.
- Evans, M. G., and Gergely, J. (1949). *Biochim. et Biophys. Acta* **3**, 188.
- Fraenkel-Conrat, H. (1954). *Biochim. et Biophys. Acta* **15**, 307.

- Frieden, E., and Alles, J. (1958). *J. Biol. Chem.* **230**, 797.
- Garner, W. E. (1957). *Advances in Catalysis* **9**, 169.
- Glass, B. (1961). In "Light and Life" (W. D. McElroy and B. Glass, eds.), p. 1911. Johns Hopkins Press, Baltimore, Maryland.
- Gordy, W. (1959). *Radiation Research, Suppl.* **1**, 491.
- Heidelberger, C. (1959). In "Carcinogens: Mechanisms of Action," Ciba Foundation Symposium, p. 179. Churchill, London.
- Hurlock, B., and Talalay, P. (1958). *J. Biol. Chem.* **234**, 886.
- Hurlock, B., and Talalay, P. (1959). *Arch. Biochem. Biophys.* **80**, 468.
- Jardetzky, C. D., and Jardetzky, O. (1960). *J. Am. Chem. Soc.* **82**, 222.
- Jones, A. S., and Letham, D. S. (1958). *Biochim. et Biophys. Acta* **14**, 438.
- Kautzmann, W. (1957). "Quantum Chemistry." Academic Press, New York.
- Lawley, P. D. (1957). *Biochim. et Biophys. Acta* **26**, 450.
- Leach, J. S. (1954). *Advances in Enzymol.* **15**, 1.
- Marmur, J., and Doty, P. (1959). *Nature* **183**, 1427.
- Oliverio, V. T., and Heidelberger, C. (1958). *Cancer Research* **18**, 1904.
- Pauling, L. (1958). In "Symposium on Protein Structure" (A. Neuberger, ed.), p. 17. Methuen, London.
- Perault, A.-M., and Pullman, B. (1960). *Biochim. et Biophys. Acta* **44**, 251.
- Perault, A.-M., and Pullman, B. (1961). *Biochim. et Biophys. Acta* **52**, 266.
- Perault, A.-M., Pullman, B., and Valdemoro, C. (1960). *Biochim. et Biophys. Acta* **46**, 555.
- Peterson, G. B., and Burton, K. (1959). *Trans. Faraday Soc.* **55**, 492.
- Pullman, A. (1945). *Compt. rend. acad. sci.* **221**, 140.
- Pullman, A., and Pullman, B. (1955a). *Advances in Cancer Research* **3**, 117.
- Pullman, A., and Pullman, B. (1955b). "Cancérisation par les Substances Chimiques et Structure Moléculaire." Masson, Paris.
- Pullman, A., and Pullman, B. (1959). *Proc. Natl. Acad. Sci. U.S.* **45**, 1572.
- Pullman, A., and Pullman, B. (1961). *Proc. Natl. Acad. Sci. U.S.* **47**, 7.
- Pullman, B. (1961). *Acad. Roy. Belg., Classe Sci.* **33**, 174.
- Pullman, B., and Pullman, A. (1952). "Les Théories Electroniques de la Chimie Organique." Masson, Paris.
- Pullman, B., and Pullman, A. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 1197.
- Pullman, B., and Pullman, A. (1959a). *Biochim. et Biophys. Acta* **36**, 343.
- Pullman, B., and Pullman, A. (1959b). *Proc. Natl. Acad. Sci. U.S.* **45**, 136.
- Pullman, B., and Pullman, A. (1960a). In "Comparative Effects of Radiation" (M. Burton, J. S. Kirby-Smith, and J. L. Magee, eds.), p. 105. Wiley, New York.
- Pullman, B., and Pullman, A. (1960b). *Rev. Modern Phys.* **32**, 428.
- Pullman, B., and Pullman, A. (1960c). *Radiation Research, Suppl.* **2**, 160.
- Pullman, B., and Pullman, A. (1961). *Nature* **189**, 725.
- Pullman, B., and Spanjaard, C. (1960). *Biochim. et Biophys. Acta* **46**, 576.
- Pullman, B., Spanjaard, C., and Berthier, G. (1960). *Proc. Natl. Acad. Sci. U.S.* **46**, 1011.
- Reiner, B., and Zamenhof, S. (1957). *J. Biol. Chem.* **228**, 475.
- Schuster, H. (1960). *Biochim. Biophys. Research Commun.* **2**, 320.
- Staehelin, M. (1958). *Biochim. et Biophys. Acta* **29**, 410.
- Stevens, M. A., and Brown, G. B. (1958). *J. Am. Chem. Soc.* **80**, 2759.

- Suard, M., Berthier, G., and Pullman, B. (1961). *Biochim. et Biophys. Acta* **52**, 254.
- Szent-Györgyi, A. (1941). *Nature* **148**, 157.
- Szent-Györgyi, A. (1957). "Bioenergetics." Academic Press, New York.
- Szent-Györgyi, A. (1960). "Introduction to a Submolecular Biology." Academic Press, New York.
- Szent-Györgyi, A. (1961). *The New York Times Magazine*, Section 6, p. 16, July 30, 1961.
- Talalay, P., and Williams-Ashman, H. G. (1958). *Proc. Natl. Acad. Sci. U.S.* **44**, 15.
- Van Vleck, J. H., and Sherman, A. (1935). *Rev. Modern Phys.* **7**, 167.
- Whitehead, C. W., and Traverso, J. J. (1960). *J. Am. Chem. Soc.* **82**, 3971.

Quantum Chemistry in Molecular Biology

MICHAEL KASHA¹

Department of Chemistry and Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida

I. The Scope of Quantum Chemistry	583
II. The Search for New Biochemical Mechanisms	584
III. Basic Theories of Energy Transfer and Electron Transfer	587
A. Energy Transfer (Excitation Transfer)—Exciton Theory	587
B. Charge-Transfer Interaction (Donor-Acceptor Interaction)	589
C. Electronic Semiconductivity in Molecular Aggregates	590
IV. Molecular Memory Models	591
V. Commentary on the Molecular Orbital Method	594
VI. Conclusion	597
References	598

I. The Scope of Quantum Chemistry

Molecular biology, in its development as a discipline seeking to interpret biological phenomena on a molecular level, may range in requirement on molecular detail from realms of pure abstract mathematics such as the topological ideas of molecular genetics, through the detailed atomic-molecular structural information of X-ray diffraction, to the ultimate electronic detail of molecular electronic structure and of intermolecular and intramolecular interactions by quantum chemical calculations.

Quantum chemistry is capable of yielding qualitative interpretations of a very wide range of molecular properties and phenomena. It is the great diversity of the problems with which quantum chemical methods can deal approximately which makes them of particular value to the molecular chemist and molecular biologist interested in detailed electronic interpretations. It is useful and important for the biochemist to realize what molecular theories are available, what range of topics can be considered, and how well they can provide answers to his problems.

Quite often the results of quantum chemical calculations give only a semiquantitative result, based on approximations and semi-

¹ Work supported by the Division of Biology and Medicine, U.S. Atomic Energy Commission.

empirical schemes. Such results may be dissatisfying to the mathematical physicist interested in the rigorous non-empirical calculation of the properties of diatomic and other simple molecules. However, if such results suggest wholly new experimental approaches and new qualitative interpretations, then no one would deny that these applications of quantum chemistry may be accepted as very satisfying and fruitful.

One of the purposes of this essay is to correct two widely held impressions regarding quantum mechanical calculations on molecules: the first being that only the energy levels and electronic structure of the hydrogen molecule can be dealt with with any useful precision; and the other, that quantum chemistry as a general discipline is fully represented by simple molecular orbital calculations. Although undoubtedly the molecular orbital method does lead to the most tractable procedure for the evaluation of electronic energies and distribution functions in large molecules, a host of intermolecular electronic interactions fall largely outside the scope of the method and yet are of the greatest interest to biochemists and molecular biologists.

Intermolecular phenomena of special interest to molecular biology include hydrogen bonding, van der Waals and dispersion forces, charge-transfer interaction, exciton interaction, and conduction band formation. Most of these phenomena will receive some analysis in this presentation.

Just as it is important for the biologist and biochemist to realize the order of complexity of molecular problems with which contemporary quantum chemical studies can deal, from amino acids, thru chlorophyll and carotenoids, to polypeptides, proteins, and nucleic acids—so also is it natural for them to desire to learn of the kinds of limitations which exist in the methods and in their results. In the face of the complexity of the problems facing the modern biochemist and molecular biologist, a judicious blend of imaginative speculation and hard science may be the order of the day, rather than an overdose of either. It is in this spirit that the quantum chemist may offer the most useful service to the development of molecular biology.

II. The Search for New Biochemical Mechanisms

In 1941 Albert Szent-Györgyi ended his Korányi lecture (Szent-Györgyi, 1941) with these words:

"Biochemistry is, at present, in a peculiar state. By means of our active substances we can produce the most astounding biological reactions, but we fail wherever a real explanation of molecular mechanisms is wanted. It looks as if some basic fact about life were still missing, without which any real understanding is impossible. It may be that the knowledge of common energy-levels will start a new period in biochemistry, taking this science into the realm of quantum mechanics."

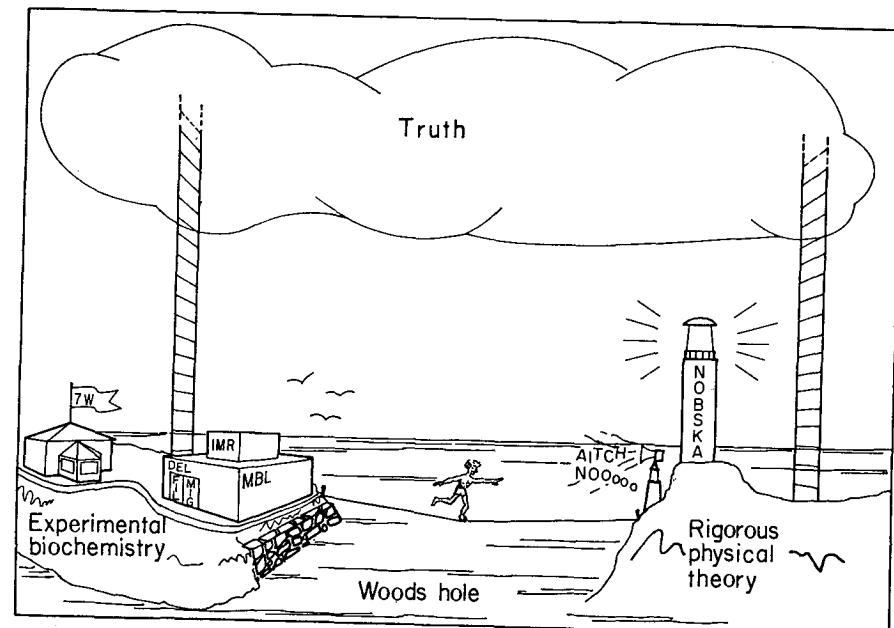


FIG. 1. Albert Szent-Györgyi's solution to the biochemist's dilemma.

Since then, Albert Szent-Györgyi has himself been a stimulating searcher for new physical approaches and possible shortcuts to the understanding of molecular mechanisms. We can picture his dilemma. Fully settled on the shores of experimental biochemistry (Fig. 1), he may well aspire to reach the shores of rigorous physical theory (although these shores may at times appear to be hidden in deep fog). But the paradox lies in the thought that the ladder to ultimate truth points straight upward, and parallel lines never meet except at infinity. Undaunted, Albert Szent-Györgyi casts a line straight

across the chasm, and feels very much at home in daring experiments serving as a short-cut to the desired contact between physical theories and experimental biochemistry. And he would be the last to worry about an occasional dip into the waters below.

One of Albert Szent-Györgyi's early attempts (Szent-Györgyi, 1941) to seek physical mechanisms for biochemical processes focused on the conduction-band model of solid state physics. In his monograph "Chemistry of Muscular Contraction" (Szent-Györgyi, 1951), he discussed this model under the title *the continuum theory*, suggesting that electron transport in proteins might be brought about through their action as semiconductors. This suggestion has inspired whole series of researches in numerous laboratories. Although even today the detailed physical bases for such properties are not well established for molecular systems, there has been a wide influence which is undoubtedly uncovering new phenomena and presenting new systems for study.

More recently *triplet states* of molecules have engaged the attention of Albert Szent-Györgyi (1956) and have been emphasized especially in his monograph "Bioenergetics" (1957) in relation to the role of water structure in biological systems. The role of metastable excited states in biochemical reactions is far from having been established. Nevertheless, the exciting discovery of the thymine photodimer formation (Beukers, Ijlstra, and Behrends, 1958) was based on Albert Szent-Györgyi's ideas on triplet-state excitation in frozen water media. The influence of Szent-Györgyi's discussion of triplet states in biochemistry will certainly continue, even though the details of the interpretation of the mechanism of triplet-state enhancement are revised as conditioned by molecular aggregation (McRae and Kasha, 1958), rather than due to water structure.

In the realm of *charge-transfer complexes* Albert Szent-Györgyi has undoubtedly found his most profitable excursion into molecular mechanisms in biochemistry. The physical interpretation and formulation of electron-donor-acceptor complexes proposed by Robert S. Mulliken (1950, 1951, 1952a,b) implies a near-universality of this type of intermolecular interaction in chemical systems. Albert Szent-Györgyi has begun an extensive study into the application of these ideas to biochemical processes, as summarized in his latest monograph "Introduction to a Submolecular Biology" (1960). A careful pursuit of these suggestions can be expected to lead to many fruitful developments.

In all of these searches Albert Szent-Györgyi has been more intent upon uncovering new points of departure for biochemical investigation, rather than settling on a mechanism as a final proof. So the experiments may be lacking in physicochemical precision, and the interpretations in detailed accuracy. But it can be said that the results are rewarding from the point of view of stimulating much new research inquiry.

We shall now turn to a critique of some areas of quantum chemistry which have potential application to molecular biology.

III. Basic Theories of Energy Transfer and Electron Transfer²

A. ENERGY TRANSFER (EXCITATION TRANSFER)— EXCITON THEORY

This theory was introduced by J. Frenkel. It involves the resonance interaction of excited states of weakly coupled composite systems. The exciton theory in its variations comprises the fundamental theory of (excitation) energy transfer. The energy interaction for electric-dipole allowed molecular excited states is a dipole-dipole interaction, which varies as $1/r^3$, and thus can be a very long range intermolecular effect (e.g., 10–50 Å). Orbital overlap is not required, and no electron transfer is involved. The energy states belong to the system as a whole, rather than to locally excited molecular units.

In the *free exciton*, or fast exciton theory, collective excitation of molecular units is involved, since the rate of energy transfer ($> 10^{12} \text{ sec}^{-1}$) is much greater than the rate of displacement of molecules from their equilibrium position. In this case the exciton interaction is evidenced experimentally by characteristic spectral shifts for the molecular aggregate relative to the spectra of the molecular units in isolation, by characteristic intensity changes, and by energy migration over the entire molecular aggregate.

In the *localized exciton*, or slow exciton theory, stepwise molecule-to-molecule energy transfer may be involved, since the rate of transfer may be much slower ($< 10^{12} \text{ sec}^{-1}$) than the rate of molecular displacement from equilibrium position in a molecular

² Taken with amendment from abstracts of the MIT Biology Department seminar series "Fast Fundamental Transfer Processes in Aqueous Biomolecular Systems" (Kasha, 1960a), Spring 1960.

array. Direct evidence of energy transfer may be demonstrable, whereas the only spectral changes observable may be in intensity of absorption, but not in spectral position.

The two cases are also describable as *strong* and *weak* coupling cases of exciton theory, respectively (Simpson and Peterson, 1957). Simpson and Peterson have developed criteria for these limiting cases in terms of the strength of interaction relative to the width of the molecular absorption band for the individual or isolated molecule. These criteria serve to distinguish between the situations in which relative strength, of intermolecular electronic interaction versus intramolecular vibrational-electronic interaction, fundamentally determines which limiting exciton model is appropriate in a given case. Intermediate coupling by the exciton model is a common possibility.

A third type of energy transfer mechanism is the vibrational-relaxation mechanism developed by Förster. This involves the quantum mechanical interaction of continua and applies specifically to energy transfer between non-identical pairs of molecules. Unlike the exciton model, an r^{-6} dependence (for dipole-dipole interaction) of the rate of transfer is observed. This is still a long-range interaction, but not as pronouncedly so as in the case of exciton interaction. The Förster mechanism is a relatively slow transfer mechanism, with rates in the range 10^9 to 10^6 sec^{-1} , comparable to spontaneous fluorescence rates. Förster (1960) has given a critical quantitative comparison of the three different mechanisms of energy transfer, labeling them strong, medium, and weak interaction.

The applications of the theory include (a) spectral changes and energy transfer in biological pigment aggregates (visual purple, chlorophyll), (b) hypochromism and energy transfer in deoxyribonucleic acids, (c) excitation energy transfer in high energy radiation experiments (scintillation counters, etc.), (d) effects on rotary dispersion through exciton interaction in helical configurations in proteins.

The exciton model has been developed as a very general quantum chemical approach to a wide range of excited state resonance interaction cases. A general discussion of the model has been given (Kasha, 1959) with reference to biological lamellar systems. The classic monograph by Davydov (1962) summarizes the detailed theory and growing literature of the field.

B. CHARGE-TRANSFER INTERACTION (DONOR-ACCEPTOR INTERACTION)

The theory dealing with charge-transfer interaction was put into a quantum mechanical framework by R. S. Mulliken (1950, 1951, 1952a,b), although a number of empirical formulations of the idea appeared in the chemical literature much earlier. The interaction between two (or more) units of a molecular system, existing either within one composite molecule or as a pair (or assembly) of molecules, can be approximated by considering a resonance interaction between a no-bond state of the isolated units, and the ionic excited state of the assembly arising from the transfer of a unit charge. Thus, an electron is formally transferred in the zeroth-order excited state, from one molecular unit to another. In the first-order description, after resonance interaction, the ground state will be stabilized by the admixture of some of the excited-state ionic character; likewise the excited state will be correspondingly destabilized or repelled by the admixture of some of the ground no-bond state.

The charge-transfer theory is an approximation to the valency interaction of two or more units of a composite system, and is a very short-range interaction. Good orbital overlap is required; other conditions set by the theory include proper symmetry relations between the overlapping orbitals, favorable energetic relations, and favorable dielectric environment. The chief applications of the theory are concerned with the attractive forces which result when a good electron-donor molecule and a good electron-acceptor molecule interact to form a complex.

Experimentally, the charge-transfer interaction is usually evidenced by the appearance of an intense, continuous new absorption band, characteristic of the composite system and not of either component taken separately. This charge-transfer absorption is predicted by the theory and is a characteristic feature; much research on this aspect has been published. There are some reports of electron spin-resonance absorption appearing in special cases of charge-transfer complexes, but this aspect is in an early research stage.

No electrical conductivity is predicted by the theory, in spite of the suggestiveness of the name of the theory. However, as the attractive force between molecules, and hence the orbital overlap, is increased, under charge-transfer interaction, there might be an

expectation of increase in electronic semiconductivity in this indirect fashion; a considerable increase in semiconductivity of charge-transfer complex crystals is actually observed, relative to crystals of the components taken separately.

Reviews of charge-transfer theory and applications have been published by several workers, including Orgel (1954) and McGlynn (1958).

C. ELECTRONIC SEMICONDUCTIVITY IN MOLECULAR AGGREGATES

A satisfactory quantum-mechanical theory for electronic semiconductivity in molecular crystals and aggregates has not yet been formulated. The property of semiconductivity requires orbital overlap between molecular units and a small energy gap between the ground state and the excited state of the molecular units, so that electron promotion (thermally) can occur. In the case of a semiconductor, closed shell ground states (which exist in most molecules normally) would preclude electron mobility under an electrical potential. However, after thermal promotion to an excited "band" of levels, arising from intermolecular overlap, hole migration or electron migration under the potential difference may occur.

The theoretical problem for molecular systems centers on the extent of intermolecular orbital overlap. The orbital overlap is certainly quite small, as evidenced by the fact that molecules retain their individuality in molecular aggregates and crystals, and by the low melting points of molecular crystals, and by the volatilization of molecular crystals as whole-molecule units. In fact, the success of the exciton theory in dealing with the spectroscopic properties of molecular crystals depends on small electron migration between molecules. The intermolecular orbital overlap of a molecular aggregate should be quantitatively related to the pre-exponential factor of the empirical semiconductivity relation, thus setting a theoretical limit on the absolute semiconductivity.

The exponential temperature-dependent factor in the empirical semiconductivity relation is related to the molecular energy gap. In the isolated molecule this gap should correspond to the lowest electronic state which can be excited in the molecule, although it is not clear at present whether this can be a triplet state, or must be singlet. In the molecular aggregate the gap should correspond to the "conduction band" arising from the individual molecule states.

The term conduction band is used advisedly, since in the face of weak intermolecular interaction, the question is just how much band development there is actually. There is less caution than enthusiasm in the field at present for lifting the ideas of the inorganic semiconductor theory and applying them directly to the molecular semiconductor cases. The values of semiconductivity of these two very different groups of materials are often many orders of magnitude apart, and doubtful progress can be expected if similarities rather than differences are explored.

Most of the experimental work naturally has been concerned with measurements of electrical semiconductivity of dried, generally crystalline, samples at temperatures from 20°C to several hundred degrees centigrade. In the case of proteins and other biological materials, the question of ionic conductivity by adsorbed salt impurities always comes up; a recent investigation has turned up an ionic conduction mechanism in a case typical of those thought to be electronic semiconductors. The whole technique of measurement of semiconductivity seems to provide information which may be very remote from actual electron mobility problems in biological macromolecule and molecular-aggregate systems.

The review by Garrett (1959) offers a critical appraisal of the status of this field.

IV. Molecular Memory Models³

In numerous biological problems the question arises of a molecular mechanism for the stabilization of transient intermediates. Photoinactivation and photoreactivation, vision, photosynthesis, radiobiological events, and others should have some interpretation on a molecular basis largely in terms of known molecular phenomena—possibly occasionally in terms of yet undiscovered ones. The mechanism of brain function ultimately must also reside in molecular electronic phenomena, even though its complexity may mask this necessary condition. It is worthwhile to examine the nature of transient molecular states and species, and especially with regard to their time constants, to see whether any of these could possibly serve as a basis for models for the biological phenomena studied.

Molecular electronic excitation principles for photoexcitation are

³ Taken with amendment from the abstracts of the MIT Biology Department seminar series "Molecular Specificity and Biological Memory" (Kasha 1962), Spring 1961.

generally understood sufficiently today to permit a description of their main features (Kasha, 1960b). The commonly observed electronic transitions of conjugated polyatomic molecules involve two broad classes of molecular orbital promotional types, $\pi \rightarrow \pi^*$ (involving fully delocalized molecular orbitals), and $n \rightarrow \pi^*$ (involving lone-pair atomic electron to π -anti-bonding molecular orbital promotion (Kasha, 1961a). Both singlet excited states and triplet excited states (Lewis and Kasha, 1944; Kasha and McGlynn, 1956), are generated with high quantum yield, in spite of the illusory paradox of the prohibition of singlet-triplet transitions. However, as much as triplet states of molecules have appealed to the biochemist as transient molecular intermediates, or as energy-storage states, such states could never be expected to persist longer than 1 to 30 seconds in most cases; in many cases mean lifetimes of much shorter than a millisecond are expected.

Intramolecular energy transfer in photoexcited metal chelates (Weissman, 1942; Crosby and Kasha, 1958; Crosby and Whan, 1960; Crosby, *et al.*, 1961) might appeal as a molecular mechanism in certain metal-organic systems. In the cases where the phenomenon is observed, however, the excited states of the metal ion, produced upon ultraviolet excitation of the aryl chelating groups, have a persistence time of only the order of milliseconds.

Molecular exciton interaction in photoexcited molecular aggregate systems (polymers, molecular laminae) lead to the production of exciton bands of levels, which for certain geometries have forbidden lowest levels. However, molecular distortions of the aggregate and singlet-triplet interaction both serve to remove the prohibition against transitions to lower states (Levinson, Simpson, and Curtis, 1957; McRae and Kasha, 1958). Consequently, metastable exciton levels generally cannot be thought of as a useful energy storage device.

In the above three broad fields, the natural mean lifetime is too short, or can be too readily shortened by external perturbation effects to permit the phenomena involved to serve as any persistent molecular memory, even if light or radiation were involved in the excitation process.

Reversible photochemical processes offer the first long-time molecular intermediate or energy storage possibility among molecular phenomena. Such cases were studied extensively by Lewis and Lipkin (1942) and more recently by Linschitz, Rennert, and Korn (1954).

The basic simplicity of the process, e.g., the photoejection of an electron from the molecule excited and its capture by the solvents, allows efficient reversibility in a liquid system in many cases; but in rigid glass systems the trapped electron and odd molecule ion may persist for extremely long times (days, years). However, such processes have been observed only under the action of light excitation, and more recently, high energy radiation (Hirshberg, 1957).

Photoreversible photochemical processes are even more interesting from the biological point of view, since they may offer a model for the various biological photoreactivation phenomena. The type systems thus far studied (Hirshberg, 1956) are principally spiropyrans and bianthrone, but the phenomena are probably more generally observable. Certainly the further study of this phenomenon is of the greatest interest.

All of the above cases involve photoexcitation or radiation excitation (which in the last stages will be analogous to photoexcitation), and as a molecular model for general biological processes, such phenomena could not be looked upon too realistically, except for the gap in our knowledge of other means of electronic excitation, e.g., excitation of electronic states of molecules during chemical reactions.

Electroluminescence of molecular vapors (Tesla discharge) is a well-known phenomenon which deserves some special study in this connection. If one could find a reversible chemical process (analogous to the photochemical cases discussed above), caused by and reversed by an electrical potential applied to a solid state system, one might then have a useful molecular model which could have biological applications to such a problem as storage and recall of information in the brain. The microwave and thermally induced fading of memory might be expected consequences of this model.

In this presentation the most general viewpoint of energy storage on the basis of molecular phenomena has been considered, with a cataloguing of molecular processes as they are known. If the brain works by nucleic acid information storage, the problem of the quantum chemist and solid state physicist and solid state chemist will be to find what molecular electronic mechanisms are operative and in what fashion, especially in the fascinating problem of recall. The search for molecular memory models, inspired by the great puzzle of the molecular basis of brain function, is certain to lead to new approaches to research on the brain. At the same time, deficiencies

in our present knowledge of molecular phenomena can be expected to be exposed.

V. Commentary on the Molecular Orbital Method⁴

The molecular orbital method as developed by Hund, and by Mulliken, and extended in simplified form by Hückel, has become the most feasible and widely used procedure for the calculation of electronic properties of polyatomic molecules. The applications of the method are very extensive and cover molecules of relatively great complexity, including large planar aromatic polycyclic hydrocarbons, and large heterocyclic molecules of biochemical significance, including amino acids, porphyrins, the DNA base-pairs, coenzymes, etc.

The biochemist not initiated into the secrets of the methods used often wonders how can such calculations be done (when some physicists are still struggling with simple diatomic molecules!) and what do the calculated predictions, which seem to flow out so easily, mean (when some chemists and biochemists still seem to have to resort to the old-fashioned method of doing experiments!)?

A flow chart of the simple semiempirical molecular orbital procedure (Fig. 2) will serve to indicate the operational nature of the steps involved. In the central band of boxes running from left to right are indicated the individual steps in the calculation. The arrows are to be read: "leads to."

Below the main band of steps are indicated the theoretical steps and approximations which are used, and above the main band are the crucial empirical elements without which the calculation could not be carried out.

The object of the molecular orbital method for π -electrons is to deduce the form of orbital for each π -electron, and to calculate its energy. Each orbital is assumed to be an independent or one-electron orbital, calculated with respect to the resultant potential field of the nuclei and remaining electrons in the molecule. Each orbital is a spatial distribution function enveloping the skeleton of the conjugated parts of the molecule, and must conform to the symmetry of this skeleton. The fact that the orbital for π -electrons is considered separately from the σ -bonding framework of the skeleton, and the fact that a one-electron orbital is used, disregarding interelectronic

⁴ This section is a supplement to the author's remarks presented at the 11^e Réunion de la Société de Chimie Physique (Kasha, 1961b).

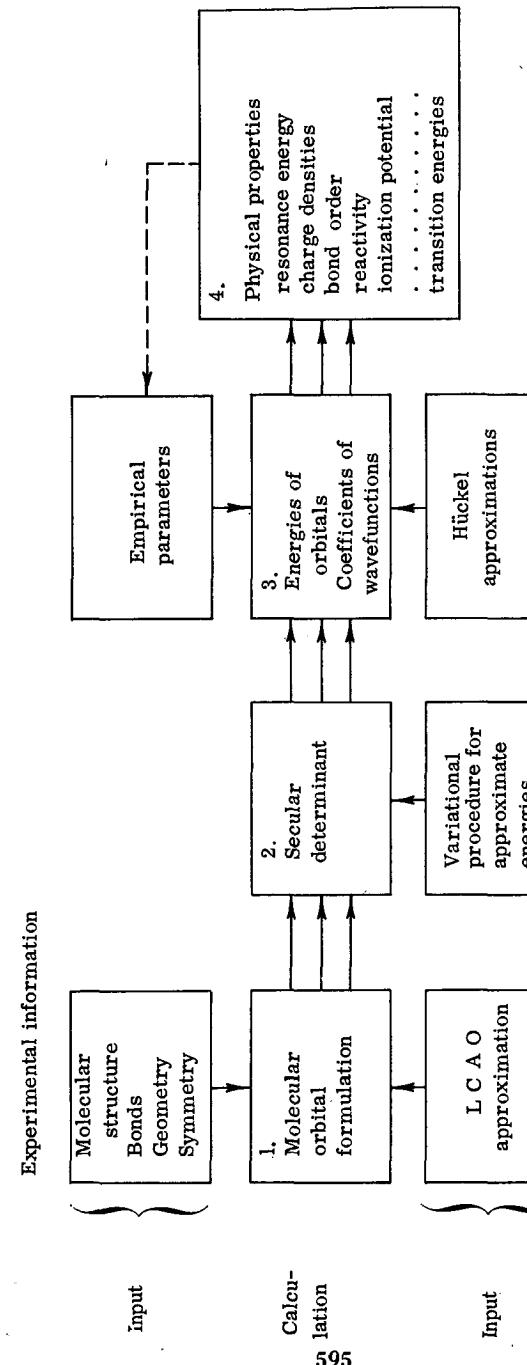


FIG. 2. Flow chart of the simple semiempirical molecular orbital calculation. (After M. Kasha, *J. chim. phys.*, 1961.)

correlation effects, are both approximations. Furthermore, the LCAO (linear combination of atomic orbitals) approximation considers that a satisfactory form for the molecular orbitals can be obtained simply by proper linear combinations of atomic orbitals. Coulson (1961) indicates the steps in the variational procedure; the better the approximation of the orbital assumed, to the correct shape an orbital would have as a one electron orbital in the real molecule, the closer will be the orbital energy calculated, to the true energy. The Hückel approximations simplify the final determinant so that the roots for the energy may be extracted. But the orbital energies are obtained as a function of two unknown parameters, α and β ,

$$\epsilon = \alpha + k\beta.$$

The parameters α and β are the coulomb integral and resonance integrals respectively. In the simple procedure, β is evaluated empirically instead of by calculation, and α is not needed for comparing relative values of the orbital energy. Heteroatom α 's and β 's also must be evaluated empirically in terms of the values for carbon. This empirical fitting is indicated by the cyclic process involving steps 3, 4, and "empirical parameters." Some physical property for some known molecules is accepted as a standard, and the values of the heteroatom parameters are adjusted so that the physical property is reproduced most closely (β for carbon having been fixed previously). Of course, up to this point nothing new has been calculated concerning molecular properties. But by using the best set of empirical parameters, properties of other molecules in principle can be now determined, or predicted. The recent book of Streitwieser (1961) may be consulted for detailed steps and refinements.

Obviously an accurate solution of the Schrödinger equation of quantum mechanics is not involved, and inherent in the simple molecular orbital procedure are many simplifying assumptions. As indicated in the foregoing, these assumptions include the idea of the existence of orbitals in molecules, the use of one-electron orbitals, the neglect of inter-electronic correlation within an orbital, and the neglect of σ - π electron interaction. In addition, inherent in the molecular orbital method is an overemphasis on highly ionic molecular orbital configurations. Although it can be expected that the empirical parameters in the simple molecular orbital method may compensate in part for the errors introduced by the neglect of

terms arising from the above-named assumptions and deficiencies, it could not be anticipated that so many qualitatively different kinds of deficiencies could be compensated by one type of empirical correction. Thus, for a particular set of semiempirical parameters, variations in calculated molecular properties of 20% or more can occur in comparison with experiment. The more closely related is a set of molecules, and the more restricted is the set of physical properties to be compared, the more precise can the fit of calculated to observed be made. The fact that these are inherently vapor state calculations, with solvation effects neglected, is a worrisome point.

Generally speaking, the value of the method is that it does give a numerical guide to the ordering of molecular properties. Thus, although the actual numbers obtained may have little absolute accuracy in some cases, they may serve to establish a rational correlation of experimentally observed quantities.

VI. Conclusion

In this essay a number of topics in quantum chemistry have been singled out for evaluation and comment as an indication of the potential which these topics offer for detailed molecular electronic mechanisms in biological phenomena. A number of important topics, such as ligand-field theory of metal ions, the intricacies of intramolecular electronic excitation, and numerous others have been omitted from discussion here. In fact, the limits of space on such an article have prevented more than a passing commentary. Nevertheless, the author hopes some illumination of the field will result.

To a researcher approaching biological phenomena from the side of physical theory, it is obvious that some pessimism, born of past experience on the part of biologists with regard to contributions from quantum mechanics, is fairly common. It is apparent, however, that a new breed of quantum chemists is in the process of evolution, with a mixture of optimism and experience in suitably developed approximation methods for large molecules, and also a humility before the wonders of the biological world. Certainly there are many indications of a new and active participation of physical scientists in biological fields.

It is most appropriate to close with the encouragement given by Coulson (1960):

"The work itself, however, will almost inevitably be crude,

and of the type that it is now fashionable to call semiempirical. . . . But in the establishment of correlations and primitive patterns of understanding it does not do to be too fussy. A rough track through the jungle precedes the construction of a metaled highway. And there is much experience possessed by professional biologists which could be linked with the deeper levels of interpretation associated with the quantum theory, to the enrichment of both. But let no one here make his claims too easily or too definitely. Biological systems are much more perverse than are laboratory chemical systems. In this field the prizes are immense—no less than the understanding and control of life itself. The future here may be far off. But, as was first said in rather different circumstances, "Ce n'est que le premier pas qui coûte"; and there are very few sensible people who would wish to deny that in this ultimate human enterprise there is to be no contribution from quantum chemistry."

REFERENCES

- Beukers, R., Ijlstra, J., and Behrends, W. (1958). *Rec. trav. chim.* **77**, 729.
 Coulson, C. A. (1960). *Rev. Modern Phys.* **32**, 170.
 Coulson, C. A. (1961). "Valence," 2nd ed. Oxford Univ. Press, London and New York.
 Crosby, G. A., and Kasha, M. (1958). *Spectrochim. Acta* **10**, 377.
 Crosby, G. A., and Whan, R. E. (1960). *J. Chem. Phys.* **32**, 614; *Naturwissenschaften* **47**, 276.
 Crosby, G. A., Whan, R. E., and Alire, R. M. (1961). *J. Chem. Phys.* **34**, 743.
 Davydov, A. S. (1962). "Theory of Molecular Excitons" (translated by M. Kasha and M. Oppenheimer, Jr.), 175 pp. McGraw-Hill, New York.
 Förster, Th. (1960). In "Comparative Effects of Radiation" (M. Burton, J. S. Kirby-Smith, and J. L. Magee, eds.), pp. 300–319. Wiley, New York.
 Garrett, C. G. B. (1959). In "Semiconductors" (N. B. Hannay, ed.). R. Reinhold, New York.
 Hirshberg, Y. (1956). *J. Am. Chem. Soc.* **78**, 2304.
 Hirshberg, Y. (1957). *J. Chem. Phys.* **27**, 758.
 Kasha, M. (1959). *Rev. Modern Phys.* **31**, 162.
 Kasha, M. (1960a). In "Fast Fundamental Transfer Processes in Aqueous Biomolecular Systems" (F. O. Schmitt, ed.), pp. 3–6. Mass. Inst. Tech., Cambridge, Massachusetts.
 Kasha, M. (1960b). *Radiation Research, Suppl.* **2**, 243.
 Kasha, M. (1961a). In "Light and Life" (W. McElroy and B. Glass, eds.), pp. 31–64. The Johns Hopkins Press, Baltimore, Maryland.
 Kasha, M. (1961b). *J. chim. phys.* **58**, 914.
 Kasha, M. (1962). In "Molecular Specificity and Biological Memory" (F. O. Schmitt, ed.). Technology Press, Cambridge, Massachusetts.
 Kasha, M., and McGlynn, S. P. (1956). *Ann. Rev. Phys. Chem.* **7**, 403.
 Levinson, G. S., Simpson, W. T., and Curtis, W. (1957). *J. Am. Chem. Soc.* **79**, 4314.
 Lewis, G. N., and Kasha, M. (1944). *J. Am. Chem. Soc.* **66**, 2100.
 Lewis, G. N., and Lipkin, D. (1942). *J. Am. Chem. Soc.* **64**, 2801.
 Linschitz, H. L., Rennert, J., and Korn, T. M. (1954). *J. Am. Chem. Soc.* **76**, 5839.
 McGlynn, S. P. (1958). *Chem. Revs.* **58**, 1113.
 McRae, E. G., and Kasha, M. (1958). *J. Chem. Phys.* **28**, 721.
 Mulliken, R. S. (1950). *J. Am. Chem. Soc.* **72**, 600, 4493.
 Mulliken, R. S. (1951). *J. Chem. Phys.* **19**, 514.
 Mulliken, R. S. (1952a). *J. Am. Chem. Soc.* **74**, 811.
 Mulliken, R. S. (1952b). *J. Phys. Chem.* **56**, 811.
 Orgel, L. (1954). *Quart. Revs. (London)*, **8**, 422.
 Simpson, W. T., and Peterson, D. L. (1957). *J. Chem. Phys.* **26**, 588.
 Streitwieser, A. (1961). "Molecular Orbital Theory for Organic Chemists," 489 pp. Wiley, New York.
 Szent-Györgyi, A. (1941). *Nature* **148**, 157.
 Szent-Györgyi, A. (1951). "Chemistry of Muscular Contraction," 162 pp. Academic Press, New York.
 Szent-Györgyi, A. (1956). *Science* **124**, 873.
 Szent-Györgyi, A. (1956). "Bioenergetics," 143 pp. Academic Press, New York.
 Szent-Györgyi, A. (1960). "Introduction to a Submolecular Biology," 135 pp. Academic Press, New York.
 Weissman, S. I. (1942). *J. Chem. Phys.* **10**, 214.

Index

- Actin, 4, 382
Active site, 275
Active transport, 422
Adenosine triphosphate, 4, 28, 30, 33, 36, 139, 257, 417, 428
Aging, 190, 217, 281
Algae, 25, 63, 79, 80
Amino acid sequence, 191, 194, 197, 214, 215
Amino acids, 120, 146
Anaesthesia, 545
Ascorbic acid, 2, 71
Ascorbic acid oxidase, 462
Asymmetry, molecular, 11
Autotrophic plants, 61
Axon, 439
- Bacteria, photosynthetic, 27, 63, 77, 80
Bacterial transformation, 153
Bacteriochlorophyll, 41
Bacteriophage, 16, 157, 174
Band structure, 296, 586, 590
Base pairing, 109, 110, 569, 570
Biochemical evolution, 12, 20, 60, 67, 113, 198, 202, 204, 208, 210
Biological evolution, 245
Bioluminescence, 91, 100, 516, 519
Biopoesis, 59
Biosphere, 92
Biosynthetic pathways, 254, 260
- Carbon, path of, in photosynthesis, 27
Carbon reduction cycle, 26
Carbonyl transition, 235
Carcinogenesis, 373, 497, 501, 505, 578, 579
Carotenes, 17, 87
Catalase, 49, 71, 244
Catalysis, 48, 65, 91, 112, 116, 230, 266, 268, 272, 489, 490, 492, 546
Cell death, 506
- Cellular chemistry, 332
Cell theory, 321
Ceruloplasmin, 464, 465, 487
Charge transfer complexes, 355, 357, 453, 586, 589
Chemical evolution, 42, 111, 244
Chemiluminescence, 91
Chlorophyll, 17, 31, 37, 39, 41, 63, 68, 82
Chloroplast, 37, 38, 372
Chromosome, 211
Cistron, 147
Coding ratio, 107
Coenzyme Q, 389, 398, 399
Coenzymes, 575
group transfer, 576
oxidation-reduction, 575
Collagen, 47, 542
Complexes, charge transfer, 357, 453, 586, 589
Conduction band model, 586, 590
Conformation, molecular, 237
Contraction factor, 428
Contraction, muscle, 3
Copper, biological, 461, 467, 478, 483, 486
Copper enzymes, 462, 470, 472, 473, 480
Copper metabolism, 469
Copper proteins, 462, 468, 471, 476, 483, 486
Cytochrome, 33, 84, 203, 364, 389
Cytochrome oxidase, 462
- Dark reactions, photosynthetic, 28
Death hormone, 281
Deletions, 148
Denaturation, 516, 541
d-orbitals, 134
Deoxyribonucleic acid, 103, 123, 153, 168, 313, 321, 335, 518
synthesis, 324, 325
Development, embryonic, 212

Differentiation, 183
Diphosphopyridine nucleotide, 29, 96

Electron acceptor properties, 577
Electron donor properties, 577
Electron gas model, 348
Electron transfer, 33, 346, 452, 528
Electron transport, 92, 94, 96, 99, 288, 389, 406, 417
Electron transport particle, 388
Electronic delocalization, 564
Electronic indices, 562
Electrophysiology, 440
Elementary particle of mitochondria, 402, 403, 404, 405
Emerson effect, 81
Energy-rich phosphates, 573
Energy transfer, 452, 587, 592
Engram, 444, 447, 450
Enzyme, 5, 16, 28, 65, 93, 96, 116, 117, 127, 153, 155, 160, 161, 168, 175, 208, 230, 251, 265, 266, 268, 270, 273, 276, 285, 423, 462, 473
Eobiontic, 14, 18
Epigenesis, 322
Evolution, biochemical, 12, 20, 60, 67, 113, 198, 202, 204, 208, 210 biological, 245 chemical, 42, 111, 244 organic, 20, 60, 61, 242
Exciton theory, 587, 592
Extraterrestrial life, 124, 248

Fast reactions, 452
Flavin, 71, 94, 390
Free electron model, 348
Free radicals, 356

Gene, 145, 153, 167, 197, 198, 206, 207, 209, 212, 219
Genetic code, 154, 169, 326
Genetic map, 168, 170
Genetics, 80, 103, 147, 148, 154, 167, 281, 322, 323
Genotype, 150
Glial cells, 442, 443
Group transfer, 132

Heme group, 194
Hemocyanin, 436, 488
Hemoglobin, 14, 145, 146, 194, 195, 214, 215, 344 fetal, 150, 195, 218 fish, 199, 216 gorilla, 199 tetrahemic, 195
High pressure, 513
Hormone, death, 281
Hormone, steroid, 498
Hydration of ions, 533
Hydrocarbons, polynuclear, 347, 497, 502
Hydrogen transfer, 3
Hydrophobic bond, 395

Inducer, 168, 169, 181
Information transfer reactions, 109, 153, 167
Ionizing radiation, 43, 243, 463, 508

Lamellae, 39, 405
L.C.A.O. method, 560
Life, definition of, 14, 117, 191, 319, 320 origin of, 12, 19, 59, 62, 92, 93, 111, 112, 113, 245, 247
Lipids, 391, 396
Luciferase, 91
Luciferin, 91

Mammary cancer, 497, 499
Mechanochemical changes, 425, 431
Messenger RNA, 107, 120, 157, 158, 170, 171
Metabolism, 12, 59, 279, 469 aerobic, 2
Meteorites, 20
Methemoglobin, 214, 216
Microgenetic map, 170
Mitochondria, 34, 94, 96, 381, 382, 383, 384, 386, 400, 414, 415, 417, 421
Molecular disease, 145, 189, 213, 220 definition, 190, 191

Molecular indices, 561, 562
Molecular language, 104
Molecular memory models, 591
Molecular orbital method, 559, 562, 594, 595
Multiple bonds, 129, 135
Muscular contraction, 3, 513
Mutation, 191, 201, 203, 204, 205, 210, 518
Myogen, 517
Myoglobin, 194, 196, 205
Myosin, 4, 279, 382

Neurofibrils, 445
Neurofilaments, 445, 446
Neurons, 439, 443, 444
Neurophysiology, 439
n-pi excited state, 82, 94, 235
Nucleic acid, 14, 65, 103, 104, 119, 335, 368, 566
Nucleotide sequence, 324

Octant rule, 237
Operon, 167, 169, 173
Optical activity, 229, 230, 231, 232, 240
Organic evolution, 20, 60, 61, 242
Oxidation, biological, 1, 285
Oxidative phosphorylation, 98, 410, 423

Periodic structures, 296, 298
Pharmacological agents, 415
Phenylketonuria, 192
Phospholipid, 397, 399, 401
Phosphorolytic cleavage, 258, 259
Phosphorus, biological, 127
Phosphorylases, 252, 258
Photoconduction, 369
Photoreactivation model, 593
Photosynthesis, 24, 32, 62, 74, 93, 94
Piezoelectricity, 335
Polymerase, DNA, 104, 155, 167
Polymerase, RNA, 106, 167
Polymers, linear, 328, 330
Polynucleotides, 113, 154, 157, 518
Polypeptides, 230, 517

Porphyrins, 48, 50, 63, 360
Primer, 104, 155, 325
Protein(s), 65, 112, 197, 336, 361, 364, 468, 568 energy bands in, 571 semiconductivity of, 572
Protein, structural, 395
Protein hydration, 366, 541, 544
Protein structure, 46, 145, 515
Protein synthesis, 106, 246, 517, 519
Protochlorophyll, 41
Proton transfer, 365, 452, 524, 525, 527, 546
Psychophysics, 437, 438, 441
Purines, 18, 159, 314, 569, 570, 577
Pyrimidines, 18, 159, 314, 569, 570, 577
Pyrophosphate, 36, 51, 93
Pyrophosphorylases, 252

Quantum chemistry, 553, 583
Quantum yield, photosynthetic, 68

Readout, 451
Replication, 67, 104, 112, 114, 122, 154, 321, 323, 324, 325, 326, 327, 329, 330, 331, 333, 450
Repressor, 168, 175, 179, 208
Respiration, 2, 86, 421, 422, 423, 433
Ribonucleic acid, 103, 123, 153, 168, 443, see also Messenger RNA, Stencil RNA, Transfer RNA soluble, 156, 180
Ribosomal particle, 105, 107, 155, 158, 444, 519
Ribulose diphosphate, 27

Satellite cells, 442
Semiconductors, 295, 304, 305, 337, 342, 343, 350, 354, 371, 590
Sequence, amino acid, 191, 194, 197, 214, 215 nucleotide, 324
Sickle cell anemia, 145, 193
Silicon, 129
Specificity, enzyme, 273

- Stencil RNA, 160, 164
Structural conservatism, 15
Sulfur, 127
Synthetase, RNA, 161
Synthetases, 253
Thalassemia, 193, 195, 213
TPNH, 29
Transducing systems, 381, 415, 448
Transfer RNA, 106, 156, 158, 180
Transition metals, 127, 408, 472
- Triplet code, 108, 120, 157, 171
Triplet state, 87, 95, 353, 586
Tyrosinase, 464
- Valence-bond method, 557
Virus, 153, 174
Vitamin C, 2
- Watson-Crick double-helix, 105, 154, 171, 178, 241, 315

KASHA
PULLMAN



LIBRARY OF THE UNIVERSITY OF TORONTO LIBRARIES

HORIZONS IN BIOCHEMISTRY

Albert Szent-Györgyi Dedicatory Volume

HORIZONS IN BIOCHEMISTRY

Albert Szent-Györgyi Dedicatory Volume

