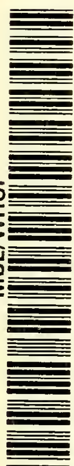
CURRENTS IN BIOCHEMICAL RESEARCH

Woods Hole, Mass.

Presented by

Estate of Dr. Otto Loewi
August, 1962


MBL/WHOI
0 0301 0010822 1

CURRENTS IN BIOCHEMICAL RESEARCH

CURRENTS IN BIOCHEMICAL RESEARCH

Edited by DAVID E. GREEN

Thirty-one essays charting the present course of Biochemical Research and considering the intimate relationship of biochemistry to medicine, agriculture and social problems

DR. OTTO LOEWI

INTERSCIENCE PUBLISHERS, INC., NEW YORK
1946

Copyright, 1946, by
INTERSCIENCE PUBLISHERS, INC.
215 Fourth Avenue, New York 3, N. Y.

Printed in the United States of America
by the MACK PRINTING COMPANY, EASTON, PA.

PREFACE

With the ever-increasing degree of specialization in scientific research and with the terrifying rate of growth of technical nomenclature, men of science are literally compelled to know more and more about less and less. The scientific literature furthers this trend, since journals, textbooks (apart from those for students), and review articles are written primarily for the specialist. There is an acute need for stripping complex subjects and getting at the simple, essential concepts which are basic to their appreciation. After all, the same scientific principles are applicable to all fields of inquiry. The art of presentation consists in the elimination of the barriers of terminology which effectively conceal these fundamental principles. *Currents in Biochemical Research* represents an attempt by some thirty research workers to describe in as simple language as possible the important developments in their own fields and to speculate a little on the most likely paths of future progress. The aim of these essays has been to excite the imagination and to provide glimpses of some of the fascinating horizons of biochemical research. However, no popularizations were intended. The various contributors were asked to write simply and provocatively but without sacrifice of scholarship. Dealing as they do on the one hand with pharmacology, chemotherapy, public health, genetics, photosynthesis, and agriculture and on the other with considerations of organic, analytical, and physical chemistry, they emphasize the focal position of biochemical research in biology, chemistry, and medicine. It is hoped that this survey from so many different points of view may assist biochemists, chemists, and medical doctors in seeing biochemistry in clearer perspective and in its proper relation to other fields of inquiry.

David E. Green

March, 1946

CONTENTS

PAGE

1. THE GENE AND BIOCHEMISTRY by *G. W. Beadle* 1
2. VIRUSES by *W. M. Stanley* 13
3. PHOTOSYNTHESIS AND THE PRODUCTION OF ORGANIC MATTER ON EARTH by *H. Gaffron* 25
4. THE BACTERIAL CELL by *René J. Dubos* 49
5. THE NUTRITION AND BIOCHEMISTRY OF PLANTS by *D. R. Hoagland* 61
6. BIOLOGICAL SIGNIFICANCE OF VITAMINS by *C. A. Elvehjem* 79
7. SOME ASPECTS OF VITAMIN RESEARCH by *Karl Folkers* 89
8. QUANTITATIVE ANALYSIS IN BIOCHEMISTRY by *Donald D. Van Slyke* 109
9. ENZYMIC HYDROLYSIS AND SYNTHESIS OF PEPTIDE BONDS by *Joseph S. Fruton* 123
10. METABOLIC PROCESS PATTERNS by *Fritz Lipmann* 137
11. BIOCHEMISTRY FROM THE STANDPOINT OF ENZYMES by *David E. Green* 149
12. ENZYMIC MECHANISMS OF CARBON DIOXIDE ASSIMILATION by *Severo Ochoa* 165
13. HORMONES by *B. A. Houssay* 187

59913

CONTENTS

		PAGE
14.	FUNDAMENTALS OF OXIDATION AND REDUCTION by *Leonor Michaelis*	207
15.	MESOMERIC CONCEPTS IN THE BIOLOGICAL SCIENCES by *Herman M. Kalckar*	229
16.	VISCOMETRY IN BIOCHEMICAL INVESTIGATIONS by *Max A. Lauffer*	241
17.	ISOTOPE TECHNIQUE IN THE STUDY OF INTERMEDIARY METABOLISM by *D. Rittenberg and David Shemin*	261
18.	MUCOLYTIC ENZYMES by *Karl Meyer*	277
19.	SOME ASPECTS OF INTERMEDIARY METABOLISM by *Konrad Bloch*	291
20.	THE STEROID HORMONES by *Gregory Pincus*	305
21.	PLANT HORMONES AND THE ANALYSIS OF GROWTH by *Kenneth V. Thimann*	321
22.	CHEMICAL MECHANISM OF NERVOUS ACTION by *David Nachmansohn*	335
23.	SOME ASPECTS OF BIOCHEMICAL ANTAGONISM by *D. W. Woolley*	357
24.	CHEMOTHERAPY: APPLIED CYTOCHEMISTRY by *Rollin D. Hotchkiss*	379
25.	BIOCHEMICAL ASPECTS OF PHARMACOLOGY by *Arnold D. Welch and Ernest Bueding*	399
26.	SOME BIOCHEMICAL PROBLEMS POSED BY A DISEASE OF MUSCLE by *Charles L. Hoagland*	413
27.	PHYSIOLOGY AND BIOCHEMISTRY by *Surgeon Captain C. H. Best*	427
28.	X-RAY DIFFRACTION AND THE STUDY OF FIBROUS PROTEINS by *I. Fankuchen and H. Mark*	439
29.	IMMUNOCHEMISTRY by *Michael Heidelberger*	453
30.	SOCIAL ASPECTS OF NUTRITION by *W. H. Sebrell*	461
31.	ORGANIZATION AND SUPPORT OF SCIENCE IN THE UNITED STATES by *L. C. Dunn*	473

THE GENE AND BIOCHEMISTRY

G. W. BEADLE, PROFESSOR OF GENETICS, SCHOOL OF BIOLOGICAL SCIENCES, STANFORD UNIVERSITY

IT IS both an accident of organic evolution and an indication of man's lack of foresight that the organisms studied in most detail by biochemists have not been those on which geneticists have concentrated. It is natural that man should have a prejudice in favor of himself, and it is therefore not remarkable that the urge of medicine on biochemistry has been in the direction of specialization on mammals, particularly on man himself. For obvious reasons, bacterial biochemistry has likewise been well nourished through medicine. Man has few inherent advantages for biochemical study while the bacteria abound in them. But both are most difficult for the geneticist—the one because of a long life cycle and social obstacles to controlled matings, the other because of the absence of a sexual cycle without which the geneticist cannot use his particular methods. The geneticist, on the other hand, has chosen to make the vinegar fly and Indian corn the classical organisms of his science. Both suffer disadvantages to the biochemist in not lending themselves readily to culture under precisely defined environmental conditions. Neither can be grown conveniently on a medium completely known from a chemical standpoint.

In spite of this situation and additional impediments arising through divergence in outlook, such persons as Garrod, Onslow (née Wheldale), Troland, Goldschmidt, Wright, Haldane, and others have

urged that the two fields have much in common and that each stands to profit through contact with the other. Through the efforts of these individuals and others of like mind there are many instances known in which the relation of genetics to biochemistry is so clear that it can no longer be disregarded by intelligent investigators in either field. In fact, from this relation there tends to emerge a new interest, known as biochemical genetics, which promises to tell us what the genes do and how they do it, on the one hand, and to lead us to further knowledge in the ways of biosynthesis on the other. In both directions there obviously lie many opportunities.

One of the earliest instances in which a Mendelian trait could be interpreted in terms of specific chemical reactions is that involving the human disease known as alcaptonuria. In individuals homozygous for the mutant gene responsible for this character, 2,5-dihydroxyphenylacetic acid (homogentisic acid or alcapton) is excreted in the urine instead of being broken down to carbon dioxide and water, as it is in persons receiving the normal form of the alcaptonuric gene from one or both parents (15). Homogentisic acid is oxidized to a black pigment on exposure to air and it is this process that is responsible for darkening of the urine, the most striking symptom of the disease. According to Gross (cited by Garrod), alcaptonurics lack a specific enzyme found in the blood of normal persons which catalyzes the degradation of homogentisic acid. Alcaptonuria therefore represents the first recorded instance in which it could be said that a particular chemical reaction is controlled by a known gene through the mediation of a specific enzyme.

Within the past dozen years, additional examples have become known in which organisms unable to carry out specific reactions differ in a single gene from their chemically more successful relatives. In flower pigment synthesis, for example, the formation of carotenoids, anthocyanins, anthoxanthins, chalcones, and flavocyanins is known to be genetically controlled in one plant or another (7,24). Specific oxidations of pelargonidin derivatives to cyanidin analogues and of cyanidin compounds to delphinidin counterparts are dependent on the activities of specific genes. The addition of sugars to anthocyanidins through glycosidal linkages and the transformation of the anthoxanthin quercetin-3-glucoside to the corresponding cyanidin-3-glucoside are likewise unable to proceed if specific genes are modified.

Fölling (14) and Penrose (35) have shown that the genetically determined failure to oxidize phenylpyruvic acid in man is invariably associated with subnormal mentality. Here again there appears to be an intimate relation between a particular gene and a specific chemical reaction. Because of its obvious importance to an understanding of the mechanisms underlying mental processes, this case is of particular interest. It is of course related metabolically to alcaptonuria in so far as phenylalanine is concerned in both. Other abnormalities in phenylalanine–tyrosine metabolism are also known (15,19).

Most remarkable progress has recently been made in understanding the genetic and chemical mechanisms of sex determination and differentiation in the green alga, *Chlamydomonas*, by Moewus, Kuhn, and co-workers. From the carotenoid pigment, protocrocin, there is derived through cleavage the motility hormone, crocin, and a female-determining hormone known as gynotermone. This cleavage is known to be genetically controlled. In genetically male individuals gynotermone is hydrolyzed to a male-determining hormone known as androtermone. Under the direction of specific genes the *cis* and *trans* forms of the motility hormone, crocin, are converted into the corresponding *cis* and *trans* dimethyl esters of crocetin. In various specific mixtures, these serve as gamones, *i. e.*, they render individuals of the specific genetic constitutions capable of conjugation. The relations between genes and chemical reactions disclosed by this work support the thesis that genes act in directing specific processes. The work on *Chlamydomonas* is so spectacular and its importance so great that independent confirmation is desirable (see 7,28,41).

The splitting of specific di- and trisaccharides by yeasts is under genetic control, as shown by Winge and Laustsen (54) and Lindegren, Spiegelman, and Lindegren (25). It appears that the genes concerned determine whether or not specific enzymes are present in active form. Somewhat similar situations are known in the rabbit, where Sawin and Glick (37) have shown that the activity of the enzyme, atropine esterase, is dependent on the presence of the normal allele of a particular gene, and in white clover, where an enzyme responsible for hydrolysis of specific cyanogenetic glucosides is known to show a similar dependence on a gene (2,9).

In the bread mold, *Neurospora*, Srb and Horowitz (44) have shown that there is present an ornithine cycle essentially similar to

that postulated by Krebs and Henseleit for the mammalian liver. In the bread mold it is known that mutation in any one of seven different genes will interrupt the synthesis of ornithine or its conversion to arginine. So far as the data go, they are consistent with the assumption that each of the seven genes is normally concerned with a different chemical reaction in the system. It is an interesting point that it was possible to establish the presence of the ornithine cycle in *Neurospora* because of the existence of the mutant strains indicated.

Tatum and Bonner (50) have shown that tryptophan is normally synthesized in *Neurospora* through the condensation of indole and serine. Evidently the indole is somehow derived from anthranilic acid, for there exist two mutant strains, one of which accumulates anthranilic acid when it is grown under suitable conditions, while the other is able to grow normally when supplied with anthranilic acid in place of indole or tryptophan (51). The gene by which the first strain differs from wild type is evidently concerned with the reaction by which anthranilic acid is converted into indole, whereas the mutant gene of the second strain appears to be concerned with failure of some reaction essential to the synthesis of anthranilic acid. It is obvious, in this case, that genetics has provided a tool of great usefulness in investigating the biosynthesis of the important amino acid, tryptophan.

Relations similar to those mentioned above are known for other biosyntheses in the bread mold and in other organisms. By following methods developed by Beadle and Tatum (8), it has been possible to obtain a series of mutant strains of *Neurospora* in each of which some particular reaction has been blocked. These are concerned with the synthesis of amino acids, vitamins, purines, pyrimidines, and other compounds of biological importance (6,21,48,49).

We can be sure from such cases as those just cited that genes function in directing biochemical reactions. We know, further, that this direction may involve enzymes as intermediates between gene and reaction. All our information is consistent with the hypothesis that in all cases in which genes control specific reactions they do so indirectly through enzymes. In other words, genes direct enzyme specificities, and enzymes control reactions. This is not a new idea. Bateson (4), Moore (29), Troland (52), Goldschmidt (16), Muller (30), Alexander and Bridges (1), Haldane (18,19), Wright (55), and others have sug-

gested it. We are only now beginning to do something definite about it from an experimental standpoint. Since the specificities of enzymes are referable to protein specificities, the hypothesis implies that genes direct protein specificities. In this case we might expect that the specificities of proteins other than those found in enzymes would show a direct relation to genes. This is indeed the situation as evidenced by the fact that in many organisms a general one-to-one relation between genes and antigens has been shown (22,47). It is true that a few deviations from this correspondence are known, but they may well represent instances in which antigens have specificities made up of two components, each corresponding to one gene.

If we knew the chemical nature of genes, we should be in a much better position than we are now to determine how they direct protein specificities. Direct chemical analyses of whole chromosomes show them to be largely nucleoprotein (27), which suggests that genes too are nucleoproteins. But since chromosomes probably contain much nongenic material, the deduction is not too satisfying. Ultraviolet radiation induces gene mutation; and its efficiency in this respect varies with wave length in the same way as does its absorption by nucleic acid (20,45), strongly indicating that the energy effective in producing mutations in genes is absorbed by nucleic acid. The simplest assumption possible is that this is so because the nucleic acid is part of the gene.

The similarity of genes and viruses constitutes a third line of evidence concerning the chemical nature of genes. Both have the property of self-duplication, which in both cases is dependent on the presence of a series of compounds such as those found in the living cell. Genes and viruses appear to be within the same size range (46). Both are capable of undergoing mutation to new forms which have altered biological activities but retain the power of self-duplication (46). Since viruses and genes have so many properties in common, it is probable that they are similar in chemical makeup. Following Stanley's isolation of crystalline tobacco mosaic virus, several other viruses have been prepared in pure form and all have been shown to be nucleoproteins (5,12,46). The circumstantial evidence that genes, too, are nucleoproteins, or at least contain nucleoproteins as essential parts, is therefore substantial.

In duplicating themselves, genes have been assumed to act as

master molecules or models from which exact copies are made (11,18, 19,30,31,55). If this is so, their action may be visualized as one of directing the construction of specific protein types plus whatever other component parts genes may have. If the specificities of proteins generally are copied from genes, the observed relations between genes and enzymes and between genes and antigens should follow. For every specific protein there should exist a gene carrying this same protein. For every enzyme and antigen type there likewise should be a gene. Because of a general tendency of mutation pressure to eliminate genes that are of no advantage to the organisms, it might be expected that for every protein type there would be only one corresponding gene. The experimental evidence appears to support this general interpretation, although it must be recognized that, in dealing with genetic traits that can be described in terms of chemical reactions, there may be an unavoidable selection of those cases in which gene action is relatively simple.

However proteins and other components of genes are synthesized —whether by an orthodox stepwise mechanism or by some as yet unknown mechanism by which many component parts are simultaneously directed into their proper places by the master molecule (11)—the precursors of proteins, nucleic acids, and whatever other parts genes may have, must be synthesized. Their synthesis will involve many enzymes and a corresponding number of genes. Thus, before one gene can determine the specificity of a new protein molecule, many other genes must have acted. This amounts to saying that, in any multigenic organism, the genes constitute a highly organized system, just as the chemical reactions they direct are integrated in time and space in a manner characteristic of a particular species. Furthermore, while a particular gene will have only one primary action in determining specificity of an enzyme or an antigen, the final physiological consequences of a change in a single gene will be manifold. This can be appreciated when one considers the consequences of depriving an organism such as a rat of thiamin. The final consequence is of course death, but before death occurs a series of changes of increasing complexity take place. These can be brought about in the rat by removing thiamin from the diet. In the bread mold, which normally synthesizes thiamin, the same end result can be effected as a result of an analogous series of changes initiated by replacing a normal gene

necessary for thiamin synthesis with a defective form of the same gene. In one particular case, the primary action of the gene is presumed to be in directing the specificity of the enzyme catalyzing the reaction by which thiazole and pyrimidine are combined (49). Viewed in this way, an understanding of gene action does not appear hopelessly difficult even though the final effects of a single gene change may involve alterations so complex as to defy complete description. Grüneberg (17) has pointed out that a similar type of interpretation in terms of one primary action of a given gene is tenable in the case of certain hereditary developmental defects in the mouse and rat that at first sight appear to involve several unrelated changes in the organism. That the gene has a functional as well as structural unity is therefore a hypothesis that has demonstrated its heuristic value. Until evidence with which it is inconsistent is presented, it will no doubt continue to play an important role in our concepts of what the gene is and how it acts.

As Troland (52), Muller (30), Alexander and Bridges (1), Oparin (33), Plunkett (36), and others have pointed out, the similarity of viruses and genes suggests that the first living structures, *i. e.*, those with the power of self-duplication, were probably somewhat similar to present-day viruses with the important difference that they were free-living. The evolution of systems of such units, each acquiring the property of directing the specificity of an enzyme or other protein, would be expected to give rise to a series of forms of increasing complexity such as we see today in the larger and more complex viruses, the rickettsias, bacteria, and higher organisms. It is probable that the present viruses and rickettsias are not relics of these ancestral forms but are forms secondarily derived through specialization in connection with parasitism (10). The true ancestral types must have been capable of multiplying outside living cells in a kind of environment which, because of the presence of many organisms, is no longer likely to exist (33). In terms of genes directing chemical reactions through their control of enzyme specificities it is possible to imagine how, in principle, these simple forms evolved in the direction of the more highly specialized and complex forms of multicellular plants and animals (56), although it is of course not easy to visualize the way in which the process occurred in detail in particular instances.

In the specialization of higher animals with respect to their nutri-

tion, it is possible to suggest a scheme of evolution that has some support at least in analogy. It has become increasingly evident that, with respect to their need for and use of vitamins of the B group, purines, pyrimidines, choline, amino acids, and other compounds, all cellular organisms are fundamentally very similar (23,26,53). To consider a specific example, carboxylase is presumably present in all protoplasm and apparently always contains thiamin as thiamin pyrophosphate. Many organisms, *e. g.*, most plants, are able to synthesize the thiamin they need, while others are dependent on an external supply of this essential compound. From an evolutionary standpoint, this difference is presumably determined by whether or not it is of advantage to a particular organism to synthesize thiamin. Evidently for *Neurospora* there is selective advantage in being able to carry out this synthesis for we find in wild strains that all essential genes concerned with it are present in active form. In mammals, on the other hand, thiamin is presumably so frequently present in the diet that the genes originally concerned with its elaboration have been permitted by natural selection to become inactive so far as thiamin synthesis goes. It may well be that they have not disappeared entirely but have been modified so as to enable the mammal to carry out chemical reactions of which the bread mold is incapable. In a similar way, mammals have become specialized through loss of ability to synthesize other vitamins, the indispensable amino acids, and other compounds. With the development of parasitism it would be expected that still further loss in synthetic ability would be encountered. As Knight (23), Lwoff (26), Schopfer (38), and others have pointed out, this is indeed the case. The work on *Neurospora* makes it most probable that the dropping out of specific chemical reactions no longer of selective advantage is the result of gene mutation. The limit of such parasitic specialization is probably represented in the molecular viruses that have lost all power of heterosynthesis and have retained only the one property essential for their continued existence in an environment in which all necessary compounds are available—the property of autosynthesis.

One may quite properly raise the question as to the course of positive evolution in terms of chemical reactions—how are new syntheses developed in the course of organic evolution? Unfortunately, the experimental evidence bearing on this is meager, which is not surprising, for obviously it should be much easier to destroy or inactivate a

complex self-duplicating unit than to modify it so as to give it a new and useful property without sacrificing its power of self-duplication. The first self-duplicating unit must have evolved from nonliving matter at some time, and more complex forms must have evolved from it—the alternative is some form of special creation. There would seem to be less difficulty in imagining a primitive "protogene" mutating to a true gene with a heterocatalytic property than its spontaneous origin in the first place, even if, as Oparin (33) supposes, it arose in a world containing preformed organic molecules of many kinds. Nor is there any apparent reason why such protogenes could not mutate in many different directions in order to give rise to many different organic catalysts. In present-day cellular organisms there exists a possible mechanism for acquiring totally new reactions. Occasionally, through accident, one or more genes become duplicated, *i. e.*, a small segment of a chromosome occurs twice in every set. The duplicated genes will be unnecessary to the organism and will be expected to disappear through loss mutations, since such mutations are not disadvantageous. But such a duplicate gene may occasionally undergo mutation in such a way that it directs the formation of an entirely new enzyme. If this new enzyme should happen to catalyze a reaction that improved the organism in competition with its relatives, the new reaction would be retained. Such new reactions might add new compounds or they might bring about the reverse of the specialization process, which leads in the direction of parasitism. In this way, as Horowitz (20a) has pointed out, the first primitive organisms might gradually have built up systems of synthesis which freed them of their dependence on preformed organic molecules originally present in the environment.

Through such advances as have been indicated we appear to be moving rapidly in the direction of a better understanding of what genes are and what they do. We are no longer content with a knowledge of the laws by which they are transmitted from one generation to the next. We see that they are basic functional units of the organism and that, by taking advantage of their tendency to mutate, we can use them as powerful tools in determining the course of biosynthesis and in understanding other aspects of metabolism. Their relations to enzymes and antigens are becoming known. Precisely how they function in duplicating themselves and in directing the specificities of proteins, nucleic acids, and possibly other large molecules is a question

for the future. But there can be no doubt that the years that lie ahead will be exciting ones in this field. The work of Avery and his co-workers (3) on the transformation of types in *Pneumococcus* and that of Emerson (13) and of others suggests that we may one day learn to direct gene mutations in predetermined ways. Work on enzymes (32) and viruses (5,46) is so closely related to the general problem of gene structure and gene action that only short steps appear to be necessary to bridge the gaps that separate them. Nucleic acid certainly plays an important role in gene action and in protein synthesis (34,39,40), and it is not too much to hope that this role will be made clear in the near future. The relation of genes to cytoplasmic elements is not well understood, but after many years in which discouragingly little progress has been made, important leads are being followed by Sonneborn (42), Spiegelman (43), and others. After half a century of growth, genetics seems to be assuming a position in the broad field of biology in which its close relations to evolution, development, physiology, and biochemistry are now more evident.

References

(1) Alexander, J., and Bridges, C. B., in *Colloid Chemistry.* Vol. II, Chemical Catalog Co., New York, 1928.
(2) Atwood, S. S., and Sullivan, J. T., *J. Heredity*, **34**, 311 (1943).
(3) Avery, O. T. F., McLeod, C. M., and McCarty, M., *J. Exptl. Med.*, **79**, 137 (1944).
(4) Bateson, W., *Mendel's Principles of Heredity.* Cambridge Univ. Press, London, 1909.
(5) Bawden, F. C., *Plant Viruses and Virus Diseases.* Chronica Botanica, Waltham, 1943.
(6) Beadle, G. W., *Physiol. Revs.*, **25**, 643 (1945).
(7) Beadle, G. W., *Chem. Revs.*, **37**, 15 (1945).
(8) Beadle, G. W., and Tatum, E. L., *Proc. Natl. Acad. Sci. U. S.*, **27**, 499 (1941).
(9) Corkill, L., *N. Zealand J. Sci. Tech.*, **B23**, 178 (1942).
(10) Darlington, C. D., *Nature*, **154**, 164 (1944).
(11) Delbrück, M., *Cold Spring Harbor Symposia Quant. Biol.*, **9**, 122 (1941).
(12) Delbrück, M., in *Advances in Enzymology*, Vol. II. Interscience, New York, 1942, p.1.
(13) Emerson, S., *Proc. Natl. Acad. Sci. U. S.*, **30**, 179 (1944).
(14) Fölling, A., *Z. physiol. Chem.*, **227**, 169 (1934).

(15) Garrod, A. E., *Inborn Errors of Metabolism.* 2nd ed., Oxford Univ. Press, London, 1923.

(16) Goldschmidt, R., *Physiological Genetics.* McGraw-Hill, New York, 1938.

(17) Grüneberg, H., *The Genetics of the Mouse.* Cambridge Univ. Press, London, 1943.

(18) Haldane, J. B. S., "The biochemistry of the individual," in *Perspectives in Biochemistry.* Cambridge Univ. Press, London, 1938.

(19) Haldane, J. B. S., *New Paths in Genetics.* Harper, New York, 1942.

(20) Hollaender, A., and Emmons, C. W., *Cold Spring Harbor Symposia Quant. Biol.*, **9**, 179 (1941).

(20a) Horowitz, N. H., *Proc. Natl. Acad. Sci. U. S.*, **31**, 153 (1945).

(21) Horowitz, N. H., Bonner, D., Mitchell, H. K., Tatum, E. L., and Beadle, G. W., *Am. Naturalist*, **79**, 304 (1945).

(22) Irwin, M. R., and Cumley, R. W., *Am. Naturalist*, **77**, 211 (1943).

(23) Knight, B. C. J. G., *Med. Research Council* (*Brit.*), *Special Rept. Series*, No. 210 (1936).

(24) Lawrence, W. J. C., and Price, J. R., *Biol. Rev. Cambridge Phil. Soc.*, **15**, 35 (1940).

(25) Lindegren, C. C., Spiegelman, S., and Lindegren, G., *Proc. Natl. Acad. Sci. U. S.*, **30**, 346 (1944).

(26) Lwoff, A., *Ann. inst. Pasteur*, **61**, 580 (1938).

(27) Mirsky, A. E., in *Advances in Enzymology*, Vol. III. Interscience, New York, 1943, p. 1.

(28) Moewus, F., *Ergeb. Biol.*, **18**, 287 (1941).

(29) Moore, A. R., *Univ. Calif. Pub. Physiol.*, **4**, 9 (1910).

(30) Muller, H. J., *Am. Naturalist*, **56**, 32 (1922).

(31) Muller, H. J., *Cold Spring Harbor Symposia Quant. Biol.*, **9**, 290 (1941).

(32) Northrop, J. H., *Crystalline Enzymes.* Columbia Univ. Press, New York, 1939.

(33) Oparin, A. I., *The Origin of Life.* Trans. by S. Morgulis. Macmillan, New York, 1938.

(34) Painter, T. S., *Texas Repts. Biol. Med.*, **2**, 206 (1944).

(35) Penrose, L. S., *Lancet*, **2**, 192 (1935).

(36) Plunkett, C. R., in *Colloid Chemistry.* Vol. V, Reinhold, New York, 1944.

(37) Sawin, P. B., and Glick, D., *Proc. Natl. Acad. Sci. U. S.*, **29**, 55 (1943).

(38) Schopfer, W. H., *Plants and Vitamins.* Chronica Botanica, Waltham, 1943.

(39) Schultz, J., *Cold Spring Harbor Symposia Quant. Biol.*, **9**, 55 (1941).

(40) Schultz, J., in *Colloid Chemistry.* Vol. V, Reinhold, New York, 1944.

(41) Sonneborn, T. M., *Cold Spring Harbor Symposia Quant. Biol.*, **10**, 111 (1942).
(42) Sonneborn, T. M., *Ann. Missouri Botan. Garden*, **32**, 213 (1945).
(43) Spiegelman, S., *Ann. Missouri Botan. Garden*, **32**, 139 (1945).
(44) Srb, A. M., and Horowitz, N. H., *J. Biol. Chem.*, **154**, 129 (1944).
(45) Stadler, L. J., and Uber, F. M., *Genetics*, **27**, 84 (1942).
(46) Stanley, W. M., "Chemical structure and the mutation of viruses," in *Virus Diseases*. Cornell Univ. Press, Ithaca, 1943.
(47) Stranskov, H. H., *Physiol. Revs.*, **24**, 445 (1944).
(48) Tatum, E. L., *Ann. Rev. Biochem.*, **13**, 667 (1944).
(49) Tatum, E. L., and Beadle, G. W., *Ann. Missouri Botan. Garden*, **32**, 125 (1945).
(50) Tatum, E. L., and Bonner, D., *Proc. Natl. Acad. Sci. U. S.*, **30**, 30 (1944).
(51) Tatum, E. L., Bonner, D., and Beadle, G. W., *Arch. Biochem.*, **3**, 477 (1944).
(52) Troland, L. T., *Am. Naturalist*, **51**, 321 (1917).
(53) Williams, R. R., *Science*, **94**, 471 (1941).
(54) Winge, Ö., and Laustsen, O., *Compt. rend. trav. lab. Carlsberg, Sér. physiol.*, **22**, 337 (1939).
(55) Wright, S., *Physiol. Revs.*, **21**, 487 (1941).
(56) Wright, S., *Biol. Symposia*, **6**, 337 (1942).

2

VIRUSES

W. M. STANLEY, MEMBER OF THE ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH, PRINCETON; MEMBER OF THE NATIONAL ACADEMY OF SCIENCES

DURING the past ten years there has converged on a group of small, infectious, disease-producing entities, known as viruses, an array of scientific talent almost as diverse in nature as the Allied forces that were brought to bear on the Axis powers. The viruses are responsible for many diseases of man, animals, plants, and bacteria. There is no single criterion by means of which viruses can be differentiated from bacteria, yet the virus group has been segregated by means of certain general characteristics. Among the most important of these are small size, the ability to reproduce or multiply when within the living cells of a given host, the ability to change or mutate during multiplication, and the inability to reproduce or grow on artificial media. The sole means of recognizing the existence of a virus is provided by the manifestations of disease which result from the growth of the virus. Although pathologists have studied viruses for over fifty years and have added much to our knowledge, the new attack on viruses has been spearheaded by chemistry, chiefly biochemistry and physical chemistry, and their allied disciplines. The impact of these diverse disciplines on viruses has been accompanied by reverberations and repercussions, which, however, bode ill for viruses and good for scientific thought. Certain dominant facts stand out in bold relief and these are understood and accepted by chemists and pathologists alike, but there is an undercurrent of indecision where neither feels

quite sure of himself. This indecision becomes apparent when one considers the mode of reproduction and mutation of viruses, whether viruses are molecules or organisms, or whether some are molecules with others being organisms, or whether viruses represent a new type of structure, hitherto unrecognized and undefined, and whether one should speak of a solution or of a suspension of a given virus. For the moment, because of the lack of precise experimental data, discussion of these questions must remain more or less philosophical in nature. But since some of these questions pertain to the very nature of life itself, and others to fundamental physicochemical problems, it is obvious that they are of great importance. During recent years the chemist, as well as the pathologist, has become aware of the limitations of his tools. It has become obvious that, if the tremendous problems posed by the viruses are to be solved, it will be necessary to forge new tools, both material and of the mind, and to carry forth the attack on a united front with a new perspective and with renewed courage and vigor. It is the purpose of this short essay to chart briefly the roads that the chemist and pathologist have already constructed into the field of viruses and to attempt to outline, in general terms, the manner in which these can be continued until they join and provide, through a common effort, a broad highway of fact and information leading from the lowly electron to the lofty heights of man.

Viruses range in size from about 10 mμ to about 300 mμ. Certain small viruses, such as alfalfa mosaic virus, are smaller than certain accepted protein molecules, such as the *Busycon hemocyanin* molecules. On the other hand, certain large viruses, such as vaccine virus, are larger than certain accepted organisms, such as the minimal reproductive units of the microorganisms of the pleuropneumonia group. With respect to size, therefore, the viruses overlap with molecules at one extreme and with organisms at the other extreme. Since the discovery of viruses by Iwanowski, a plant pathologist, in 1892, and by Beijerinck, a chemist and botanist, in 1898, many investigations on the sizes of different viruses have been carried out. Ultraviolet-light microscopes and ultrafiltration were used in most of these studies. Following the isolation of essentially pure virus preparations, methods involving ultracentrifugation, diffusion, x-ray diffraction, viscosity, osmotic pressure, and stream double refraction measurements were used with great success. Recently the chemist and pathologist have

joined in the extensive use of the newly developed electron microscope. This instrument covers the entire range of sizes occupied by viruses and has proved, and will doubtless continue to prove, of the greatest value in the estimation of the sizes of viruses. In those cases in which more than one method has been employed, good agreement has usually been secured. The occasional discrepancies have been found to be due to errors or to a failure to appreciate the limitations of a given method, and generally have been resolved. At present the sizes of several viruses are well established and the values are accepted by both chemists and pathologists.

Tobacco mosaic, the first virus to be discovered, was also the first virus to be prepared in essentially pure form and subjected to extensive chemical study. No difference was noted in virus samples prepared from a wide variety of hosts or from the same kind of host at different times of the year. The virus particles were found to consist of about 6% nucleic acid of the ribose type and about 94% protein. The exact nature of the linkage between protein and nucleic acid is unknown, but it appears to be considerably stronger than that which exists in the sperm nucleoproteins. The protein component contains definite and reproducible amounts of over thirteen amino acids. It is of interest that, in contrast to the sperm nucleoproteins, there does not appear to be an excess of basic amino acids. The single virus particles are about 280 mμ in length and 15 mμ in diameter. Tobacco mosaic virus activity has never been demonstrated to be associated with smaller particles. However, there is good evidence that a single virus particle is built up from similar subunits fitted together in a hexagonal lattice to yield the final structure which possesses virus activity. The nucleic acid of this final structure appears to exist in the form of eight threadlike units laid down along the length of the particle. Because of the repeat pattern within a single virus particle, it can be regarded as a submicroscopic crystal. In addition, these single virus particles can aggregate with a two-dimensional regularity to form large needlelike crystals that are readily visible with a low-power hand lens. Of especial interest and significance is the fact that the virus particles appear to be utterly devoid of water and of any enzymic activity other than virus activity. The complete lack of water and the crystal-like inner structure of the individual virus particles would appear to preclude the existence of metabolism of the type usually associated with

organisms. Yet, when introduced into the cells of susceptible hosts, these particles can direct or enter into the metabolic chain of events of the cell.

The rod-shaped anhydrous tobacco mosaic virus particle is representative of a small group of viruses, but is certainly not representative of all viruses, for most viruses that have been studied adequately appear to be essentially spherical in shape and hydrated. Among the viruses that have been obtained in essentially pure form and studied in some detail are alfalfa mosaic virus with a diameter of about 17 mμ, tobacco ring spot virus with a diameter of about 19 mμ, tomato bushy stunt virus with a diameter of about 26 mμ, equine encephalomyelitis and rabbit papilloma viruses with diameters near 40 mμ, influenza virus with a diameter of about 100 mμ and vaccine virus with a diameter of about 225 mμ. Of these, tomato bushy stunt virus is the only one that has been obtained in crystalline form. This virus, which contains about 17% nucleic acid, and about 83% protein, crystallizes in the form of large, beautiful, rhombic dodecahedra. The particles of bushy stunt virus appear to be strictly homogeneous with respect to size, shape, and density; hence the case for regarding these particles as molecules is as good as for any protein.

The papilloma virus appears to be a nucleoprotein containing about 8% nucleic acid and little or no lipid. The particles of the equine encephalomyelitis virus appear to be a liponucleoprotein complex containing about 48% lipid, about 5% nucleic acid, and protein plus a small amount of excess carbohydrate. Data on influenza virus indicate that the 100 mμ particle has a water content of about 60% by weight, with the solid portion being composed of about 65% protein, about 25% lipid, about 7% carbohydrate, and a very small amount of nucleic acid. Vaccine virus is the largest and most complex virus that has been subjected to chemical investigation. The preparations were found to contain protein, lipid, carbohydrate, and thymus nucleic acid in concentrations not materially different from those found in bacterial cells. The vaccine virus preparations were also found to contain phosphatase, catalase, lipase, biotin, riboflavin, flavin-adenine-dinucleotide, and apparently significant and reproducible amounts of copper. It is exceedingly difficult to prove that all of these represent integral components of vaccine virus, but it must be regarded as significant that this large and complicated virus appears to retain

certain enzymic activities quite tenaciously, whereas tobacco mosaic and bushy stunt viruses appear to possess no enzymic activities other than that of virus activity. It is also of interest that, in marked contrast to bacteria and other cells, tobacco mosaic and influenza viruses contain negligible amounts of the B vitamins. Electron micrographs of vaccine virus, as well as of certain bacteriophages, have revealed the presence within individual particles of an internal structure consisting of a pattern of granules.

As a whole, the data now available on viruses indicate that, as one goes from the small to the large viruses, there is, with increase in mass, an increase in complexity of composition, structure, and function. The viruses appear to provide, in truth, a bridge between proteins and organisms. Indeed, if one wishes to regard the transformation agent of the pneumococcus as a virus, the bridge could be extended to nucleic acid. If one were starting out anew to construct a link between the molecules of the chemist and the organisms of the pathologist or bacteriologist, it is difficult to visualize how it would be possible to improve upon what Nature has already provided. The structural complexity encountered in tomato bushy stunt virus nucleoprotein is but little more complex than that of hemoglobin and no more complex than that of the hemocyanins. Physically and chemically these behave as molecules; and if it were not for the virus activity of the bushy stunt nucleoprotein it would not be given a second thought. Between bushy stunt virus and vaccine virus, Nature has provided a continuous series of structures of gradually increasing mass and complexity, all linked by a common biological property, virus activity. Vaccine virus is as large as some organisms that can be grown on artificial media; and if vaccine virus could be grown on artificial media it would be accepted generally as an organism. Even larger structures exist which cannot be grown on artificial media and which are just as fastidious as vaccine virus with respect to growth requirements; yet these are accepted as organisms. Nature has provided us with an accomplished fact, and it is time for all to brush away the barriers of the mind and to recognize the possibilities that are provided by the viruses. Although viruses are disease-producing agents and have caused untold suffering, the complete acceptance of Nature's dubious gift may not be without recompense. For its acceptance and exploitation may provide the key to broad and wonderful vistas.

Pathologists and bacteriologists have labored long and arduously with viruses. They have found that a given virus will reproduce only within the cells of certain specific living hosts, and that, although some viruses will reproduce within the cells of several different hosts, other viruses will multiply only within the cells of one given host or sometimes only within certain specialized cells of that host. They have shown that the primary pathological changes produced in cells by viruses are either proliferative or degenerative in character. In some virus diseases, such as yellow fever, poliomyelitis, and tobacco necrosis, degenerative changes predominate; but in many, as in smallpox and fowl pox, both proliferation and necrosis occur. Still other virus diseases, such as Rous chicken sarcoma, Shope rabbit papilloma, and tobacco enation mosaic, are characterized by a rapid and unorganized cellular proliferation.

Long before the discovery of viruses, a means was recognized of protecting man against the virus disease, smallpox. This was achieved by vaccination with active virus, presumably altered by passage in an unnatural host. However, it has only been in recent years that there has come a full realization of the great benefits, both with respect to methods of protection against virus diseases and to the study of the viruses themselves, that can be achieved through the use of new virus hosts. Yellow fever has been eliminated as a major health problem, and much has been learned of the virus because the virus was taken from man and grown in monkeys, in mice, and in chick embryos. The possibility of a recurrence of the 1918 influenza epidemic, which killed more people than have died from combat activities in World Wars I plus II (to date of writing), has been reduced and perhaps eliminated because of the production of a vaccine which was made possible by the growth of this virus in the chick embryo. Truly remarkable progress has already been made in the study of viruses and in the prevention of virus diseases, and it is to be expected that this progress will continue. Yet withal, the one fundamental and all-important problem posed by the viruses—that of the mode of virus reproduction—remains unsolved. For many years little hope for a solution of this problem was held, for viruses were generally regarded as living organisms and the nature of life was considered to be a hallowed, insoluble secret. However, chemists recognized in the virus activity of certain crystallizable nucleoproteins a type of biological

activity somewhat akin to that possessed by certain protein hormones and enzymes. To them, virus activity appeared no less wonderful and possibly no more complicated than the changes that can be induced within cells by enzymes and hormones. The mental barrier of the living state is being eliminated gradually and chemists are recognizing and accepting Nature's gift of the viruses. The true issue is only beclouded by the insistence in some quarters for a decision as to whether a given structure is living or inanimate. The fundamental meaninglessness of such terms has been commented on before and is becoming ever more apparent. In the meantime, the requirements for the solution of the riddle of virus reproduction—perhaps the most important problem in all of biochemistry—are becoming clearer.

Many studies on the nature and mode of virus reproduction are in progress, and it seems certain that such studies will increase in the future, not only in volume, but also in scope and in diversity. It is possible that the solution of this very important and fundamental problem will not be realized until after new weapons of the hand and mind have been brought into play. So far, the most spectacular advances have been made along three lines, each of which merits considerable further attention: (a) studies on the very favorable bacterial cell–bacteriophage system; (b) studies on the changing of the chemical structure of a virus by means of known chemical reactions; and (c) studies on the nature of the differences in chemical structure that are responsible for the existence of virus strains. The bacterial cell–bacteriophage system provides an extraordinary opportunity to follow the interaction of a virus with its host cell. The host cells can be grown *in vitro* on artificial media in large quantities and under constant and reproducible conditions. The metabolism of the host can be controlled to a certain extent and analyses can be made on the system throughout the reaction. The bacterial and phage or virus materials can be differentiated up to the entrance of the virus into the cells and following the lysis of the cells. Studies on this system have permitted the conclusion that multiple infection of a bacterial cell with several virus particles of the same type has the same effect, both qualitatively and quantitatively, as infection with a single virus particle. It was also found that infection of a bacterial cell with two kinds of virus particles resulted in the reproduction of only one kind and the suppression of growth of the other kind. However, more than one virus

can multiply in a single cell if the viruses are sufficiently distinct in their requirements. These and similar results have permitted many stimulating inferences regarding the mode of virus reproduction and have suggested innumerable approaches for future experimentation.

Work on the changing of the chemical structure of a virus by means of known chemical reactions has been both encouraging and discouraging. It has, in fact, proved possible actually to change the chemical structure of a virus; but so far no change has been found to be perpetuated in the virus particles produced as a result of infection of a host with the altered structures. Thus, the abolishment of the sulfhydryl groups in tobacco mosaic virus or the introduction into the structure of this virus of several thousand acetyl, phenylureido, carbobenzoxy, benzene sulfonyl, or malonyl groupings yields diverse altered virus structures. Although these are infectious, the disease which they produce is the ordinary tobacco mosaic disease, and is accompanied by the production of particles, not of the respective altered structures, but of ordinary tobacco mosaic virus. However, encouraging results were obtained in a study of the specific virus activity of these chemical derivatives on different hosts. It was found that a property of the virus, which perhaps can best be described as virulence, can remain constant for one host but be modified with respect to a different host, upon formation of a given chemical derivative of the virus. This result lends encouragement to the belief that, eventually, heritable structural changes in a virus will be achieved in the chemical laboratory by means of known chemical reactions. Contemplation of the implications that would accompany the actual accomplishment of this feat tends to stagger the imagination.

Spectacular progress has attended studies on the nature of the differences in chemical structure that are responsible for the existence of virus strains. It was indeed fortunate that Nature provided, and the pathologist recognized and separated, two or more strains of each of several different viruses. Strains of a virus appear to arise during the reproduction of a virus by a process which can be regarded as similar to that of gene mutation. It can be presumed that the strains of a virus have arisen from some parent strain, one by one, during the course of the years. Each strain thus probably bears a definite relationship to the parent strain and to each of the similar strains. Each strain causes a more or less different disease; because of this fact it

has been possible for the pathologist to recognize, separate, and grow individual virus strains. Some of these have been found of great use as vaccines for the prevention of certain virus diseases.

In view of the large size and complexity of structure of viruses, it may appear that chemists were somewhat optimistic in expecting to be able to detect differences in the chemical structure of different strains of a virus. Two similar large mountains may appear identical when viewed from a distance, but close inspection will reveal differences. So, too, with the viruses. The over-all structures of different strains of tobacco mosaic virus were found to be very similar; yet in several instances it has been possible to demonstrate definite differences in chemical structure. These consist of differences in the amount of one or more amino acids, in the presence of an entirely new amino acid, or in the complete elimination from the virus structure of a given amino acid. These changes represent deep seated and fundamental alterations in the virus structure, and it seems unlikely that they could have resulted from alterations of fully formed virus particles. It appears more likely that these changes occurred as a result of a diversion of the synthetic process by means of which a virus reproduces. Since new strains tend to appear or to become dominant when a virus is grown in an unnatural host, it is possible that the altered environment of this host provides a somewhat different supply of amino acids and enzyme systems, and in the effort to adhere to some basic pattern it becomes necessary to build into the virus structure amino acids that would not be used normally. The drive to follow a basic pattern and the aberrations that result bear a certain kinship to the forces of heredity. As a matter of fact, there is a striking similarity between the properties of viruses and those that have been ascribed to genes. Both may be regarded as large nucleoprotein structures that have the ability to perpetuate themselves within, and only within, certain specific living cells. Both can undergo sudden changes, apparently either spontaneously or as a result of external factors and those changes are then reproduced in subsequent generations. Within limits, the concentration of both in cells can be changed by proper treatment of the host. Some viruses appear to be concentrated in the cytoplasm of cells and others in the nuclei.

The similarity between viruses and genes may not be without significance, for the abode of genes is the cell and no virus has been

proved to arise *de novo*—they are always first found in cells. Viewed in this light, the differences in chemical structure that have been demonstrated to exist between different virus strains, and especially the changing of the chemical structure of a virus in a definite manner by means of known chemical reactions, take on a new and perhaps startling significance. There are many virus diseases in which a symbiotic relationship is set up between virus and host. Although the virus enters into and alters the metabolic activity of the cell, the cell survives and continues to divide. The virus is carried continuously within the cells, and in some cases could almost be regarded as a normal component of the cells because of the lack of obvious damage. In fact, one virus has come to be known as the "healthy potato virus" because it is present in almost all potatoes grown in this country and yet the potato plants appear healthy. It is as if one had introduced, from without, a nucleoprotein which was accepted by the cell as a part of its own germ plasm. The fact that different strains of a virus, which can cause different manifestations of disease, are characterized by different nucleoproteins, and especially the fact that the structure of these nucleoproteins can be changed in the test tube by means of known chemical reactions, could be interpreted to mean that eventually the germ plasm of cells may prove to be susceptible to similar chemical manipulation. The viruses have assuredly provided a link between molecules and organisms, and there now exists a pathway leading from simple structures, such as the electron, to massive, highly complex structures, such as man. This pathway is broad and well established in some places and narrow and difficult to traverse in others. But as the latter are broadened and placed on a firm foundation, through the common effort of chemists and pathologists, it is possible that information will be acquired which could affect the future destiny of every living being in the world.

Selected References

Anson, M. L., and Stanley, W. M., *J. Gen. Physiol.*, **24**, 679 (1941).

Bawden, F. C., *Plant Viruses and Virus Diseases*. Chronica Botanica, Waltham, 1943.

Beard, J. W., Bryan, W. R., and Wyckoff, R. W. G., *J. Infectious Diseases*, **65**, 43 (1939).

Bernal, J. D., and Fankuchen, I., *J. Gen. Physiol.*, **25**, 111, 147 (1941).

Cohen, S. S., and Stanley, W. M., *J. Biol. Chem.*, **144,** 589 (1942).
Doerr, R., and Hallauer, C., *Handbuch der Virusforschung.* Springer, Berlin, 1938–39.
Green, R. H., Anderson, T. F., and Smadel, J. E., *J. Exptl. Med.*, **75,** 651 (1942).
Hoagland, C. L., *Ann. Rev. Biochem.*, **12,** 615 (1943).
Hoagland, C. L., Ward, S. M., Smadel, J. E., and Rivers, T. M., *J. Exptl. Med.*, **74,** 69, 133 (1941).
Hoagland, C. L., Ward, S. M., Smadel, J. E., and Rivers, T. M., *J. Exptl. Med.*, **76,** 163 (1942).
Knight, C. A., *J. Biol. Chem.*, **145,** 11 (1942).
Knight, C. A., and Stanley, W. M., *J. Biol. Chem.*, **141,** 29 (1941).
Lauffer, M. A., *J. Biol. Chem.*, **143,** 99 (1942).
Lauffer, M. A., *J. Am. Chem. Soc.*, **66,** 1188 (1944).
Lauffer, M. A., and Ross, A. F., *J. Am. Chem. Soc.*, **62,** 3296 (1940).
Lauffer, M. A., and Stanley, W. M., *J. Exptl. Med.*, **80,** 531 (1944).
Luria, S. E., and Anderson, T. F., *Proc. Natl. Acad. Sci. U. S.*, **28,** 127 (1942).
Luria, S. E., and Delbrück, M., *Arch. Biochem.*, **1,** 207 (1942).
Miller, G. L., and Stanley, W. M., *J. Biol. Chem.*, **141,** 905 (1941).
Rivers, T. M., *Viruses and Virus Diseases.* Stanford Univ. Press, Stanford University, 1939.
Rivers, T. M., *et al.*, *Virus Diseases.* Cornell Univ. Press, Ithaca, 1943.
Stanley, W. M., *Physiol. Revs.*, **19,** 524 (1939).
Stanley, W. M., *Ann. Rev. Biochem.*, **9,** 545 (1940).
Stanley, W. M., *J. Biol. Chem.*, **135,** 437 (1940).
Stanley, W. M., *Sci. Monthly*, **53,** 197 (1941).
Stanley, W. M., *J. Exptl. Med.*, **79,** 267 (1944).
Stanley, W. M., and Anderson, T. F., *J. Biol. Chem.*, **139,** 325 (1941).
Stanley, W. M., and Knight, C. A., *Cold Spring Harbor Symposia Quant. Biol.*, **9,** 255 (1941).
Taylor, A. R., Sharp, D. G., Beard, D., and Beard, J. W., *J. Infectious Diseases*, **72,** 31 (1943).
Williams, R. J., Schlenk, F., and Eppright, M. A., *J. Am. Chem. Soc.*, **66,** 896 (1944).

3

PHOTOSYNTHESIS AND THE PRODUCTION OF ORGANIC MATTER ON EARTH

H. GAFFRON, RESEARCH ASSOCIATE, PROFESSOR, DEPARTMENTS OF BIOCHEMISTRY AND CHEMISTRY (FELS FUND), UNIVERSITY OF CHICAGO

Practical Importance of the Assimilation of Carbon Dioxide

We know, for example, that if we abuse the soil, it will lose its fertility, that if we massacre the forests, our children will lack timber and see their uplands eroded, their valleys swept by floods. Nevertheless, we continue to abuse the soil and massacre the forests . . . in the simple affairs of nature, where we know quite well what is likely to happen we immolate the future to the present. "Those whom the gods would destroy they first make mad."

ALDOUS HUXLEY, *Time Must Have a Stop*, p. 298.

EVEN the less informed layman is aware of the reaction which makes possible the abundance of living things on earth—the conversion in daylight by plants of carbon dioxide and water, the waste products of plant and animal life, back into food. This reconversion is called the assimilation of carbon, or photosynthesis. Hardly ever does the layman spend another thought on this fundamental and quite spectacular achievement of living cells. If, however, he suddenly realizes that, without this reaction, life as we know it would perish quickly and completely (except for a few species of bacteria), he is likely to place false hopes on its artificial reproduction by the toiling scientist. He probably believes that mankind's in-

adequate supplies of food and dwindling stores of energy will turn into a surplus the moment photosynthesis has been duplicated in the test tube. Popular articles and radio talks—spreading the lamentable misconception that scientific research represents nothing but the first step in technology—foster this notion by stressing the practical importance of artificial food production from sunlight and carbon dioxide. The error is easily demonstrable. The solar energy flux per acre and year is a constant. Unless our devices are to be much more efficient than the green plants and unless we are willing to spoil with ugly machinery an acreage equivalent to that now covered by beautiful forests and pastures, artificial photosynthesis is not only utopian but impractical. Even assuming we were to discover some sort of artificial photochemical reduction of carbon dioxide into a digestible carbohydrate with an over-all efficiency surpassing that of the plants which use on the average two per cent of the incident radiation, it would not help us much, since we need, not one product of plant metabolism, but a thousand (16). Let us mention only the proteins among foodstuffs, and wood and rubber among industrial raw materials.

True, chemists have learned to make ammonia and nitrates from atmospheric nitrogen. In that respect man is no longer dependent on other natural sources. But up to now only the plant converts these simple nitrogen compounds into proteins,* and so we are forced in a second way, equally fundamental, to rely on the growth of plants for our continued existence. In order to stretch our limited supply of organic substances already formed, such as coal and petroleum, the production of most compounds which plants can synthesize and which we need in large quantities should be reserved for natural photosynthesis, regardless of whether the chemist can duplicate the synthesis *in vitro*. Oil, for instance, is too versatile a material to be converted into rubber as long as plants are capable of continuously producing essential raw materials like alcohol, or better still, rubber of excellent quality.†

Denying that artificial photosynthesis will be the solution to a serious and fascinating problem is not equivalent to saying that sun-

* Nonphotosynthetic plants like yeasts are a good source of protein, provided they are fed with products of photosynthesis.

† Guayule plants grown in California produce rubber far superior in automobile tires to the synthetic compounds now available.

light will not be used at all in the technology of the future. In selected places, such as roof tops in sunny countries, bare rock, and deserts inaccessible to irrigation, sunlight may soon be utilized to produce heat or electric power or even to drive some photochemical process like that of hydrolysis. These possibilities have exactly as much and as little relation to the problem under discussion as exists between any other common source of energy and a filled granary. There is every indication that products of plant photosynthesis will be needed ever more urgently, not only as food but also as fuel and building material. Recent technological inventions indicate that wood, strengthened with wood-derived chemicals, will be in even greater demand than now.

One and one-half billion of the two billion humans on earth are illnourished or permanently hungry. "There has never been enough food for the health of all people" (18). Since this globe offers only a certain area of habitable and tillable ground, it is obvious that populations will have to be adjusted to a certain optimum density determined by the general standard of living that man is capable of attaining or willing to endure.

We have ample testimony of the improvident way in which the ancients exhausted their supply of wood. The mountains of Persia, Syria, Greece, Dalmatia, Italy, and Tripolitania are now to a large extent barren and infertile. The goats of the Arabs are said to have done away with the last traces of vegetation in North Africa, thus allowing the desert sands to advance to the shores of the Mediterranean. The changes in climate brought about by the denuding and erosion of the sites of the most important early civilizations make it difficult or impossible to retrieve the lost fertility of the land.

Are we wiser today? The rate of consumption of coal and petroleum in the world today (1.75×10^9 tons carbon per year) is roughly one-tenth of the rate of the total carbon assimilation achieved by the land plants on earth (1.6×10^{10} tons carbon per year) (14). The coming industrialization of Asia will soon diminish the gap between the demand for, and the supply of, organic material. The established reserves of oil in the United States, according to official figures, are about twenty billion barrels. They are being used up at a rate of one and one-half billion barrels per year. The industry insists upon a more optimistic outlook, based on the (ever-declining) rate of new discoveries. True or not, considering also the certainly

limited supply of coal, it is quite clear that we are spending mainly as fuel irreplaceable organic material.

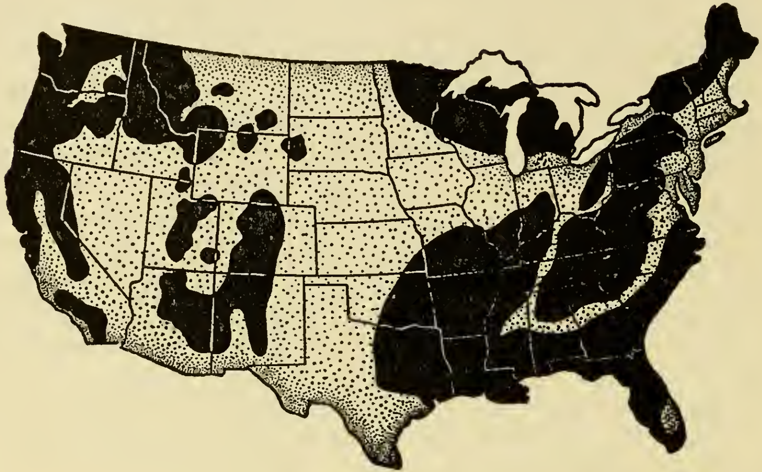


Fig. 1.—Virgin forests in 1850.

Can these losses of stored products of photosynthesis be offset by the assimilation of carbon going on today? Within the cultivated areas of industrialized countries this is not the case, and foresters agree that virgin forests are more or less stationary. New growth balances natural decay. Only well-planned agriculture and expert forestry may perhaps furnish all that we need in the future provided the increasing demand for fuel and energy is eventually met by the general development of atomic power. At present, the products of agriculture are consumed within a few years after harvesting. There is no increase of our reserve of organic carbon due to this source. On the contrary, the improvident exploitation of the soil in many places leads to diminishing returns. In the United States, fifty million acres now under cultivation are so badly eroded as to invite abandonment. These sad circumstances have received widespread attention, and effective measures are being taken to check further losses and to regain the lost fertility of the soil (13). Not so well known are the conditions regarding the forests in this country (1). The following figures speak for themselves:

	Billion board feet
Original stand of forests...........	5,200 on 850 million acres
Present (1927) stand of forests......	2,500 on 550 million acres
Rate of annual consumption........	35

We are cutting down our forests three times as fast as they grow. Certain species of trees are in danger of extinction. War conditions have aggravated the situation. Figures referring to the years just before the war for the states of Washington, Oregon, Montana, and Idaho, which now supply half of the timber cut in the United States, are as follows (3):

	Billion board feet
Annual cut and losses in the Northwest...........	12.5
Current annual growth.........................	4.2
Potential annual growth after universal adoption of best forest practices.........................	23

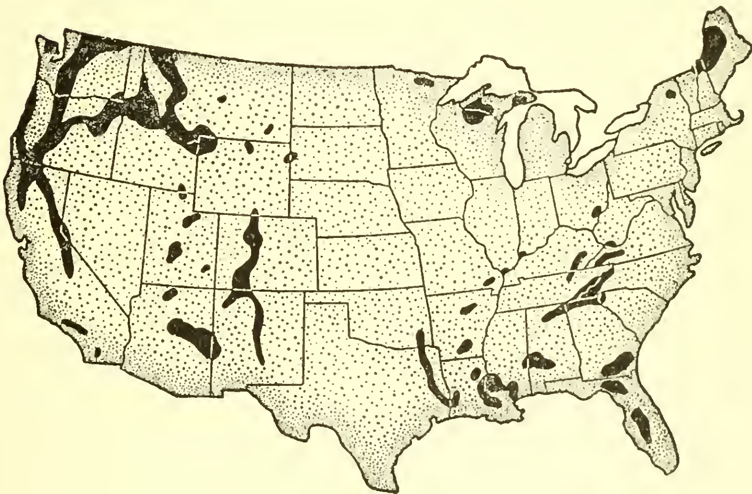


Fig. 2.—Virgin forests in 1945.

The last figure shows that we should not abandon all hope. There is a well-tested way to prevent the exhaustion of our wood reserves. But it is not enough for the government to administer the national forests providently. Industrialists and private owners hold

not only the larger area of commercial forest land but also the more valuable timber from the standpoint of accessibility and of quality. "The main objective of this group is to market their lumber as speedily as possible" (3). Management according to the principles of far-sighted modern forestry is encountered only on an estimated two per cent of the total area. The practices of commercial competition are such that in general only the State can afford to look ahead two or three generations.

Meanwhile, the waste continues. According to N. C. Brown (1) less than half of the total amount of wood cut in this country is ultimately utilized. Hence it is imperative that the efforts of the Departments of Agriculture and of the Interior to remedy the situation should have the conscious support of every citizen. Unless we plan for a permanent forestry everywhere, not only our supplies of oil but also those of wood will be gone in about seventy years and the country dependent upon foreign sources. No wonder that scientists begin to turn toward the oceans as a source of products of photosynthesis; it is estimated that the amount of carbon assimilated by marine algae surpasses by four or five times the yield of the land plants (14). But when can we expect to replace wood by plankton? The journals of chemical industry often display an advertisement showing a magnificent mountain forest with a caption from which we quote: "Ever see a forest through a chemist's glasses? Do you see . . . plastics . . . plywood . . . laminated beams . . . silk . . . smokeless powder . . . rolls of newsprint . . . movie film . . . All these and many more items of beauty, strength, and utility the chemist makes from wood, holding out brighter hopes for a better future." If man continues to consume the products of photosynthesis in the way he does at the present time this kind of better future may not last very long.*

The understanding of photosynthesis becomes essential if we are to solve part of these rapidly approaching difficulties, not in order to reproduce the process technically on a big scale but rather in order to

* Two publications have recently appeared on the subject of our dwindling supply of wood: "Forestry and the public welfare," in *Proc. Am. Phil. Soc.*, **89**, 399 (1945); and "What's happening to the timber," by R. A. H. Thompson, in *Harper's Magazine*, issue of August, 1945, page 125. The instructive maps (Figs. 1 and 2) from the latter article are reproduced through the courtesy of *Harper's Magazine*.

increase its natural efficiency. As everybody knows, different plants grow in different locations. The efficiency of the chloroplasts or of the photochemical mechanism proper may be equal in all plants, yet some plants may not have attained the optimum in utilizing their own photosynthetic products. Here the knowledge of internal factors influencing photosynthesis will show which plants to breed. Investigations on the photosynthetic behavior responsible for the over-all efficiency of plants in producing the countless substances we need will increase in practical importance. Much has been done and will be done empirically by the gardener, agriculturist, and forester, but the shortest way to success lies in systematic investigations of the correlation between growth or fruition of plants and the rate of photosynthesis. Such work is in progress, for instance, at the California Institute of Technology (22). The intimate connection with studies concerning such factors as soil, climate, and inheritance are obvious.

In contrast, it appears doubtful at this moment whether the analysis of the nature of the photosynthetic process can produce immediate practical results surpassing those obtainable by the kind of studies mentioned above. True, pure research remains the only source of sudden technical advances. At the present time, however, the question as to the nature of photosynthesis is a problem of science and not of so-called "applied science," that is, technology. The reward in joining the few who insist upon spending their time in tackling this problem will be only the pleasure of knowing a little more and seeing a little farther than those who worked on the same problem twenty years ago.

Partial Reactions in Photosynthesis

> "*The enormous amount of research upon the process of photosynthesis during the past half century has thrown little or no additional light upon the subject. The problem is evidently too complex for any specialists in any one field of science to solve. . . .*"
>
> E. C. MILLER, *Plant Physiology*, 2nd ed., 1938

During the past decade a new laboratory technique—that of enzyme chemistry—has been developed and is at present taught to every student in cell physiology. The isolation of an enzyme is an advance that requires no mental effort to appreciate. Equally obvious are the advantages of using traceable isotopes for elucidating the course

of obscure metabolic reactions. Among younger biochemists the attitude toward a field like photosynthesis is one of waiting until disagreeable obstacles have been cleared away by hand, so to speak, so that they may roll in with modern methods. During the past twenty years the moment when this will happen has quietly drawn nearer. How near we shall try to demonstrate in the rest of this essay.

It is mostly forgotten that, after Buchner had demonstrated fermentation outside the living cell, some thirty years passed before enzymes were isolated and the mechanism of organic catalysis could be said to be clearly understood. Progress in the field of cellular breakdown reactions, slow in the beginning of this century, nevertheless encouraged a handful of students to devote their lives to the investigation of respiration and fermentation. In contrast, the return for the toil directed toward solving the problem of photosynthesis was so small that even the leading biochemists, after establishing a few important new facts such as the existence of a photochemical reaction distinguishable from an enzymic one, saw no point in pursuing the analysis any further. In 1918, Willstätter wrote that it was evidently too early to try to elucidate the mechanism of carbon dioxide assimilation in living cells. We now know that he was right. At that time photosynthesis appeared as an absolutely unique process showing no connection or analogy with other metabolic processes in the cell. It was looked upon as a direct photochemical decomposition of carbon dioxide and mysteriously connected with the phenomenon of life on earth.

The absolute ignorance of the kind of reactions photosynthesis might involve and, consequently, the lack of a theoretical framework accurate enough to direct a reasonable approach barred further progress. And as to purely empirical experimental attempts, they all ended with the destruction of the intact living cell. Photosynthesis stops the moment the cell is hurt. While this is still true, of course, the great difference is that today we know more or less why. Hence there is hope we shall be able to overcome this difficulty.

Despite Miller's verdict quoted above, there has been very decisive progress between 1918 and 1938, not merely by the addition of several important observations to a hundred earlier ones but, mainly, by the change in our conception of what constitutes the problem of photosynthesis. This change was due, on the one hand, to the understanding of the nature of other metabolic processes such as respiration

and fermentation and, on the other, to a clearer knowledge of the essence of photochemical reactions in complex molecules. Twenty-five years ago there was an overabundance of uncorrelated general observations and only a few very simple pictures of the mechanism of photosynthesis. The latter were hardly supported by any straightforward experiments or permissible analogies. I need mention only the concepts of a chlorophyll carbonic acid complex, of the release of oxygen from this complex, and of the formation and polymerization of formaldehyde, which are still faithfully reproduced in most textbooks of botany. About 1930, van Niel (19) pointed out that photosynthesis should be considered as a coupled oxidoreduction comparable to other reactions of this kind. This approach to the problem proved very fruitful and was soon generally adopted. Today, in 1945, we can divide the process of photosynthesis into several partial reactions, each with its particular problems, none of which seems to present insurmountable conceptual difficulties. By drawing upon analogies with other metabolic processes and from the results of direct experiments it is possible to build up a theoretical picture which, though incomplete in many places, satisfies the essential requirement that the main observations can be correlated. The mystery of photosynthesis is mainly gone, and this in a rather fundamental sense. We are now quite certain that photosynthesis promoted by chlorophyll in visible light has nothing to do with the origin of life. Instead, it must be regarded as a rather late achievement of the living cell. It is a unique combination of a few reactions found only in the green plant, with important devices characteristic of any kind of metabolism in living cells.

With the acquisition of pigments, the early living cells became able to accelerate the reaction between hydrogen donors and acceptors by absorbing radiant energy. The biological currency, H and OH, the constituents of water, became available in larger quantities because of the interaction between the irradiated dye and water. Despite the photochemical reduction of carbon dioxide no true gain in free energy of the complete system was yet possible, since the resulting "hydroxylated" counterparts (*cf.* reference 20) could only re-form water with some valuable hydrogen donor. Any increase in the amount of organic matter depended on the presence of inorganic hydrogen donors such as free hydrogen, hydrogen sulfide, sulfur, ferrous iron,

etc., and hence could not proceed very rapidly. The picture changed radically, however, the moment some cells happened to combine the photochemically accelerated utilization of water as hydrogen donor with a reaction allowing for the elimination of oxidation products by liberating free oxygen.

As long as there was an abundant supply of carbon dioxide, of water, and of light energy, the accumulation of organic matter could continue unchecked. Oxygen appeared in quantities as free gas in the atmosphere. This was followed by the enormous multiplication of organisms capable of using the oxidation of organic compounds for synthetic reactions.* Carbon dioxide released by respiration was again utilized in the photosynthetic reaction. Some million years later, the cycle of carbon arrived at a steady state. The concentration of carbon dioxide in the air is now 0.03% and barely sufficient to support maximum photosynthesis in full sunlight. Most plants do better when carbon dioxide is added artificially.

This picture of photosynthesis as a process that developed gradually from less complicated reactions is supported by the following observations. In the plant, the catalytic systems bringing about the assimilation of carbon dioxide can be distinguished by the differentiating effects of metabolic poisons. The whole process is specific for carbon dioxide, just as other metabolic reactions are specific for their substrates. The various catalysts seem to consist of proteins combined with special reaction groups. Chlorophyll itself is bound to protein in the cell. The reduction of carbon dioxide is not an exclusive privilege of the green plants. Besides normal photosynthesis we distinguish at present the following types of carbon dioxide reduction:

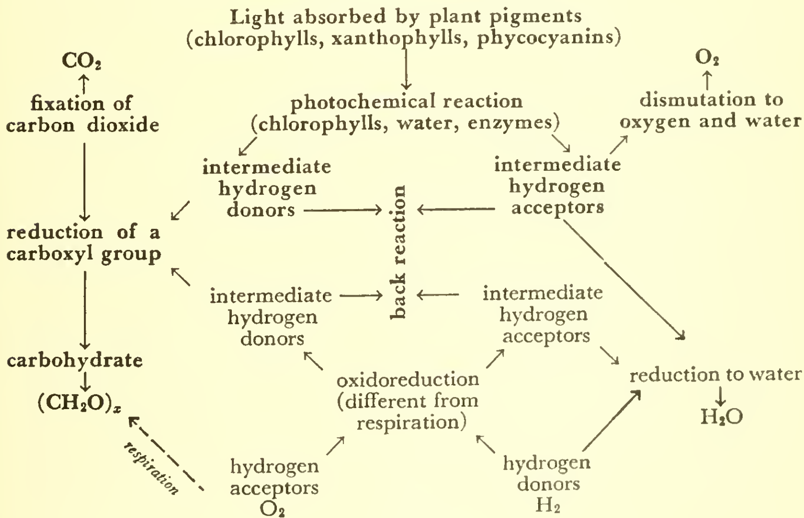
(*a*) In the dark (10,23) coupled with
(*1*) bacterial fermentations (methane bacteria, propionic acid bacteria),
(*2*) bacterial and plant oxidations (sulfur bacteria, "Knallgas" bacteria, unicellular algae),
(*3*) metabolic reactions in animal tissues.
(*b*) In the light (6,7,19), with the simultaneous consumption

* Van Niel has demonstrated a gradual adaptation to aerobic conditions with some strains of purple bacteria, which at first grew only anaerobically in the light (20).

of inorganic or organic hydrogen donors (purple bacteria, unicellular algae).

Unique to the green plants is the coupling of photochemical reduction with an evolution of free oxygen. Such an evolution of oxygen can be obtained by illuminating isolated chloroplasts in the presence of oxidizing substances like ferric salts or *p*-benzoquinone (5,9,21). The existence of several more or less autonomous enzyme systems working together can best be demonstrated in unicellular algae belonging to the *Scenedesmaceae*. In these algae it is possible to interchange at will all three known types of carbon dioxide reduction; a chemoreduction promoted in the dark by the burning of hydrogen with oxygen to water; a photoreduction in which two hydrogen molecules disappear together with a molecule of carbon dioxide; and a complete photosynthesis with the liberation of an equivalent amount of oxygen (7).

A rough summary of the way in which the problems of carbon dioxide reduction in an alga like *Scenedesmus* can be subdivided according to present knowledge is given in the scheme below. More complete and complicated schemes can be found in a recent review (7) and in Rabinowitch's new book (14).



In addition to what has been said in the preceding paragraphs about the probable evolution of the photosynthetic system and about the three main types of carbon dioxide reduction, the scheme tries to coordinate the following fairly well-established facts (13):

(*1*) Chlorophyll is necessary even in those cases in which the light is effectively absorbed by other plant pigments such as xanthophyll and phycocyanin.

(*2*) The oxygen liberated does not originate from carbon dioxide but from molecules of water entering somewhere as ultimate hydrogen donors into the reaction.

(*3*) Carbon dioxide is fixed initially by way of a reaction which is reversible and nonphotochemical, probably in the form of a carboxyl group.

(*4*) There are several intermediate steps between the photochemical reaction proper and the appearance of free oxygen. The proof for this is found in experiments with specific inhibitors and in the fact that the reversible switching from oxygen evolution to an equivalent consumption of hydrogen ("photoreduction") does not change the quantum yield.

(*5*) The chemoreduction in the dark has many traits in common with photoreduction. Yet we should expect differences between some of the intermediates produced and utilized in the dark and those made under the impact of a light quantum with an energy content of 40,000 calories.

Since neither the reactions of chlorophyll or of any other catalyst involved nor the nature of any one of the latter has been clearly established, the scheme will convince the reader that there is a host of problems to be solved. The interested investigator is likely to look upon *that* partial problem as the most exciting and, hence, important one to whose solution he believes he can contribute something significant. Since each of the partial reactions of which the whole of photosynthesis consists is indispensable, truly none can be more important than the other. However, the utilization of radiant energy and the evolution of free oxygen sets photosynthesis apart from other better analyzed metabolic reactions. It is here, therefore, where we must expect the greatest deviations from the common types of cellular catalyses and where we may find the unpredicted.

Though the picture given in the scheme could be seen to

emerge gradually, since about 1930, from the discussions accompanying the newly gained experimental data (van Niel, Stoll, Emerson, Wohl, Franck, and others), the interrelation between the different types of carbon dioxide assimilation has struck some investigators quite suddenly. In their enthusiasm over this fact they are likely to overlook the peculiar problems which photosynthesis offers in contrast to, and distinction from, the other metabolic processes. To them, now, all is very simple. The light absorbed by chlorophyll is used to mobilize hydrogen. Water is decomposed into H and OH. And once hydrogen is available the reduction of carbon dioxide proceeds as a dark reaction exactly as in the cases mentioned above. This has led to some strange expressions and statements, such as "photosynthesis in the dark" or "light *per se* is not essential for photosynthesis."

Lately, the analogy between thermal and photochemical reactions has been pushed still further, and attempts have been made to bring the so-called energy-rich phosphate bond into the picture. Before we analyze these attempts we must say a few words about intermediates.

In Search of Intermediates

In taking apart the mechanism of photosynthesis, we shall in all probability find two types of intermediates: on the one hand, the enzymes necessary for the transfer of hydrogen and the removal of oxygen; on the other, the final acceptors, the precursors of carbohydrates and of oxygen. The former we may compare to the pyridine nucleotides, flavoproteins, and cytochromes, the latter, to the degradation products of glucose, with no clear counterpart to what has been called the "photoperoxide" or "peroxide" or "moloxide" or "hydroxylated compounds" (van Niel) in photosynthesis, that is, the hypothetical substance which decomposes with the evolution of oxygen (we are pretty certain that it is not hydrogen peroxide and perhaps it is no peroxide at all).

Most of the intermediate catalytic steps in respiration and fermentation have been shown to be reversible—even the decarboxylation of pyruvate, if it proceeds with the formation of acetyl phosphate (11). The only requirement is the coupling with another reaction, furnishing the energy to reverse the particular step. Without such a coupling the breakdown processes continue unchecked until the specific

substrates are exhausted. If such an easy reversibility existed everywhere in the reaction chain leading from carbon dioxide to carbohydrate, an accumulation of photosynthetic products would be unlikely. In fact, it is quite essential for the efficiency of the synthetic process that the carbohydrate finally formed is not broken down again in the dark by the same catalytic systems which helped to build it up. One of the difficulties we shall encounter in trying to reconstruct the photosynthetic mechanism *in vitro* with enzymes isolated from the plant cell will probably be the reversibility of partial reactions, unless we succeed in removing and thereby stabilizing the intermediate just formed. At least, the way "down" is the normal course of events as demonstrated in respiration and fermentation. Now, save for attack by the respiratory or glycolytic systems, the carbohydrate synthesized in the light is stable (a very slow reversion is postulated for theoretical reasons not discussed here). The solution of the problem, therefore, may be hidden at either or both ends of the photosynthetic system. The primary unstable carbohydrate may polymerize with the release of some free energy. One might even think of a reduction up to the alcohol (or hydrocarbon) level and an oxidation by another path back to an aldehyde (or alcohol). At the other end of the line, the oxygen-liberating system should have properties making it rather ineffective as an oxidase. At any rate, all attempts to cause an accumulation of intermediates by mistreating the irradiated cell with narcotics or specific poisons have failed. Hydrogen donors capable of reducing carbon dioxide afterwards in the dark do not survive a period of illumination. There is no easily detectable formation of partly reduced substances. Nevertheless, there must be at least three intermediate steps in the process of reduction, since four hydrogen atoms have to be transferred from water to carbon dioxide. This particular problem will probably be solved with the aid of carbon isotopes in continuation of the work of Ruben, Kamen, and Hassid (*cf.* 14).

Several independent investigations have shown that nine or ten light quanta are necessary for this process. The maximum number of initial steps, hence, could be ten. For reasons of stoichiometry it is sensible to assume that there are only eight photochemical reactions, a pair for each release and recovery of a hydrogen atom by the molecule transforming electronic into chemical energy (4). This molecule may be chlorophyll itself. The difference between the theoretical

value of eight and the observed one of ten we attribute to losses due to inactive chlorophyll. The simplest explanation, then, for the fact that photosynthesis proceeds either all the way from carbon dioxide to carbohydrate with a permanent gain of 120 kcal. per mole, or not at all, is that the photochemically produced hydrogen donors and acceptors disappear by back reactions whenever the process becomes artificially inhibited. Despite the most drastic effects of certain poisons on the rate or the type of the reaction, there is no permanent shift of the assimilatory quotient.* Hence, all but the final reaction products, oxygen and starch or sucrose, must react back to form water.

It is very unlikely that this disappearance of intermediates should proceed by an exact reversion of the photochemical process. A chemiluminescence with a 100% yield can certainly not occur. In other words, we may expect a special kind of oxidoreduction within the assimilatory system (7) (compare the last part of this article, pages 43 *et seq.*).

One way of avoiding immediate back reactions lies in the use of radiant instead of thermal energy. A light quantum heats up, as it were, a single molecule in an otherwise cool environment. Once the structure of the absorbing molecule allows for a spontaneous conversion of electronic excitation energy into chemical energy (that is, a change of structure to a configuration with a higher free energy content), the new tautomer has a good chance to survive until some catalytic action makes use of its potential energy. The new tautomer will live longer the higher the wall of activation energy between the original and the photochemically changed structure. In other words, its lifetime depends on how much of the available light energy (in the case of chlorophyll, this is always *ca.* 40,000 calories and is independent of the size of the light quantum absorbed) is expended in forming the new structure. Actually, the very first photochemical product appears to be short-lived. It probably contains still much of the original 40,000 calories. J. Franck assumes, further, that a reduced intermediate hydrogen donor comparable to, let us say, the reduced pyridine nucleotides is not formed. Rather, the carboxyl group becomes reduced directly as a consequence of the photochemical reaction taking place in the chlorophyll–protein complex to which the essential car-

* Assimilatory quotient = (oxygen liberated)/(carbon dioxide consumed) or (hydrogen absorbed)/(carbon dioxide consumed).

boxyl group seems to be attached. Franck considers the influence which the presence of carbon dioxide exerts on the intensity of chlorophyll fluorescence *in vivo* as the most cogent evidence supporting this view. The scheme shown on page 35 does not do justice to these important considerations.

In case, however, the fluorescence experiments could be explained in another way, there would be no objection against assuming the existence of an intermediate hydrogen donor. Perhaps catalyst B of Franck and Herzfeld might play this part (4). Its concentration should be about two thousand times smaller than that of chlorophyll (the evidence can be found in the discussion about the photosynthetic unit).

Summing up, we may state that the heart of the photosynthesis problem is the effective utilization of energy which has to be accepted in a few big lumps. Franck and Herzfeld (4) calculate an immediate and permanent gain of 21,000 calories for each quantum absorbed. Does comparison with chemosynthesis point to an orthodox solution of this problem?

Possible Role of Energy-Rich Phosphate

Recent research on the utilization of metabolic energy in living cells had led to the discovery that this energy is handled in parcels not surpassing 12,000 calories packed away in so-called energy-rich phosphate bonds. Such phosphate bonds occur, for instance, in substances like Lipmann's acetyl phosphate (11) which, for the purposes of our discussion, may be regarded as stable for an indefinite period.

Recently, Emerson, Stauffer, and Umbreit (2) suggested "that for each quantum of light absorbed *one* 'energy-rich' phosphate bond (of *ca.* 10,000 cal./mole) is formed." There is no experiment or observation new or old which requires such an assumption. The only merit of this proposal, therefore, would lie in the complete analogy of photosynthesis with synthesis by thermal reactions. Part of the energy contained in the unstable first product formed in the photochemical reaction, the tautomer mentioned above (4), would be stabilized by conversion into a phosphate bond. The loss involved would be two-thirds of the chlorophyll excitation energy. There is general agreement among those who have considered carefully the theoretical

energy requirements of photosynthesis, that a minimum of 160,000 calories is needed per mole of carbon dioxide. Ten to twelve light quanta (let alone eight) transformed into phosphate bonds at 12,000 calories each would mean that the latter would have to be used with 120% efficiency. Actually, the efficiency with which the energy-rich phosphate bond can be used for synthetic reactions appears to be around 60%. If we postulate two phosphate bonds per quantum (that is, twenty bonds altogether), there would be enough energy. However, a complicated mechanism must be provided to divide the energy of the excited molecule between the two phosphate bonds. Finally, the resulting phosphorylated compounds should be stable enough to survive the end of an illumination period and cause the reduction of carbon dioxide for some time afterward in the dark. As said above, nothing of that kind has ever been observed, though many an investigator has looked for it. Hence, it is simpler to assume that Nature makes use of the particular advantages inherent in photochemical reactions and produces intermediate hydrogen donors which are capable of a one-step gain in free energy larger than those possible by way of energy-rich phosphate bonds.

A question quite different from that discussed is whether carbon dioxide becomes reduced to carbohydrate in the form of a phosphorylated compound. It is known that plant constituents combining with carbon dioxide in the dark are half saturated at a carbon dioxide partial pressure of less than 0.1 mm. mercury. The nature of the very first fixation of carbon dioxide preceding the reaction with excited chlorophyll is still under investigation. Suggestions like "reversal of a decarboxylation" are no more helpful than the dictum that the liberation of oxygen is "the reversal of respiration." What is needed is a working model fulfilling the energy requirements. A carboxylation reaction as effective as that taking place in the first step of photosynthesis needs the coupling with a reaction releasing some free energy (*ca.* 10,000 cal.). Ruben (15) suggests that the carboxylation occurs with the aid of an energy-rich phosphate bond present in the molecule which takes up carbon dioxide. Lipmann (11) recently has shown that, in the presence of certain enzyme preparations, acetyl phosphate, carbon dioxide, and molecular hydrogen can condense to pyruvate. By proposing that the pyruvate is removed quickly through further reduction and a new energy-rich phosphate furnished by an oxida-

tive side reaction, he has provided the first plausible model for a continuous chemoreduction of carbon dioxide.

As an experimental approach to tie up the phosphorylation reaction and photosynthesis, it is obviously insufficient to demonstrate that phosphorylations take part in the metabolism of plant cells and that the photosynthetic production of carbohydrates changes the dark equilibrium of various phosphate esters. It would be strange if it were not so. In a respiring or fermenting cell, the model for the chemical reduction of carbon dioxide as presented by Lipmann should work with any available metabolic hydrogen, particularly with free hydrogen activated by hydrogenase. Since, under the latter circumstances, the over-all energy requirements are very small, there is no reason why fermenting algae should not exhibit a clear-cut reduction of carbon dioxide in the dark. This is not the case, at least not at a noticeable rate. Apparently some extra energy for activation is necessary. If so, we have still to explain why, as found by Rieke in unpublished experiments, the number of quanta for the reduction of one carbon dioxide molecule remains ten when the algae begin to consume hydrogen instead of liberating oxygen. Energetically, one or two quanta would now suffice. The simplest explanation would involve the assumption of the existence of a rigid coupling of the reduction of carbon dioxide with the utilization of water as hydrogen donor. This would mean the exclusion of energetically "cheaper" hydrogen from a direct participation in the photochemical reaction during anaerobic photoreductions in algae and, by analogy, also in the photosynthetic purple bacteria. We cannot present in detail the points for and against this explanation. It is also not necessary, for in van Niel's comprehensive reviews (20) and Rabinowitch's book (14) the reader will find much stimulating discussion.

Interrelation of Carbon Assimilation with Oxidation Reactions in Plants and Purple Bacteria

We have seen above that there is a possibility that substrates, intermediates, and final products in photosynthesis may undergo reduction and polymerization as phosphorylated compounds. Experiments supporting Ruben's or Lipmann's hypotheses (11,15) are still missing, and Ochoa (12) has just reported the formation of oxalo-

succinic acid from α-ketoglutaric acid and carbon dioxide. This type of fixation reaction does not depend on energy-rich acyl phosphate. Our knowledge of phosphorylation and the "energy-rich" phosphate bond is based entirely on the carbohydrate metabolism involved in respiration and fermentation. We know that both types of catabolic reactions make use of the same phosphorylated compounds. If phosphorylated substances participate in photosynthesis, are these intermediates related to, or identical with, those occurring in general cell metabolism? Some special reaction must provide for the energy-rich phosphate bond assumed to promote the initial fixation of carbon dioxide in the dark. Is this special reaction part of the normal respiratory system or something different? In this regard Ruben (15) mentioned as a possible source of energy the dismutation to oxygen or reduction to water of the intermediate hydrogen acceptors (hydroxylated substances). The reduction of carbon dioxide by a purely thermal oxidoreduction in green algae occurs under circumstances which exclude normal respiration (7). Here the energy-rich phosphate bond would present a welcome means of explaining not only the fixation or carboxylation but also the coupling between the oxidation of hydrogen and the reduction of carbon dioxide. Experimentally, one should try therefore to establish the existence of an independent phosphorylation cycle serving exclusively the oxyhydrogen or "Knallgas" reaction.

While the oxyhydrogen reaction may be considered to be a sort of respiration and consequently invites comparison with the normal cell respiration in air, conditions become more unfamiliar when we turn to photoreduction. We mentioned earlier in this article why special back reactions in photosynthesis must be assumed. The question arises whether such (hypothetical) back reactions contribute to the formation of energy-rich phosphate bonds. The autonomy of the mechanism of photochemical reduction becomes quite apparent in the following observation (8). A concentration of 0.001 *M* phthiocol [also of methylnaphthoquinone (vitamin K) or of *o*-phenanthroline] inhibits strongly the respiration of algae like *Scenedesmus*, prevents completely normal aerobic photosynthesis, hinders the adaptation to and the reversion from photoreduction under anaerobic conditions, and severs the coupling between the oxyhydrogen reaction and the reduction of carbon dioxide in the dark. But if the inhibitor is added

to the plants after their adaptation to photoreduction it does not interfere with a reduction of carbon dioxide. The photochemical reaction proceeds with a normal quotient of two. Two volumes of hydrogen are absorbed together with one volume of carbon dioxide. The only change of importance concerns the quantum yield. It is exactly one-half of that found with the unpoisoned algae. Figure 3 shows how, with increasing concentrations of phthiocol at 560 lux,

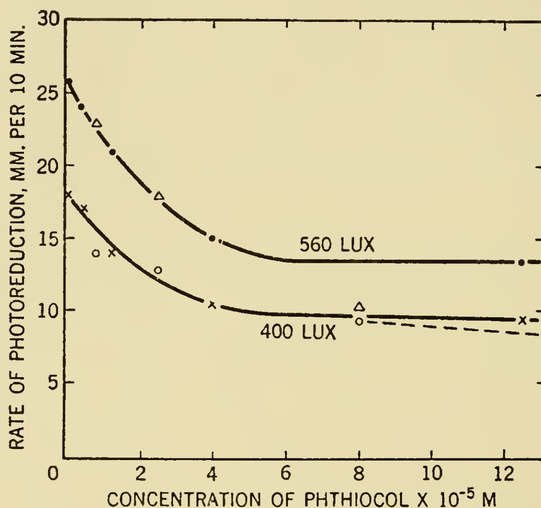


Fig. 3.—The inhibition and stabilization of photoreduction in *Scenedesmus* by increasing concentrations of phthiocol: ●—●, rates at 560 lux; Δ—Δ, same a day later; X—X, rates at 400 lux; O--O, same a day later.

the rate of the reaction falls from 26 to 13 mm. per 10 min. (or at 400 lux from 18 to 9 mm. per 10 min.) and then stays constant. What this means is quite obscure. A possible solution of the riddle may be found in the following considerations. The probable occurrence was mentioned above of back reactions between the reduced and oxidized products whenever photosynthesis is artificially inhibited. Now we postulate that, in the poisoned algae, all intermediates react back, and thereby activate a reduction of carbon dioxide with hydrogen in a manner similar to that brought about by the "Knallgas" reaction in the dark. Apart from any specific explanation we can state that, if

phosphorylated compounds participate at all in the reduction of carbon dioxide, they must be either drawn from a reservoir filled during preceding anaerobic periods or produced by a cycle belonging to the assimilatory system. No indication in favor of the first alternative has been observed. Special experiments to test it, however, have yet to be performed.

Another approach to the question of whether phosphorylations occur within the assimilatory mechanism would be by way of investigating further the metabolism of the photosynthetic purple bacteria (6,19,20). Studies on purple bacteria have proved extremely useful in the past for elucidating the similarities between photoreductions and the photosynthesis of green plants (*cf.* references 19 and 20). We now arrive at the point at which it would be of interest to analyze in detail the differences in their mechanisms.

An important difference between the green plants and the purple bacteria from the point of view of the possible role of phosphorylated compounds is the fact that the metabolism of the plant centers around carbohydrates, whereas most of the purple bacteria decline to utilize them in any way either in light or dark. They do not respire and they do not ferment glucose and consequently cannot grow in glucose media. The metabolism of purple bacteria revolves, if we consider only organic substrates, mainly around aliphatic acids and, in a few cases, simple alcohols. Acetate, propionate, butyrate, crotonate, malate, etc., are favorite substrates. The elementary analysis of an entire green plant yields, in general, data indicating the predominance of compounds of carbohydrate nature. The available elementary analyses of some purple bacteria have yielded figures indicating the presence of more hydrogen and of less oxygen than is found in carbohydrates and a composition very close to that of an extracted substance having the formula $(C_4H_6O_2)_n$. The latter proved to be a polymer of crotonic acid. The green plants store carbohydrates in more or less pure form whenever photosynthesis has lasted for a while in strong light, for respiration proceeds at a much slower pace. Some purple bacteria (*Athiorhodaceae*) do not seem to accumulate photoproducts in excess of what can be used for immediate synthesis of integrated cell material, that is to say, for growth of more bacteria. Others (*Thiorhodaceae*) accumulate unknown photosynthetic products that break down anaerobically in the dark by a sort of back

reaction which yields again hydrogen donors capable of being utilized in the light (20).

A second difference between plants and purple bacteria lies in the relation of cell multiplication to the assimilation of carbon. Green plants can, as a rule, grow normally in air on a heterotrophic diet without photosynthesis. Spoehr (17) succeeded in obtaining growth of hereditary chlorophyll-free albino corn plants by feeding them with only one organic compound, sucrose. On the other hand, we have not observed true growth of algae, during a photochemical or thermal reduction of carbon dioxide, under anaerobic conditions, that is, in the absence of respiration.

By contrast, some strains of purple bacteria multiply exclusively under anaerobic conditions and only during periods of illumination, and never in the dark. They are not capable of linking a synthetic process to either oxidative reactions (which do occur in the presence of oxygen) or fermentations (which seem not to occur at all). They depend for synthesis and growth upon photoreduction with special hydrogen donors like fatty acids, hydrogen sulfide, or molecular hydrogen. A few species among the nonsulfur purple bacteria (*Athiorhodaceae*) appear to grow in ways similar to that of the green plants in that they do not depend upon the photochemical reaction alone. As van Niel has shown, they can also grow aerobically in the dark by oxidizing the same substrates which they use as hydrogen donors in the light. But the fact that no carbohydrates are attacked points to a deviation from the metabolism of the green plants.

The third interesting difference between plants and purple bacteria concerns a direct interrelation between the respiratory and the assimilatory systems. In the ordinary green plant, respiration and photosynthesis can go on simultaneously. Metabolic measurements appear most consistent if we assume that both reactions run independently of one another and that any conspicuous increase of the rate of respiration in the light is caused in an indirect way. Photosynthesis appears to provide only reserve material, while the synthetic reactions leading to cell multiplication are coupled exclusively with respiration (and perhaps with fermentation). In those purple bacteria capable of growing at the expense of oxidation reactions, the utilization of oxygen must compete with the utilization of carbon dioxide plus light for the same hydrogen donors. The reactions do not occur

independently; they supplement or exclude each other. In a very interesting quantitative experiment, van Niel (20) found that the bacteria simply cease to take up oxygen when exposed to a sufficiently intense radiation. Here we have to search for a direct interrelation, an intermediate metabolic link.

We may sum up as follows. Whether phosphorylated compounds participate in photosynthesis must be considered in the light of two sets of observations. First, the assimilatory mechanism in green plants and in purple bacteria must be "self-supporting" as far as phosphorylations are concerned, since it functions under conditions in which neither respiration nor fermentation of carbohydrates seem capable of providing enough ready-made phosphorylated compounds. Second, in purple bacteria a carbohydrate appears perhaps as the primary product of the photochemical reaction, but instead of being stored it is converted into cell material of different elementary composition. These organisms are incapable of oxidizing or fermenting ordinary plant carbohydrates.

From the evolutionary point of view it is interesting that the "respiration" of the purple bacteria corresponds to the oxyhydrogen reaction and related oxidations in anaerobically adapted algae and not to the "normal" respiration in plants. We may speculate that the liberation of oxygen from "hydroxylated" compounds could be combined effectively with the reduction of carbon dioxide only after the synthesis and the utilization of sugars had become separated. Under aerobic conditions, the back reactions with free oxygen in the assimilatory system had to be prevented. We know that in adaptable algae this is brought about by the oxidative inactivation of one or two of the catalysts involved. Apparently we have here a parallel to the well-known case in which anaerobic fermentations are prevented from continuing in air by a special oxidation, the so-called "Pasteur reaction."

References

(1) Brown, N. C., *Forest Products*. 2nd ed., Wiley, New York, 1927.

(2) Emerson, R. L., Stauffer, T. F., and Umbreit, W. W., "Relationship between phosphorylation and photosynthesis in *Chlorella*," *Am. J. Botany*, **31**, 107 (1944).

(3) *Forestry Depletion in Outline*. Northwest Regional Council, Portland, 1940 (25¢).

(4) Franck, J., and Herzfeld, K. F., "Contribution to a theory of photosynthesis," *J. Phys. Chem.*, **45**, 978 (1941).
(5) French, C. S., and Rabideau, G. S., "The quantum yield of oxygen production by chloroplasts suspended in solutions containing ferric oxalate," *J. Gen. Physiol.*, **28**, 329 (1945).
(6) Gaffron, H., "Über den Stoffwechsel der Purpurbakterien," Part I, *Biochem. Z.*, **260**, 1 (1933); Part II, *ibid.*, **275**, 351 (1935).
(7) Gaffron, H., "Photosynthesis, photoreduction and dark reduction of carbon dioxide in certain algae," *Biol. Rev.*, **19**, 1–20 (1944).
(8) Gaffron, H., "*o*-Phenanthroline and derivatives of vitamin K as stabilizers of photoreduction in *Scenedesmus*," *J. Gen. Physiol.*, **28**, 259 (1945).
(9) Hill, R., and Scarisbrick, R., *Proc. Roy. Soc. London*, **B129**, supplement, 39 (1940).
(10) Krebs, H. A., "Carbon dioxide assimilation in heterotrophic organisms," *Ann. Rev. Biochem.*, **12**, 529 (1943).
(11) Lipmann, F., *J. Biol. Chem.*, **158**, 515 (1945).
(12) Ochoa, S., "Isocitric dehydrogenase and carbon dioxide fixation," *J. Biol. Chem.*, **159**, 243 (1945).
(13) *Planning for a Permanent Agriculture.* U. S. Dept. Agr., *Misc. Pub.* No. 351 (1939).
(14) Rabinowitch, E., *Photosynthesis*, Vol. I. Interscience, New York, 1945.
(15) Ruben, S., "Photosynthesis and phosphorylation," *J. Am. Chem. Soc.*, **65**, 279 (1943).
(16) Sinnott, E. W., "Plants and the material basis of civilization," *Am. Naturalist*, **79**, 28 (1945).
(17) Spoehr, H. A., "The culture of albino maize," *Plant Physiol.*, **17**, 397 (1942).
(18) United Nations Conference on Food and Agriculture, Report, Hot Springs, May, 1943.
(19) van Niel, C. B., "On the morphology and physiology of the purple and green sulphur bacteria," *Arch. Mikrobiol.*, **3**, 1 (1931).
(20) van Niel, C. B., "The bacterial photosyntheses and their importance for the general problems of photosynthesis," in *Advances in Enzymology*, Vol. I. Interscience, New York, 1941, p. 263.
"The culture, general physiology, morphology, and classification of the non-sulfur purple and brown bacteria," *Bact. Revs.*, **8**, 1 (1944).
(21) Warburg, O., and Lüttgens, W., "Weitere Experimente zur Kohlensäureassimilation," *Naturwissenschaften*, **32**, 301 (1945).
(22) Went, F. W., *et al.*, "Plant growth under controlled conditions," several articles in *Am. J. Botany*, **31–32**, 1944–1945.
(23) Werkman, C. H., and Wood, H. G., "Heterotrophic assimilation of carbon dioxide," in *Advances in Enzymology*, Vol. II, 1942, p. 135.

4

THE BACTERIAL CELL

RENÉ J. DUBOS, MEMBER OF THE ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH, NEW YORK; MEMBER OF THE NATIONAL ACADEMY OF SCIENCES; JOHN PHILLIPS MEMORIAL AWARD; MEAD JOHNSON AWARD

Only such substances can be anchored at any particular part of the organism which fit into the molecule of the recipient combination as a piece of mosaic fits into a certain pattern.

PAUL EHRLICH

BACTERIA appeared to the nineteenth century biologist as a type of protoplasmic material devoid of any organization, almost as a link between the animate and the inanimate world. "They are," said Ferdinand Cohn, "the simplest and lowest of all living forms—beyond them, life does not exist." The compound microscope failed to reveal any structure within their cellular boundaries, and the biochemist was inclined to consider the bacterial cell as a mere bag of enzymes which owed its enormous biochemical activity to its colloidal dimensions. The primitiveness of bacterial life appeared to be confirmed in chemical terms when Winogradsky demonstrated in 1887 that certain autotrophic species can grow in purely inorganic media and can synthesize their protoplasm from mineral salts and carbon dioxide, utilizing for the reduction of the latter the energy released by the oxidation of sulfur, iron, ammonia, nitrite, etc. (21). Was it not permissible to consider this production of organic matter from inorganic elements as the most primitive biochemical expression of life, as the beginning of life on earth?

Advances on the diverse fronts of bacteriology were quick to dispel these early illusions concerning the biochemical primitiveness and the simplicity of organization of the bacterial cell. Analysis of the chemical activities of bacteria soon revealed that microbial life takes place through the agency of the same type of reactions, the same

metabolic channels and products, and the same biocatalysts which constitute the mechanism of life in the highest and most evolved organisms. For example, the oxidation of sulfur by the autotrophic bacterium *Thiobacillus thiooxidans* depends upon an intimate linking between oxidation and phosphate turnover; the oxidative phase is accompanied by phosphate fixation and the reductive phase of carbon dioxide fixation is accompanied by a release of phosphate (22). Moreover, *Thiobacillus thiooxidans* is fully equipped with the regular complement of water-soluble vitamins found in other living organisms: thiamin, riboflavin, nicotinic acid, pantothenic acid, pyridoxine, biotin, etc. (17). In other words, the mechanisms of energy transfer and of intermediary metabolism are essentially as complex in the least exacting bacteria as they are in the most fastidious organism.

The chemistry of *Thiobacillus thiooxidans* is not unique among bacteria; most of the known water-soluble vitamins—with the possible exception of ascorbic acid—have now been found to be either produced by, or required for the growth of, all the microbial species so far studied. Fat-soluble vitamins also probably play a part in microbial metabolism since at least one of them, vitamin K, is an essential growth factor for Johne's bacillus (*Mycobacterium paratuberculosis*) and the other mycobacteria produce biologically active naphthoquinones during growth (24).

Thus, bacteria utilize the multiple and complex biocatalysts which govern and integrate the metabolism of all living cells. Moreover, many bacteria, the autotrophic species for example, possess in addition the ability to synthesize these same biocatalysts from inorganic elements in the course of their growth in synthetic media, a property which most plant and animal cells never possessed or have lost entirely. The high degree of cellular organization required for the performance and for the integration of these complex syntheses need not be emphasized; at the biochemical level, at least, there is no ground to consider that bacteria represent primitive forms of life. The growth requirements of autotrophic bacteria are extremely simple indeed, but how complex their vital machinery, their performance, and their products! If they are truly the first representatives of life on earth, they sprang, like Minerva, fully armed from the forehead of Jove.

Simultaneously with the realization that the biochemical processes of bacteria are no simpler than those of other organisms came

the recognition of the existence in the bacterial cell of a number of structures which, although often ill defined in nature and function, obviously express a morphological complexity parallel to the biochemical complexity of all known forms of life. Utilization of the classical methods of cytology soon revealed, for example, the existence in bacteria of flagella, spores, different kinds of membranes and capsules etc., which give to each bacterial type a fairly characteristic morphological individuality (12). Little by little, bacteriological staining techniques are gaining the dignity of cytochemical reactions and give chemical definition to the cellular objects which they reveal; the more skillful utilization of Feulgen's reagent for instance, permits the identification among the other basophilic constituents of the bacterial cell of discrete bodies rich in desoxyribonucleic acid which are almost certainly the equivalent of the vesicular nucleus in larger cells (12,19). Photography in the ultraviolet and electron microscopy have permitted the optical resolution of cellular structures—intracellular granules, membranes, individual components of flagella, etc.—which are below the limit of resolution by ordinary microscopy. These classical cytological techniques aim at the *direct* visualization of the constituents of the cell. On the other hand, the analysis of the response of the cell to the effect of certain reagents and procedures provides an *indirect* approach to cytological problems, by suggesting the existence and often the chemical nature of important cellular components which cannot be seen by any of the known methods of microscopy. Interestingly enough, it is the study of pathogenic bacteria which has been the most fruitful from the point of view of this indirect approach to cytology.

In order to analyze the host–parasite relationship, the student of infection must concern himself with those structures and products of bacteria—the cellular antigens and toxins—which affect the course of the infectious process and against which are directed the reactions of immunity. Similarly, the attempts to understand the mode of action of antiseptics on bacteria led to the study of the structures through which antiseptic agent and susceptible cell come in contact. Thus, many constituents of the bacterial cell have been recognized first by biological reactions; and the analysis of the phenomena of infection, immunity, and chemotherapy has provided important information concerning the biochemical architecture of bacteria. Paul

Ehrlich first stated clearly the possibility of describing these reactions in terms of cellular structure. He postulated that the living cell possesses a number of chemically reactive groups which he called "receptors" and with which dyes, bactericidal substances, and immune antibodies react selectively. Ehrlich regarded these "receptors" as definite chemical entities capable of entering into union with dyes, antiseptics, and antibodies. According to his theory, characteristic staining reactions, differential susceptibilities to toxic substances, and specific reactions with corresponding antibodies could all be explained by assuming the existence of a sufficient number of receptors in the bacterial cell. These phenomena can consequently serve as tests to facilitate the recognition of the receptors and their isolation in pure form. During the past decades, immunochemists and students of the theory of chemotherapy have gone far toward identifying, and in several cases separating in a purified state, several of the cellular components with which antibodies and antibacterial agents react selectively.

The specific chemical relationships involved in the chemotherapeutic reactions are discussed in other essays of this volume and need not be considered here. There are certain aspects of the problem, however, which are too ill-defined to warrant discussion in terms of a chemical theory, but which deserve mention at this time since they bid fair to help in the elucidation of some interesting details of cellular structure. Empirical staining reactions led very early to the division of the bacterial world into three broad groups, the Gram-positive, the Gram-negative, and the acid-fast, which are defined by their behavior toward two staining techniques (the Gram and Ziehl-Neelsen methods). These three bacterial groups not only differ in their staining properties, but also exhibit striking differential susceptibilities to the different types of antiseptics and antibacterial agents. By way of illustration, most Gram-negative bacilli grow readily in the presence of basic dyes, penicillin, or gramicidin, whereas these organisms are extremely susceptible to the bactericidal and lytic effect of immune serum. On the contrary, many Gram-positive species are completely resistant to lysis by immune serum, but are extremely susceptible to the bacteriostatic and bactericidal effect of small concentrations of dyes, penicillin, and gramicidin. As for the tubercle bacilli (acid-fast), they are remarkably resistant to all the classical antiseptics and to a great variety

of other toxic agents, retaining their viability, for example, after exposure to 5% sodium hydroxide or sulfuric acid. In order to account for these strikingly different susceptibilities, many theories have been proposed which assume that the bacterial groups vary with reference to cellular permeability, presence of certain lipids in the cell membranes, acid-base properties of the cell body, etc. Thus, observations of electrokinetic behavior and of affinity for dyes at different *p*H levels suggest that the cell material in the colon-dysentery-typhoid group (Gram-negative) is less acidic than in the Gram-positive bacterial species (pneumococci, staphylococci, streptococci, anthrax bacilli, etc.) (1,20). The acid-fast bacilli (*e. g.*, tubercle bacilli) produce astonishingly high concentrations of a variety of lipids which gives them marked hydrophobic properties (2,23). All the Gram-negative species readily yield in solution phospholipid–protein–polysaccharide complexes which constitute 5–10% of the total cell weight; similar complexes have not been obtained from the Gram-positive organisms (16,18). Recent observations have established a correlation between ability to retain the Gram stain (correlated with greater susceptibility to many antiseptics) and the presence around the cell of a magnesium ribonucleate complex (5,10). All these facts provide examples of the type of chemical information which results from the analysis of the biological behavior of the cell and which will undoubtedly reveal important differences of structure between the different bacterial groups.

The specific antibodies produced as the result of the injection of bacterial antigens into the animal body have provided another set of reagents that have yielded important information of a cytochemical nature. The immune reaction to any one type of bacterial cell is not a simple phenomenon, since bacteria are made up of a multiplicity of chemical constituents many of which elicit the production of specific antibodies. In other words, the injection of one type of bacterial cell usually results in the production of several antibodies, each one of which is directed against one particular cellular component. These different cellular constituents obviously bear a definite spatial morphological relationship to each other in the intact cell. Some are masked by membranes and become exposed only as a result of cellular disintegration; others are peripherally disposed and in direct contact with the environment. This stratification of cellular structures affects the im-

mune response of the animal host to the whole bacterium and is reflected in the type and amount of antibody produced. It also conditions the reaction of the bacterial cell with a given antibody, since the intact microorganism unites much more readily, if not solely, with the antibodies which are specifically directed against those of its constituents exposed at or near the surface. Analysis of antibody production and of antigen–antibody reaction can therefore help in formulating an approximate picture of the arrangement of the different antigens in the architecture of the cell.

To summarize, the use of immunological procedures for the study of cellular structure involves a number of successive steps: (a) the preparation and separation of antibodies specific for each one of the chemical constituents of the cell; (b) the utilization of these antibodies as specific reagents for the detection and preparation in pure form of the cellular constituents; and finally, (c) the interpretation of antigen–antibody reaction in an attempt to define the relative positions occupied by these chemical constituents in the living cell (7,9,11,14).

The results obtained by immunochemical analysis have led to the recognition that the different bacterial groups (pneumococci, streptococci, sporulating aerobic bacilli, organisms of the colon-typhoid-dysentery group, etc.) are characterized by a general pattern of antigenic organization which is common to the different members of each group. On the other hand, the various species and immunochemical types within any general group differ from each other by virtue of the chemical specificity of the different components of their antigenic mosaic. Thus, all virulent pneumococci possess a capsule which is polysaccharide in nature, but the polysaccharide varies chemically and antigenically from type to type (3). The virulent forms of human streptococci (group A) can also produce a capsule made up of hyaluronic acid; moreover, they all possess as surface constituents peculiar proteins (the M substances), and other substances of unknown chemical nature (the T antigens), which vary in immunochemical specificity from type to type (13). Several of the aerobic sporulating bacilli have been found to produce a capsule consisting essentially, if not solely, of polypeptides; the polypeptide, in the case of the anthrax bacillus, appears to be made up exclusively of *d*-glutamic acid (6). The virulent coliform bacilli (typhoid, dysentery, etc.) all produce the

lipid–protein–polysaccharide complexes mentioned above. The polysaccharide components determine the immunochemical specificity of the different species. The protein component, on the contrary, appears to be essentially the same in all the members of the group; in fact, it can be made to combine with the different specific polysaccharides to reconstitute complexes similar to those which normally occur in the cell (16).

In general, the immunochemist has concerned himself primarily with the constituents which are present at the cell surface because these elements appear to play the most important part in the phenomena of immunity. However, other proteins, polysaccharides, etc., which certainly occupy less superficial positions in the cell have also been recognized by immunochemical analysis. A striking illustration of the potentialities of immunochemical methods in cytology is given by the case of the typhoid bacillus. Specific antibodies have been prepared for the following cellular components of this organism: the flagella; the O and Vi antigens of the cell surface; the R, ρ, and Q antigens, which are intracellular components. Living typhoid bacilli resuspended in solutions of these different antibodies exhibit a characteristic behavior which is determined by the relative position of the corresponding antigen in the cellular architecture. Loss of motility, different patterns of agglutination, bacteriolysis, etc., are phenomena which are characteristic for each antigen–antibody reaction and which can be interpreted in terms of cellular organization of the different antigens.

Dyes, antiseptics, and antibodies are not the only reagents which can be used to recognize and identify the cellular receptors. If it be found, for example, that a given enzyme attacks the cells of a certain microbial species causing death or the alteration of a characteristic cellular property, it can be surmised that the chemical substrate which is susceptible to the enzyme is present in the cell under consideration and that it plays some part in the function altered by the enzyme. Thus, the fact that lysozyme causes the death and lysis of the cells of several bacterial species indicates that the mucopolysaccharides hydrolyzed by this enzyme are essential components of the cellular structure of the susceptible species (15). It has been shown also that other polysaccharidases decompose the capsular polysaccharides of pneumococci and of streptococci, and that proteolytic

enzymes inactivate the specific M protein antigens of group A streptococci (8,13). Contrary to what is observed with lysozyme, however, the polysaccharidases and proteases do not affect in any way the viability of the treated cells, even when they rid it entirely by hydrolysis of the specific polysaccharides or proteins. It is likely, therefore, that, instead of being considered as structural constituents of the bacterial bodies, the capsular polysaccharides and the M proteins should be regarded as excretion products which accumulate around the cell, since they can be removed or destroyed without interfering with the essential living processes.

There are many other instances of biological reactions which, because of the specific relationships they bear to certain cellular components of bacteria, can be used as indirect methods for the analysis of cellular structures. Let us mention, for example, the remarkable selectivity of pure lines of bacteriophage with reference to the strains of bacteria which they can attack. The relationship between specificity and cellular structure is illustrated by the fact that the bacteriophage can be absorbed specifically by the cells, living or dead, of the susceptible bacterial cultures. It has also been found that, in certain cases, soluble fractions extracted from the susceptible bacterial cells inhibit specifically the lysis of the homologous organisms by the bacteriophage. It is likely, therefore, that the phenomena of bacteriophage lysis will also yield a number of new specific reactions which by revealing the existence and nature of new types of receptors could serve in the analysis of cellular structure.

The biological phenomena which we have considered have in common the following characteristics permitting their utilization as indirect cytological methods. They are all the result of a reaction between a given reagent (antiseptic, enzyme, antibody, bacteriophage) and a specific cellular receptor, this reaction manifesting itself by inhibition of growth, enzymic destruction of cellular component, agglutination or lysis by antibody, or lysis by bacteriophage. In many cases the reagent can be absorbed on the homologous cell substrate, and the reaction can be inhibited by the addition to the system of the specific substrate which constitutes the cellular receptor. Inhibition of growth, enzymic decomposition, agglutination, lysis, etc,. are only the secondary manifestations of primary reactions which depend upon the union between the cellular receptors, on the one hand,

and the biological reagents, be they antiseptics, antibodies, enzymes, bacteriophages, on the other.

One of the most intriguing applications of the indirect cytological methods discussed in the preceding pages has been the analysis of the phenomena of bacteria variability. It has long been known that most bacterial cultures—even those arising from single cells—often undergo profound transmissible modifications of their morphological, biochemical, and physiological properties. Immunochemical studies have revealed, in particular, a type of variation, now recognized in practically all bacterial species, which involves the loss of the specific surface components of the cell (the capsular polysaccharides of pneumococci, the M proteins of streptococci, the capsular polypeptide of anthrax, the lipid-protein-polysaccharide complexes of the dysentery and typhoid bacilli, etc.). These transmissible modifications of the surface of the bacterial cell have attracted particular attention because they are in many cases correlated with alteration of the virulence of the organism concerned. They constitute, however, only a very narrow aspect of the total problem of bacterial variability. One can observe within one given bacterial culture transmissible modifications of many unrelated properties: ability to attack sugars or proteins, to synthesize amino acids or pigments, to resist antiseptics or other injurious procedures, to produce flagella, spores, capsules, and so on. All these variations occur independently of each other, thus giving to each bacterium the possibility of manifesting its existence under a great diversity of forms and properties. The production of these large numbers of variant forms deficient in one or another of the cellular components has greatly helped in the analysis of many immunochemical problems.

From a more general point of view, it is of the greatest interest that a given organism can successfully continue to exist and to multiply as an independent living object after having lost a great variety of structures and functions which had appeared to constitute important components and attributes of the "normal" parent form. As already stated, these structures and functions can be lost and regained independently of each other, without altering the essential nature of the germ or the potentialities of the cell. Even more striking is the fact that it is possible to substitute experimentally one character for another. Thus, by adding to a strain of pneumococcus which has lost

the ability to produce its specific capsular polysaccharide an extract of the cell of another type of encapsulated pneumococcus, one can convert the former organism into the type from which the extract was made. From then on, the cell can produce, and transfer to its progeny the ability to produce, a polysaccharide different from the one it had been known to synthesize heretofore. All available evidence indicates that the substance which is capable of inducing the transformation is a form of desoxyribonucleic acid—specific for each pneumococcus type (4). Of equal interest is the unavoidable conclusion that the bacterial cell is not only an integrated complex of independent characters, but that it is possible to substitute for one of these characters another one, homologous but different, without interfering essentially with cellular organization.

Thus, a large body of knowledge concerning cytology is slowly emerging from the study of bacterial variability and of the behavior of the cell in the presence of a number of biological reagents. It is becoming possible to recognize and to define in chemical terms a number of structures not yet detectable by any microscopic technique. Furthermore, by bold, even though admittedly dangerous, extrapolation, one can guess at the approximate position of these cellular constituents in the architecture of bacteria. The history of science provides, of course, many examples of the fruitfulness of indirect methods, and in particular of the utilization of chemical and biological manifestations as indices and guides for the recognition and identification of morphological structures. Claude Bernard stated as early as 1855 that "anatomical localization is often revealed first through the analysis of the physiological processes." Much of the morphology indirectly revealed by antibodies, enzymes, and cytotoxic substances lies beyond the microscopic range and in fact often reaches the molecular level. It concerns the organization of those molecular groupings which, because of their chemical reactivity, condition the behavior of the cell both as an independent functioning unit and in its relation to the environment. This knowledge is of obvious interest to the cytologist. It also forms the fundamental basis upon which is being erected the theory of immunity, since all the phenomena of host–parasite relationship are essentially a reflection of the biochemical architecture of the cell.

*References**

(1) Albert, A., "Chemistry and physics of antiseptics in relation to mode of action ," *Lancet*, **1942,** II, 633–636.

(2) Anderson, R. J., "The chemistry of the lipids of the tubercle bacilli," *Harvey Lectures*, **35,** 271–313 (1939–1940).

(3) Avery, O. T., "The role of specific carbohydrates in pneumococcus infection and immunity," *Ann. Internal Med.*, **6,** 1–9 (1932–1933).

(4) Avery, O. T., MacLeod, C. M., and McCarty, M., "Studies on the chemical nature of the substance inducing transformation of pneumococcal types," *J. Exptl. Med.*, **79,** 137–158 (1944).

(5) Bartholomew, J. W., and Umbreit, W. W., "Ribonucleic acid and the Gram stain," *J. Bact.*, **48,** 567–578 (1944).

(6) Bovarnick, M., "The formation of extracellular *d*(–)glutamic acid polypeptide by *Bacillus subtilis*," *J. Biol. Chem.*, **145,** 415–424 (1942).

(7) Boyd, Wm. C., *Fundamentals of Immunology.* Interscience, New York, 1943.

(8) Dubos, R. J., "Enzymatic analysis of the antigenic structure of pneumococci," *Ergeb. Enzymforsch.*, **8,** 135–148 (1939).

(9) Heidelberger, M., "Immunology as a tool in biological research," *Am. Naturalist*, **77,** 193–198 (1943).

(10) Henry, H., and Stacey, M., "Histochemistry of the Gram-staining reaction for micro-organisms," *Nature*, **151,** 671 (1943).

(11) Kabat, E. A., "Immunochemistry of proteins," *J. Immunol.*, **47,** 513–587 (1943).

(12) Knaysi, G., *Elements of Bacterial Cytology.* Comstock, Ithaca, 1944.

(13) Lancefield, R. C., *et al.*, "Studies on the antigenic composition of group A hemolytic streptococci," *J. Exptl. Med.*, **78,** 465–476 (1943); **79,** 79–114 (1944). "Specific relationship of cell composition to biological activity of hemolytic streptococci," *Harvey Lectures*, **36,** 251–290 (1940–1941).

(14) Landsteiner, K., *The Specificity of Serological Reactions.* Rev. ed., Harvard Univ. Press, Cambridge, 1944.

(15) Meyer, K., Palmer, J. W., Thompson, R., and Khorazo, D., "On the mechanism of lysozyme action," *J. Biol. Chem.*, **113,** 479–486 (1936).

(16) Morgan, W. T. J., and Partridge, S. M., "An examination of the O antigenic complex of *Bact. typhosum*," *Brit. J. Exptl. Path.*, **23,** 151–165 (1942). "Studies in immunochemistry. 6. The use of phenol and of

* The material discussed in this essay is presented in a more complete manner and with extensive documentation in a monograph, *The Bacterial Cell*, by R. J. Dubos and C. Robinow, Harvard University Press, Cambridge, 1945.

alkali in the degradation of antigenic material isolated from *Bact. dysenteriae* (Shiga)," *Biochem. J.*, **35**, 1140–1163 (1941).

(17) O'Kane, D. F., "The presence of growth factors in the cells of the autotrophic sulphur bacteria," *J. Bact.*, **43**, 7 (1942).

(18) Partridge, S. M., and Morgan, W. T. J., "Immunization experiments with artificial complexes formed from substances isolated from the antigen of *Bact. Shigae*," *Brit. J. Exptl. Path.*, **21**, 180–195 (1940).

(19) Ribonow, C. F., "A study of the nuclear apparatus of bacteria," *Proc. Roy. Soc. London*, **130**, 299–328 (1942). "Cytological observations on *Bact. coli*, *Proteus vulgaris* and various aerobic spore-forming bacteria with special reference to the nuclear structures," *J. Hyg.*, **43**, 413–423 (1942).

(20) Stearn, A. E., and Stearn, E. W., "Metathetic staining reactions with special reference to bacterial systems," *Protoplasma*, **12**, 435–464, 580–600 (1931).

(21) van Niel, C. B., "Biochemical problems of the chemo-autotrophic bacteria," *Physiol. Revs.*, **23**, 338–354 (1943).

(22) Vogler, K. G., and Umbreit, W. W., "Studies on the metabolism of the autotrophic bacteria," *J. Gen. Physiol.*, **26**, 157–167 (1942).

(23) Wells, H. G., and Long, E. R., *The Chemistry of Tuberculosis*, 2nd ed. rev., Williams & Wilkins, Baltimore, 1932.

(24) Woolley, D. W., and McCarter, J. R., "Antihemorrhagic compounds as growth factors for the Johne's Bacillus," *Proc. Soc. Exptl. Biol. Med.*, **45**, 357–360 (1940).

5

THE NUTRITION AND BIOCHEMISTRY OF PLANTS

D. R. HOAGLAND, PROFESSOR OF PLANT NUTRITION, COLLEGE OF AGRICULTURE; PLANT PHYSIOLOGIST, AGRICULTURAL EXPERIMENT STATION, UNIVERSITY OF CALIFORNIA

OTHER writers for this volume will discuss the biochemistry of plants in relation to photosynthesis, plant hormones, and the activities of microorganisms. The present article is, therefore, devoted primarily to an attempt to indicate the need and the opportunities for research on the biochemistry of higher plants, especially plants of agricultural interest, as a foundation for the adequate understanding of many problems of plant nutrition and of plant physiology. This field of inquiry is relatively undeveloped in the modern period in comparison with the biochemistry of higher animals and of microorganisms, with its remarkable record of achievement during the past quarter of a century. The importance of the biochemistry of the higher plants for the cycles of living organisms in general, and for the basic occupation of agriculture is too obvious to require analysis.

The disparity of achievement in fundamental biochemical research dealing with higher plants, on the one hand, and with the higher forms of animal organisms and some groups of microorganisms on the other, becomes apparent on examination of recent monographs dealing with advances in biochemistry. They are predominantly concerned with experiments on animal tissues, or on microorganisms, and the majority of contributors are associated with medical research in-

stitutes. Recent texts on biochemistry give but little specific attention to the biochemistry of higher plants. Certain earlier treatises devoted largely to this latter subject have not appeared in new editions for a good many years. This general appraisal on a comparative basis appears to be justified even when the noteworthy contributions of individual workers or of certain groups of workers on the biochemistry of higher plants are kept in view. It receives support from some of the comments of Vickery, who is well known as an extensive contributor to several phases of the biochemistry of plants.

The situation as described may seem surprising when one recalls the vast programs of research carried on by agricultural experiment stations. It is true that, in these stations, a great number of studies on plants have been made which are to some degree biochemical in nature. But it is rare to find groups of investigators assigned the definite objective of developing the knowledge of the fundamental biochemistry of crop plants. Generally, biochemical studies are encouraged in so far as they throw light on particular questions of agricultural importance as related to plant nutrition, horticulture, agronomy, or perhaps general plant physiology. In terms of crop production, the success of the coordinated attacks on plant problems through the application of the agricultural sciences and arts is well demonstrated by the enlarged production of crops under the difficulties of war conditions. This, however, does not meet the point we have under discussion. Further, there is reason to assume that, even from a utilitarian point of view, a more intensive development of research on the biochemistry of crop plants would in due course contribute to the basic knowledge essential to the control of plant production and supplement and guide the interpretation of results of practical experimentation.

Ramifications of the biochemistry of plants into the applied field are manifold. The growth of crops needs to be appraised by the criteria not only of total yield but also of quality. This latter aspect is currently receiving much attention through consideration of plant composition in relation to the value of the plant product for animal nutrition, a point illustrated by the research program of the Federal Soil, Plant, and Nutrition Laboratory at Cornell. For example, the general problem of the synthesis of vitamins by plants might be cited, as well as the studies made to gain information on the

relative influence of climate and mineral nutrition of the plant on its vitamin content. Also, the biochemical mechanisms in the plant that result in the synthesis of essential amino acids, carbohydrates, and higher fatty acids are of interest to students of both plant and animal metabolism. Knowledge of the biochemistry of the plant is one of the sources of information that constitutes a valuable asset to the plant pathologist, the plant geneticist, the horticulturist, the specialist in forestry, and indeed to all those who must of necessity come into contact with plant systems of biochemical reactions, whether or not this is consciously recognized.

From some points of view, the plant offers, as compared with the animal, methods of study with marked advantages. On the other hand, there are disadvantages in the use of the higher plant for biochemical investigation. The nature of a plant's growth is such that most of its living cells perform the most diversified functions, a fact which renders plant cells less suitable for research on specific biochemical reactions. Tissues from animal organs often provide material with specialized and highly intense activities suitable for the investigation of enzyme reactions. It is doubtful that the plant in general is as favorable as the animal for similar investigations, but this view may be held only because fewer attempts have been made to exploit the possibilities of plant material. There is no opportunity to carry on with the plant the kind of experiments that are rendered possible by the presence of a blood stream and organs of elimination. The introduction of specific organic metabolites into the plant cell and the study of their transformations within the cell involve special complications. In the examination of biochemical reactions taking place in excised plant tissues, avoidance of bacterial and fungal contamination often presents great difficulties. Further, the correlating effects of various plant hormones and of normal translocation of metabolites make especially difficult the interpretation of the results of studies of plant tissues in terms of the intact growing plant.

Plants of the kind under consideration may be regarded as normally complete synthetic systems, building up or breaking down compounds representing an extraordinary array of organic structures of biological interest, all derived from the simple substances required for plant growth, namely, carbon dioxide, water, and inorganic elements to the number of fifteen or more. It might be postulated that

certain species of plants growing in some types of soil high in organic matter may have lost the power of synthesizing at an adequate rate vitamins or other essential organic units, and that these species depend in part on the absorption of these units from an environment in which they have been synthesized by microorganisms. But most or all species of higher green plants so far intensively studied (these are mainly plants of economic importance) can go through their cycles of growth by virtue of their own synthetic powers, at least so far as can be ascertained by the use of purified inorganic media, although usually without complete exclusion of all microorganisms. It follows that the range and diversity of biochemical reactions that need to be investigated is enormous. Correspondingly great are the opportunities for the study of different synthetic processes in a living organism, to the extent that adequate methods can be devised for attacking such complex systems.

The preoccupation of plant physiologists engaged in agricultural research with the inorganic elements absorbed by crop plants from the soil, and the frequent designation of these elements as "plant foods," tend to subordinate appreciation of the biochemical aspects of plant nutrition. The inorganic elements derived from the soil constitute only a small percentage of the dry weight of a plant, and one of the most important objectives of research in plant nutrition should be an understanding of the mechanisms by which these inorganic components become directly incorporated into the organic compounds synthesized by the plant, or activate the enzymes which catalyze the syntheses and breakdown of organic compounds.

We have as a foundation the knowledge that plants of the kind in question absorb from an inorganic medium, and have an essential need for, the elements nitrogen, phosphorus, potassium, calcium, magnesium, sulfur, and iron. Also, as a result of research in plant nutrition during the past decade or two, other elements have been shown to be equally essential though required in only minute quantities. These additional elements include boron, manganese, copper, and zinc, with strong but still limited evidence that molybdenum is also essential. There may be, and probably are, still other chemical elements indispensable to plant growth, but conclusive proof of the general indispensability of other elements, over a wide range of plant species, has not yet been obtained. Limitations of technique are soon

encountered as the effort is made to go further in the exclusion of impurities from the nutrient medium. It should be noted, however, that various chemical elements not indispensable for growth of the plant may modify its biochemical reactions either beneficially or adversely, under the conditions of a natural environment.

Many of these facts have been established primarily through the use of artificial culture methods, among which the so-called water-culture method is especially useful for studying the effects of a deficiency of an element needed by the plant in minute quantity. Sometimes, however, special care in the selection and purification of a solid inert medium, to which a purified nutrient solution is applied, provides an alternative technique, with certain advantages.

Innumerable experiments, some of them with meticulous care, have been made on many species of plants with these artificial culture techniques. The control of the inorganic nutrient medium represents only a partial control of the environment; and frequently there remains the necessity or desirability of control of the atmospheric factors to which the plant is subjected: light, temperature, humidity, carbon dioxide concentration, and air movement. The control of these factors obviously demands costly equipment and usually is not attempted in nutritional experiments with plants. But some laboratories have had the opportunity to grow plants under conditions of controlled air temperature and controlled artificial illumination. In recent years, fluorescent lamps have proved especially valuable for this purpose.

Most artificial culture experiments have not been designed for the primary purpose of obtaining information about the metabolic mechanisms of the plant and the specific enzyme systems concerned. Observations on plants grown in the presence of selected combinations of inorganic nutrients are likely to be confined to measurements of rate of growth of the plants, total yields of the tissue produced on a fresh or dry weight basis, or on yields of some part of the plant of special interest from an agricultural point of view. Chemical studies are often limited to the determination at the end of a selected growth period of certain chemical elements absorbed by the plant, or of well-known organic compounds formed as a net result of innumerable biochemical processes that have proceeded perhaps for a considerable period of time during which the plant has increased in size and differentiated its tissues. In other cases, the purpose may be to record the

pathological symptoms resulting from a marked deficiency in the medium of some one of the essential inorganic elements. Valuable as this information is for the purpose in view, it does not advance our understanding of the biochemistry of plants in a manner at all comparable with the advances made in the study of animal tissues or of some microorganisms, in which definite steps in a series of chemical reactions are identified or reasonably deduced from experimental data.

The use of the artificial culture methods of plant nutrition makes feasible the growing of plants of many species with any desired combination of inorganic nutrients, or with a given nutrient available in graduated quantities. Controlled modifications in the inorganic composition of the plant are thereby induced, although generally in no simple relation to the composition of the nutrient solution. As already stated, possibilities exist for control of illumination and temperature and, thus, to some degree for control of carbon assimilation and rates of metabolic reactions.

It is tempting to propose that these methods of controlled culture afford techniques for endless rewarding studies on biochemical mechanisms in the plant. To what extent this is a realistic view is difficult to say. The extraordinary complexity of the growing plant and of the conditions of its nutrition may set narrow limits to what can be done of fundamental biochemical importance, but whatever opportunities do exist have yet to be adequately explored. There are, of course, other methods of experimentation on plants which can be adapted to biochemical research, such as embryo culture, culture of root tips, and experiments with excised leaves, roots, or other parts of the plant immersed in solutions of known composition. A technique has been recently described (13) for physiological and chemical studies on albino plants, whereby the transformations of a known carbohydrate supplied to the plant might be followed, without being complicated or obscured by reactions which are associated with photosynthesis.

It may be noted again that from the standpoint of plant nutrition the main biochemical problem is the fate of the chemical elements derived from the nutrient medium and the way in which they interact with the products of photosynthesis. Most of the elements essential to plants are also essential to all organisms; research on their metabolic functions has therefore a wide biochemical interest. For some of

the essential inorganic elements, and especially certain "trace" elements, we have little or no guidance from studies on other organisms. Boron is indispensable for all the higher plants so far properly investigated. It has not been shown to be indispensable to the animal, although little work has been done on this point. Boron is one of the elements plants require in minute amounts, yet deficiencies even under some soil conditions have assumed first-rate agricultural importance, a fact which accentuates interest in the biochemical functions of boron. Research by plant physiologists indicates that deficiency of boron often results in a pathological state in the plant nearly the same as that caused by a deficiency of calcium. One view is that an inadequate supply of boron may limit the maintenance of an effective level of calcium in a soluble or active form within the tissues of plants. Some workers think that formation of pectin compounds does not proceed normally when boron is deficient. Clearly these are questions which need the attention of skilled biochemists. Possibly productive leads might come from comparative biochemistry. Certain groups of fungi, and perhaps some algae, appear not to require boron. The same fungi also can grow without calcium, or at least the amounts needed are too small to be removed from the media by present methods of purification. Further study of certain phases of organic metabolism in plant organisms with different boron or calcium requirements might conceivably point to biochemical reactions for which boron or calcium, or both, may be indispensable.

Another among the chemical elements effective in micro quantities to which an indispensable function in the growth of higher green plants must be assigned is zinc. This element, like boron, is not always adequately supplied by the soil, and the deficient plant becomes diseased ("little-leaf," "mottle-leaf" of trees, and pathological conditions shown by other crop plants as a result of zinc deficiency). Physiological studies under the control of artificial culture disclose some of the effects of zinc deficiency. One such study in this laboratory yielded evidence that, without an adequate supply of zinc, plant growth substances of the auxin type are either not synthesized at a rate sufficient for normal growth or else are destroyed too rapidly (12). But it has also been learned that protein synthesis is retarded when the zinc concentration in the plant falls below a critical level. Tomato plants were grown with graduated supplies of zinc, so that some of the plants

at the time of the experiment gave no visible symptomatic deficiency response, yet addition of zinc to the nutrient medium rapidly induced an increased rate of protein synthesis (1). Observations of this kind, however, only show that some unknown link in a chain of reactions is broken. The nature of the enzyme systems existing in the plant, of which zinc is an essential component, or activator, is an unanswered question. From investigations on blood cells comes evidence that zinc is a component of carbonic anhydrase; and there is a suggestion from studies on yeast that zinc is one of the activators of the enzyme aldolase. We have no positive evidence of this character derived from experiments performed directly on enzyme systems of higher plants.

Zinc is often successfully applied in curing zinc deficiency disease by spraying the plant or treating the soil. This discovery, valuable as it is in agricultural practice, does not satisfy the curious investigator who seeks enlightenment on the function of zinc in plant metabolism. Knowledge of why zinc is needed by the plant might improve existing agricultural practice by providing a rational basis for the treatment of zinc deficiency disease; but clearly, as in other fields of research, sound progress in the study of plant biochemistry cannot be hoped for if the direction of research is to be governed by the degree of probability that a given investigation will in itself have a practical outcome.

These are only illustrations of the wide gaps in fundamental biochemical insight into the functions of the so-called plant foods. One could write of the lack of the kind of experimentation which might elucidate the role of potassium, one of the most important fertilizer elements in enzymic reactions in the plant. At the present time, it is possible to cite data from one source or another that could be interpreted in terms of an effect of potassium on almost every general biochemical process of which the plant is capable. The net conclusion is that potassium is an essential element for plant growth and that its deficiency may impair growth in various ways or alter the composition of the plant, depending upon the factors such as degree of potassium deficiency, the concentration in the media of calcium, sodium, or other ions, the species of plant studied, and the physiological age of the plant. The desirability of further basic information on the role of potassium in biochemical processes in the plant is evident. For example, a suggestion has been advanced that potassium has an im-

portant effect on certain of the phosphorylation processes in muscle, with an antagonistic action by calcium. Studies of this type, applied to the enzyme systems of higher plants, would obviously be of great significance for plant nutrition.

If we turn to the essential element phosphorus, great encouragement is gained for the view that biochemical research on animal tissues and on microorganisms can serve as a guide for extension of research on the metabolism of the plant. In fact, current research on phosphorus metabolism of various organisms has far-reaching implications for problems of plant nutrition from both a theoretical and practical standpoint.

Following the work of Cori on the phosphorolytic system in animal tissues which brings about the synthesis of glycogen from glucose-1-phosphate came the discovery by Hanes (5) that an enzyme system can be prepared from plant tissues (potato, peas) which catalyzes reversible phosphorolytic reactions by which starch can be synthesized *in vitro*. Physical and chemical studies have been made of the synthesized starch, and its structure compared with that of natural plant starch (7,11).

The conclusion is that the artificially synthesized starch represents only the amylose component of natural starch, that is, the one made up of long chains of glucose units, whereas the natural starch also includes a component characterized by a branched-chain structure. Recently, the failure to reproduce artificially the complete natural starch was apparently overcome. A preliminary report has been made of the isolation of another enzyme system from potato which can accomplish the synthesis *in vitro* of the amylopectin component of starch, with the branched-chain structure (8). The investigations as a whole on this question therefore represent clarification of the role of an essential inorganic element in synthesis by the plant of one of its most important carbohydrates. It is not difficult to appreciate the way in which this addition to plant biochemistry may aid in the guidance of researches in plant nutrition with reference to the utilization of phosphate. From the agricultural point of view, the information now available may well become of value in appraising the adequacy of the phosphate supply for high yields or starch content of a crop, especially as further research provides more data on concentrations of various forms of phosphate in the plant under diverse nutrient and atmospheric

environments. To what extent, if at all, the relative proportions of amylose and amylopectin may be subject to modification by physiological conditions remains to be studied.

The phosphorolytic mechanisms probably will supply the key to an understanding of the synthesis of another carbohydrate almost universally synthesized by higher plants, namely, sucrose. This is also the dominant sugar of commerce and, as one authority has pointed out, is manufactured commercially in far greater quantity than any other pure chemical product. It is natural, therefore, that many attempts have been made to analyze the mechanism of sucrose synthesis in the plant. Recently the enzymic synthesis of sucrose *in vitro* was accomplished (7); and, while the enzyme system responsible is derived from a bacterial organism (*Pseudomonas saccarophila*), the achievement has such great suggestive interest for the investigator of the nutrition of higher plants that it seems appropriate to mention it in this connection. The substrates utilized in the synthesis were glucose-1-phosphate and fructose. Sucrose is broken down into these components in a phosphorolytic reaction catalyzed by an enzyme system in the bacterial cell. By starting with glucose-1-phosphate and fructose in the presence of the enzyme, pure crystalline sucrose was prepared which was identical with the natural product. The identity was established by all available physical and chemical criteria. This is the first well-authenticated synthesis of this sugar.

It is true that a similar enzyme system has not yet been isolated from the tissues of higher plants, despite various attempts to do so in this laboratory. Nevertheless, biochemical studies on various species of plants strongly support the view that the synthesis of sucrose does proceed by chemical reactions in which glucose or fructose phosphate esters, or both, serve as substrates, although the mechanism is probably not identical with that of the bacterial enzyme system. Certain studies on sugar cane leaves suggest that, in the higher plants, fructose diphosphate takes part in the synthetic reaction (6). It is of fundamental importance that the experimental evidence now available shows that, for the synthesis of sucrose from glucose and fructose in the plant, aerobic metabolism is indispensable. Possibly aerobic oxidations are essential to the phosphorylation of one of the substrates involved in the synthesis of the sucrose. The question is complicated by the observation that various substrates other than glucose and fructose

may result in sucrose formation by plant tissues. For example, in experiments on barley shoots by infiltration procedures, galactose could be utilized for this purpose, as well as various other carbohydrates or related compounds. In the leaves of the sugar cane, as studied by Hartt, the oxidative system involved was not inhibited by cyanide, although in some experiments on other plant tissues in this laboratory the synthesis was cyanide sensitive. While the mechanisms of sucrose synthesis operating in the higher plant are by no means sufficiently elucidated as yet, there is reason for an optimistic view that further developments along the general lines of attack already pursued will eventually lead to a satisfactory biochemical solution of this important problem of plant metabolism and plant nutrition.

Closely related to the investigations just outlined is the long-standing question of the biochemical nature of the interconversion of starch and sucrose in the plant. As an illustration, the well-known sweetening of potatoes at low temperatures may be cited. In this process, starch is converted to sucrose. This conversion is also an aerobic process and is inhibited by cyanide and some other respiratory poisons. A mixture of hexose-6-phosphates has been isolated from potato juice, and also phosphatases capable of hydrolyzing these compounds. A tentative scheme to explain the conversion of starch to sucrose has been advanced on the basis of reactions for which hexose phosphate esters are requisite; and, according to the explanation offered, both glucose-1-phosphate ester and fructose diphosphate are essential (11). In the earlier experiments in Hawaii on sucrose synthesis by sugar cane leaves, fructose diphosphate was likewise regarded as an essential substrate. On the other hand, in the bacterial enzyme synthesis referred to above, only glucose-1-phosphate and fructose could be converted to sucrose.

The great problem of cellulose synthesis remains without a biochemical explanation. Whether this synthesis can take place only from activities of the organized protoplasm, and whether phosphorolytic processes are involved, are in the realm of speculation at the present time.

The point to be emphasized by the foregoing remarks is that biochemical research, by contributing to the basic knowledge of carbohydrate transformations, can influence profoundly the study of plant nutrition and physiology. There is, of course, open for further research

the immense problem of the origin of polysaccharides other than cellulose and starch, which comprise a large fraction of many tissues of higher plants, such as hexosans, galactans, pentosans, and related compounds. There is need for more information about the synthesis of the important pectin compounds.

As a brief digression from the main theme, it is of general biochemical interest to refer to the specificity of the enzyme system from *Pseudomonas saccharophila* which catalyzes the synthesis of sucrose. This enzyme system was not restricted in its catalytic potentiality to the synthesis of sucrose. It was also effective in bringing about the synthesis of two new disaccharides. One was synthesized *in vitro* from glucose-1-phosphate and *l*-sorbose, the other from the glucose ester and a ketoxylose (3). The versatility of enzyme systems of this type was thereby demonstrated.

To the student of practical plant nutrition interested in the application of fertilizers to soil, the utilization of simple nitrogen compounds by crop plants is always a topic of dominant interest. This interest is accentuated today because of the enormous expansion of industries for the fixation of atmospheric nitrogen. Greatly increased quantities of fixed nitrogen could be made available after the war for agriculture. Field and pot experiments on the effects of nitrogen fertilizers on crop growth are of course legion, but this type of investigation discloses little of the biochemical processes by which the nitrate or ammonia absorbed by the plant is elaborated into organic compounds such as the amino acids.

Fortunately, this subject has attracted the efforts of a number of able investigators whose primary concern has been that of biochemistry; for reviews on the history of research on nitrogen metabolism of plants and recent trends, the monograph by Chibnall (2) and reports by Vickery *et al.* (15–18) may be consulted. In the consideration of this aspect of plant biochemistry, it is apparent once more that guidance in the interpretation of data and in the design of experiments is greatly influenced by previous research concerned with the biochemistry of muscle and other animal tissues.

The importance of organic acids and their cycles of metabolism, including the tricarboxylic acid cycle, have been stressed by some investigators. Prominent among the organic constituents of plants are malic and citric acids. Oxalic acid also occurs very frequently, and

succinic acid has been established as a component of the tissues of several plant species whose content of organic acids has been examined. Frequently all the organic acid content of the plant is not accounted for, and unknown organic acids require identification and quantitative estimation. But it seems that all the organic acids postulated as components of organic acid cycles may be present in plant tissues.

A general theory of protein metabolism based on experiments with seedlings and detached leaves has been evolved which assigns significant roles to the amides, asparagine and glutamine, and to various keto acids. As already noted, the mechanisms postulated draw heavily on the explanations advanced to account for transformations of nitrogen compounds in animal tissues; but caution is needed in applying mechanisms based on the study of animal tissues to plant processes, without adequate confirmatory evidence. Efforts have been made, however, to integrate available data on changes in the organic composition of excised plant leaves under experimental conditions, as well as data on the changes that occur in the intact growing plant, into schemes of protein synthesis and breakdown correlated with catalytic cycles. The possibilities of applying to this field of study in the plant the new tool of isotopic nitrogen have been opened by Vickery and his collaborators in a preliminary experiment with the tobacco plant, following the well-known research of the Schoenheimer group on animal metabolism.

Another aspect of the problem of nitrogen metabolism is concerned with the symbiotic fixation of nitrogen by leguminous plants. Much more will have to be learned about protein metabolism before the biochemical reactions of the nodule organism can be properly understood. Some progress has been made—compare the review by Wilson (19)—but the theories and evaluation of evidence now available are apparently subject to controversy. The extraordinary practical importance of nitrogen fixation and its scientific interest invites further efforts in research by biochemists.

The brilliant study of oxidation systems in living organisms rests primarily on the experiments and insight of those who have been concerned with muscle and other animal tissues, or with yeast. No comparable achievements by investigators of higher plants come to mind. Some investigators of higher plants, however, have sought to apply fundamental knowledge gained by studies on other organisms

to the study of plant respiration. Goddard, for example, considered the role of the cytochrome system (4). It appears that this system is in fact active in some plant tissues, but not in all. Thus, cytochrome oxidase activity was demonstrated in the embryos and roots of certain species of plants, also in immature, but not in adult, leaves. A report is made of the successful isolation of cytochrome C from wheat germ. Wheat cytochrome C has the same absorption spectrum as heart cytochrome C. Its reduced form is oxidized by heart or wheat cytochrome oxidase. Succinic dehydrogenase was found to be present in wheat germ in small amounts.

W. O. James and his associates (10) have undertaken a series of investigations on the enzyme systems extracted from the sap of the barley plant. Here, an ascorbic acid system appeared to be active in normal aerobic respiration. This oxidation system is thought to be characteristic of higher plants. The degradation of sugars, however, seems to proceed by way of well-known reactions of phosphorylation and hydrogen transfer, as described for other kinds of tissue. Various plant storage tissues have been selected by a considerable number of workers as suitable material for the study of respiratory processes. The effects of respiratory inhibitors on these and other plant tissues have also received much attention.

A comprehensive survey of the literature of plant respiration from the point of view of modern concepts is greatly to be desired, but these examples may perhaps serve to illustrate the point that enzyme chemists can find broad opportunities in the higher plants. This material provides an important field of research in which the number of workers is inadequate to cope effectively with the many and formidable problems presented. It is reasonable to suppose that a sufficiently intensive effort directed at plant research might be calculated to advance the subject of respiratory systems in general, as well as supply much needed basic knowledge for plant nutrition and all its ramifications in agriculture.

The growth of plants in soil and its dependence on the absorption of inorganic salts or their ions by root cells, with its implications for soil and plant interrelations and for fertilizer practices, brings into the foreground the phenomenon of solute absorption and translocation by living cells—compare the review by Hoagland (9). While these phenomena are of general significance to the study of physiological

processes in all living organisms, they possess a peculiar importance in the consideration of the nutrition of higher plants for the reasons already suggested. At one time, the absorption of salts by plant cells would usually not have seemed to belong to the domain of biochemistry. The intake of solutes was regarded as a passive process, in which the permeability of protoplasmic membranes was chiefly stressed. It is now well recognized that solutes may move into plant cells against concentration or activity gradients at the expense of metabolic energy. This kind of absorption of salts by living cells of the root, often referred to as salt accumulation, is dependent on aerobic respiration. The accumulation process is inhibited by many respiratory poisons, such as cyanide and iodoacetate, not only in the initial absorption of salts or ions by the root, but also in their polarized movements into the plant's upward conducting system, and their subsequent accumulation in the cells of the leaf or reproductive organs. There are various theories of the mechanisms by which solutes move through the living conducting system of the plant. It is agreed that simple diffusion cannot explain the movement, and the conclusion cannot be escaped that at some point solutes move against gradients or are accelerated in their movement through coupling with some energy-yielding process.

Concurrently with the accumulation of some ions by plant cells, various metabolic processes are stimulated, such as synthesis of organic acids, or proteins, and oxidation of sugars. In fact, the biochemistry of salt absorption by plant cells, in its various aspects, appears to offer a profitable branch of research in the plant field. Researches on storage tissues of plants provide eloquent testimony of the value of studying biochemical transformations as part of an investigation of salt accumulation—compare the review by Steward (14).

In the development of studies on salt absorption or movement, the tool of radioactive isotopes may become of great value. Thus certain metabolic reactions may be related to the movement of a selected radioactive ion, which can be detected with extreme sensitivity of measurement. So far, the use of radioactive isotopes in research on higher plants has been limited in the main to simple tracer studies designed to obtain information on rates or direction of movement and to identify tissues through which translocation or accumulation takes place. A much wider field for the application of isotopes, stable and radioactive, awaits development, and with it may come knowledge of

the biochemical aspects of salt absorption and particularly of the energy-yielding reactions which are inextricably bound up with this process.

While it is generally accepted now that the accumulation of salt by plant cells is in some way closely linked with metabolism, and particularly with aerobic respiration, the incompleteness of this knowledge should be clearly recognized. Even if the steps in the particular respiratory cycles, and the active enzymes, coenzymes, and activators participating, should be identified to the extent that they sometimes have been in biochemical studies, the mechanisms by which metabolic reactions are coupled to the movement of solutes into the living plant cell, or to its polarized translocation from one tissue to another, would still be obscure. These are questions that have not been answered for any living cells, and the plant organisms may offer an especially favorable system for further research on solute movement.

It is hoped that these remarks may serve to call attention to the abundant opportunities for fundamental biochemical research on higher plants, including the great groups of plants of economic importance. There is a place in this field for more workers of the kind who have done so splendidly in advancing biochemistry, particularly in its relation to animal nutrition, medicine, and the metabolism of microorganisms.

References

(1) Bean, R. S., *Ph.D. thesis*, University of California, 1943.

(2) Chibnall, A. C., *Protein Metabolism in the Plant.* Yale Univ. Press, New Haven, 1939.

(3) Doudoroff, M., Hassid, W. Z., and Barker, H. A., *Science*, **100**, 315 (1944).

(4) Goddard, D. R., *Am. J. Botany*, **31**, 270 (1944).

(5) Hanes, C. S., *Proc. Roy. Soc. London* **B128**, 1421 (1940); **B129**, 174 (1940).

(6) Hartt, C. E., *Hawaiian Planters' Record*, **48**, 31 (1944).

(7) Hassid, W. Z., Doudoroff, M., and Barker, H. A., *J. Am. Chem. Soc.*, **66**, 1416 (1944).

(8) Haworth, W. N., Peat, S., and Bourne, E. J., *Nature*, **154**, 236 (1944).

(9) Hoagland, D. R., *Inorganic Nutrition of Plants.* Chronica Botanica, Waltham, 1944.

(10) James, W. O., Heard, C. R. C., and James, G. M., *New Phytologist*, **43**, 62 (1944).

(11) McCready, R. M., *Ph.D. thesis*, University of California, 1944.

(12) Skoog, F., *Am. J. Botany*, **27**, 939 (1940).

(13) Spoehr, H., *Plant Physiol.*, **17**, 397 (1942).

(14) Steward, F. C., *Trans. Faraday Soc.*, **33**, 1006 (1937).

(15) Vickery, H. B., Leavenworth, C. S., and Wakeman, A. J., *J. Biol. Chem.*, **125**, 527 (1938).

(16) Vickery, H. B., Leavenworth, C. S., and Wakeman, A. J., *Conn. Agr. Expt. Sta. Bull.*, No. 422 (1940).

(17) Vickery, H. B., and Pucher, G. W., *J. Biol. Chem.*, **128**, 703 (1939).

(18) Vickery, H. B., Pucher, G. W., Schoenheimer, R., and Rittenberg, D., *J. Biol. Chem.*, **135**, 531 (1940).

(19) Wilson, P. W., *The Biochemistry of Symbiotic Nitrogen Fixation.* Univ. Wisconsin Press, Madison, 1940.

6

BIOLOGICAL SIGNIFICANCE OF VITAMINS

C. A. ELVEHJEM, PROFESSOR OF BIOCHEMISTRY, COLLEGE OF AGRICULTURE, UNIVERSITY OF WISCONSIN; WILLARD GIBBS MEDALIST

PREVIOUS to the 20th century, thousands, and more likely millions, of people suffered and died because of a lack of scientific knowledge about vitamins or of an insufficient supply of foods rich in these essential nutrients. During the first three decades of this century, about a dozen vitamins not only have been identified, isolated, and synthesized, but manufacturing methods have been perfected to the point at which some vitamins can be supplied at a cost relatively lower than the cost of calories and proteins. Most nutritionists agree that there are more vitamins to be isolated and better methods of synthesis to be developed, but I wonder how many have given thought to the possibility that the uncontrolled production of synthetic nutrients may lead to sufficient economic disturbances in agricultural production to affect the health of the people of the world adversely.

It is unwise, especially for a biochemist, to make any predictions of future developments. The field of vitamins, however, has now developed to the point at which it is possible to look ahead in light of past experiences.

The early workers on vitamins had a point of view or philosophy quite different from that held by present investigators. Many of the pioneers were motivated by the single purpose of alleviating human

suffering. When liver was given to relieve night blindness and fresh vegetables were used to cure scurvy, the practitioner knew nothing about vitamins—he was interested in healing the patient. Eijkman, I am sure, carried a mental picture of the severe cases of beriberi which he encountered in Java during all his attempts to relate this disease to a specific essential nutrient. R. R. Williams referred to his early contact with the disease in his Willard Gibbs award address as follows: "In short, beri-beri was a principal topic of conversation in scientific and medical circles in Manila during those early years of my enlistment with Vedder in the Philippines." When Goldberger was called upon in 1914 to undertake studies on the cause of pellagra, he knew very little about the disease, but on December 13, 1915, he wrote as follows to Dr. Milton Rosenau: "I can hardly describe the feeling that I experience as I go through our wards at the asylum and see the poor insane women who a year ago had pellagra but who this year are perfectly well—so far as pellagra is concerned."

Regardless of the satisfaction experienced by the individual workers, these phenomenal results did not captivate the interest of administrators of research funds. More support was given for studies on animal nutrition, since a premium was placed on production. Few recognized that the development of strong human bodies would also pay dividends. Perhaps we can now explain this difference in reaction. The animal husbandrymen took great interest in judging and selecting fine stock. If better nutrition produced better stock they were interested. On the other hand, medical students have always been given sick people to study rather than the ultrahealthy. For example, Sir Robert McCarrison of England went to India to study disease but his most important contributions originated because he was impressed by the perfect physique of the Hunza race. At present many are interested in expanding our conception of the relation of nutrition to optimum health, but we still are not too certain about the procedure; some talk about extra quantities of vitamins, others advocate physical training.

It was not surprising, therefore, that between 1910 and 1920 the following laboratory findings continued to attract widespread interest. (1) Calves maintained on diets balanced according to the recognized standards failed to grow on rations made entirely from products of the wheat plant but thrived on rations made from products of the corn plant. (2) Rats placed on purified diets developed nor-

mally when the fats of the diet consisted largely of butterfat but failed when certain vegetable oils devoid of vitamin A were used; this observation was especially significant at a time when Danish children given skim milk plus vegetable fats were developing the same symptoms as those observed in rats. (3) Chicks grew normally if allowed access to sunlight shortly after hatching, but developed severe leg weakness when hatched early in the spring; at that time attempts were being made to start the chicks during the winter months so that broilers would be available when the demand was heavy. When the deficiency agent was found to be vitamin D, not only was the poultry industry saved but a program was initiated which led to the eradication of human rickets.

By 1925, many laboratories were undertaking systematic nutritional studies, and further attempts were being made to produce specific beriberi in rats and chicks, vitamin A deficiency in rats and dogs, vitamin C deficiency in guinea pigs and monkeys, vitamin D deficiency in chicks, dogs and rats, and pellagra in dogs. Rough assay procedures were developed; foods richest in each of these vitamins were designated "protective" foods. Theoretically, this knowledge was all that was needed to prevent the deficiency disease resulting from the lack of each respective vitamin. Carlson, a few years ago, stated that we had sufficient knowledge to prevent all beriberi in the world long before vitamin B_1 was synthesized, a statement which is true in a limited sense. Thus, Chamberlain and Vedder reduced the incidence of beriberi in the Philippine Scouts by issuing unpolished rice, but the problems of producing high-quality unpolished rice and of educating people to use this type of rice product as part of their diet are still with us.

All these studies were most intriguing to ever-increasing numbers of research workers. The bars were now down to the progress of nutrition research and the studies gained momentum each year.

Some wanted to know what happened during each deficiency and how the vitamin functioned in producing normal animals. One of the first attempts involved histological studies of tissues from vitamin-deficient animals; and it was soon recognized that the absence of a minute trace of a nutrient led to extensive structural changes in certain tissues. Even greater progress was made when the function of vitamins was related to the dynamics of the living cell. In 1921, Seidell stated that, aside from a possible significant difference in the degree of

dialyzability, there are no grounds for not classifying vitamins with enzymes. Prior to this, Harden and Young, in 1906, emphasized the importance of organic dialyzable substances in yeast fermentation and called these substances coenzymes. In 1918, Meyerhof found the coenzymes of yeast to be present in a number of animal tissues; but animal workers were too busy compounding rations to pay any attention to this finding. R. J. Williams, in 1919, concluded that the substance or substances which stimulate the growth of yeast are identical with the substance or substances which in animal nutrition prevent beriberi or polyneuritis.

Then, in 1932, Warburg and Christian found that the so-called yellow enzyme which they had shown to be active in a reconstructed oxidation system contained a derivative of riboflavin, the second member of the B complex, as the prosthetic group. A short time later, it was found that the coenzyme used in the same system and prepared from red blood cells contained nicotinic acid. In 1937, nicotinic acid was shown to be the antipellagra factor and thus became identified with the third member of the B complex. In 1932, Auhagen split carboxylase, the enzyme necessary for the metabolism of pyruvic acid, into a protein component and a thermostable part called cocarboxylase. Soon cocarboxylase was identified as the pyrophosphoric acid ester of vitamin B_1.

The fourth decade of the 20th century will undoubtedly be recognized as the period of greatest advance in our knowledge of the mechanism of action of the vitamins. We must recognize that much is still unknown and that some of the most difficult problems lie ahead. However, the results so far obtained have had a much broader influence—they have given new impetus to the study of enzyme chemistry. R. R. Williams states: "These enzyme molecules are too vast and complex for the chemist to decipher completely today, but we can now say that the prosthetic group or business ends of these molecules are in many instances what we earlier came to call vitamins. These vitamins are, therefore, the bits, the working ends, of the keys which unlock stores of vital energy from glucose and other foods." The enzyme approach has emphasized the close relationship of all types of life. Vitamins have turned out to be growth factors and metabolic regulators for plants, bacteria, protozoa and yeast, as well as for animals. See W. H. Peterson, *Biol. Symposia*, **5**, 31 (1941).

Others wanted to know what a vitamin looked like. The first crystalline material to be isolated from a natural concentrate having vitamin activity was probably nicotinic acid. It was obtained between 1912 and 1914 from rice bran and yeast; but unfortunately its biological activity was tested for antineuritic activity rather than for antipellagra activity. The successful establishment of the chemical constitution of several of the vitamins depended upon an enormous amount of work and true chemical ability. The first of the vitamins to be given serious chemical consideration was undoubtedly vitamin D. In 1925, Steenbock and Black, and Hess and co-workers showed that crude cholesterol could be activated by ultraviolet light, and the following year ergosterol was recognized as the actual provitamin. These observations stimulated the interest of organic chemists in the structure of sterols, a problem that had received only sporadic attention. A little later, the chemical basis for the relationship between carotene and vitamin A was established. It is interesting to note that, although the structure of vitamins A and D received early attention, these are the only well-known vitamins still not available in synthetic form. Vitamins C and B_1 were the first to be made synthetically, but only about ten years ago. During the past decade, methods for the synthesis of ten different vitamins have been perfected.

I have merely recorded the final results without paying tribute to the individual workers for their years of study of the details of chemical structure. It is true that some of the work was stimulated by commercial interest, but in many cases the individual workers were rewarded only by the satisfaction obtained from the successful proof of structure of "their" vitamin. That chemical industry did become interested in the production of vitamins was indeed fortunate. The availability of each new vitamin facilitated progress on the remaining vitamins.

As far as I am aware, no one has formally expressed the gratitude of laboratory workers for the large quantities of vitamins supplied gratis by industry for experimental purposes. Although it is true that many papers carry a footnote indicating indebtedness to a particular firm "for a generous supply of crystalline vitamins," such an acknowledgment is so common today that it is often taken for granted. I have no way of estimating the total expenditure involved, for the value of these gifts cannot be calculated merely by multiplying the number

of pounds supplied by the current price. The crystalline material was most valuable to the investigator when the supply was still in limited production. For example, our work on the newer members of the B complex with the chick was directly dependent upon our ability to obtain adequate supplies of pure biotin. Currently, everyone is interested in feeding his animals purified rations containing only the synthetic vitamins; and the vitamin requirements become rather large when dogs, monkeys, pigs, and even human subjects are used. Many of us today would be willing to pay a fancy price for even a few milligrams of pure folic acid. If work on the chemistry of this and related compounds had not been limited by the war, sufficient quantities of it would undoubtedly now be available for experimental purposes.

As the methods of synthesis improved and the demand for the compound increased, very substantial decreases in the wholesale prices of most of the vitamins were made. The cost of riboflavin has decreased from $17.50 per gram in April, 1938, to 30 cents per gram in gram lots in October, 1944. In January, 1934, vitamin C cost $213 per ounce; today, one ounce may be purchased for 95 cents.

The reduced prices made these vitamins available for many purposes other than for the manufacture of elixirs, tablets, and capsules: synthetic ascorbic acid is added to the lemon powder used in army rations; B vitamins are added to flour, bread, and corn grits; and several of the vitamins were supplied to other countries through "Lend-Lease." In 1944, production of certain of the individual vitamins ranged from 100,000 to 1,000,000 pounds.

Vitamins are no longer limited to the laboratory and the doctor's office—they are now part of big business: more extensive use of vitamins means greater dividends to the stockholders of many industries; therefore, large advertising campaigns have been instituted; and as profits increase, more funds become available for research. A few years ago some of us were highly pleased if we received $500 to support a favorite project. Today, yearly grants as high as $50,000 are made for nutrition research—a magnificent start. We have the interest of the public and the support of industry, and we should have many well-trained and energetic investigators in the postwar period.

But what of the future? First we must realize that many of the workers will be interested in research merely for the sake of research. Many will have had little contact with extensive deficiency diseases.

This should not be disturbing if the findings are properly applied. The practical problems will involve economics as well as chemistry and physiology. In the field of medicine, the doctor will continue to prescribe vitamins for deficiencies which are clearly diagnosed, and in some cases he will try vitamins to determine if any beneficial effects can be obtained. If the patient recognizes some benefit, the use of vitamin supplements will be continued for some time; if no benefits are recognized, the box of capsules will probably remain on the shelf of the medicine cabinet. A certain group of people will buy vitamin preparations on their own initiative but few will take capsules continually for any length of time.

Extension of the use of vitamins will probably have to come by way of the addition of vitamins to widely used foods. The supply of niacin did not become critical immediately after the discovery of its role in curing pellagra; but the supply was critical within a few weeks after the introduction of flour enrichment. If my calculations are correct, 1,000,000 pounds of niacin is almost sufficient to supply the minimum requirement of all the people in the United States for one year, an amount which the annual production is now reaching. What will happen now that war is over and synthetic niacin will be in direct competition with niacin in meat, and synthetic ascorbic acid will be in competition with vitamin C in oranges, tomatoes, etc.? Will industry be willing to control the production of synthetic vitamins in relation to the true demand for these products? I would be greatly disturbed by an extensive advertising campaign advocating greater use of synthetic vitamin C by the public at a time when the orange crop is rotting in orchards. On the other hand, if the synthetic vitamins are used to supplement rather than replace our food supply, we can plan for this country—yes, even the world—a continuous supply of nutrients which will be little affected by crop failures. I hope such a plan can be made and executed before the problem becomes so acute that government may have to step in. At present, the use of vitamins is promoted largely among people who are rather adequately fed. What results we could expect from the proper use of vitamins in the control of famines in India! It is true that vitamins cannot replace other food nutrients but certain vitamins at least increase the efficiency of utilization of the total nutrients and may also help in the synthesis of other vitamins in the intestinal tract.

I believe there is another important angle which applies not only to the use of synthetic vitamins but also to synthetic amino acids, which undoubtedly will be produced in the postwar period. Since, in general, the synthetics have no taste appeal, and since mankind will continue to consume food for reasons other than that of mere nutrition, it may be more important to use a larger part of the synthetics in animal feeding. As we learn more about nutrition, we are finding that many of the more expensive animal feeds can be replaced by cheaper substitutes. For example, riboflavin can be used in poultry feeding without relying upon more expensive milk products. Thus, the cost of animal production can be reduced to such an extent that animal products can be used more widely for human consumption.

The necessity of fortifying certain human foods may continue for some time. Although new types of food fortifications may be introduced, we must recognize that any enrichment program is not necessarily permanent, and we should be willing to discontinue any one program when and if scientific evidence indicates that it is no longer necessary. It will be the duty of nutritionists to give careful consideration to these programs; proper decisions can be made only if we have extensive knowledge of the vitamin content of all foods. Food industries have generously supported such programs but the work has certainly not reached completion. Plans should be made to set aside funds which are now easily obtainable so that work of this kind can be carried out when personnel become available. There are two important lines of approach: one deals with the production of food products high in vitamins and can be accomplished by improved breeding, cultivation, and fertilization; the other deals with improved methods of handling the food products after harvesting and slaughtering.

Because fundamental research must continue in the field of vitamins, it will be fortunate to have young men interested in pure research. During the past few years, the practical problems have received greatest emphasis, but we have now reached the point at which fundamental research again becomes the limiting factor in further progress. We must study cellular mechanisms within the body and the relation of bacterial cells to the vitamins within the intestinal tract. The relation of vitamins to enzymes has already been discussed. After all, sturdy bodies are largely dependent upon properly functioning enzymes in all the cells of the body. Vitamins are only a small part

of the enzyme molecule and we need to know more about the rest of the molecule. Perhaps proper exercise with limited amounts of vitamins may be more conducive to rigorous enzyme systems than the consumption of vitamin cocktails while reclining in an easy chair. Studies on the enzyme systems will help to relate nutrition to such important problems as resistance to infection, prevention of cancer, and resistance to the process of aging. Biochemists have learned how to disorganize cells into parts, but greater integration of the parts into a whole must be attained.

More attention also must be paid to the bacteria in the intestinal tract. Bacteriologists have studied the bacteria in soils, in milk, in foods, and in disease, but have largely disregarded the bacteria of our intestines. There are still nutritional disturbances which must be related indirectly to the changes in the digestive tract. Very recently, a most interesting report was presented on the nutritional status of about 800 individuals living in Newfoundland: certain symptoms observed were ascribed to riboflavin and niacin deficiencies, and yet a rough estimate indicated that the intake of these vitamins was not seriously inadequate. I believe some of the changes, at least, are due to a lack of as yet unknown vitamins which were not synthesized in sufficient amounts because of the type of dietary regime. Pellagra has always been associated with a large consumption of corn. Preliminary evidence in our laboratory indicates that an extensive corn intake may adversely affect the synthesis of vitamins in the intestinal tract, an effect which is overcome by high levels of nicotinic acid. High levels of protein also have a counteracting effect, which may explain why milk has been found to have antipellagra activity although it is known to contain little nicotinic acid.

Intestinal synthesis is not limited to the production of the vitamins which are the last to be discovered. Obviously, the degree of synthesis in the intestine must be less in the case of the older vitamins, or we would have had more difficulty in producing the deficiency state of these vitamins. Some time ago, we showed that the fat content of the diet had a marked effect on the riboflavin requirement of the rat. Diets which contained dextrin and low levels of riboflavin produced a much more severe riboflavin deficiency when a large portion of the dextrin was isocalorically replaced by fat. Later work has shown that this effect is directly dependent upon a decreased synthesis of riboflavin

in the presence of fat. There is much evidence to show that fat, carbohydrate, protein, and vitamins are all interrelated in their effect on the production of both known and unknown vitamins in the digestive tract of all animals; but the results for one animal cannot be predicted from the results obtained with another species. Just where the human fits into the picture is impossible to say. It is encouraging to find that several groups of workers are engaged in studying this problem on human subjects, a project which will undoubtedly clear up many of the difficulties now encountered in attempting to establish quantitative requirements for each of the vitamins in human subjects. The final answer can probably be made only when animals are rendered bacteria-free and their requirements are studied under these conditions. Although this will answer the questions from an academic point of view, for practical purposes we must continue to recognize the interrelationship of food and intestinal bacteria. No one need feel that further work in the field of vitamins will not be productive. Much human suffering has been alleviated through our knowledge of vitamins and we can expect much success in the future if we learn more about these interesting compounds and if we apply what we learn in a sensible manner.

Selected References

Addinall, C. R., "Synthesis and production of vitamins," *Chem. Eng. News*, **22**, 2174 (1944).

Black, J. D., ed., "Nutrition and food supply: The war and after," *Ann. Am. Acad. Political Social Sci.*, **225** (1943).

"Enrichment of flour and bread. A history of the movement," *Bull. Natl. Research Council*, No. 110 (1944).

Evans, E. A., Jr., ed., *Biological Action of the Vitamins.* Univ. Chicago Press, Chicago, 1942.

Major, R. T., "Industrial development of synthetic vitamins," *Chem. Eng. News*, **20**, 517 (1942).

"Medical survey of nutrition in Newfoundland by a group of investigators," *Can. Med. Assoc. J.*, **52**, 227 (1945).

Rosenberg, H. R., *Chemistry and Physiology of the Vitamins.* Interscience, New York, 1945.

Schultz, T. W., ed., *Food for the World.* Univ. Chicago Press, Chicago, 1945.

Williams, R. R., and Williams, R. J., "Vitamins in the future," *Science*, **95**, 335–344 (1942).

SOME ASPECTS OF VITAMIN RESEARCH

KARL FOLKERS, DIRECTOR OF ORGANIC AND BIOCHEMICAL RESEARCH, MERCK & CO., INC.; AMERICAN CHEMICAL SOCIETY AWARD IN PURE CHEMISTRY; CORECIPIENT OF THE MEAD JOHNSON AWARD

So little is known about the chemistry of vitamins—not a single one has been isolated with absolute certainty—that I have hesitated to include this subject among the applications of organic chemistry. The very extensive contemporary literature on vitamins which takes up much space in journals devoted to biochemistry, contains few chemical facts, and very few that are thoroughly well established.

THIS statement was made in 1928 at Cornell University by George Barger (1) during his lectures on some applications of organic chemistry to biology and medicine. These lectures were concerned with hormones, vitamins, chemical constitution and physiological action, chemotherapy, and blue adsorption compounds of iodine. The number of well-established chemical facts on the chemistry of the vitamins developed so enormously during 1928 to 1945 that an entire university course could justifiably be devoted now to the organic chemical aspects of the vitamins. The companion developments on the biochemistry of the vitamins and on the application of the vitamins in clinical medicine might also require a course each for adequate presentation to students. The industrial production of vitamins on a ton basis, a subject recently reviewed by Major (30), is no less amazing in the rapidity and magnitude of the development.

Today there are so many excellent books and review articles on the chemistry of the vitamins available, that no effort will be made in the following sections to cover any topic completely. Instead, a

few observations, facts, and results which may have unique interest have been selected for comment.

On the Discovery of Vitamins

The discovery of the "major vitamins"* has been based upon observations which related the syndrome of a human disease to constituents of natural materials used in nutrition. The discovery of the "lesser vitamins"* has been based upon observations which related biological reactions generally produced experimentally with animals or microorganisms to constituents of natural materials. Although the "lesser vitamins" are at present not known to correspond to any historically recognized human disease, they probably are essential for the human being. They may be considered "lesser vitamins" today because their absence in human diets is less frequent statistically or because the signs of their absence are not yet fully recognized, as was the case for riboflavin deficiency until 1938, when Sebrell and Butler (43) characterized ariboflavinosis. Oden, Oden, and Sebrell (37) concluded a little later that ariboflavinosis is "a common dietary-deficiency disease in the southern United States."

It seems not unlikely that certain of the "lesser vitamins" will ascend to the class of "major vitamins" after further clinical research. Some of this future clinical research might lie in the borderline fields between biochemistry and psychology, according to R. J. Williams (63). His observations on "personality differences" among animals in nutrition experiments, and the fact that hallucinations and mental symptoms of pellagra are known to be eliminated by administration of nicotinic acid, helped stimulate this interesting thought. Certainly, if other vitamins were found to benefit mental disease or psychological disturbances, the rank of importance of these vitamins would be elevated. R. R. Williams (69) believed it was probable that any

---

* R. R. Williams, in his stimulating address on the occasion of the presentation of the Chandler Medal in 1942, defined the "major vitamins" as thiamin, riboflavin, nicotinic acid, and vitamins A, D, and C. Five of these vitamins are related to ancient and widespread diseases. Ariboflavinosis, which is cured by riboflavin, had been confused with and masked by pellagra and was not recognized *per se* until recently. The "lesser vitamins" include choline, vitamin B_6, pantothenic acid, biotin, inositol, etc. See reference (69).

vitamins yet to be discovered are destined to have lesser nutritional significance for human welfare and that the vitamins which are required to check the nutritional plagues of mankind have already been discovered and produced. Nevertheless, he recognized that there may be exceptions, particularly in the case of obscure diseases. Undoubtedly, chemists and nutritionists must cooperate scientifically for many years on the problems of the discovery of new vitamins.

On the Isolation of Vitamins

The isolation of a vitamin from the natural material in which it exists is essentially a chemical problem of the same nature as the older problems on the isolation of an alkaloid or a glycoside, but with at least two important differences. One of these differences is that vitamins generally occur in quantities amounting to a few parts per million of the natural material, whereas the common alkaloids and glycosides are found frequently in quantities amounting to a few parts per hundred. The greatest difficulties in vitamin isolation might be said to lie in the region of converting the natural materials with a few parts per million of the substance to a concentrate containing a few parts of the substance per hundred. New techniques and new procedures are frequently devised to surmount the difficulties in making such a purification. The isolation of trace substances in milligram or gram quantities requires the processing of hundreds of pounds of the natural material. A second important difference in the isolation procedure is the necessity for countless biological assays throughout the whole isolation work to show the investigator the location (and loss!) of the vitamin in the fractionation.

Pioneering researches on the isolation of a vitamin are very costly and time-consuming. It has been said (69) that the first gram of pure natural thiamin must have cost an aggregate of several hundred thousand dollars. Eight years transpired between the first success in isolating this vitamin in 1926 by Jansen and Donath (21) and the work of Williams, Waterman, and Keresztesy (70) in 1934, which resulted in greatly improved yields, so that a sufficient quantity of the pure vitamin could be made available for its structure determination. It took five years in Kögl's laboratory at the University of Utrecht in Holland to work out the pioneering methods which yielded seventy

milligrams of crystalline biotin (26). The difficult and tedious scheme of fractionation involved a three million-fold purification. Kögl (26) estimated that to produce one gram of their biotin from ordinary yeast would have required 360 tons of the yeast as starting material. Although he found egg yolk to contain ten times as much biotin as yeast, the number of fresh eggs required for the production of one gram of biotin would have cost about $165,000 in 1937.

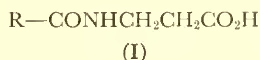
Special chemical steps or new techniques often have to be devised or applied to the isolation process before the vitamin can be obtained in sufficient amounts for the complete elucidation of its structure. The improved yields (70) in the process for the isolation of thiamin depended upon the elution of the vitamin from fuller's earth with quinine acid sulfate instead of barium hydroxide and the introduction of a benzoylation step for purification. Probably one of the most important factors contributing to the success of the isolation of additional quantities of crystalline biotin was the application of the chromatographic adsorption technique by du Vigneaud, Hofmann, Melville, and György (59) to concentrates of the vitamin which had been obtained from beef liver (13). The concentrate contained 0.1% biotin, and, after esterification, the material was chromatographed twice over aluminum oxide. After the final sublimation and crystallization steps, pure crystalline biotin methyl ester was obtained in 38% yield based on the amount of the vitamin in the concentrate.

On the Structure Determination of Vitamins

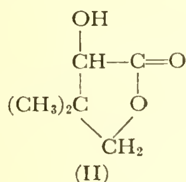
Ordinarily, it is desirable to isolate any unknown natural product in a state of complete purity before the carrying out of chemical reactions for establishment of molecular formula, identification of functional groups, and finally the determination of structure. Special attention to the question of purity is often justified, because natural products are frequently isolated which are extremely difficult to separate from final impurities of unknown but of allied properties.

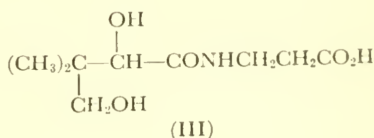
There are, however, exceptions to the purity requirement. The isolation of calcium pantothenate in pure form was found by Williams' group (67) to involve extraordinary difficulties, and it was necessary to conduct the structure studies with a highly purified concentrate estimated to be about 90% pure (34,67). The establishment

of the structure of pantothenic acid under these conditions was accomplished by a rather unique series of developments. β-Alanine was first reported by Williams, Weinstock, and Mitchell (61,68) to be formed from this calcium salt in acid or alkaline medium. The other hydrolytic product was found (34) to be an α-hydroxy acid capable of spontaneous transformation to a lactone. The crude lactone fraction was recombined (66,72) with β-alanine to give material which possessed the physiological activity of pantothenic acid and, of course, actually was this acid. This condensation reaction constituted resynthesis, and the results supported the earlier observation (34) that β-alanine was combined as an amide through its β-amino group, as in structure I. At this stage of the investigation it was clear that it



would be unnecessary to isolate pure pantothenic acid if the hydroxy acid fragment or its lactone could be isolated in pure form. The structure determination of the hydroxy acid would give also the structure of pantothenic acid. Further research (34,53) showed how concentrates containing only 3 to 40% of pantothenic acid could be purified and treated so as to yield the pure lactone. Once having obtained the pure lactone, Stiller, Keresztesy, and Finkelstein (53) applied degradation reactions which showed the lactone to be α-hydroxy-β,β-dimethyl-γ-butyrolactone (II):





Obviously, pantothenic acid was α,γ-dihydroxy-β,β-dimethylbutyryl-β'-alanide (III).

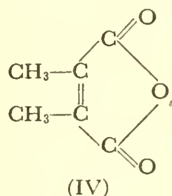
Because of the great cost and technical difficulties in the isolation of new vitamins, it has been necessary on occasion to open the program of structural research with a series of studies which are designed to characterize the functional groups of the substance and which can be

carried out with micro quantities of the substance. For example, early evidence concerning the non-β-alanine portion of pantothenic acid was obtained from the results of a series of new micro procedures (34). Some of these new micro procedures involved determination of active hydrogen atoms with deuterium oxide and of hydroxyl groups with hydriodic acid, selective oxidation with iodic acid, oxidation equivalent analysis, determination of α- and β-hydroxy acids, and estimation of microorganisms in suspension. The use of such techniques represented an unconventional but fruitful approach to the study of pantothenic acid and often required only one or two milligrams of the compound for each determination.

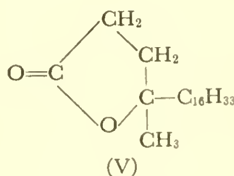
Another type of study has been employed to open a program of structural research upon a costly vitamin available in only very limited quantities. This study, involving a series of "inactivation" experiments, requires only a milligram or less of the substance and yields information on functional groups and constitution. These experiments involve adding to the micro sample of the vitamin the chemical reagent(s) required for a given chemical reaction, such as nitrosation or hydrolysis, and following with a microbiological assay to test whether chemical reaction took place as judged by a change or lack of change of activity. These "inactivation" experiments yield valuable results, but they must nevertheless be interpreted with considerable caution. They do aid in guiding the exploratory efforts in direct chemical studies. An example of such inactivation experiments may be found in certain biotin studies. Brown and du Vigneaud (2) described the effect of certain reagents on the activity of biotin. They obtained this preliminary information with experiments on 1- or 2-cc. aliquots of solution containing only 12.5 γ of biotin per cc., and the criterion of reaction was the effect upon yeast growth activity. Such reagents as 5% hydrogen peroxide solution, aqueous bromine, hydrochloric acid, potassium hydroxide, formaldehyde, and nitrous acid caused "inactivation," indicating that a change in the structure had undoubtedly been brought about by the reagent. On the other hand, such reagents as acetic anhydride–sodium hydroxide, ketene, benzoyl chloride–pyridine, sodium ethoxide–methyl iodide, and ninhydrin caused no "inactivation," indicating that a change in structure had not been brought about by these reagents. The results of these and related experiments indicated to Brown and du Vigneaud that biotin

is inactivated by vigorous treatment with acid or alkali, is not an α-amino acid, is not destroyed by acylating or alkylating reagents, and is easily oxidized.

Not all of the well-known vitamins have had their constitution elucidated by an application and development of micro methods. The correct structure of α-tocopherol of the vitamin E group was determined largely by interpretation of the results of three chemical reactions, which were carried out on 2.1, 4.3, and 25 grams of the vitamin in each case. While repeating some of the isolation work of α-tocopherol from cottonseed oil described by Emerson, Emerson, and Evans (7), the late Dr. E. Fernholz, working in the laboratory adjoining that of the author, intimated that he desired to accumulate enough of the α-tocopherol to permit reactions on a gram-scale basis. It was feasible to isolate α-tocopherol on this scale. Subsequently, Fernholz described (8) the thermal decomposition of α-tocopherol which yielded durohydroquinone, and the details (9) of the experiment reveal that 2.1 g. of the vitamin had been heated for six hours at 355°. The crystalline sublimate yielded 257 mg., or 67%, of durohydroquinone and a hydrocarbon of the composition $C_{18-19}H_{36}$. This characterization of durohydroquinone shattered the frequently discussed idea that α-tocopherol is related to the sterols. The absorption spectra and chemical properties of some synthetic monoethers of durohydroquinone, when considered in conjunction with other suggestive evidence, led to the hypothesis that α-tocopherol was derived from chroman or coumaran. This hypothesis of a heterocyclic ring structure was justified by the results of an oxidation reaction with chromic acid (9). When 4.3 g. of α-tocopherol was dissolved in glacial acetic acid and oxidized with chromic acid, dimethylmaleic anhydride (IV) and



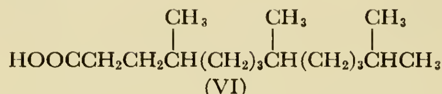
(IV)



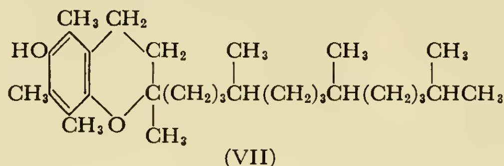
(V)

a lactone, $C_{21}H_{40}O_2$, were isolated. Structure V was proposed for the lactone after a study of its properties and after a study of the oxidation of the acetate of α-tocopherol obtained from 25 g. of α-

tocopheryl allophanate. The chromic acid oxidation of this acetate yielded diacetyl, acetone, an acid $C_{16}H_{32}O_2$ of probable structure VI, and a ketone $C_{18}H_{36}O$.

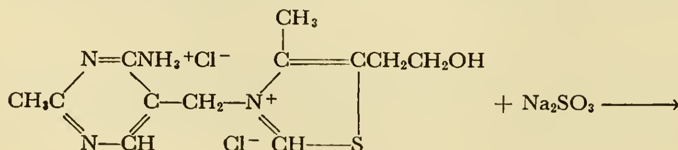


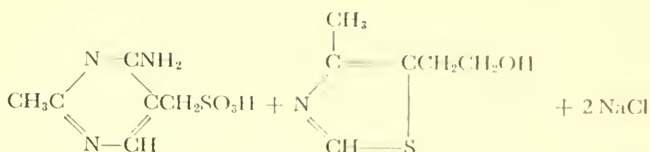
The interpretation of these degradation products and related evidence led to the proposal by Fernholz (9) of structure VII for α-



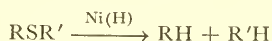
tocopherol. Although the related coumaran structure was considered further by Karrer, Fritzsche, Ringier, and Salomon (22,23) all the results of the investigations—refer to review paper by Smith (44)—were soon interpreted in favor of the chroman structure VII, which was proposed after a study of a few relatively large-scale degradation reactions.

An interesting aspect of the structure studies on thiamin and biotin was the application of new organic structural reactions. It had been observed during the isolation work that an attempt to use sulfurous acid as a preservative against the bacterial decay of extracts of vitamin B_1 from rice polish caused a prompt and complete loss of vitamin activity. This observation led to the development of a procedure by Williams, Waterman, Keresztesy, and Buchman (71) for the quantitative cleavage by sulfite of pure crystalline vitamin B_1 at *p*H 5 and room temperature into two products of the composition $C_6H_9N_3SO_3$ and C_6H_9NSO. There are obvious advantages of this reaction over the usual oxidative and hydrolytic reactions which appeared to give a miscellany of substances. The reaction is expressed in terms of the structural formulas by the following equation:





The new organic structural reaction used in the investigation of biotin was applied in the belief that the results would lead to a final selection between two alternative structures for biotin. It was believed that organic sulfides could be cleaved by the Raney nickel catalyst according to the equation:



$$RSR' \xrightarrow{Ni(H)} RH + R'H$$

This sulfur hydrogenolysis reaction was developed first on several "model" compounds. It was found by Mozingo, Wolf, Harris, and Folkers (35) that representative sulfides could be cleaved to their corresponding sulfur-free products in yields of 65 to 95% on both a macro and semimicro scale. As described by du Vigneaud, Melville, Folkers, Wolf, Mozingo, Keresztesy, and Harris (60), the application of this hydrogenolysis reaction to biotin methyl ester yielded desthiobiotin methyl ester, and the subsequent study of this hydrogenolysis product was the first of two independent methods which led to the final proof of the structure of biotin. It was considered originally (35) that this reaction would be of general value in investigations on the structures of natural products containing sulfur.

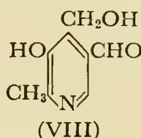
On the Use of Microorganisms in Vitamin Research

It is beyond the scope of this article to present all of the interesting aspects of the application of microorganisms to vitamin research. Williams recently discussed the importance of microorganisms in vitamin research (62) and reviewed microbiological tests (64). György reviewed further developments in the use of microorganisms in vitamin research (12). There are, however, certain recent developments in the chemistry of vitamin B_6 which were originally provoked by the results of microbiological experiments, and recent results on new vitamins which seem to merit comment.

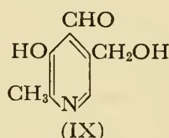
Snell, Guirard, and Williams (50) found that assays with *Streptococcus lactis* gave values for the pyridoxine content of certain

natural materials which were several hundred to several thousand times the values obtained by different biological or chemical methods. Other experiments showed that the factor responsible for the increased activity was very similar in properties to pyridoxine, and they provisionally named the factor "pseudopyridoxine." Subsequent to these studies, Williams stated (65) in his review on the water-soluble vitamins for 1942: "Contrary to the common impression, the chemistry of vitamin B_6 is in a highly unsatisfactory state. There is no question, of course, regarding the fundamental chemistry of pyridoxine, that pyridoxine occurs naturally, or that pyridoxine has vitamin properties. The serious question is whether the vitamin B_6 activity of tissues is due solely to pyridoxine or whether there are other substances (probably closely related) which serve equally well or better."

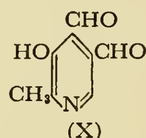
After about two years of further research at the University of Texas on the one or more substances in natural materials tentatively called "pseudopyridoxine," Snell (45) concluded that one of these substances was probably an aldehyde and another was probably an amine. These studies were concerned essentially with the effect of certain selected chemical treatments on the biological activity of pyridoxine for *Streptococcus faecalis* R, *Lactobacillus casei*, and *Saccharomyces cerevisiae*. After consideration of the functional groups of pyridoxine which might be involved in the reactions, it was believed that the biologically active aldehyde would have one of three structures, VIII, IX, or X, and the biologically active amine would have the corresponding structure, XI, XII, or XIII.



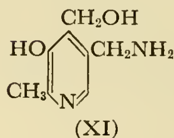
(VIII)



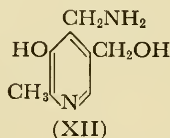
(IX)



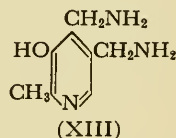
(X)



(XI)



(XII)



(XIII)

Collaborative studies by Harris, Heyl, and Folkers (15) on the structure and synthesis of the active aldehyde and active amine resulted

in the synthesis of aldehydes VIII and IX and amines XI and XII. Biological tests on these synthetic compounds by Snell (46) showed that the biologically active aldehyde was 2-methyl-3-hydroxy-4-formyl-5-hydroxymethylpyridine (IX), and that the biologically active amine was 2-methyl-3-hydroxy-4-aminomethyl-5-hydroxymethylpyridine (see XII). The active aldehyde and amine were given the trivial names "pyridoxal" and "pyridoxamine," respectively. The microbiological assays showed that pyridoxal was about 1400 times more active, and pyridoxamine about 10 times more active, than pyridoxine hydrochloride for promoting the growth of *L. casei*. Pyridoxamine was about 8000 times more active, pyridoxal was about 5500 times more active, in promoting the growth of *S. faecalis* R, than was pyridoxine hydrochloride. The comparative activity of pyridoxal, pyridoxamine, and pyridoxine for *Saccharomyces carlsbergensis* was of the same order of magnitude.

Evidence for the occurrence of pyridoxal and pyridoxamine in natural extracts was secured by Snell (47) by development of a differential microbiological assay technique with the three organisms mentioned above and application of the assay to extracts of natural materials. Further evidence for the existence of pyridoxal and pyridoxamine in nature was secured by studying the effect of certain chemical treatments upon the "pyridoxine, pyridoxal, and pyridoxamine fractions" in comparison with the effect of these treatments upon the synthetic vitamins.

Vitamin B_6 was originally considered to be a single pyridine derivative, pyridoxine. It may now be considered, as a result of these combined microbiological and organic chemical studies, as a name which designates a group of vitamins, *i. e.*, the "vitamin B_6 group." Pyridoxal and pyridoxamine may occupy a place of equal or greater importance in this group as compared with that of pyridoxine.

In retrospect, it is interesting to note that, in the original isolation work, Keresztesy and Stevens (25) and Lepkovsky (29) used rice bran as the source of their vitamin B_6, while Kuhn and Wendt (28) and György (11) used yeast as their source of the crystalline vitamin B_6 (pyridoxine). Snell's microbiological differential assays showed (47) that a rice-bran concentrate contained far more pyridoxine fraction than pyridoxal or pyridoxamine fractions, whereas

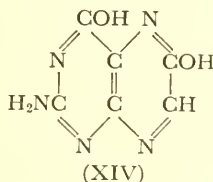
yeast and liver extracts contained an excess of the pyridoxamine fraction, and animal assays showed no marked difference in the activity of the three substances. By assuming that the yeast supply used by Kuhn and Wendt and by György contained an excess of pyridoxamine also, it is evident that at one or more of the steps in the isolation process the pyridoxamine was lost. Williams (65), in commenting on György's communication (11) on isolation, noted that the yield of active substance in the first few steps of the concentration was only 10 to 30% of the original activity. If the isolation of vitamin B_6 from yeast had been guided by the results of microbiological assays with *S. faecalis* R instead of rats, one might predict today that it is quite probable that pyridoxamine would have been isolated instead of pyridoxine and the pyridoxine present would have been lost at some step of the isolation procedure.

In the field of new vitamins of unknown structure, several substances are currently of great interest and papers concerning them are appearing frequently in the literature. Although studies of biological activities in animals are being made, the use of microorganisms for tests of biological activities is resulting in the rapid accumulation of much valuable data on the differentiation of these substances. The following citations may exemplify the importance of the role of microorganisms in the study of these new growth factors.

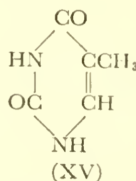
Snell and Peterson (52) and Hutchings, Bohonos, and Peterson (19) have described the preparation and some properties of a concentrate of a norite eluate factor from liver and yeast which resulted from a study of the nutrient requirements of *L. casei* and related lactic acid bacteria. Mitchell, Snell, and Williams (33) reported on the preparation of a highly purified nutrilite from spinach which they designated folic acid and defined as the material responsible for the growth stimulation of *S. lactis* R. Crystalline vitamin B_c from liver was highly active in growth activity for *L. casei* according to Pfiffner, Binkley, Bloom, Brown, Bird, Emmett, Hogan, and O'Dell (39). Stokstad (56) has described some properties of two crystalline preparations, one from liver and one from yeast, which had somewhat different activities for promoting the growth of *L. casei* and *S. lactis* R. Keresztesy, Rickes, and Stokes (24) have reported the isolation of a different factor which was highly active for the growth of *S. lactis* R but relatively inactive for the growth of *L. casei*. Another new compound

which is active for the growth of *L. casei* was described by Hutchings, Stokstad, Bohonos, and Slobodkin (20).

The similarity of the biological activities of these several substances suggests that they also may have a similarity in chemical structure. Concerning this chemical structure, Mitchell (32) presented evidence which showed that the ultraviolet absorption spectrum of folic acid resembled that of xanthopterin (XIV). Addition of large amounts of thymine (XV) was found to substitute (biological pre-



(XIV)



(XV)

cursor?) for folic acid according to Snell and Mitchell (51). Thus, chemical relationships to the pterins and pyrimidines are suggestive possibilities.

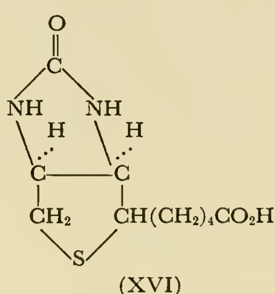
Precise determination of the chemical structure of these several biological factors is probably necessary before the identity of any two or more of them, and before exact similarities in chemical structure, can be established with certainty. It is interesting that much of the microbiological characterization of folic acid, vitamin B_c, and related factors is being developed before the elucidation of their chemical structures. This situation is somewhat the reverse of that for the vitamin B_6 group since much of the chemical characterization of these factors was developed before the clarification of the microbiological aspects.

On the Synthesis of Vitamins

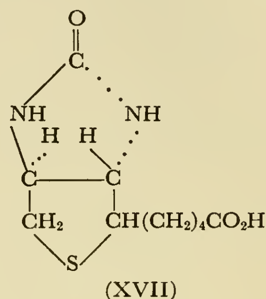
The methods of the organic syntheses of the synthetic vitamins have been amply covered in appropriate review articles. These methods exemplify the adaptation and modification of classical organic laboratory reactions to the synthesis of the desired structure.

The three asymmetric carbon atoms and the two fused five-membered saturated heterocyclic nuclei of biotin present interesting stereochemical features to the studies on the synthesis of this vitamin. There are two racemates which have *cis* forms of the rings and two

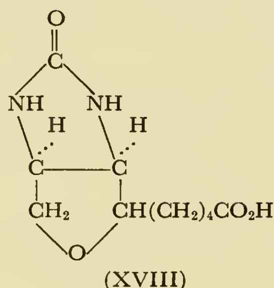
racemates which have *trans* forms of the rings. The *cis* and *trans* relationship of the rings is shown by structures XVI and XVII, re-



(XVI)



(XVII)



(XVIII)

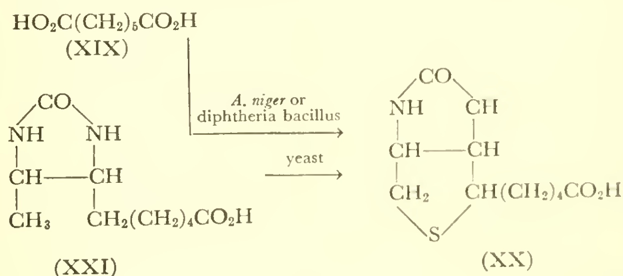
spectively. Biotin is one of the eight stereoisomeric forms. *dl*-Biotin and related *cis* and *trans* forms were obtained by Harris, Mozingo, Wolf, Wilson, Arth, and Folkers (16). These other forms were designated *dl*-allobiotin and *dl*-epiallobiotin. Grussner, Bourquin, and Schnider (10) have described *dl*-pseudo-*β*-biotin and *dl*-iso-*β*-biotin in addition to *dl*-*β*-biotin. Each group may have certain racemates in common. Further work on the comparison of the properties and degradation products of *α*-biotin and *β*-biotin is apparently in progress, according to a recent paper by Kögl and Borg (27).

Studies on synthesis provide interesting information concerning the specificity of structure or the relationship between the constitution and vitamin activity. The specificity of biotin has been discussed in an excellent review article by Melville (31). Recent additions may be made to the subject of specificity of biotin. The described *cis* and *trans* forms related to *dl*-biotin appear to be biologically inactive according to Emerson (6), Ott (38), and Stokes and Gunness (54); see

also Grussner *et al.* (10). An interesting oxygen analogue has been described by Hofmann (17). It has significant biological activity according to Pilgrim, Axelrod, Winnick, and Hofmann (40); an oxygen analogue of about the same melting point was also reported as biologically active by Duschinsky, Dolan, Flower, and Rubin (4). Subsequent studies by Hofmann (18) suggest that *dl*-oxybiotin (XVIII) and biotin have identical spatial configurations and differ only in the nature of one of the hetero atoms.

Biosynthesis of vitamins has received little attention. The pioneering experiments on the synthesis of alkaloids under so-called physiological conditions, as described by Robinson (41), Schöpf (42), and Hahn and Schales (14), involved the isolation of the alkaloid synthesized as the criterion of the success of the experiments. Experiments on biological precursors and on the biosynthesis of vitamins have involved assays with microorganisms, and the production of vitamin activity as the criterion of success. Once again, the utility of microorganisms in vitamin research may be noted. The following examples may serve to illustrate the trends of such vitamin synthesis research.

Mueller's studies (36) on the nutritional requirements of certain pathogenic bacteria led to the isolation of pimelic acid (XIX) from cow's urine and the establishment of pimelic acid as a growth accessory for the diphtheria bacillus. Eakin and Eakin (5) recognized that this growth activity of pimelic acid, when considered in conjunction with the tentative structural formulas for biotin, as published by du Vigneaud, Hofmann, and Melville (58), might mean that pimelic acid was being utilized as a biological precursor in the synthesis of biotin (XX). They selected *Aspergillus niger* as a satisfactory mold to test this biosynthesis and obtained data which demonstrated the activity



of pimelic acid in promoting the biosynthesis of biotin. Cysteine or cystine, as sources of organic sulfur, enhanced the effect of pimelic acid. The studies by du Vigneaud, Dittmer, Hague, and Long (57) on the growth-stimulating effect of biotin for the diphtheria bacillus in the presence and absence of pimelic acid led to the interpretation that pimelic acid was being utilized as a precursor by the diphtheria bacillus for the biosynthesis of biotin.

Experiments described by Dittmer, Melville, and du Vigneaud (3) on the activity of desthiobiotin (XXI) for stimulating the growth of *S. cerevisiae* showed that desthiobiotin disappeared from the incubating yeast cultures and was replaced by an equivalent amount of a substance possessing growth activity for *L. casei.* To these investigators, the most logical interpretation was that desthiobiotin (XXI) was transformed to biotin (XX) by the growing yeast cell. Further support for this interpretation was supplied by the experiments of Stokes and Gunness (54), which showed that, when extracts of yeast grown with desthiobiotin were treated with Raney nickel, the process (60) for converting biotin to desthiobiotin, the activity of the yeast-formed substance for *L. casei* was destroyed and activity for the yeast was retained. Avidin also neutralized the yeast-formed substance, as it does biotin.

Another example of biosynthesis was found by Stokes, Keresztesy, and Foster (55), who reported that the S.L.R. factor was converted by *S. lactis* R into a substance which was active for the growth stimulation of *L. casei.*

The possibility that alanine might be a biological precursor of vitamin B_6 was recognized two years ago by Snell and Guirard (49) when it was found that vitamin B_6 was not required for the growth of *S. faecalis* R if sufficient alanine was added to the medium. Subsequent studies by Snell (48) showed that an enzymic digest of casein contained an unknown biological precursor which, together with *dl*-alanine, permitted the growth of *L. casei* in the absence of vitamin B_6. Since *d*(−)alanine was active and *l*(+)alanine was almost inactive, these data seem to be the first which indicate that the "unnatural" amino acids may be essential for normal metabolic processes.

It is possible that future research on biological precursors and biosyntheses of the vitamins will make these substances available by new methods of startling simplicity.

References

(1) Barger, G., *Some Applications of Organic Chemistry to Biology and Medicine.* McGraw-Hill, New York, 1930, p. 62.

(2) Brown, G. B., and du Vigneaud, V., *J. Biol. Chem.*, **141**, 85 (1941).

(3) Dittmer, K., Melville, D. B., and du Vigneaud, V., *Science*, **99**, 203 (1944).

(4) Duschinsky, R., Dolan, L. A., Flower, D., and Rubin, S. A., *Arch. Biochem.*, **6**, 480 (1945).

(5) Eakin, R. E., and Eakin, E. A., *Science*, **96**, 187 (1942).

(6) Emerson, G., *J. Biol. Chem.*, **157**, 127 (1945).

(7) Emerson, O. H., Emerson, G. A., and Evans, H. M., *Science*, **83**, 421 (1936).

(8) Fernholz, E., *J. Am. Chem. Soc.*, **59**, 1154 (1937).

(9) Fernholz, E., *J. Am. Chem. Soc.*, **60**, 700 (1938).

(10) Grussner, A., Bourquin, J. P., and Schnider, O., *Helv. Chim. Acta*, **28**, 517 (1945).

(11) György, P., *J. Am. Chem. Soc.*, **60**, 983 (1938).

(12) György, P., *Ann. Rev. Biochem.*, **11**, 310 (1942).

(13) György, P., Kuhn, R., and Lederer, E., *J. Biol. Chem.*, **131**, 745 (1939).

(14) Hahn, G., and Schales, O., *Ber.*, **68**, 24 (1935).

(15) Harris, S. A., Heyl, D., and Folkers, K., *J. Biol. Chem.*, **154**, 315 (1944); *J. Am. Chem. Soc.*, **66**, 2088 (1944).

(16) Harris, S. A., Mozingo, R., Wolf, D. E., Wilson, A. N., Arth, G. E., and Folkers, K., *J. Am. Chem. Soc.*, **66**, 1800 (1944).

(17) Hofmann, K., *J. Am. Chem. Soc.*, **67**, 694 (1945).

(18) Hofmann, K., *J. Am. Chem. Soc.*, **67**, 1459 (1945).

(19) Hutchings, B. L., Bohonos, N., and Peterson, W. H., *J. Biol. Chem.*, **141**, 521 (1941).

(20) Hutchings, B. L., Stokstad, E. L. R., Bohonos, N., and Slobodkin, N. H., *Science*, **99**, 371 (1944).

(21) Jansen, B. C. P., and Donath, W. F., *Mededeel. Dienst. Volksgezondheid Nederland-Indië*, **16**, 186 (1926).

(22) Karrer, P., Fritzsche, H., Ringier, B. H., and Salomon, H., *Helv. Chim. Acta*, **21**, 820 (1938).

(23) Karrer, P., Salomon, H., and Fritzsche, H., *Helv. Chim. Acta*, **21**, 309 (1938).

(24) Keresztesy, J. C., Rickes, E. R., and Stokes, J. L., *Science*, **97**, 465 (1943).

(25) Keresztesy, J. C., and Stevens, J. R., *Proc. Soc. Exptl. Biol. Med.*, **38**, 64 (1938); Stiller, E. T., Keresztesy, J. C., and Stevens, J. R., *J. Am. Chem. Soc.*, **61**, 1237 (1939).

(26) Kögl, F., *J. Soc. Chem. Ind.*, **57**, 49 (1938).
(27) Kögl, F., and Borg, W. A. J., *Z. physiol. Chem.*, **281**, 65 (1944).
(28) Kuhn, R., and Wendt, G., *Ber.*, **71**, 780, 1118 (1938).
(29) Lepkovsky, S., *Science*, **87**, 169 (1938); *J. Biol. Chem.*, **124**, 123 (1938).
(30) Major, R. T., *Chem. Eng. News*, **20**, 517 (1942).
(31) Melville, D. B., in *Vitamins and Hormones.* Vol. II, Academic Press, New York, 1944, p. 29.
(32) Mitchell, H. K., *J. Am. Chem. Soc.*, **66**, 274 (1944).
(33) Mitchell, H. K., Snell, E. E., and Williams, R. J., *J. Am. Chem. Soc.*, **63**, 2284 (1941); **66**, 267 (1944).
(34) Mitchell, H. K., Weinstock, H. H., Jr., Snell, E. E., Stanbery, S. R., and Williams, R. J., *J. Am. Chem. Soc.*, **62**, 1776 (1940).
(35) Mozingo, R., Wolf, D. E., Harris, S. A., and Folkers, K., *J. Am. Chem. Soc.*, **65**, 1013 (1943).
(36) Mueller, J. H., *J. Biol. Chem.*, **119**, 121 (1937).
(37) Oden, J. W., Oden, L. H., and Sebrell, W. H., *U. S. Pub. Health Repts.*, **54**, 790 (1939).
(38) Ott, W. H., *J. Biol. Chem.*, **157**, 131 (1945).
(39) Pfiffner, J. J., Binkley, S. B., Bloom, E. S., Brown, R. A., Bird, O. D., Emmett, A. D., Hogan, A. G., and O'Dell, B. L., *Science*, **97**, 404 (1943).
(40) Pilgrim, F. J., Axelrod, A. E., Winnick, T., and Hofmann, K., *Science*, **102**, 35 (1945).
(41) Robinson, R., *J. Chem. Soc.*, **111**, 876 (1917).
(42) Schöpf, C., *Ann.*, **497**, 1 (1932).
(43) Sebrell, W. H., and Butler, R. E., *U. S. Pub. Health Repts.*, **53**, 2282 (1938).
(44) Smith, L. I., *Chem. Revs.*, **27**, 287 (1940).
(45) Snell, E. E., *J. Am. Chem. Soc.*, **66**, 2082 (1944).
(46) Snell, E. E., *J. Biol. Chem.*, **154**, 313 (1944); **157**, 475 (1945).
(47) Snell, E. E., *J. Biol. Chem.*, **157**, 491 (1945).
(48) Snell, E. E., *J. Biol. Chem.*, **158**, 497 (1945).
(49) Snell, E. E., and Guirard, B. M., *Proc. Natl. Acad. Sci. U. S.*, **29**, 66 (1943).
(50) Snell, E. E., Guirard, B. M., and Williams, R. J., *J. Biol. Chem.*, **143**, 519 (1942).
(51) Snell, E. E., and Mitchell, H. K., *Proc. Natl. Acad. Sci. U. S.*, **27**, 1–7 (1941).
(52) Snell, E. E., and Peterson, W. H., *J. Bact.*, **39**, 273 (1940).
(53) Stiller, E. T., Keresztesy, J. C., and Finkelstein, J., *J. Am. Chem. Soc.*, **62**, 1779 (1940).
(54) Stokes, J. L., and Gunness, M., *J. Biol. Chem.*, **157**, 121 (1945).

(55) Stokes, J. L., Keresztesy, J. C., and Foster, J. W., *Science*, **100**, 522 (1944).
(56) Stokstad, E. L. R., *J. Biol. Chem.*, **149**, 573 (1943).
(57) du Vigneaud, V., Dittmer, K., Hague, E., and Long, B., *Science*, 96, 186 (1942).
(58) du Vigneaud, V., Hofmann, K., and Melville, D. B., *J. Am. Chem. Soc.*, **64**, 188 (1942).
(59) du Vigneaud, V., Hofmann, K., Melville, D. B., and György, P., *J. Biol. Chem.*, **140**, 643 (1941).
(60) du Vigneaud, V., Melville, D. B., Folkers, K., Wolf, D. E., Mozingo, R., Keresztesy, J. C., and Harris, S. A., *J. Biol. Chem.*, **146**, 475 (1942).
(61) Weinstock, H. H., Jr., Mitchell, H. K., Pratt, E. F., and Williams, R. J., *J. Am. Chem. Soc.*, **61**, 1421 (1939).
(62) Williams, R. J., *Science*, **93**, 412 (1941).
(63) Williams, R. J., *Science*, **95**, 340 (1942).
(64) Williams, R. J., *Ann. Rev. Biochem.*, **12**, 309 (1943).
(65) Wiliams, R. J., *Ann. Rev. Biochem.*, **12**, 331 (1943).
(66) Williams, R. J., Mitchell, H. K., Weinstock, H. H., Jr., and Snell, E. E., *J. Am. Chem. Soc.*, **62**, 1784 (1940).
(67) Williams, R. J., Truesdail, J. H., Weinstock, H. H., Jr., Rohrmann, E., Lyman, C. M., and McBurney, C. H., *J. Am. Chem. Soc.*, **60**, 2719 (1938).
(68) Williams, R. J., Weinstock, H. H., Jr., and Mitchell, H. K., *Abstracts, 96th Meeting, Am. Chem. Soc., Sept., 1938*, Division of Organic Chemistry, p. 34.
(69) Williams, R. R., *Science*, **95**, 335 (1942).
(70) Williams, R. R., Waterman, R. E., and Keresztesy, J. C., *J. Am. Chem. Soc.*, **56**, 1187 (1934).
(71) Williams, R. R., Waterman, R. E., Keresztesy, J. C., and Buchman, E. R., *J. Am. Chem. Soc.*, **57**, 536 (1935).
(72) Woolley, D. W., Waisman, H. A., and Elvehjem, C. A., *J. Am. Chem. Soc.*, **61**, 977 (1939); *J. Biol. Chem.*, **129**, 673 (1939).

QUANTITATIVE ANALYSIS IN BIOCHEMISTRY

DONALD D. VAN SLYKE, MEMBER OF THE ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH, NEW YORK; WILLARD GIBBS MEDALIST

THE BIOCHEMISTRY of today is based to a large extent on quantitative analyses by means of micro methods. It has not only applied the methods developed in laboratories of organic and inorganic chemistry, but has also rapidly developed new procedures to meet the demands of its expanding range of research. While Pregl's system of elementary organic microanalysis, published in 1911, was gaining application, Folin and Wu in 1919 and Bang in 1916 published quite different systems adapted to blood analyses. The ultra in microanalyses is represented by the methods of capillary colorimetry developed by A. N. Richards and his collaborators for the analysis of glomerular urine, and the extraordinary combination of physical and chemical procedures applied by Linderstrøm-Lang in his studies of the enzymes of cells.

In a brief survey of the field it will be possible only to mention some of the different types of analytical procedure that have been applied in biochemistry, with examples of a few applications, and references to reviews in which descriptions and bibliographies can be found.

Gravimetric Analysis

The appearance in 1911 of the Kuhlman micro balance, and of Pregl's system of micro elementary analyses decreased the size of

samples usually taken for analyses from 100–150 mg. to 4 or 5 mg., and opened a new epoch, not only in organic analysis, but also in organic chemistry. For it became possible to solve problems when amounts of material were available that would have been inadequate even for preliminary analyses by the old macro methods. It is questionable, for example, that the rapid development of knowledge concerning the structure of the vitamins would have been possible without these micro methods. The techniques of micro analysis developed in Austria by Pregl and by Emich (1) were brought to this country, especially by J. B. and V. Niederl (3), and are available in a recent volume by these authors. For ultra micro gravimetric work, Lowry (2) has recently described a simple quartz fiber balance which will weigh 200 gammas to 0.03 gamma.

Volumetric Analysis

Bang's work (4) introduced micro titrations into biochemistry, utilizing the principle of keeping the volume of titrated liquid small, to maintain the sharpness of the end point. Rehberg (9) followed shortly with his micro burette, in which a thread of standard solution in a calibrated capillary of about 1 mm. bore is expelled into the titrated solution by mercury moved by a plunger advanced by turning a steel screw. With this burette, the solution could be measured to within ±0.1 cu. mm., so that 50 cu. mm. was ample for a good titration. Linderstrøm-Lang (7) moved the error down to ±0.01 cu. mm. He used still narrower capillaries, and stirred the titrated solution in a minute glass thimble by a magnetic stirrer consisting of a bit of iron, enclosed in a glass droplet, which was lifted and lowered in the titrated solution by the automatic opening and closing of the circuit to a magnet. Scholander (10,11) has applied the metal micrometer screw that is readily available because of its general use in accurate industrial mechanical work. The column of standard solution in a fine capillary is moved, as in the Rehberg burette, by pressure of a mercury column advanced by a screw and plunger; but in the Scholander burette the extent to which the screw is turned serves as a measure of the volume of solution delivered, and calibration of the capillary is not necessary.

A limitation of the Rehberg and Scholander burettes is that they cannot be used to deliver solutions which react with mercury. For

such solutions, Longwell and Hill (8) have modified the Rehberg burette by introducing an elastic rubber diaphragm between the mercury and the solution. Clark, Levitan, Gleason, and Greenberg (5) have applied Scholander's micrometer principle, but employ the micrometer screw to push air (instead of mercury) from a hypodermic syringe into a capillary which delivers the solution.

The general principles of volumetric microanalyses have been clearly elucidated by Conway (6).

Gasometric Analysis

One of the oldest quantitative analyses in biochemistry is the determination of urea by measurement of the nitrogen gas liberated by reaction with alkaline bromine solution. It exemplifies procedures in which a substance is measured by the amount of gas that it liberates when it reacts with properly chosen reagents. In such analyses, the measurement is based, as in gravimetric methods, on direct observation of the amount of substance obtained, independent of comparison with standard solutions, such as are required in titration and colorimetry. Combined with this independence are the advantages of a quick measurement and easy adaptation to micro quantities. Historically, micro gasometric procedures were introduced into biochemistry for determination of the blood gases, and were then adapted to more general analyses.

The first of such procedures was the blood gas analysis of Barcroft and Haldane (12) in which the oxygen liberated from blood by ferricyanide, or the carbon dioxide liberated by acid, was measured by the gas displaced into a capillary tube. The procedure requires accurate temperature control and constant shaking until equilibrium between dissolved and supernatant gases is reached. The method was elaborated by Warburg (19), and has been used for a great variety of purposes by Warburg and others, particularly in following the course of enzymic reactions by measurement of the oxygen absorbed or the carbon dioxide evolved. The ease with which the course of a reaction can be followed by observing the increase in gas volume particularly adapts the procedure to the observation of comparative reaction velocities (13). A recent refinement of the apparatus, and the principles of its use, are described by Summerson (15).

The manometric apparatus of Van Slyke and Neill (18), like the Barcroft-Haldane apparatus, was first developed for determination of the blood gases, and its use then spread to other micro analyses, including the determination of urea, reducing and fermentable sugars, the ammonia yielded by Kjeldahl digestions, amino nitrogen by measurement of the nitrogen yielded by reaction with nitrous acid (18), free alpha-amino acids by measurement of the carbon dioxide yielded by reaction with ninhydrin [$RCH(NH_2)COOH \rightarrow RCHO + CO_2 + NH_3$] (16), organic carbon by measurement of the carbon dioxide evolved by a wet combustion completed in two minutes (17), and various other determinations. In these procedures, the gases are either evolved in, or transferred to, a 50-cc. chamber, provided at the top with small bulbs for measuring 0.5 and 2.0 cc. of gas, and connected at the bottom with a mercury manometer. The evolved gas is brought to 0.5 or 2.0 cc. volume, and its pressure is read on the manometer. In a mixture of gases, each gas can be measured separately by measuring the pressure before and after the absorption of each gas by introduction of a proper reagent. The carbon combustion method permits micro determination of any organic substance, such as the blood fats, that can be isolated by extraction with volatile solvents; the combustion also provides a micro measurement of any substance that can be isolated as a carbon-containing precipitate, *e. g.*, sulfate as benzidine sulfate, magnesium as hydroxyquinolate, phosphorus as strychnine phosphomolybdate (14).

While the Barcroft-Haldane-Warburg apparatus is adapted to following the course of time reactions, the Van Slyke-Neill apparatus is fitted for quick determination of the total amounts of gases evolved by rapid quantitative reactions. Hence the Haldane apparatus has found its chief application in following enzymic time reactions, while the Van Slyke-Neill apparatus is the one usually employed in quantitative micro determinations of specific substances.

Photometric Analysis

Under the chief initial stimulus of Folin and of S. R. Benedict, during the past forty years chromogenic reactions have been developed for estimating numerous biological substances by producing colored products from them (25). In some cases the colored products are

defined; in other cases neither the reactions nor the colored products are defined with certainty; but empirical conditions have been fixed which relate the color quantitatively to the amount of the substance under analysis. Thus Folin (21) used the color produced by Nessler's reagent as the means for quantitative estimation of ammonia, and hence of nitrogen, through the ammonia obtained by Kjeldahl digestion, and of urea through the ammonia obtained by urea hydrolysis. The constitution of the colored compound of ammonia and potassium mercuri-iodide is still under dispute, but the colorimetric results are accurate. Sugars reducing Cu^{++} to Cu^{+} were determined by Folin and Wu (21) by letting the cuprous ion thus formed act on a molybdate solution, with reduction of the colorless hexavalent molybdenum to a lower valence, which shows an intense blue color; the reactions do not appear to be stoichiometric, but quantitative relations can be obtained between colored molybdenum products and the initial sugar. For almost every substance of interest in quantitative biochemical analysis, chromogenic reactions have been devised which can be used for more or less accurate estimation. As a rule these procedures are rapid, and are adapted to minute amounts of material.

During the first years of the colorimetric epoch, the instrument in general use was the familiar Duboscq colorimeter, in which the depths of colored solution layers in two parallel columns, one of the unknown solution and the other of a standard, are varied until the two fields viewed with the eye appear equal. The simplicity and versatility of the Duboscq colorimeter assisted greatly in the rapid adoption of colorimetric procedures.

Photometers, in which the percentage transmission of light could be measured without standard solutions for direct comparison, were known long before this period, but were too complicated and expensive for ordinary routine in biochemical laboratories. During the past two decades, however, photometers have been progressively made more adaptable to such routine, and have been gradually displacing colorimeters of the Duboscq type. In the photometer, the concentration of light-absorbing solute is related to the optical density according to the simple linear formula of Beer's law (20,22,23):

$$C = k D \qquad (1)$$

where C is the concentration and D is the optical density (or extinction).

D is the logarithm of $1/T$, T being the fraction of light transmitted by the substance measured.

The value of k for a given colored solute varies with the wave length of light. Hence Beer's law holds exactly only for monochromatic light; and the accuracy with which the formula applies to a given photometer depends partly on how narrow a spectral band can be given by the analyzing device interposed between the source of light and the solution, a limitation under which the Duboscq colorimeter does not suffer. Some solutes do not exactly follow Beer's law, even with the narrowest spectral bands. Most solutions, however, do follow the law over the concentration ranges used for analysis, and with the spectral bands provided by the instruments now available for routine analytical work. Solutions for which Beer's law does not hold can be analyzed by using empirical calculation curves of optical density *vs.* concentration.

The photometer has several advantages over the Duboscq colorimeter. The validity of equation (1) makes it possible to measure the concentration of one colored solute in the presence of others, since the total optical density is additive:

$$D = C_1/k_1 + C_2/k_2 \ldots \tag{2}$$

Hence, if the medium in which the concentration of a solute is to be measured is itself colored or turbid, the *increase* in optical density due to the presence of the specific solute can be used as a measure of its concentration. Correction for nonspecific color is much less simple in a Duboscq colorimeter.

Because the k value for each solute changes with the wave length, it is possible to determine two colored solutes in the same solution by measuring the optical densities at two different wave lengths. Two densities, D_a and D_b, are thus measured:

$$D_a = C_1/k_1 + C_2/k_2 \tag{2a}$$

$$D_b = C_1/k_3 + C_2/k_4 \tag{2b}$$

If k_1, k_2, k_3, and k_4 are known, C_1 and C_2 can be calculated by simultaneous equations from the observed D_a and D_b.

The concentration of a turbid suspension can be estimated from its optical density, in the same way as the concentration of a colored solute.

The eye strain and subjective error accompanying the use of a visual instrument are obviated in most of the modern photometers by using photocells to measure the intensity of the transmitted light. The rapidity with which a series of observations can be carried out is also much greater when the eye is replaced by the photocell.

A further advantage of the photometer is that, with photocell measuring devices, it can be used with light waves extending into the ranges of the ultraviolet and infrared, broadening the range of accessible analyses.

In the ultra micro colorimetric methods which Richards and his collaborators (24) developed for analyses of samples of 1 cu. mm. of glomerular filtrate, the sample is drawn into a small glass capillary in which it is mixed with chromogenic reagent, and the color is estimated by comparing the capillary under a microscope with a series of standards similarly prepared. The comparison with graded standards is a return to simplest first principles, but the technique of the application gave results to 1%.

Fluorimetric Analysis

The diffuse fluorescent light developed by passing light rays into solutions of fluorescent substances (28) can be measured by the same visual or electrical means employed in colorimeters and photometers, and serves in some cases, such as determination of riboflavin (27), quinine, atabrine (26), and related compounds, to measure substances in more dilute solutions than can be handled with a photometer. In fluorimetry, the concentration of fluorescent substance is directly proportional to the intensity of the light measured, instead of being inversely proportional to the log of the transmission, as in photometry.

Polarographic Analysis

This procedure, which has gained rapid utility during the past few years, was introduced by Heyrovský in Prague in 1925, and has been applied to a multiplicity of analyses of substances, both organic and inorganic. It is based on measurement of the amperages obtained at observed voltages applied to solutions of electroreducible or electro-

oxidizable substances in a cell in which one electrode consists of mercury falling from a fine capillary.

When a current at gradually increasing voltage is passed to a small mercury electrode through a solution containing a solute capable of giving or receiving electrons ("redox solute") at a given potential, relatively little current is obtained until the decomposition potential of the redox solute is reached. Then the current rapidly rises, with further increase in voltage until a plateau is reached at which the redox solute is providing its maximal flow of electrons. This flow, and the resultant current, are proportional to the concentration of the redox solute, which determines the rate at which its molecules diffuse to the electrode and discharge or receive electrons. Since diffusion is a process independent of the voltage, increase of voltage above that at which electrolytic decomposition of the redox solute equals its rate of diffusion to the mercury electrode does not further increase the flow of electricity; the concentration of the active solute thus forms a bottleneck which limits the current. The curve of current *vs.* voltage then reaches a plateau, the height of which, in milliamperes, is proportional to the concentration of active redox solute (32–34).

Maintenance of the proportionality requires a continually renewed surface of the mercury electrode; both renewal of the surface and setting its area at a small size are obtained by using mercury dropping from a fine capillary, of about 0.03 mm. diameter, as the electrode; a drop is delivered about once in two to four seconds.

The current fluctuates somewhat as each drop of mercury expands and falls, but the average current, i, in microamperes is given by the "Ilkovič equation":

$$i = 605\, n\, D^{1/2}\, C\, m^{2/3}\, t^{1/6} \qquad (3)$$

C is the millimoles of redox solute per liter, n is the number of electrons exchanged per molecule of redox solute in the electrolytic decomposition, m is the weight of mercury flowing from the capillary per second, and t is the time required for formation of one drop of mercury. The procedure is adapted to analysis of highly dilute solutions, 0.001 molar and lower concentrations.

Measurement of the current in such a system provides a measure of the concentration of the redox solute. Furthermore, if two different redox solutes are present with different decomposition potentials, they

can be determined one after another by using voltages related to their respective decomposition potentials.

The setup also can be used for electrometric titrations. If a reagent which reacts with the redox solute is added while the current is measured at proper voltage, a drop in current will accompany the disappearance of the redox solute, and the end point will be indicated when the current has fallen to a residual value, representing conductance by factors other than the redox solute.

For literature, theoretical discussion, description of the different types of analysis to which the procedure has already been applied, and the precautions that must be observed, the reader is referred to the bibliography (35), in particular to Kolthoff (32,33) and Müller (34). The various inorganic cations and anions are determinable; also reducible organic compounds, such as aldehydes, ketones, and nitro compounds. Eisenbrand and Picher (31) found that the sex hormones with the O═C—C═C— group are reducible at the mercury electrode and can be determined polarographically: these hormones include testosterone, progesterone, and desoxycorticosterone, but not androsterone nor dehydroandrosterone. Application to solve a hitherto difficult biochemical problem is illustrated by the work of Berggren (30) and of Beecher *et al.* (29) in determining the oxygen tension of arterial blood plasma and other body fluids. These authors also describe their apparatus in detail.

Spectrographic Analysis

Measurement of the intensity of light of characteristic wave length emitted by incandescent elements has been long used in both qualitative and quantitative analysis, and in special problems for quantitative or semiquantitative estimation of minute amounts of specific elements. Lundegårdh (39) in 1929 applied the principle in a manner which makes it applicable to micro determination of mineral bases in biological fluids. The solution is sprayed from an atomizer at a constant rate, and the stream of air with suspended fluid is mixed with acetylene, which is burned, heating the mineral bases in the suspension to such a temperature that they emit their characteristic light bands, the intensity of which is measured by an electrophotometer (37). The procedure is reviewed by Ells (38) and by

Cholak and Hubbard (36). The latter describe the application of the procedure to determination of minute amounts of cadmium in blood and urine, and compare it with polarographic and photometric procedures for this purpose. An apparatus devised by the American Cyanamid Company, not yet in general production, determines potassium in serum in a few minutes with an error not over $\pm 5\%$.

Micro Diffusion Analysis

Conway (41) has devised a simple chamber for the determination, primarily, of ammonia, but applicable also to estimation of other volatile substances that can be set free by quantitative reactions and transferred by diffusion at atmospheric pressure to absorbing solutions in which the diffused substance can be measured, by titration, photometry, or otherwise. The apparatus consists of a flat, cylindrical dish, of 60-mm. diameter and 10 mm. high (inner measurements), from the inner bottom of which rises a ring of 33-mm. diameter and 5 mm. high. The top of the dish is ground accurately flat, so that when lubricated with vaseline or other proper material it can be closed gastight by a flat glass cover. The chamber consists, therefore, of an outer ring, about 60 mm. wide, surrounding an inner low cylinder; when the chamber is covered a free space of 5 mm. is left open between the covering plate and the wall about the inner cylinder, permitting free diffusion of volatile substances from the outer compartment to the inner. To determine ammonia, the solution containing it is alkalinized in the outer compartment, and acid is placed in the inner compartment. In one or more hours, depending on temperature and the other conditions, the ammonia from the former diffuses through the air space of the chamber into the acid, where it can be measured in amounts of a few micrograms. Conway applied this procedure to determination of the minute amounts of ammonia in blood, to estimation of the ammonia formed by micro Kjeldahl nitrogen digestion, and of the ammonia formed by decomposition of urea with urease; also to chloride and bromide, which were oxidized to chlorine and bromine and diffused into potassium iodide solution for iodometric titration. Conway also applied the diffusion chamber to the determination of carbon dioxide, which was caused to diffuse into a barium hydroxide solution, where the excess alkali was titrated. Borsook (40)

developed applications of the ammonia procedures. Winnick (42) applied the diffusion apparatus to the determination of: alcohol, with chromic acid in the inner chamber; lactic acid, which was oxidized in the outer chamber to acetaldehyde, the latter diffusing to a bisulfite solution in the inner chamber; acetone, which diffused to bisulfite; threonine, which was oxidized by periodate to acetaldehyde in the outer chamber, with diffusion of the aldehyde to bisulfite.

References

GRAVIMETRIC ANALYSIS

(1) Emich, F., *Microchemical Laboratory Manual.* Trans. by F. Schneider. Wiley, New York, 1932.

(2) Lowry, O. H., "A quartz fiber balance," *J. Biol. Chem.*, **140**, 183 (1941). "A simple quartz torsion balance," *ibid.*, **152**, 293 (1944).

(3) Niederl, J. B., and Niederl, V., *Micromethods of Quantitative Organic Analysis.* Wiley, New York, 1942.

VOLUMETRIC ANALYSIS

(4) Bang, I., *Methoden zur Mikrobestimmung einiger Blutbestandteile.* Bergmann, Wiesbaden, 1916.

(5) Clark, W. G., Levitan, N. I., Gleason, D. F., and Greenberg, G., "Titrimetric microdetermination of chloride, sodium, and potassium in a single tissue or blood sample," *J. Biol. Chem.*, **145**, 85 (1942).

(6) Conway, E. S., *Micro-diffusion Analysis and Volumetric Error.* Van Nostrand, New York, 1942.

(7) Linderstrøm-Lang, K., "Distribution of enzymes in tissues and cells," *Harvey Lectures*, **34**, 214 (1939).

(8) Longwell, B., and Hill, R. M., "A modified Rehberg burette for use with titrating solutions which react with mercury," *J. Biol. Chem.*, **112**, 319 (1935).

(9) Rehberg, P. B., "A method of microtitration," *Biochem. J.*, **19**, 270 (1925).

(10) Scholander, P. F., "Microburette," *Science*, **95**, 177 (1942).

(11) Scholander, P. F., and Edwards, G. A., and Irving, L., "Improved microburette," *J. Biol. Chem.*, **148**, 495 (1943).

GASOMETRIC ANALYSIS

(12) Barcroft, J., and Haldane, J. S., "A method of estimating the oxygen and carbonic acid in small quantities of blood," *J. Physiol.*, **28**, 232 (1902).

(13) Dixon, M., *Manometric Methods.* Cambridge Univ. Press, London, 1934.

(14) Hoagland, C. L., "Microdetermination of sulfate and phosphate by manometric combustion of their organic precipitates," *J. Biol. Chem.*, **136,** 543 (1940). "Micromanometric determination of magnesium," *ibid.*, 553.

(15) Summerson, W. H., "A combination simple manometer and constant differential manometer for studies in metabolism," *J. Biol. Chem.*, **131,** 579 (1939).

(16) Van Slyke, D. D., Dillon, R. T., MacFadyen, D. A., and Hamilton, P. B., "Gasometric determination of carboxyl groups in free amino acids," *J. Biol. Chem.*, **141,** 627 (1941). Hamilton, P. B., and Van Slyke, D. D., "The gasometric determination of free amino acids in blood filtrates by the ninhydrin–carbon dioxide method," *ibid.*, **150,** 231 (1943).

(17) Van Slyke, D. D., and Folch, J., "Manometric carbon determination," *J. Biol. Chem.*, **136,** 509 (1940).

(18) Van Slyke, D. D., and Neill, J. M., "The determination of gases in blood and other solutions by vacuum extraction and manometric measurement," *J. Biol. Chem.*, **61,** 523 (1924); **83,** 449 (1929). "Applications to other analyses," in Peters, J. P., and Van Slyke, D. D., *Quantitative Clinical Chemistry. Methods.* Williams & Wilkins, Baltimore, 1932; rev., 1943.

(19) Warburg, O., "Verbesserte Methode zur Messung der Atmung und Glykolyse," *Biochem. Z.*, **152,** 51 (1924).

PHOTOMETRIC ANALYSIS

(20) Ashley, S. E. Q., "Spectrophotometric methods in modern analytical chemistry," *Ind. Eng. Chem., Anal. Ed.*, **11,** 72 (1939).

(21) Folin, O., and Wu, H., "A system of blood analysis," *J. Biol. Chem.*, **38,** 81 (1919).

(22) Hamilton, R. H., "Photoelectric photometry. An analysis of errors at high and low absorption," *Ind. Eng. Chem., Anal. Ed.*, **16,** 123 (1944).

(23) Müller, Ralph H., "Photoelectric methods in analytical chemistry," *Ind. Eng. Chem., Anal. Ed.*, **7,** 223 (1935); **11,** 1 (1939).

(24) Richards, A. N., Bordley, J., 3rd, and Walker, A. M., *J. Biol. Chem.*, **101,** 179, 193, 223, 229 (1933).

(25) Snell, F. D., and Snell, C. T., *Colorimetric Methods of Analysis, including Some Turbidimetric and Nephelometric Methods.* 2nd ed., 2 vols., Van Nostrand, New York, 1936–1937.

FLUORIMETRIC ANALYSIS

(26) Brodie, B. B., and Udenfriend, S., "The estimation of atabrine in biological fluids," *J. Biol. Chem.*, **151,** 299 (1943).

(27) Hand, D. B., "Determination of riboflavine in milk by photoelectric fluorescence measurements," *Ind. Eng. Chem., Anal. Ed.*, **11**, 306 (1939).

(28) Kavanagh, F., "New photoelectric fluorimeter and some applications," *Ind. Eng. Chem., Anal. Ed.*, **13**, 108 (1941).

POLAROGRAPHIC ANALYSIS

(29) Beecher, H. K., Follansbee, R., Murphy, A. J., and Craig, F. N., "Determination of the oxygen content of small quantities of body fluids by polarographic analysis," *J. Biol. Chem.*, **146**, 197 (1942).

(30) Berggren, S. M., "The oxygen deficit of arterial blood caused by non-ventilating parts of the lung," *Acta Physiol. Scand. Suppl.*, **4**, XI (1942).

(31) Eisenbrand, J., and Picher, H., "Über den polarographischen Nachweis von biologisch wichtigen Ketonen der Steringruppe," *Z. physiol. Chem.*, **260**, 83 (1939).

(32) Kolthoff, I. M., "Factors to be considered in quantitative polarography," *Ind. Eng. Chem., Anal. Ed.*, **14**, 195 (1942).

(33) Kolthoff, I. M., and Lingane, J. J., *Polarography.* Interscience, New York, 1941.

(34) Müller, O. H., *The Polarographic Method of Analysis.* J. Chem. Education, Easton, 1941.

(35) Sand, H. J. S., *Bibliography of the Dropping-Mercury Electrode.* Leeds & Northrup, Philadelphia, 1941.

SPECTROGRAPHIC ANALYSIS

(36) Cholak, J., and Hubbard, D. M., "Spectrochemical analysis with the air–acetylene flame," *Ind. Eng. Chem., Anal. Ed.*, **16**, 728 (1944).

(37) Churchill, J. R., "Techniques of quantitative spectrographic analysis," *Ind. Eng. Chem., Anal. Ed.*, **16**, 653 (1944).

(38) Ells, V. R., "The Lundegårdh flame method of spectrographic analysis," *J. Optical Soc. Am.*, **31**, 534 (1941).

(39) Lundegårdh, H., *Die Quantitative Spektralanalyse der Elemente.* Fischer, Jena. Part I, 1929; Part II, 1934.

MICRO DIFFUSION ANALYSIS

(40) Borsook, H., "Micromethods for determination of ammonia, urea, and total nitrogen," *J. Biol. Chem.*, **110**, 481 (1935).

(41) Conway, E. J., *Micro-diffusion Analysis and Volumetric Error.* Van Nostrand, New York, 1940.

(42) Winnick, T., "Micro-diffusion methods. Alcohol," *Ind. Eng. Chem., Anal. Ed.*, **14**, 523 (1942). "Acetone," *J. Biol. Chem.*, **141**, 115 (1941). "Lactic acid," *ibid.*, **142**, 451 (1942). "Threonine," *ibid.*, **142**, 461 (1942).

9

ENZYMIC HYDROLYSIS AND SYNTHESIS OF PEPTIDE BONDS

JOSEPH S. FRUTON, ASSOCIATE PROFESSOR OF PHYSIOLOGICAL CHEMISTRY, YALE UNIVERSITY; LILLY AWARD IN BIOLOGICAL CHEMISTRY

Specificity of Proteolytic Enzymes

THE ACTION of proteolytic enzymes on peptide linkages involves a high degree of specificity. No proteolytic enzyme acts on peptide bonds indiscriminately, and each enzyme hydrolyzes only such peptide bonds as are present in the substrate in a certain structural setting. Thus, the nature of the requisite structural attributes of the substrate is an expression of the specificity of the enzyme which hydrolyzes the substrate. In recent years, much attention has been devoted to the determination of those structural elements in the substrate molecule which are essential for the action of various proteolytic enzymes. These studies have permitted the formulation of several hypotheses concerning the specific action of the proteolytic enzymes.

Modern theories concerning the specificity of proteolytic enzymes are based on the assumption, made by von Euler and Josephsohn (17) in their "dual-affinity" theory, that ereptic peptidase—it was not recognized at that time that "erepsin" represents a mixture of many peptidases—combines with two atomic groupings of the substrate molecule. Subsequent work of Balls and Köhler (2) presented

further evidence for the "dual-affinity" hypothesis. More recently, the finding of synthetic substrates for the protein-splitting enzymes (pepsin, trypsin, chymotrypsin, papain, etc.) has led to the extension of this hypothesis to all representative proteolytic enzymes (10).

Of the two essential points in the substrate, one, of necessity, must be the sensitive CO—NH group or some part thereof. The other requisite point of contact lies in the "backbone"* of the substrate and varies with the nature of the enzyme. The nature of this second group and its position in the backbone relative to the sensitive peptide bond provide the basis for the classification of proteolytic enzymes into four groups as given in Table I (the requisite groups are italicized and the sensitive peptide linkage is indicated by means of a dotted line).

TABLE I

CLASSIFICATION OF PROTEOLYTIC ENZYMES

Group No.	Linkage attacked (R on the CH)	Classification	
I	*NH₂*·CH·*CO*⋮*NH*·	Aminopeptidases	Exopeptidases
II	·*CO*⋮*NH*·CH·*COOH*	Carboxypeptidases	Exopeptidases
III	·*CO*—*NH*·CH·*CO*⋮*NH*·	Proteinases	Endopeptidases
IV	·*CO*⋮*NH*·CH·*CO*—*NH*·	Proteinases	Endopeptidases

In groups I and II are included the enzymes restricted in their action to peptide bonds at the end of a peptide chain. The peptidases belonging to group I (aminopeptidases) selectively attack the chain at the peptide linkage adjacent to the amino end of the chain while the peptidases of group II (carboxypeptidases) attack the chain at the

* In order to describe the structural setting of a peptide bond, it is desirable to speak of a "backbone" of the substrate, *i. e.*, the sequence of —NH—CH—CO— groupings linked through peptide bonds, and the "side chains," *i. e.*, the groups attached to the CH groups of the backbone.

peptide linkage adjacent to the carboxyl end of the chain. The amino- and carboxypeptidases cannot split linkages that are centrally located in the peptide chain; for this reason they are referred to as exopeptidases.

All the protein-splitting enzymes whose backbone requirements have been determined belong to group III. These enzymes are capable of hydrolyzing central peptide bonds and, therefore, are referred to as endopeptidases. They were found to require, in their substrates, a peptide bond in close proximity to the carbonyl group of the peptide bond which is hydrolyzed by the enzyme. Although no known proteolytic enzyme has been identified as belonging to group IV, the suggestion was made recently that an enzyme which hydrolyzes leucylglycylglycine and which is found in intestinal mucosa may belong to this group (31).

The presence of the indispensable groups in the backbone of a substrate in itself is insufficient to render the substrate susceptible to the action of an enzyme. It has been found that each of the proteolytic enzymes tested thus far also requires the presence, in the substrate, of a certain type of side chain (R) in a precisely defined location. In the second column of Table II, the required location of side chain R is indicated for each of the enzymes mentioned; and the chemical nature of these R groups is given in the third column. The enzymes which have been listed require in their substrates one of the following side chains: isobutyl as in leucine, benzyl or *p*-hydroxybenzyl as in phenylalanine or tyrosine, aminobutyl or guanidopropyl as in lysine or arginine. The side chains mentioned here represent only a few of those which jut out from the peptide chain of proteins. It will be a task of the future to determine precisely the specificity of proteolytic enzymes which require, in their substrates, the side chains of amino acids such as glycine, glutamic acid, histidine, tryptophane, etc.

Homospecific Proteolytic Enzymes

In the course of the systematic study of the specificity of proteolytic enzymes, it was noted that several different enzymes exhibited the same backbone and side-chain requirements in their substrates. For example, as will be seen from Table II, an enzymic component of papain and an enzymic component of beef spleen cathepsin have the

same type of specificity and split the same synthetic substrates as does crystalline pancreatic trypsin. Because of this relationship, the members of this group of enzymes have been designated "trypsinases." Similarly, evidence has been obtained for the existence of groups of enzymes related in specificity to pepsin, to leucine aminopeptidase,

TABLE II

SPECIFICITY OF PROTEOLYTIC ENZYMES

Enzyme	Requisite groups in substrate backbone	Requisite groups in substrate side chain
Peptidases (Exopeptidases)		
Leucine aminopeptidase from intestinal mucosa, beef spleen, beef kidney, and swine kidney	$NH_2 \cdot CH(R) \cdot CO \vdots NH \ldots$	$(CH_3)_2 > CH \cdot CH_2 \ldots$
Chymotrypsin aminopeptidase	Same	$HO \cdot C_6H_4 \cdot CH_2 \ldots$ or $C_6H_5 \cdot CH_2 \ldots$
Other aminopeptidases	Same	
Carboxypeptidase from pancreas, beef spleen, beef kidney, and swine kidney	$\ldots CO \vdots NH \cdot CH(R) \cdot COOH$	$HO \cdot C_6H_4 \cdot CH_2 \ldots$ or $C_6H_5 \cdot CH_2 \ldots$
Other carboxypeptidases	Same	
Proteinases (Endopeptidases)		
Pepsin Pepsinases from beef spleen, beef kidney, and swine kidney	$\ldots CO—NH \cdot CHX \cdot CO \vdots NH \cdot CH(R) \ldots$	$HO \cdot C_6H_4 \cdot CH_2 \ldots$ or $C_6H_5 \cdot CH_2 \ldots$
Trypsin Tripsinases from beef spleen, beef kidney, swine kidney, and papain	$\ldots CO—NH \cdot CH(R) \cdot CO \vdots NH \ldots$	$NH_2 \cdot CH_2 \cdot (CH_2)_3 \ldots$ or $NH_2(NH{=})C \cdot NH \cdot (CH_2)_3 \ldots$
Chymotrypsin	$\ldots CO—NH \cdot CH(R) \cdot CO \vdots NH \ldots$	$HO \cdot C_6H_4 \cdot CH_2 \ldots$ or $C_6H_5 \cdot CH_2 \ldots$

and to crystalline pancreatic carboxypeptidase. These are referred to as "pepsinases," "leucine aminopeptidases," and "carboxypeptidases," respectively. Such groups of enzymes possessing identical backbone and side-chain requirements are termed homospecific enzymes (6). On the other hand, two enzymes which differ from one another with respect to their backbone or side-chain requirements, or both, are designated heterospecific enzymes.

To our knowledge, the proteolytic enzymes represent the first class of enzymes for which the property of homospecificity has been demonstrated. It is not unlikely, however, that similar relationships may exist in other classes of enzymes. Some years ago, the question was raised whether the hydrolysis of the various β-glucosides by emulsin is to be attributed to a single β-glucosidase or to several glucosidases of slightly different specificity, and also whether the emulsins of various plants contain identical or different β-glucosidases. In particular, Weidenhagen (33) advocated the theory that the emulsins from various sources contain the same β-D-glucosidase, and that this enzyme splits not only all β-D-glucosides containing various aglucones, but also all oligosaccharides in which the sugar components are linked through β-D-glucosidic linkage. More recently, Pigman (28), in his classification of carbohydrases, suggested that the individual enzymes of Weidenhagen's system be considered as classes of enzymes acting on the same substrates but with different specificities. The finding of homospecific proteolytic enzymes now raises the question whether similar groups of carbohydrases of identical specificity type exist. It may be added that Bisseger and Zeller (13) have applied the concept of homospecificity to choline esterases obtained from various tissues.

Mechanism of Enzymic Proteolysis

It is inviting to speculate about the structural factors in the enzyme which give rise to the phenomenon of homospecificity. It was mentioned earlier that each proteolytic enzyme requires, for its action, certain atomic groupings in the backbone of its substrates. This backbone specificity may perhaps best be explained by the hypothesis that each enzyme molecule contains an essential center, composed of several distinct atomic groupings in a definite arrangement, and that the first step of enzymic action consists in a combination of several atomic groupings of the essential center of the enzyme with the indispensable backbone groups of the substrate. As in the Michaelis concept of the enzyme–substrate compound, it is assumed that this combination would result in the activation, and subsequent hydrolysis, of an adjacent peptide bond of the substrate. To explain the homospecificity phenomenon, it seems necessary to conclude that the enzymic action and its specificity originate from a rather restricted area

of the enzyme–substrate complex—the reacting nucleus—while the rest of the complex may be expected to have some influence only on the rate of the enzymic action. Two homospecific enzymes thus may differ in many respects but are assumed to contain identical essential centers, and therefore yield, with the same substrate, two enzyme–substrate compounds containing identical reacting nuclei.

Support for this hypothesis comes from the finding that, if each of two homospecific enzymes is allowed to act on two substrates, the quotient of the rates of hydrolysis of each substrate by the two enzymes is independent of the nature of the substrate. For example, papain trypsinase was found to hydrolyze benzoylargininamide with a proteolytic coefficient* of 167, while beef spleen trypsinase splits the same substrate with a coefficient of 8.3. The quotient of these two values is 20.1. If the same enzymes are tested with benzoyllysinamide as the substrate, 78 and 3.8 are the coefficients found, giving a quotient of 20.5. Similar examples may be cited for other pairs of homospecific enzymes; for further data, *cf.* Bergmann (6). The fact that two homospecific enzymes, E_1 and E_2, give similar proteolytic quotients for the hydrolysis of the substrates S_1 and S_2, may be explained by the assumption that the enzyme substrate compounds, E_1S_1, E_2S_1, E_1S_2, and E_2S_2, all contain identical reacting nuclei.

Furthermore, if this concept is correct, it may be expected that parts of the enzyme molecule may be split off without destruction of the enzymic activity, and without alteration of the enzymic specificity. Indeed, Kunitz (21) has shown that α-chymotrypsin may be transformed into γ-chymotrypsin, and that in the course of this transformation about one-third of the α-chymotrypsin molecule is removed. However, neither the proteolytic activity toward proteins nor the specificity of action on synthetic substrates was altered.

It may be added that both α- and γ-chymotrypsin have been found to exhibit two distinct proteolytic specificities, one of the aminopeptidase type and another of the proteinase (endopeptidase) type (18). Since all efforts to alter the ratio of the two specificities were unsuccessful, and also since α- and γ-chymotrypsin both conform to the

* The proteolytic coefficient (C) is defined as the value of the reaction velocity constant (K) for the hydrolysis of a peptide bond in the presence of an amount of enzyme corresponding to one milligram of protein nitrogen per cubic centimeter of the test solution.

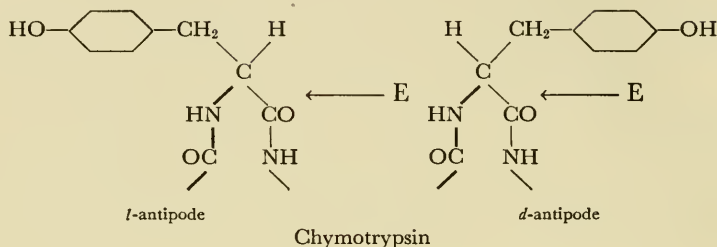
phase rule criteria for a pure protein (15), it would appear that each of these two proteins exhibits more than one distinct enzymic specificity. The twofold specificity of the chymotrypsins may be explained on the basis of the hypothesis presented above by assuming that each of the protein molecules contains two distinct and different essential centers.

The hypothesis of the predominant role of the essential center as against that of the remainder of the enzyme molecule cannot be considered the same as the well-known theory of Willstätter (34), who assumed that an enzyme molecule consists of a colloidal carrier and a prosthetic group, the latter being responsible for the enzymic activity and specificity. On the basis of this theory, Kraut distinguishes between the pheron and the agon as the two components of enzymes (20). Numerous workers refer to the prosthetic group as the coenzyme and the carrier as the apoenzyme (1), the dissociable flavoproteins frequently being regarded as the prototype of this postulated dual structure. It should be recalled, however, that the flavin part of the flavoproteins usually represents one of the partners in the chemical reaction and that the essential catalytic activity resides in the protein moiety; *cf.* also Parnas (27). In the case of proteolytic enzymes that do not require activation by sulfhydryl compounds (pepsin, trypsin, etc.), no evidence of a dual structure is available. For the activatable proteolytic enzymes, our present knowledge indicates that the active enzyme represents a dissociable combination of a protein with one of several activators (19). However, in the case of these latter enzymes, it cannot be claimed that the protein part of the enzyme acts merely as a colloidal carrier for another active part of the enzyme. On the contrary, it is the protein part of the activated enzyme which contains the essential center and which determines the specificity. No proteolytic enzyme is known in which the nature of the activator determines the specificity type of the enzyme.

Antipodal Specificity of Proteolytic Enzymes

The majority of the known proteolytic enzymes of higher plants and animals has been found to be adapted to the hydrolysis of substrates in which the essential side chain belongs to an *l*-amino acid. For example, chymotrypsin endopeptidase rapidly hydrolyzes the substrate benzoyl-*l*-tyrosylglycinamide, but does not hydrolyze benzoyl-

d-tyrosylglycinamide (9). The hypothesis of the essential center offers an explanation for such antipodal specificity. In order to act



Chymotrypsin

upon the substrate, the enzyme must approach the substrate closely so that the groups in the essential center of the enzyme can combine with the indispensable groups in the backbone of the substrate. If we consider the tetrahedral arrangement of the groups about the asymmetric carbon atom, it becomes evident that the enzyme must approach the substrate from the right side in order to combine with the essential backbone groups of the substrate. In the case of benzoyl-*d*-tyrosylglycinamide, the side chain prevents the enzyme from approaching the backbone groups, and therefore the enzyme cannot split the substrate.

When this theory of antipodal specificity was first proposed (5), it was suggested that the size of the side chain of *d*-alanine should be sufficiently small so that the combination of the essential center of the enzyme with the backbone groups of the substrate would not be prevented completely, but would be made more difficult. This steric hindrance should result in a retardation of the enzymic action on peptides of *d*-alanine as compared with its action on the *l*-form. In fact, it was found that crystalline pancreatic carboxypeptidase was able to hydrolyze carbobenzoxyglycyl-*d*-alanine, albeit much more slowly than the *l*-form (8).

It must be emphasized that the above conclusion can be correct only when it has been established that it is the same enzyme which hydrolyzes the two substrates containing *d*- and *l*-alanine. This reservation applies particularly to earlier experiments in which a study was made of the antipodal specificity of so-called dipeptidase and of aminopeptidase (12), since the existence of individual dipeptidases as well as the homogeneity of "aminopeptidase" have become doubtful.

Finally it should be mentioned that recent years have witnessed the discovery, in ever-increasing number, of proteolytic enzymes adapted to the hydrolysis of peptides of *d*-amino acids (3,25,29). The existence of such enzymes presents us with an interesting problem: if the antipodal specificity of a peptidase has its basis in the nature of the essential center, then the essential center of a *d*-peptidase should possess a configuration antipodal to that of the corresponding *l*-peptidase; *cf.* also Lettré (22).

Role of Proteolytic Enzymes in Peptide Synthesis

In recent years, especial emphasis has been given to the dynamic character of protein metabolism. The studies of Whipple (24), Schoenheimer (30), and others have given dramatic evidence for the view that, in the tissues of animals and plants, protein molecules are rapidly and continuously broken down and new protein molecules built up. The recognition of the "dynamic equilibrium" of proteins *in vivo* has brought to the fore the question of the nature of the enzymes that selectively catalyze the sequences of chemical reactions in protein synthesis and breakdown. More particularly, much attention has been given to the nature of the enzymes that mediate the synthesis of peptide bonds between the individual amino acids. One view, which is held widely, is that the biosynthesis of peptide bonds is catalyzed by the same enzymes that are responsible for the cleavage of peptide bonds; in other words, biological peptide synthesis is thought to represent a reversal of the degradative action of the proteolytic enzymes. The tissue proteolytic enzymes which are presumed to perform *in vivo* synthesis are those frequently designated "cathepsins" (in the case of animal tissues) and "papainases" (in the case of plant tissues).

The view that the hydrolytic action of proteolytic enzymes might be reversed was advanced by several workers at the start of the century and was later championed by Wasteneys and Borsook (32). More recently, it was shown unequivocally (7) that numerous proteolytic enzymes can catalyze, in model experiments, the synthesis of peptide bonds. One of many examples of such synthesis is the catalysis, by activated papain, of the following reaction: benzoyl-*l*-leucine + *l*-leucinanilide → benzoyl-*l*-leucyl-*l*-leucinanilide.

The fact that, in model experiments of this type, compounds of known and relatively simple structure are involved, in contrast to

the heterogeneous character of the protein hydrolyzates employed by previous workers, has permitted a closer study of various factors which play a role in the synthesis of peptide bonds by proteolytic enzymes. The most important of these factors are: (a) the specificity of the enzyme action; (b) the role of activators in the enzyme action; and (c) the energy relationships involved in peptide synthesis.

With respect to the specificity of peptide synthesis, it may be sufficient, at this point, to recall that the action of a proteolytic enzyme is to catalyze the attainment of equilibrium between a peptide and its hydrolytic products. Consequently, one should expect the specificity of synthesis to be the same as that of hydrolysis. Indeed, it has been found experimentally that, if the chemical nature of one of the groups near a peptide linkage is altered so that a given enzyme no longer is able to hydrolyze that linkage, then a similar structural change in the components for the enzymic synthesis will also prevent the formation of the peptide.

As is well known, several of the intracellular proteolytic enzymes of animals and plants require, for their full catalytic activity, the addition of sulfhydryl compounds as activators. The view was expressed some years ago (26) that, by oxidation of the sulfhydryl groups of a proteolytic enzyme into disulfide groups, the enzyme would cause peptide synthesis instead of hydrolysis. Experiments with model substrates soon showed, however, that the activation requirements were the same for the synthetic as for the hydrolytic reaction (7).

Turning now to the energy relationships involved in peptide synthesis, we should recall that the hydrolysis of peptide bonds in proteins and peptides proceeds spontaneously in the presence of a suitable enzyme and that the equilibrium which is established is very far on the side of hydrolysis. Thus, in order to reverse the hydrolytic reaction, energy is required. Borsook (14) has calculated, from thermal data, that the energy needed for the synthesis of a peptide bond is approximately 3000 calories per mole. This energy may be obtained in a variety of ways. In the case of the synthesis of benzoylleucylleucinanilide, mentioned earlier, the driving force for synthesis comes from the removal, by crystallization, of the synthetic product from the solution. In order to restore the balance of the equilibrium reaction, synthesis occurs, which, in turn, causes more of the synthetic product to crystallize.

Another mechanism for favoring the formation of peptide bonds is coupling the synthetic reaction with another energy-yielding chemical reaction. At the present writing, few experimentally demonstrated examples of such coupling can be cited (4). Surely the careful study of the coupling between peptide synthesis and energy-yielding systems represents one of the most interesting directions for future research; *cf.* Bergmann and Fruton (11).

Before leaving the question of the role of the proteolytic enzymes in peptide synthesis, it should be emphasized that the available experimental knowledge does not yet permit the conclusion that reversal of proteolysis is actually the process employed in biological systems for the synthesis of peptide bonds. It is well to remember that in the metabolic transformation of the polysaccharides, for example, synthesis is not effected by the reversal of the hydrolytic action of the amylases, but rather through a different chemical pathway, namely, the synthetic action of the phosphorylases (16). The intervention of phosphate in the biosynthesis of peptides also has been suggested (23), but no experimental evidence for this view has as yet been brought forward. Other speculations concerning the biological mechanisms for peptide synthesis have been advanced, largely on the basis of *in vitro* reactions (11). Clearly a greater fund of data on coupled reactions in metabolic systems is required before it will be possible to decide which, if any, of these theories is correct.

Perhaps the strongest reason for assuming, as a working hypothesis, the view that proteolytic enzymes do play an important role in protein synthesis is the fact they are the only known biocatalysts which, by virtue of their sharp specificity, could direct, precisely and reproducibly, the coupled sequence of successive peptide syntheses required for the formation of a protein. The considerations concerning specificity which have been discussed earlier in this article cannot fail to modify our picture of the possible role of the proteolytic enzymes in the biological synthesis of proteins. Until a few years ago, intracellular proteolytic enzymes, such as papain or cathepsin, were regarded either as single enzymes or mixtures of very few enzymes. On this basis it was concluded that the specificity of a single enzyme can predetermine the molecular pattern of a protein. Thus it was assumed that the specificity range of an intracellular proteinase would be sufficiently broad to comprise all the peptide bonds present in a protein

molecule. The demonstration of the extremely precise side-chain specificity of the proteolytic enzymes suggests that the synthesis of a protein from amino acids or small peptides could be accomplished only by the cooperative successive action of many enzymes of different specificity. In offering this suggestion, it must be emphasized again, however, that the view that the proteolytic enzymes mediate the biosynthesis of proteins is not supported as yet by unequivocal experimental evidence. There can be no doubt that the efforts of numerous biochemists will be directed in the future to the elucidation of this fundamental problem.

References

(1) Albers, H., *Angew. Chem.*, **49**, 448 (1936).
(2) Balls, A. K., and Köhler, F., *Ber.*, **64**, 34 (1931).
(3) Bamann, E., and Schimke, O., *Naturwissenschaften*, **29**, 558 (1941).
(4) Behrens, O. K., and Bergmann, M., *J. Biol. Chem.*, **129**, 587 (1939).
(5) Bergmann, M., *Science*, **79**, 439 (1934).
(6) Bergmann, M., in *Advances in Enzymology*, Vol. II. Interscience, New York, 1942, p. 49.
(7) Bergmann, M., and Fraenkel-Conrat, H., *J. Biol. Chem.*, **119**, 707 (1937).
(8) Bergmann, M., and Fruton, J. S., *J. Biol. Chem.*, **117**, 189 (1937).
(9) Bergmann, M., and Fruton, J. S., *J. Biol. Chem.*, **124**, 321 (1938).
(10) Bergmann, M., and Fruton, J. S., in *Advances in Enzymology*, Vol. I. Interscience, New York, 1941, p. 63.
(11) Bergmann, M., and Fruton, J. S., *Ann. N. Y. Acad. Sci.*, **45**, 357 (1944).
(12) Bergmann, M., Zervas, L., Fruton, J. S., Schneider, F., and Schleich, H., *J. Biol. Chem.*, **109**, 325 (1935).
(13) Bisseger, A., and Zeller, E. A., *Helv. Physiol. Pharm. Acta*, **1**, C86 (1943).
(14) Borsook, H., and Dubnoff, J. W., *J. Biol. Chem.*, **132**, 307 (1940).
(15) Butler, J. A. V., *J. Gen. Physiol.*, **24**, 189 (1940).
(16) Cori, C. F., *Biol. Symposia*, **5**, 131 (1941).
(17) Euler, H. v., and Josephsohn, K., *Z. physiol. Chem.*, **162**, 85 (1926).
(18) Fruton, J. S., and Bergmann, M., *J. Biol. Chem.*, **145**, 253 (1942).
(19) Irving, G. W., Fruton, J. S., and Bergmann, M., *J. Biol. Chem.*, **139**, 569 (1941).
(20) Kraut, H., and Pantschenko-Jurewicz, W. v., *Biochem. Z.*, **275**, 114 (1924).
(21) Kunitz, M., *J. Gen. Physiol.*, **22**, 207 (1938).

(22) Lettré, H., *Angew. Chem.*, **50,** 581 (1937).
(23) Lipmann, F., in *Advances in Enzymology*, Vol. I. Interscience, New York, 1941, p. 154.
(24) Madden, S. C., and Whipple, G. H., *Physiol. Revs.*, **20,** 194 (1940).
(25) Maschmann, E., *Biochem. Z.*, **313,** 129 (1942).
(26) Maver, M. E., and Voegtlin, C., *Enzymologia*, **6,** 219 (1939).
(27) Parnas, J., *Am. Rev. Soviet Med.*, **1,** 485 (1944).
(28) Pigman, W. W., *J. Research. Natl. Bur. Standards*, **30,** 257 (1943).
(29) Schmitz, A., and Merten, R., *Z. physiol. Chem.*, **278,** 43 (1943).
(30) Schoenheimer, R., *Dynamic State of Body Constituents.* Harvard Univ. Press, Cambridge, 1942.
(31) Smith, E. L., and Bergmann, M., *J. Biol. Chem.*, **153,** 627 (1944).
(32) Wasteneys, H., and Borsook, H., *Physiol. Revs.*, **10,** 110 (1930).
(33) Weidenhagen, R., *Ergeb. Enzymforsch.*, **1,** 205 (1932).
(34) Willstätter, R., *Ber.*, **59,** 1 (1926).

10

METABOLIC PROCESS PATTERNS

FRITZ LIPMANN, RESEARCH CHEMIST, MASSACHUSETTS GENERAL HOSPITAL, BOSTON; RESEARCH FELLOW IN BIOCHEMISTRY AND SURGERY, HARVARD MEDICAL SCHOOL

We have hitherto failed in our comprehension of life mainly because we have been involved in the absolute method of dealing with things.

E. NOBLE (14)

THE CEASELESS occurrence of metabolic processes in a living cell has long been understood to imply at large a need of energy for maintenance of active life. There was, however, and still is, only a vague realization of the tasks for which uninterruptible flux of energy is needed. A first opening here appeared when, through fuller chemical resolution, recently, reaction chains unfolded which, rather unexpectedly, were found to involve a multitude of substances containing phosphate in peculiar linkages. When the way in which these phosphate intermediates are manipulated in the cell was understood, it became possible to see clearly the connection between phosphate cycles and transformation and transport of energy. In all cells studied, a chemical network of energy distribution, the adenylic acid system, was found to be present and able to carry in the form of special energy-rich phosphate bonds standard portions of energy, amounting to about one-fiftieth of that liberated by total combustion of a mole of

carbohydrate. Therewith the view developed that catabolism consists to a considerable extent of a conversion of potential energy of foodstuffs into directly utilizable phosphate bond energy (7), and that, through alternate attachment and release of energy-rich phosphate bonds, catabolism and anabolism are knit together into a largely reversible reaction continuum.

This new appreciation of aspects of the metabolic apparatus which have hitherto been well concealed is beginning to affect our general attitude toward problems of metabolic chemistry. The more we recognize transformation of energy as a primary problem in metabolic processes, the more are we compelled to treat metabolic procedures for what they really are, namely, technical devices. In detail, the manner in which the living organism solves the problem of energy conversion is rather different from the technological methods employed by man. But whether in the case of the organism or man, the ultimate objective of energy conversion is the generation of energy in a utilizable form. A fundamental analogy appears, indeed, between the increasingly close dependence of our own daily life on electric current, gas pipes, and a variety of motors and that of our body cells on food and oxygen. In both instances the supply of energy is necessary to maintain an organization, although most of the energy is ultimately dissipated in the form of heat.

In many respects a living cell is comparable to a chemical factory. The design of chemical factories, from the standpoint of a technologist, is based on a variety of technical principles (5). Only the process proper remains chemistry, but its technical execution is effected wholly by physicomechanical devices. This predominantly mechanical manipulation of unit processes represents a most significant difference between organismic chemistry and chemistry practiced by man, for, in living cells, both process design and process execution are based on chemical principles. Instead of the material being manipulated successively in spatially separated compartments, cellular chemistry involves a harmonious series of consecutive reaction steps which are brought about on a molecular scale by a host of catalysts, all present together in the same reaction fluid. This difference in type of operation tends to obscure the basic analogy of both procedures.

Most metabolic processes classify among what the chemical engineer calls "flow processes" (5), that is, procedures whereby streams

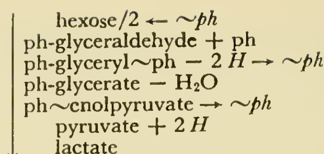
of material enter and leave a reaction system uninterruptedly. The technical flow process is based on flow charts which map the route along which a compound is driven through a series of operations. "An ideal flow process is characterized by steady states of flow, temperature and composition at any point of the process." This characterization holds likewise for almost any metabolic process.

Process Characteristics of a Fermentation

The physicomechanical environment in which we live has influenced our thinking to the point at which we must overcome certain mental inhibitions in order to comprehend the almost exclusive reliance of the living organism on chemical operations. This is particularly the case with processes of power generation which we habitually associate with highly mechanical machinery, though most of the power ultimately derives, as in our bodies, from chemical combustion of carbon and hydrogen. There has been an additional, more incidental, obstacle to a ready understanding of biochemical energy transformation. In fermentation—we are just gathering the elementary facts—human interest has centered long on manufacturing aspects, like the production of alcohol and other valued substances. From the biological point of view, a fermentation or a respiration is designed to produce power and the nature of the end product is more or less secondary and accidental.

In the simpler forms of anaerobic carbohydrate utilization, *e. g.*, lactic and alcoholic fermentation, the mapping of the sequence of reactions is now completed. But it will take some time until their pattern and design are duly comprehended. However, life and multiplication of a large variety of organisms are maintained exclusively through fermentative transformation of energy, frequently involving simple organic and nitrogenous compounds as starting materials. Therefore principles derived from the chemical mechanics of fermentation allow a fair amount of generalization.

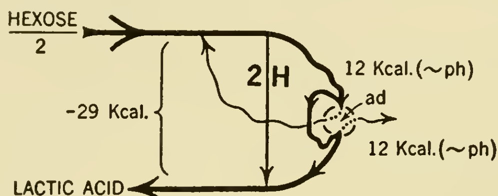
In scheme I, the simplest fermentative process, the conversion of glucose into lactic acid and phosphate bond energy is represented. To emphasize process characteristics, the now rather well-known intermediaries are omitted in the scheme; the flow chart represents the reaction sequence:



The terms ~ph, -ph, and ph characterize, respectively, the energy rich phosphate bond (12 kcal.), the ester phosphate bond (3 kcal.) and inorganic phosphate (7).

SCHEME I

The Process Pattern of Lactic Acid Fermentation



The flow line represents a projection into space of the catalytic pathway a hexose molecule travels to reach the inert end product, lactic acid. Initial fission in the middle of the six-carbon chain seems to be prompted by the introduction of phosphate groups at both ends. This phosphorylation is a rather costly investment absorbing just one-half of the gross yield of energy and thus reducing the net yield by about fifty per cent. An initial investment of part of the ultimate energy yield in the operation of the process is a notable feature. It is this need of induction energy which makes the fermentative process autocatalytic. The misleading statement is often made that, in fermentation, one half of the hexose molecule oxidizes the other half, suggesting a dismutative process. What happens, rather, is that a hydrogen donor, after unloading of phosphate bond energy, is transformed into a hydrogen acceptor. This manner of manipulation is expressed in the characteristic shape of the flow lines which, in all fermentations, fold back on themselves. The bending back, to accept a pair of hydrogens released in a previous stage on the molecular flow line, together with the initial expenditure of energy to start the process, may be considered as general and dominant characteristics of anaerobic metabolism.

Technologically there are numerous disadvantages in operation of the anaerobic type of metabolism. Most prominent is low efficiency. The probable upper limit of the gross yield is about ten per cent, whereas ninety per cent of the potential energy of combustion remains unused in the waste products. A piling up of waste products, frequently of strongly acidic character, presents a further serious technical problem. In the more highly organized living systems, therefore, we find the aerobic type predominant. If in the higher organisms we meet, as we do sporadically, a well-developed system of anaerobic energy conversion, it is in places or stages of development at which, for structural or topographical reasons, the oxygen supply is poor or unsafe: in the embryo, in cancer tissue, in parts of the placenta, parts of the retina, in muscle, etc. An anaerobic energy supply grants greater independence (8).

Process Characteristics of Respiration

The introduction of oxygen as hydrogen acceptor increases considerably the complexity of the energy-yielding process. Our present insight in this case is spotty and far removed from the completeness achieved in understanding simpler anaerobic fermentations. An attempt has been made here to coordinate the available data into a coherent process scheme, and at the same time to point out those stretches in the flow lines for which information is still missing.

When, as in Figure 1, the progressive catabolism of a substrate molecule in the manner of a flow chart is projected onto an energy–time coordinate system, some representative features emerge. The particulars of this scheme refer to degradation of half a glucose unit through the citric acid cycle. The gross energy available from this process, calculated by summation of the areas above the six consecutive steps of dehydrogenation is $6 \times 57, = 342$ kcal., a quantity practically identical with the theoretical yield for carbohydrate combustion. The dehydrogenation potentials of the intermediaries oscillate almost symmetrically around the hydrogen potential.

Contrary to present convention, the hydrogen potential at *p*H 7 is made the reference potential, which coincides with the potential of the "average" respiratory hydrogen donor. By plotting in a conventional manner oxidation-reduction potentials, E_0', as the difference between the normal potential of the system at *p*H 7 and the potential of atmospheric hydrogen gas at *p*H 0, a meaningless zero

line cuts arbitrarily through at the succinate/fumarate potential not far below the middle between hydrogen and oxygen potentials at *p*H 7.

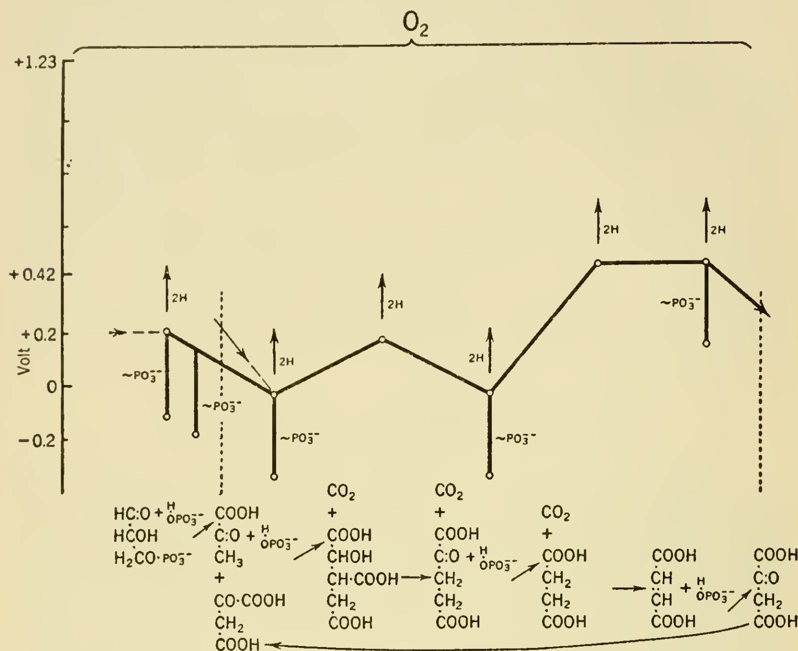


Fig. 1.—Citric Acid Cycle.

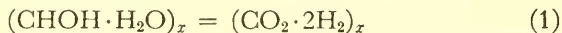
The dotted lines mark off the constantly repeating process unit. Each turn—from condensation to oxalacetate regeneration—oxidizes a two-carbon unit of carbohydrate level. The two-carbon unit is fed into the system in the form of acetate radicals.

Hydrogen donator[a]	Oxidation-reduction potential, volts[b]		Absolute potential difference between oxygen and water system, volts[c]
	Water system[c]	Phosphate system[c]	
Phosphoglyceraldehyde	−0.1	0.2	1.33 volts
Pyruvate	−0.35	−0.05	1.58
Isocitrate	0.13	—	1.10
Ketoglutarate	−0.35	−0.05	1.58
Succinate	0.43	—	0.80
Fumarate–malate	0.25	0.55	0.98
			Sum 7.37 volts
			Theory 7.46[d]

(*Continued on following page*)

Unfortunately, all schemes following the well-established rule of assigning to oxidation-reduction potentials values increasingly positive toward oxygen depict respiratory processes ambiguously. In respiration, the chemical potential gradually diminishes, of course, toward oxygen, being eventually spent with oxygen reduction, when the listed potential attains the most positive value.

As a characteristic of the process pattern, a separability of two main flow lines appears. The line representing catabolism of the substrate glides along on an approximately equipotential path; and at right angles to it the hydrogens which have been released by dehydrogenation journey to oxygen. So far, the greatest progress has been made in the field of substrate catabolism in which workable schemes have emerged. Schemes like that of the citric acid cycle, however, do not supply information about the chemical pathways of respiratory transformations of energy beyond the stage of hydrogen donation. The pathway of the substrate supplies merely the level from which electrons are emitted, loaded with potential energy and ready to be used. We learn, thus, from our map that the potential difference between oxygen and each of the six dehydrogenation steps which sum up to complete oxidation of a triose averages very closely to the value of the oxyhydrogen potential. In other words, carbohydrate reacts grossly like a mixture of carbon dioxide and molecular hydrogen:



$$(CHOH \cdot H_2O)_x = (CO_2 \cdot 2H_2)_x \quad (1)$$

At first approximation, it seems justified therefore to consider the respiratory process as a repeating series of rather uniform process

[a] For isocitric and ketoglutaric acids, tentatively, the potentials of hydroxybutyric (4) and pyruvic acid (11), respectively, were used here. For other values of the oxidation-reduction potentials, *cf.* Green (4).

[b] Reference potential: hydrogen electrode at *p*H 7; *cf.* page 141.

[c] The terms, *water* system and *phosphate* system, refer to the hydrated and the corresponding phosphorylated double bonds, as, for example, in phosphoglyceraldehyde hydrate and phosphoglyceraldehyde phosphate (9). The difference of the oxidation-reduction potentials between the water series and the phosphate series is approximately constant and equal to the volt equivalent of the energy-rich phosphate. For acetyl phosphate, recently, a bond energy of approximately 15 kcal. was calculated (10, 12) which corresponds to roughly 0.3 v. This value is appropriate for calculations primarily concerned with bond generation. The average energy is somewhat lower, 12 kcal. or 0.25 v., which value is preferred for turnover calculations. In cases in which more or less arbitrarily the actually reacting dehydrogenation system is assumed to be of the phosphate type, the connecting line is drawn through the phosphate system. In these cases, a vertical line in the graph connects the potential points of water and phosphate system, indicating the energy transformation. ***The wide empty area above the line connecting the potentials indicates the large part of the energy which remains here unaccounted for*** (*cf.* Fig. 2).

[d] Calculated from the combustion heat of one-half mole of glucose, 343 kcal. The agreement between the rather roughly approximated voltage and the combustion heat is noteworthy. To be accurate, 0.25 v., the equivalent of the nonoxidative phosphate bond in phosphopyruvate, should be added to the sum of the oxidation-reduction potentials.

units. Such a unit is essentially an oxygen–hydrogen cell. This unit is further broken down in the scheme of Figure 2, which may be representative of the respiration not only of carbohydrate but also of other substrates.

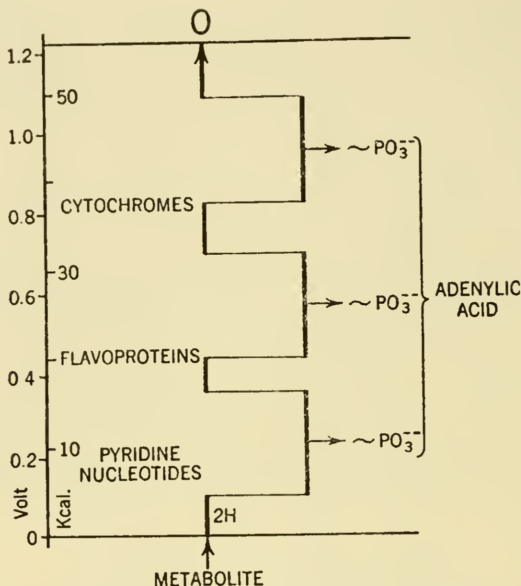


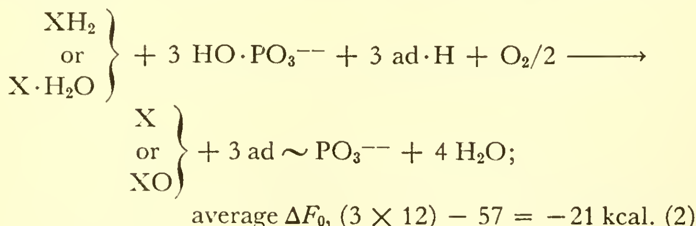
Figure 2.—Transformation of Electron Potential into Phosphate Bond Energy.

This graph accounts for the energy deficit which appeared in Figure 1 as empty space above the substrate flow line. A more detailed picture is obtained by fitting into the detours cycles of the type depicted in scheme II, page 146.

Projection of the potential gradient onto a space scale enables us to fix approximately on the map those regions where transformation of electron potential into phosphate bond energy is to be expected. It is known from a gross comparison between oxygen consumption and the resulting phosphorylation (1,6,15) that, by transport of one pair of hydrogen electrons from substrate to oxygen, 3+ energy-rich phosphate bonds may be generated. Energy-rich phosphate bonds average 12 kcal. per bond, which is equivalent to a span of 0.25 v. for a two-

electron system. The available potential between oxygen and a pair of average substrate hydrogens is 1.2 v., a potential span which may accommodate theoretically 1.2 divided by 0.25, or 4+ energy-rich bonds. This calculation fixes the upper limit of yield and shows the experimental values to be well within this limit.

In the particular scheme of Figure 2, a generation of the three phosphate bonds established by experiment is rationalized by merely cutting out, from the potential gradient, three 0.25-v. portions in succession. Such a mapping leads to the conclusion—unambiguously, it would appear—that at potential levels of around +0.1, +0.5, and +0.9 v., with reference to the hydrogen electrode of *p*H 7, the pair of hydrogen electrons is intercepted three times in succession by chemical devices which transform catalytically 0.25-v. portions into energy-rich phosphate bonds. This breaks the process of hydrogen transfer up into three smaller units; and here we probably meet the smallest units of the catalytic system designed for respiratory transformation of energy. The three, or perhaps four, transformers which are built into the pathway of the hydrogen electrons should be considered as the actual power generators of the living organism. It is very significant that these transformer systems appear largely independent of the particular hydrogen donor. Such operation of substrate-independent catalysts for transformation may explain how phosphate bonds are generated in a constantly increasing variety of oxidations, such as sulfur oxidation in *Thiobacillus* (16), the oxyhydrogen reaction (3), and fatty acid oxidation (13). To summarize, we may say that generation of phosphate bonds, regardless of the type of hydrogen donor, may be represented by the following equation:



$$\left.\begin{matrix} XH_2 \\ \text{or} \\ X\cdot H_2O \end{matrix}\right\} + 3\ HO\cdot PO_3^{--} + 3\ ad\cdot H + O_2/2 \longrightarrow$$

$$\left.\begin{matrix} X \\ \text{or} \\ XO \end{matrix}\right\} + 3\ ad \sim PO_3^{--} + 4\ H_2O;$$

$$\text{average } \Delta F_0,\ (3 \times 12) - 57 = -21 \text{ kcal.} \quad (2)$$

Equation (2) shows that a participation of phosphate and adenylic acid frequently is a reflection of general hydrogen transfer catalysis rather than a particular case of dehydrogenation.

The chemical nature of the transformer catalysts which operate on three 0.25-v. spans between the limits 0 and +1.2 v. remains to be considered. In analogy to what is known about the carbonyl compounds, which in part could have the function of transformers on the lowest potential level (9), the following generalization may be illustrative: The catalyst should offer opportunity to a phosphate molecule to add onto a double bond. From the addition product a pair of electrons is removed wherewith the energy-rich phosphate bond is generated. This step should involve little or no energy loss. Exchange of adenylic acid hydrogen for the phosphate group now delivers 12 kcal. into the cell, and the residual product may be rehydrogenated. This phase involves the refund of the delivered 12 kcal. (0.25 v.) and must ultimately lead to a regeneration of the double bond. Tentatively, the potential level around 0.1 v. may be assigned to a carbonyl double bond. For the subsequent level around 0.5 v. a C:C double bond would be suitable. Finally, a double bond of the ascorbic acid type appears as a possibility for the upper level. A transforming unit of the type outlined is charted arbitrarily for a C:C double bond in scheme II.

SCHEME II

A Catalytic Transformer of Electron Potential into Phosphate Bond Energy

O/R-CATALYST

HC···H

HC···OPO$_3^{--}$

$-H_2O$ $-2H$

POSPHATE SYSTEM E_0=0.45 V.

HCH | HCOH

HC ‖ CO~PO$_3^{--}$ H·ad +

WATER SYSTEM E_0=0.2 V.

+2H METABOLITE

H_2C | C:O

ad~PO$_3^{--}$ → UTILIZATION

=0.25 V.

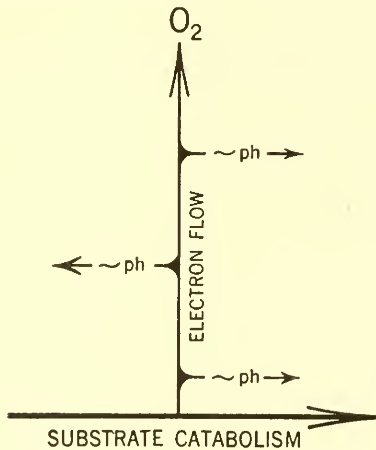
With slight modification it is applicable to any double bond of a C:X type, X being, for example, :O or :NH. The essential feature of the system is a shuttling between addition of water and of phosphate to the double bond. Starting from the bottom and moving clockwise, we distinguish four steps of the cyclic process. First, by *hydrogenation* of the *water* system the transformer is loaded with about 0.25 v. per mole, at least. Second, the reduced water

system is then converted into a phosphate system by exchange reactions. Third, in the key reaction, a *dehydrogenation* of the *phosphate* system transforms with little loss electron potential into phosphate bond energy. And fourth, the bond equivalent of 0.25 v. is unloaded to adenylic acid, thus returning the system to the original state.

In scheme III, a condensed scheme of respiration is drawn.

Scheme III

Summarizing Flow Scheme of Respiratory Energy Turnover



In the space projection the two energy fields occupy two dimensions, the third being assigned to the equipotential flow of primary metabolites. In reality, these fields of electron and phosphorylation potential are not homogeneous, but are canalized through chemical specificity of hydrogen and phosphate transfer, with enzyme specificity acting in the manner of connection plugs.

The underlying theme of this article is a reminder that, having pulled apart the chemical continuity of the living organism, we are challenged to reintegrate the scattered pieces into a whole. "There is more in a transition than a series of states or possible cuts, more in a movement than a sequence of positions or possible stops. We have

to place ourselves along the transition and, from within, to cut across in thought, in order to appreciate the successive states (2)."

References

(1) Belitzer, V. A., and Tsibakova, E. T., *Biokhimiya*, **4**, 518 (1939).
(2) Bergson, H. L., *Creative Evolution.* Modern Library, New York, 1944.
(3) Gaffron, H., *J. Gen. Physiol.*, **26**, 241 (1942).
(4) Green, D. E., *Mechanisms of Biological Oxidations.* Cambridge, Univ. Press, London, 1940.
(5) Hougen, O. A., and Watson, K. M., *Chemical Process Principles.* Wiley, New York, 1943.
(6) Kalckar, H., *Biochem. J.*, **33**, 631 (1939).
(7) Lipmann, F., in *Advances in Enzymology*, Vol. I. Interscience, New York, 1941, p. 99.
(8) Lipmann, F., "Pasteur effect" in *A Symposium on Respiratory Enzymes.* Univ. Wisconsin Press, Madison, 1942.
(9) Lipmann, F., "Biological oxidations and reductions," *Ann. Rev. Biochem.*, **7**, 1 (1943).
(10) Lipmann, F., *J. Biol. Chem.*, **155**, 55 (1944).
(11) Lipmann, F., and Tuttle, L. C., *J. Biol. Chem.*, **154**, 725 (1944).
(12) Meyerhof, O., *Ann. N. Y. Acad. Sci.*, **45**, 357 (1944).
(13) Muñoz, J. M., and Leloir, L. F., *J. Biol. Chem.*, **147**, 355 (1942).
(14) Noble, E., *Purposive Evolution.* Holt, New York, 1929.
(15) Ochoa, S., *J. Biol. Chem.*, **151**, 493 (1943); **155**, 87 (1944).
(16) Vogler, K. G., and Umbreit, W. W., *J. Gen. Physiol.*, **26**, 157 (1942).

11

BIOCHEMISTRY FROM THE STANDPOINT OF ENZYMES

DAVID E. GREEN, CHIEF OF THE ENZYME RESEARCH LABORATORY
DEPARTMENT OF MEDICINE, COLLEGE OF PHYSICIANS AND SURGEONS,
COLUMBIA UNIVERSITY; PAUL-LEWIS AWARD FOR ENZYME CHEMISTRY

IT WOULD be in the nature of a platitude to say that there is hardly a branch of biochemistry which cannot be analyzed or at least interpreted in terms of enzymes or enzymic phenomena. Yet few would maintain that enzymes represent more than an intellectual liqueur in the teaching of biochemistry in our graduate schools. Most modern textbooks of biochemistry often treat the subject of enzymes much in the manner that feathers and scales are dealt with in textbooks of anatomy. If this were merely another instance of the lag of textbooks behind developments in research, there would be no cause for concern. But more appears to be involved than the traditional lag. The implications of enzyme chemistry have yet to be more generally understood; and until that time arrives the textbooks will continue to regard enzymes as chemical oddities, if not to ignore them altogether.

During the past fifty years, biochemistry has developed from the ugly duckling of physiology to a science in its own right. During this period, interest has been focused largely on methodological and structural problems: how to estimate and what the constitution is of the innumerable compounds which make up the living cell. In this phase in the development of biochemistry, the analytical chemist and struc-

tural organic chemist played the leading roles, and indeed they laid the foundations for an exact science. Hence, it is hardly surprising that dynamic problems such as those of intermediary metabolism were approached more from the direction of what may be called chemical morphology than from the point of view of physiology. Perhaps the best way of illustrating the point is to recall the sensation which was produced by Schoenheimer's (5) early isotope experiments. His conception of an organism as a chemical system in a constant state of flux, in dynamic as opposed to static equilibrium, bore the same relation to the classical biochemical conception that the Schrodinger-Heisenberg conception of the atom bears to the rigid atom of the late 19th century. This is not to imply that the groundwork for Schoenheimer's conception was not already laid in the literature. The students of enzymes have long been aware of reversible equilibrium systems; but biochemists generally were unable to project the implications of these reversible systems in terms of intermediary metabolism. The outstanding researches of Schoenheimer and Krebs have done much to orient research on intermediary metabolism along more peculiarly functional lines; but, nonetheless, not more than the fringe of biochemical thinking has been disturbed. Intermediary metabolism is still being taught and discussed without any reference to the catalysts responsible for each of the transformations. Discussing the behavior of a car without reference to the motor represents an analogous situation. This state of affairs provides some reason for attempting an interpretation of biochemistry in terms of enzymes and enzymic phenomena. What follows will be rather sketchy, not only because of the limitations of space, but also in some cases because of the inadequacy of available information. However, the purpose of this essay is more to show that enzymes provide a logical and rational approach to many fields of biochemistry and medicine rather than to attempt a comprehensive survey of the enzyme field.

Living systems carry on their activities by virtue of myriads of chemical reactions which collectively are referred to as intermediary metabolism. Physiological functions such as growth, reproduction, secretion, nerve conduction, muscular contraction, etc., are integrations of whole series of chemical events in intermediary metabolism. These chemical events with few exceptions are not spontaneous processes. They require the presence of highly specialized protein cata-

lysts known as enzymes. Apparently, there is a different enzyme for practically every reaction, or at least every reaction type of intermediary metabolism. This can only mean that some several thousand enzymes must exist. For a long time fears were expressed that the cramped quarters of the average small sized cell could not possibly accommodate so many different enzymes. But that bogey has been laid low by the isolation from one small cell such as the yeast cell of literally hundreds of different enzymes. Of course, not all types of cells have the same complement of enzymes. The number and amount of enzymes vary from one cell type to another, and in fact determine the individuality of each cell.

There are many instantaneous ionic reactions which occur in the course of intermediary metabolism and which do not require enzymes, *e. g.*, neutralization of acids or bases, and deposition of salts such as calcium phosphate. But even in the province of reactions which occur spontaneously, a surprise is occasionally in store. Thus, the decomposition of carbonic acid into carbon dioxide and water is catalyzed by a special enzyme known as carbonic anhydrase, which can speed up the rate of the reaction far beyond the spontaneous rate.

The study of intermediary metabolism represents one of the oldest lines of biochemical investigation. It is not surprising, therefore, that of the total number of reactions known to occur in intermediary metabolism only a very small proportion have been reconstructed with isolated enzyme systems. In fact, whole chapters of intermediary metabolism such as the metabolism of steroids, porphyrins, carotenoids, sulfur compounds, bile acids, fatty acids, etc., are practically virgin territory as far as knowledge of enzymes is concerned. Enzyme chemistry has made its greatest strides in the field of fermentation or glycolysis of sugar. Here, the entire process from glucose to glycogen or from glycogen to lactic acid has been reconstructed *in vitro* with some twenty odd enzymes, each prepared in pure or largely pure state. This *in vitro* reconstruction is not to be regarded as a stunt or merely as a triumph of biological engineering. In order to effect a successful reconstruction, it becomes necessary to understand precisely the way in which certain enzyme systems are linked together and the way in which different chemical reactions are synchronized. A successful reconstruction, therefore, implies mastery of most of the chemical details and a complete knowledge of the constituent enzyme

systems. But the success in the field of glycolysis has been to some extent at the expense of progress in other fields.

Nutrition in essence deals with the relative amounts and nature of the materials which have to be supplied ultimately to the enzyme systems, and with the resyntheses and replacement of enzymes. In other words, the study of nutrition and the study of enzymes represent two sides of the same coin. Since the nature, number, and amounts of enzymes vary from one living system to another, the nutritional problem varies in the same way. If we knew all the enzymes present in a particular organism and the special components present in each of the enzymes, theoretically, we would have all the necessary data for determining the complete nutritional requirements. But since we are largely in the dark about the vast majority of enzymes, we use the data and knowledge of nutrition to ferret out information about enzymes. It has long been known, for example, that traces of certain metals such as manganese, iron, zinc, copper, and magnesium are essential in the diets of many animals. The indispensability of these metals in the diet has been correlated with the presence of these metals as structural elements of important enzyme systems. Thus, manganese has been shown to be an essential component of arginase; magnesium an essential component of carboxylase, chlorophyll, etc.; zinc of carbonic anhydrase; copper of phenolases; and iron of catalase, peroxidase, cytochromes, and lactic dehydrogenase. Traces of cobalt are known to be essential in the diets of sheep particularly. No doubt cobalt also will be identified as an essential component of some as yet unknown system. Among plants, boron and molybdenum are essential trace elements, and one must presume special enzyme systems in plants requiring these elements. The fact that, in most instances, small quantities of the metals are necessary correlates with the extraordinary activity of enzymes at high dilutions. In other words, only traces of metals are necessary for incorporation into the enzymes, since the enzymes also occur only in trace amounts.

The identification of the P-P factor with nicotinamide, of vitamin B_1 with thiamin, of vitamin B_2 with riboflavin, and of vitamin B_6 with pyridoxal provides a moral for those who prefer to study one side of a coin without reference to the other. Cozymase, a dinucleotide of nicotinamide and adenine, has long been studied as the coenzyme of fermentation since its discovery by Harden and Young (4) in 1906.

Yet it was not until 1937, some three years after Warburg had shown nicotinamide to be an essential component of the coenzyme, that Elvehjem and his colleagues (1) established a connection between the antipellagra vitamin and nicotinamide. Elvehjem's discovery was consistent with poetic justice because he had been trained both as a nutritionist and as an enzyme chemist. Vitamin B_2 was identified in 1935 by Kuhn and Karrer with riboflavin, the prosthetic group of the so-called yellow enzyme which Warburg had isolated from yeast three years earlier. The time relations were reversed in the cases of vitamins B_1 and B_6, since the identification of the vitamins preceded knowledge of their participation in enzymic reactions. The chemical identification of vitamin B_1 with thiamin by Williams and Cline in 1936 preceded by one year the demonstration by Lohmann and Schuster that the prosthetic group of yeast carboxylase is a diphosphoric ester of thiamine. Peters and his group at Oxford had established the role of vitamin B_1 in the oxidation of pyruvic acid long before the chemical nature of the vitamin was established. Vitamin B_6 was identified with pyridoxine in 1938 by Folkers, Keresteszy *et al.*, and Kuhn *et al.*, but it was not until 1944 that a more active form of the vitamin, *viz.*, pyridoxal, was discovered by Snell and that the phosphoric ester of pyridoxal was shown to be the coenzyme of tyrosine decarboxylase by Gunsalus. There are thus four authenticated identifications of vitamins with prosthetic groups. In the case of vitamin A, Wald has shown that it is an essential part of a photosensitive pigment in the eye known as visual purple. Many will not concede that visual purple has the properties of an enzyme but, whether or not that point is conceded, it is at any rate admissible that vitamin A fulfills the role of prosthetic group of a chromoprotein with an important physiological function.

This relation between vitamins and prosthetic groups makes it easy to understand the basis for the body's continuous requirement of vitamins. Since enzymes have a limited life period in consequence of their destruction during activity, there is constant need for more of all enzymes. The minimum amount of any of the vitamin compatible with viability is therefore an approximate measure of the total amount of enzymes in the body whose prosthetic groups contain that vitamin. Here, again, if we knew precisely which enzymes, for example, require flavin, what the normal levels of these enzymes are in different organs,

and the rates at which they are synthesized, we would have all the necessary data for determining the flavin requirements of that organism. There are much easier methods of getting at that data at the moment, but it is of considerable theoretical interest to arrive at the data from the enzyme side.

There are some well-informed workers in the field of nutrition who are unwilling to concede that all vitamins must have a catalytic function. While they admit that the relationship has been established in at least four and possibly six instances, they do not regard these necessarily as precedents. In 1941 the enzyme–trace substance theory (3) was developed which predicted that any substance necessary in the diet in trace amounts must be an essential part of some enzyme system. The theory was not meant to imply that substances required in higher concentrations cannot be essential parts of enzymes. That may or may not be the case. However, amounts of the order of 10 μg. per kg. per day were regarded as conclusive evidence of a catalytic role for that substance. On the basis of the enzyme–trace substance theory we may confidently expect in the near future the identification of biotin, pantothenic acid, folic acid, vitamin A, vitamin K, and vitamin D with essential parts of new prosthetic groups. The daily requirements of vitamin C are about a thousand times greater than those for most of the other vitamins, and, certainly, are well beyond the trace level. The possibility is therefore open that vitamin C may have no catalytic function whatsoever. The recent investigations of Sealock on the role of vitamin C in the oxidation of tyrosine in the liver however, do not encourage the view that vitamin C is an exception to the vitamin–enzyme relation.

In recent years the exact nutritional requirements of many bacteria and molds have been carefully investigated. A study of any of the synthetic diets proves very instructive from the standpoint of enzyme chemistry. The list of required substances can be easily divided into two categories: (*1*) substrates of enzyme systems, *e. g.*, amino acids, glutamine, dextrose, fatty acids, purines, etc.; and (*2*) precursors of prosthetic groups, *e. g.*, vitamins such as riboflavin, thiamin, etc. or the building stones thereof, hemin, and trace metals. The members of the first category can be classified not only on the basis of what we know about their intermediary metabolism but also on the basis of the amounts required, which are vastly in excess of the

minimal amounts necessary of substances in the second category. The concentrations of the first category are in milligrams per cubic centimeter, whereas those of the second are in fractions of a microgram per cubic centimeter.

The enzyme–trace substance theory in the form stated above is applicable only to naturally occurring substances in the diet or growth medium. There is an alternative form of the theory which applies to all substances regardless of their origin, and it may be stated in the following terms: Any substance which in trace amounts induces profound biological effects does so either by participating in or by specifically affecting some enzyme system. If valid, this theory must be regarded as one of the cornerstones of the science of pharmacology. An extensive list can now be compiled of substances, part if not all of whose pharmacological actions can be explained in terms of enzyme effects (*cf.* Table I).

TABLE I

Pharmacological agent	Enzyme inhibited by the agent
Fluoride	Enolase
One of the anterior pituitary hormones	Hexokinase
Cyanide	Cytochrome oxidase
Eserine	Choline exterase
Prostigmine	Choline esterase
Chlorinating agents	Triosephosphoric dehydrogenase
Iodoacetic acid	Triosephosphoric dehydrogenase
Benzedrine	Amine oxidase
Phlorhizin	Glucose phosphorylase
Gramicidin	One of the fermentation enzymes involved in phosphorylation
	Enzyme identical with the agent
α-Toxin of *Clostridium welchii*	A lecithinase
Lytic factor of cobra venom	A lecithinase
Spreading factor	Hyaluronidase
One of the toxins of *Cl. welchii*	Collogenase

One cannot but be impressed by the vindication of the enzyme hypothesis in at least fourteen instances, most of them reported within the last few years, in contrast to the absence of a single authenticated case in which any other principle of mechanism has been shown to operate. At present, the enzyme theory is really not much more than a good working hypothesis. But while it may be premature to regard all the biological effects of trace substances exclusively in terms of

enzymic effects at the present time, there is no alternative explanation that merits serious consideration. Alternative explanations usually amount to substituting an obscure phrase for an obscure phenomenon. Thus, some pharmacologists talk about trace substances upsetting an "active patch" of some important cellular membrane and thereby exerting their action. When interrogated about the properties of the "active patch," the pharmacologist usually admits that he has in mind some specific combining group and in effect admits a somewhat watered-down version of an enzyme reaction. Certainly there is no way of testing the "active patch" hypothesis as commonly stated nor is there any evidence that it has been productive either in explaining or predicting new phenomena. One cannot resist the conclusion that the "active patch" concept is a terminological device for cloaking ignorance. Another variant of the "active patch" concept is the so-called "active surface" which is sensitive to pharmacological agents and which controls certain key biological functions. The effect of agents on these surfaces is said to be exclusively a physical one, *i. e.*, the "active surface" becomes covered by the pharmacological agent and consequently is inactivated. The extraordinary specificity of pharmacological effects and the high dilutions at which these effects occur render this interpretation in purely physical terms unlikely.

The discovery by Woods (6) that the antibacterial action of the sulfonamides could be explained in terms of the resemblance of the sulfonamides to *p*-aminobenzoic acid was an important milestone in our understanding of the mechanism of the action of drugs. It became at once clear that some enzymic process was at the bottom of the chemotherapeutic action of the sulfonamides. There is still no clue as to the nature of this enzymic process but there can be little doubt that the process is enzymic. *p*-Aminobenzoic acid is a naturally occurring substance in yeast and animal tissues. Many bacteria and molds are unable to grow unless it is present in the medium. The trace concentrations in which it must be present for optimum growth exclude all but a catalytic role. In fact, *p*-aminobenzoic acid has been shown to exist in yeast largely in the form of a polypeptide, which may well be its active catalytic form in the intact cell. One theory of sulfonamide action assumes that the sulfa drugs displace *p*-aminobenzoic acid from its combination with specific proteins and thereby inactivate enzymes important in the growth of certain microorganisms. In

other words, the natural substance, *viz.*, *p*-aminobenzoic acid, competes with the drugs for the protein partner with which it forms the enzyme complex. The degree of inhibition is determined by the relative concentrations of *p*-aminobenzoic acid and drug and by their relative affinities for the protein partner. The phenomenon of competitive inhibition is well known in the enzyme literature and is perhaps the most characteristic hallmark of enzymic phenomena.

The sulfonamides inhibit the growth of susceptible bacteria only when they are actively growing. Once active growth has taken place, the sulfonamides are without effect even on susceptible bacteria. This observation suggests that the sulfonamides interfere with the process of synthesizing the *p*-aminobenzoic acid enzyme complex rather than with the activity or function of *p*-aminobenzoic acid in its final catalytic form. Unquestionably, there are many alternative pathways in bacteria by which *p*-aminobenzoic acid can be coupled with other substances to form the final catalytic complex. Only one pathway may be sulfonamide-sensitive, and only those bacteria which share that method of synthesizing the *p*-aminobenzoic acid catalytic complex will be susceptible to the action of the sulfonamides. These considerations must be borne in mind in evaluating the enzymic theory of sulfonamide action. Chemotherapeutic drugs may interfere either with the working of a key enzyme or with some stage in the process by which the key enzyme is synthesized. Since the synthesis of enzymes is also enzymic in nature, the primary action of chemotherapeutic drugs must still be considered as one of interference with enzymes.

The interpretation in terms of enzymes of the mode of action of sulfonamides is by no means in general currency. The so-called "essential metabolite" theory of Fildes (2) has gathered many adherents and is generally accepted among workers in the field of chemotherapy. This theory assumes that *p*-aminobenzoic acid is an essential metabolite for the growth of certain microorganisms rather than a part of an essential enzyme system. A metabolite is usually defined as a substance which undergoes chemical transformation; and the term is usually applied to substances like amino acids, fatty acids, etc. which are present in considerable concentration, and which are degraded or converted into more complex substances by enzyme systems. Since the amount of *p*-aminobenzoic acid present in microorganisms is

scarcely detectable by the most delicate chemical methods, *p*-aminobenzoic acid can hardly be classified as a typical metabolite. Quite clearly, the term metabolite as applied to *p*-aminobenzoic acid must imply merely that it is involved in metabolism. Furthermore, since traces of essential substances are known to participate in metabolism only in the capacity of catalysts, the "essential metabolite" theory boils down to a disguised enzyme theory. In the present state of ignorance, there is some merit in talking of *p*-aminobenzoic acid as an essential metabolite until such time as its precise enzymic role is clarified. But when the essential metabolite theory is seriously proposed as an alternative to the enzyme theory, it becomes important to recognize what the concept of essential metabolite really means. As applied to *p*-aminobenzoic acid, "essential metabolite" is just a term of caution to indicate by implication a catalytic role, without stating it in so many words. But as occasionally happens with terms of caution, their neutrality defeats their purpose. Essential metabolite has become confused by many with metabolite, and its real significance has become lost among those who see only the letter and not the spirit of the term.

The relation between the sulfonamides and *p*-aminobenzoic acid provided a blueprint for designing other chemotherapeutic agents. Every vitamin or prosthetic group theoretically should find its nemesis in some antivitamin. So the hunt began, and not without success. The sulfonic acid analogue of pantothenic acid (pantoyltaurine) was found to inhibit the growth of some bacteria which required pantothenic acid. The ratio of pantothenic acid to the sulfonic acid analogue determined whether growth or inhibition would take place. Much of the same kind of results were obtained with pyrithiamin, as antithiamin agent, with pyridine β-sulfonic acid, as antinicotinic acid agent, etc. None of these antivitamins is comparable to the sulfonamides in their efficiency as chemotherapeutic agents; but in these antivitamins, at least, we have the hopeful beginnings of a rational program of chemotherapy.

From the standpoint of enzymes, chemotherapy would appear to be the science of compounds which go for the enzymic Achilles' heel of an infectious organism without at the same time damaging the host unduly. In other words, the objective is first to find an enzyme which is present or important in the infectious organism and not in the host, and second to find a drug which specifically inhibits this

enzyme. Clearly, little progress will be made along rational lines in chemotherapy unless progress in the enzyme chemistry of infectious organisms is stimulated. Many are not convinced that progress in chemotherapy can come from that direction. "Suppose," they say, "you find the enzymic Achilles' heel, how are you going to find the drug to inhibit that enzyme?" As an example of the direction from which the solution might come, the chemical nature of the prosthetic group or active groups of the enzyme which turns out to be the weak point might provide the necessary clue for the synthesis of specific inhibitors. None will maintain that the solution will be easy, but experience has taught us that the possibilities of solving a problem are greatly increased when the nature of the problem can be accurately defined.

There is a school of thought well represented in our large pharmaceutical firms and also in the councils of our government scientific agencies which prefers to advance chemotherapy exclusively by the method of trial-and-error organic synthesis. In effect, the program followed is merely that of permutation and combination of the few effective chemotherapeutic agents we have as models. Instead of looking for new models, the old models are varied over and over again. The very limited success of this program of chemotherapy is hardly surprising. In the first place, not all pharmacologically active substances tolerate any considerable structural change. Thus, no one has been able to prepare a more active form of the vitamin than thiamin. In fact, the slightest alteration of the molecule involves partial or complete loss of activity. The same considerations apply to the vast majority of biologically active substances such as acetylcholine, histamine, flavin, ascorbic acid, etc. It is certainly true in other instances that a given effect is produced by large numbers of substances sharing a common structure, *e. g.*, adrenalin, and the hundreds of adrenalin-like bases which have been tested, sulfanilamide and the hordes of other sulfa drugs, etc. But one must keep the objective clearly in mind. In the case of adrenalin, the analogues merely imitate the adrenalin effect. They accomplish nothing that adrenalin cannot do. Their virtue lies either in their greater stability or in their greater resistance to deterioration in the animal body. All the known sulfa drugs act in exactly the same way, though they are effective at different concentration levels. The same organisms can be inhibited both by the weakest and strongest sulfa drugs provided the weakest is

sufficiently soluble. In other words, no new principle emerges. A more effective sulfa drug does not extend the range of action of sulfa drugs in the sense of inhibiting organisms which are otherwise insensitive to the sulfa drugs. From the standpoint of enzyme chemistry, the direction which much of chemotherapy research has taken does not appear to be either profitable or rational. Chemotherapy cannot be attacked intelligently without a detailed knowledge of intermediary metabolism and enzyme chemistry. We may make allowances for the element of urgency in wartime, but after the war, there ought to be a better balance between the sums spent on sheer trial-and-error organic synthesis and the sums spent on fundamental investigations.

Woolley (7) has pioneered in providing a framework for a rational pharmacology based on the antivitamin concept. He showed that certain antivitamins can produce a state of avitaminosis, in some cases in a matter of hours, merely by displacing the vitamins competitively from their catalytic complexes. Since profound pharmacological effects attend the syndrome of avitaminosis, antivitamins have to be regarded as potential pharmacological agents. Thus, pyrithiamin rapidly induces the disorders of the central nervous system which are characteristic of thiamin deficiency. The lesion is of course righted at once by addition of large enough amounts of thiamin. Since the quantitative importance of the catalytic reaction in which vitamins participate varies depending upon the organ or part of the organ, it does not follow that all antivitamins will exhibit similar pharmacological effects. On the contrary, it would appear that each antivitamin would selectively poison only a particular portion of the nervous system as well as only particular organs. A complete series of antivitamins should provide a wide range of specific pharmacological agents, all of which are reversible by addition of the vitamins which they imitate. The beginnings in this new field of exploration are still modest but the horizons seem immense.

The hormones represent a class of substance which, according to the enzyme–trace substance theory, ought unequivocally to qualify as enzymes or essential parts of enzymes. Yet no one has conclusively demonstrated that any one of this large class is either an enzyme or an essential part of an enzyme. Do we have in this class a notable exception to the theory? There is no basis for answering this question definitely one way or the other. There is a possibility that renin, the

hormone elaborated by the kidney, hydrolyzes hypertensinogen, one of the plasma proteins, with formation of a pressor substance. If this possibility is confirmed, we would have the first identification of a hormone with an enzyme function. Whatever the uncertainty about hormones as enzymes, the evidence leaves no doubt that hormones influence enzymic phenomena, *e. g.*, the dramatic effect of insulin* or adrenalin on carbohydrate metabolism. In practically every instance, hormone action has been boiled down to the regulation of some phase of intermediary metabolism. Do hormones regulate by being enzymes themselves, by influencing enzymes, or perhaps by controlling the synthesis of enzymes? There are many indications which point to some of the hormones controlling the synthesis of enzymes. But since the synthesis of proteins represents one of the most obscure corners of enzyme chemistry, there is little hope of any early clarification of the precise role which hormones might play in synthesis.

Oddly enough, our most precise knowledge of the way in which the synthesis of enzymes is regulated has been acquired from the field of genetics. The experimentation which has led to this knowledge constitutes one of the most brilliant chapters of modern biology. Geneticists have succeeded in demonstrating that single genes control the syntheses of single enzymes. This implies that genes, like hormones, are regulators of intermediary metabolism. Genes accomplish this regulation by controlling the synthesis of enzymes, whereas hormones operate in ways as yet not classified. Genes and hormones are distinguishable in another fundamental respect—hormones must be synthesized by the respective endocrine glands, while genes are autocatalytic and hence self-perpetuating. This autocatalytic property of the gene resolves the dilemma that, if enzymes are needed to synthesize other enzymes, there must then be an infinite series of enzymes making enzymes. But recognition of autocatalysis as a phenomenon is a far cry from understanding the mechanism. As a matter of fact, guesswork constitutes the sum and total of our knowledge of the way in which certain protein molecules are able to reproduce themselves.

* When this essay was in the proof stage, Cori and his group announced the epoch-making discovery that insulin reverses the inhibition of hexokinase produced by one of the anterior pituitary hormones. In these two instances, at any rate, hormones must be regarded as regulators of enzymes by virtue of their inhibiting or releasing the inhibition of key enzymic processes.

The problem of the autocatalysis of genes is essentially similar to that of viruses. What is the starting material for the synthesis and how is the synthesis accomplished? No concrete answer is possible as yet, but there are indications that the solution to the problem will be in terms of enzyme chemistry. One intriguing development in recent years has been the successful application of the concept of competitive inhibition drawn from the field of enzymes to the virus problem. Two strains of the same virus are found to antagonize one another's growth in the same host, presumably by competing for the same pabulum. In chemotherapy, two substrates, one natural, the other "fraudulent," compete for the same enzyme. In the example above, two viruses are made to compete for the same substrate. This virus interference phenomenon occurs only when the two strains are closely related, just as competitive inhibition occurs only when the "fraudulent" substrate closely resembles the natural.

The past decade has seen not only the extension of our knowledge of enzymes to other fields of biochemistry and medicine but also the extension in part of the enzyme concept to noncatalytic proteins. Let us compare, for example, two proteins, catalase and hemoglobin. Both contain iron protoporphyrin as prosthetic group and both are highly specific. Catalase catalyzes the decomposition of hydrogen peroxide into oxygen and water, whereas hemoglobin combines reversibly with molecular oxygen. Catalase cannot function as hemoglobin and conversely hemoglobin for all practical purposes does not catalyze the decomposition of hydrogen peroxide. Catalase forms a compound with hydrogen peroxide and then the enzyme–substrate compound undergoes decomposition. This cyclical process repeats itself more than two million times per minute at 0°. In a similar way, hemoglobin forms a compound with molecular oxygen over a certain range of oxygen tension, and the complex is dissociable by lowering the oxygen tension. The speed of the combination is of about the same order of magnitude as for the combination of catalase with hydrogen peroxide. This comparison is not being made to infer that hemoglobin is an enzyme. There is little to be gained by a redefinition of the classical concept of an enzyme to include a proteinlike hemoglobin, because after all there is a real distinction between a catalyzed reaction and a noncatalyzed reaction. However, it is important to recognize the many properties in common which catalase and hemoglobin share.

While enzymes must be differentiated from noncatalytic proteins, nonetheless a broader classification of proteins is conceivable in which enzymes represent a special case of what we may call the functional type of protein. We may define the functional protein as one which performs a specific physiological function. Thus, catalase decomposes hydrogen peroxide; hemoglobin combines reversibly with molecular oxygen; cytochrome C is reduced by the reduced forms of certain enzymes and in turn its reduced form is oxidized by cytochrome oxidase; prothrombin plays a specific role in blood clotting; visual purple acts as a photoreceptor, etc. Limitation of space precludes further development of the concept of functional proteins. Suffice to say that, in the author's opinion, processes like those of blood coagulation, complement fixation, and antibody formation, are phenomena which have much in common with enzymic phenomena, and that the highly specific functional proteins responsible for these processes have more in common with enzymes than with purely structural proteins. The concept of functional proteins has the virtue of opening new horizons in the form of novel types of proteins. Just as myosin is a protein of muscle specialized to convert chemical energy to mechanical energy or as visual purple is a protein in the retina specialized to convert light energy, presumably, ultimately to the electrical energy of nerve conduction, so there may be analogous proteins in nerve, cellular membranes, etc. specialized to carry out the particular physiological functions of these organs. The trend in biochemistry would appear to be toward the inclusion of more and more proteins in the category of catalysts. In fact, it is conceivable that eventually all proteins apart from purely structural proteins will be found to perform in a highly specific way some physiological catalysis, and the currently prevalent idea of storage and inert proteins will soon be as outmoded as the so-called endogenous nitrogen metabolism.

There is another aspect of myosin and visual purple that well merits consideration. Biochemists have long exercised themselves over the problem of the means by which the organism converts energy from one form to another. In one or two instances, the curtain surrounding these interconversions has been pierced. Myosin and visual purple may well be considered as examples of energy transformers. Thus, myosin in effect converts the chemical energy of hydrolysis of adenosine triphosphate into mechanical energy. The protein itself

acts as the transforming agent by combining two physiological functions. In the same way, visual purple becomes a vehicle for transforming light energy into chemical energy and, we must assume, eventually reacts in some way with nervous elements of the retina. Instances of enzymes with multiple functions and multiple active groups have been known in the literature, but the tendency hitherto has been to regard them as biological curiosities. Thus pyruvic oxidase of *Lactobacillus delbrueckii* contains two prosthetic groups, *viz.* flavin dinucleotide and diphosphothiamin. Milk flavoprotein contains two prosthetic groups (flavin dinucleotide and another as yet unidentified) and catalyzes the oxidation of purines, aldehydes, and dihydrocoenzyme I. The *l*-amino acid oxidase of rat kidney has two enzymic functions. These enzymes with multifunctions may be involved in the transfer of chemical energy from exergonic to endergonic processes. In other words, enzymes may ultimately turn out to be the energy transformers and converters of the cell.

References

(1) Elvehjem, C. A., Madden, R. J., Strong, F. M., and Woolley, D. W., *J. Am. Chem. Soc.*, **59**, 1767 (1937).

(2) Fildes, P., *Lancet*, **I**, 955 (1940).

(3) Green, D. E., *Advances in Enzymology*, Vol. I. Interscience, New York, 1941, p. 177.

(4) Harden, A., and Young, W. J., *Proc. Roy. Soc. London*, **B77**, 405 (1906).

(5) Schoenheimer, R., *The Dynamic State of Body Constituents*. Harvard Univ. Press, Cambridge, 1942.

(6) Woods, D. D., *Brit. J. Exptl. Path.*, **21**, 74 (1940).

(7) Woolley, D. W., *Science*, **100**, 579 (1945).

12

ENZYMIC MECHANISMS OF CARBON DIOXIDE ASSIMILATION

SEVERO OCHOA, ASSISTANT PROFESSOR OF BIOCHEMISTRY, NEW YORK UNIVERSITY COLLEGE OF MEDICINE

IT IS well known that animal cells depend on a supply of ready-made organic materials (such as carbohydrates, fats, and proteins) or of their building stones (such as simple sugars, fatty acids, and amino acids) for either building up cell substance or replenishing their stores of energy-yielding foodstuffs. Green plants, however, are able to fix atmospheric carbon dioxide under the influence of light and use it to synthesize the organic constituents they need; by means of this process—photosynthesis—they "assimilate" carbon dioxide. In its final over-all results, photosynthesis is essentially a reversal of respiration: In respiration, foodstuffs are oxidized to carbon dioxide and water with absorption of oxygen, the energy thereby released being utilized by the cell to carry out its activities; in photosynthesis, the chlorophyll-containing chloroplasts utilize radiant energy to build up organic substance from carbon dioxide and water, and oxygen is liberated in the process.

Some microorganisms, such as green algae and both green and purple bacteria, are photosynthetic. Certain bacteria do not possess the capacity to utilize radiant energy for carbon dioxide assimilation but are able to use, for the same purpose, the energy derived from oxidation of inorganic substances like hydrogen sulfide, thiosulfate,

sulfite, selenite, nitrite, elementary sulfur, ammonia, and molecular hydrogen. This process is known as chemosynthesis.

Both photosynthetic and chemosynthetic organisms are referred to as autotrophic because they can grow in media composed of inorganic substances exclusively. However, a number of bacteria, including most of the pathogenic species, can live only in media that contain one or more organic components, and are known as heterotrophic (24,25,28).

The importance of photosynthesis and, in general, of carbon dioxide assimilation can hardly be overemphasized, since animals depend for their subsistence on the materials formed through carbon dioxide assimilation by autotrophic organisms. Thus, carbon dioxide assimilation is one of the most fundamental of all life processes. Although it was known for some time that heterotrophic bacteria and some animal cells could utilize carbon dioxide to synthesize carbon-to-hydrogen bonds or carbon-to-nitrogen bonds (as is the case in the synthesis of formic acid from carbon dioxide and hydrogen by *Escherichia coli*, or in the synthesis of urea by the liver), the belief was current that such organisms lacked the capacity to utilize carbon dioxide for the synthesis of carbon-to-carbon bonds until the pioneer work of Wood and Werkman demonstrated that heterotrophic bacteria can fix carbon dioxide in this manner (28). The process is now known to occur also in animal cells.

The synthesis of organic material from carbon dioxide is an endergonic reaction, which results in an increase of the free energy content of the system, and thus requires energy in order to proceed. In other words, such a synthesis must be coupled with exergonic* reactions involving a decrease in free energy. For this purpose, photo- and chemosynthetic organisms can use either radiant energy or the energy derived from oxidation of inorganic compounds. Heterotrophs, on the other hand, can assimilate carbon dioxide only at the expense of oxidizing organic foodstuffs, so that no net gain in organic cell constituents can result from carbon dioxide assimilation under these conditions. It is, therefore, difficult to decide whether carbon dioxide fixation is

* C. D. Coryell, in *Science*, **92**, 380 (1940), introduced the terms "exergonic" and "endergonic" to characterize negative and positive changes in free energy (ΔF), respectively, and suggested that the use of the terms "exothermic" and "endothermic" be restricted to designate changes in heat (ΔH).

essential or even important for heterotrophic organisms and animal cells. One could conceive, however, that fixation might be used for the synthesis of special cell constituents such as growth factors or the like. It is well known that most heterotrophic bacteria require the presence of carbon dioxide in the medium for optimal growth.

The cellular mechanisms of carbon dioxide fixation have been obscure for a long time. It is only recently that light has been shed on the mechanism by which fixation occurs in heterotrophic bacteria and animal cells. As it now appears, the fundamental process in all types of carbon dioxide fixation is a reversal of the decarboxylation of some keto acids—a process catalyzed by enzymes. These reactions are reversible, but their equilibrium lies very far to the side of decarboxylation, *i. e.*, liberation of carbon dioxide. Thus, the problem faced by the cell is to shift the equilibrium as far as possible in the opposite or uphill direction, and this requires expenditure of energy.

It has been established (15) that the free energy change of a reversible chemical reaction is related to the equilibrium position in a manner expressed by the equation:

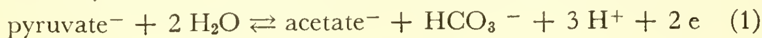
$$\Delta F = -RT \ln K$$

where ΔF represents the change in free energy (expressed in gram calories) and K is the equilibrium constant. If the equilibrium constant is expressed as:

$$K = \frac{\text{(decarboxylation product) } (CO_2)}{\text{(carboxylated product)}}$$

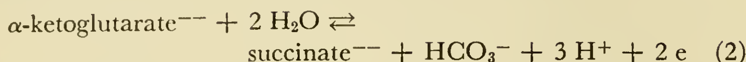
The equilibrium constant of some of the reversible decarboxylations is of the order of 10^3, so that they proceed with a decrease in free energy of about -4000 to -5000 calories. Here are included the enzymic decarboxylation of oxalacetic acid to pyruvic acid and carbon dioxide, and that of oxalosuccinic acid to α-ketoglutaric acid and carbon dioxide. In both cases the reaction involves a carboxyl in position β relative to the carbonyl group; this reaction type will be referred to here as β-carboxylation.

There is another group of enzymic decarboxylations involving simultaneous decarboxylation and dehydrogenation of α-keto acids that proceed with a much larger decrease in free energy than do β-decarboxylations. Thus, the free energy change of reaction (1):



$$\text{pyruvate}^- + 2\,H_2O \rightleftarrows \text{acetate}^- + HCO_3^- + 3\,H^+ + 2\,e \quad (1)$$

has been estimated to be −9400 cal., and that of reaction (2) −8000 cal. (8):



$$\alpha\text{-ketoglutarate}^{--} + 2\,H_2O \rightleftarrows \text{succinate}^{--} + HCO_3^- + 3\,H^+ + 2\,e \quad (2)$$

Reactions of this type are often referred to as oxidative decarboxylations and the reverse as reductive carboxylations.

From the relationship between free energy change and equilibrium constant discussed above, it is clear that shifts of equilibrium toward carboxylation, *i. e.*, carbon dioxide fixation, can only be accomplished by an input of energy into the system. We shall consider in some detail the enzymic mechanisms used by the cell for this purpose.

The fundamental pattern of biological carbon dioxide fixation can be visualized in terms of the following steps:

(*1*) *Primary Fixation Reaction. Carboxylation.* Since the equilibrium of the reversible reaction involved is very unfavorable, only small amounts of keto acid are formed at one time.

(*2*) *Reduction.* Step *1* is followed by enzymic reduction of the keto acid to the corresponding hydroxy acid. This shifts the equilibrium and more keto acid can be formed by step *1*.

(*3*) *Reduction of the Pyridine Nucleotide Oxidized in Step 2.* The second and third steps will be discussed below.

(*4*) *Dehydration, Hydration, Isomerization.* Further equilibrium shifts can be brought about by secondary enzymic transformations of the hydroxy acid formed by step *2*. Thus, the hydroxy acid may be dehydrated to the corresponding unsaturated fatty acid. The latter, in turn, may be hydrated in a different position of the molecule to form a new hydroxy acid isomeric with the one first formed.

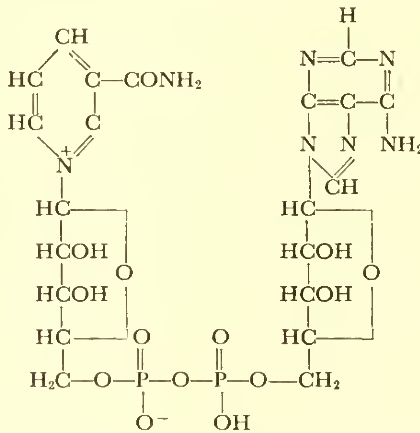
(*5*) *Further Reduction.* The unsaturated fatty acid formed in step 3 may undergo further reduction to the corresponding saturated acid.

All the reactions concerned in the various steps just outlined are reversible.

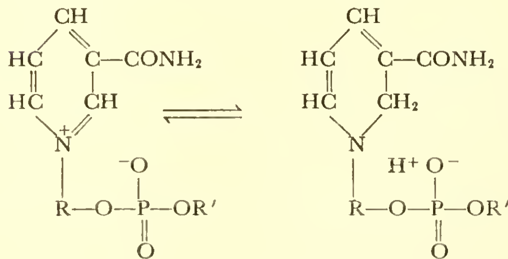
The reduction of the keto acid in step *2* is catalyzed by a specific pyridine nucleotide dehydrogenase. These dehydrogenases (27) consist of a protein and a prosthetic group or coenzyme that combine with

one another, although the equilibrium constant is preponderantly in favor of dissociation. The active group of the coenzyme is pyridine. Two such prosthetic groups are known at present, *viz.*, diphosphopyridine nucleotide (abbreviated DPN, also known as cozymase and coenzyme I), and triphosphopyridine nucleotide (abbreviated TPN, also known as Warburg's coenzyme and coenzyme II). The coenzymes are dinucleotides, each containing two bases, adenine and nicotinamide, two molecules of ribose, and either two (DPN) or three (TPN) molecules of phosphoric acid. The structural formula of DPN is shown in scheme I. The point of attachment of the third phosphoric acid residue in TPN is as yet unknown. Most of the pyridine nucleotide dehydrogenases have DPN as the prosthetic group; only two are defi-

SCHEME I



Structural formula of diphosphopyridine nucleotide



Reversible reduction of diphosphopyridine nucleotide

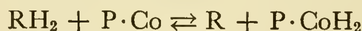
nitely known to function with TPN, and one at least can function with either coenzyme (Table I).

TABLE I

PYRIDINE NUCLEOTIDE DEHYDROGENASES

System	Ox-redox potential (E_0'; pH 7.0), volts	Prosthetic group	Occurrence of dehydrogenase
Glutamic acid ⇄ α-ketoglutaric acid + NH_3	−0.03	DPN or TPN	Yeast, bacteria, animal tissues
Malic acid ⇄ oxalacetic acid	−0.10	DPN	Bacteria, plants, animal tissues
Ethyl alcohol ⇄ acetaldehyde	−0.16	DPN	Yeast, bacteria
Lactic acid ⇄ pyruvic acid	−0.18	DPN	Bacteria, animal tissues
3-Phosphoglyceraldehyde + phosphate ⇄ 1,3-diphosphoglyceric acid	−0.28	DPN	Yeast, animal tissues
β-Hydroxybutyric acid ⇄ acetoacetic acid	−0.29	DPN	Animal tissues
Isocitric acid ⇄ oxalosuccinic acid	−0.30	TPN	Yeast, plants, animal tissues
Glucose-6-phosphate ⇄ 6-phosphogluconic acid	−0.43	TPN	Yeast, red blood cells
Glucose ⇄ gluconic acid	−0.45	DPN	Animal tissues

The pyridine in the coenzymes acts by cyclic addition and removal of hydrogen (see scheme I). Through the reversible change pyridine ⇄ dihydropyridine, it functions as a hydrogen carrier. The protein component of the dehydrogenase, upon whose nature depends the specificity for its substrate, binds both substrate and coenzyme, and in this complex hydrogen from the substrate is transferred to the pyridine of the coenzyme; the substrate is thus oxidized and the pyridine reduced. This transfer of hydrogen is reversible, so that dihydropyridine and oxidized substrate when bound by the protein can react to form pyridine and reduced substrate. Thus, the action of a typical pyridine nucleotide dehydrogenase can be represented as follows:

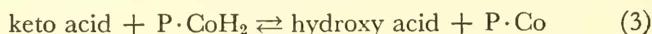


$$RH_2 + P \cdot Co \rightleftarrows R + P \cdot CoH_2$$

where RH_2 and R represent reduced and oxidized substrate and $P \cdot Co$ and $P \cdot CoH_2$ represent oxidized and reduced protein–coenzyme complex, respectively.

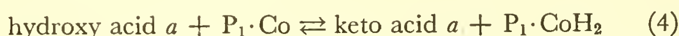
The reduced coenzymes have a sharp absorption band at 340 mμ, whereas there is no absorption at this wave length by the oxidized form. This distinction is important because the action of the pyridine nucleotide dehydrogenases can be conveniently followed by changes in the absorption of light of wave length 340 mμ (27).

It is thus clear that the biological reduction of the keto acid formed in step *1* of carbon dioxide assimilation requires the presence of the specific dehydrogenase and of the reduced form of its prosthetic group. Step *2* can then be represented as follows:



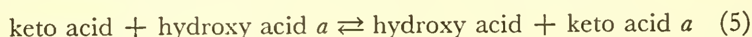
$$\text{keto acid} + P \cdot CoH_2 \rightleftarrows \text{hydroxy acid} + P \cdot Co \qquad (3)$$

Since the complex formed by the coenzyme with the protein component of the dehydrogenases is dissociable to a relatively large degree, there are always small amounts of free coenzymes present in the cell. When equilibrium has been reached in a reaction of the type represented by reaction (3), it will stop unless provision is made for a reduction of the oxidized coenzyme formed so as to displace the equilibrium to the right. Step *3* occurs here. It is carried out through the action of another dehydrogenase which functions with the same prosthetic group as that acting in step *2*. Such a reaction may be, for instance:



$$\text{hydroxy acid } a + P_1 \cdot Co \rightleftarrows \text{keto acid } a + P_1 \cdot CoH_2 \qquad (4)$$

and is possible because of the dissociable nature of the complex formed by the coenzyme with the protein components of the dehydrogenases, so that the coenzyme can alternatively be bound by either protein. Thus, some of the oxidized coenzyme dissociating from $P \cdot Co$ can be bound by the second protein, P_1, to form $P_1 \cdot Co$, as in reaction (4). Since the coenzyme oxidized in reaction (3) is reduced in reaction (4), the net result of these two reactions can be expressed by:



$$\text{keto acid} + \text{hydroxy acid } a \rightleftarrows \text{hydroxy acid} + \text{keto acid } a \qquad (5)$$

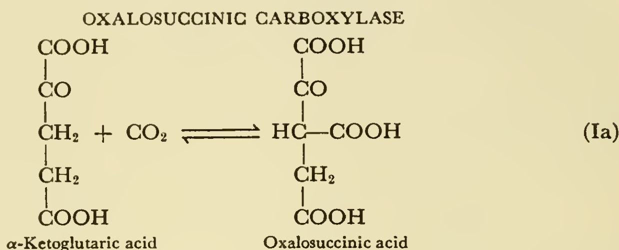
This type of reaction is known as a coenzyme-linked dismutation, and is, in general, reversible. The extent to which it will proceed in a given direction depends on the equilibrium constants of the two dehydrogenase systems involved and on the concentration of reactants. Thus, in our case, a high concentration of hydroxy acid *a* will favor carbon dioxide fixation.

We shall now consider the individual carbon dioxide fixation systems.

β-Carboxylation

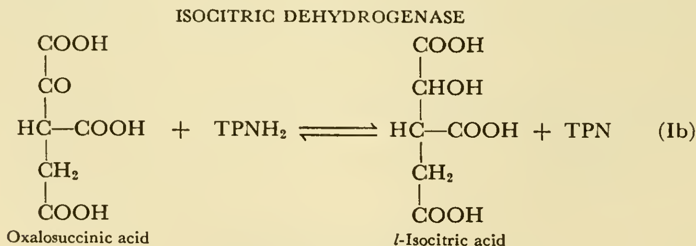
Carbon dioxide fixation by α-ketoglutaric acid. Although this type of carbon dioxide fixation is the most recently discovered

(20,21), it will be convenient to discuss it first because the methods used in the study of this system permit a clearer picture of the pattern into which the cellular mechanisms of carbon dioxide fixation can be fitted. The primary fixation reaction involves the reversal of the decarboxylation of oxalosuccinic acid to α-ketoglutaric acid and carbon dioxide (reaction Ia). This reaction is catalyzed by an



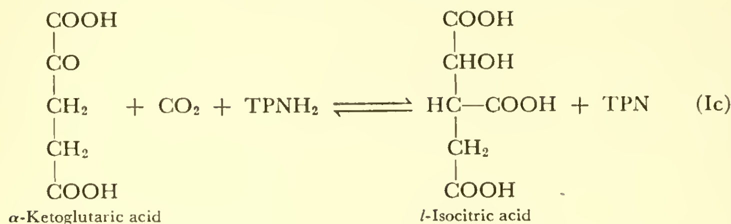
enzyme, oxalosuccinic carboxylase, present in heart muscle and probably in other animal and plant tissues. The enzyme requires either magnesium or manganese ions for activity.

The equilibrium of reaction Ia is so far to the left that the available analytical methods would fail to show a formation of oxalosuccinic acid even when starting with very high concentrations of α-ketoglutaric acid and carbon dioxide in the presence of the enzyme. However, reversibility can easily be demonstrated by adding isocitric dehydrogenase (see Table I) and reduced triphosphopyridine nucleotide. When this is done, the oxalosuccinic acid formed by carboxylation is reduced to *l*-isocitric acid by $TPNH_2$ which, in turn, is oxidized to TPN (reaction Ib):



Reaction Ib is the second step of the series by which carbon dioxide is fixed in this system. The combined result of reactions Ia and Ib is reaction Ic.

OXALOSUCCINIC CARBOXYLASE AND ISOCITRIC DEHYDROGENASE



Since both reactions Ib and Ic involve conversion of $TPNH_2$ to TPN and vice versa, they can be followed spectrophotometrically in either direction by allowing the reaction to take place in a quartz cell and measuring the absorption of light at wave length 340 mμ by the test solution. As mentioned above, reduced pyridine nucleotides strongly absorb light of this wave length. The molar extinction coefficient, which is defined by the equation:

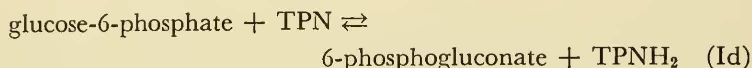
$$\alpha = \frac{\log I_0/I}{cl}$$

is 0.5644×10^7 (cm.^{2}/mole). For a transmittance of 95% and when $l = 1$ cm., the concentration of reduced pyridine nucleotide would be 0.04×10^{-7} moles per cc. In the case of TPN with a molecular weight of 743, 0.04×10^{-7} moles per cc. corresponds to 3 μg. of TPN per cc., or 0.8 μg. of isocitric acid per cc. This indicates the great sensitivity of the optical method and how suited it is for the study of reactions of this type.

By using this method, it has been possible to determine the equilibrium constants of reactions Ia, Ib, and Ic. The equilibrium constant of reaction Ib, K_b = (*l*-isocitrate) (TPN)/(oxalosuccinate) ($TPNH_2$), at *p*H 7.0 and 22° C., is approximately 0.3. That of reaction Ic, K_c = (*l*-isocitrate)(TPN)/(α-ketoglutarate)(CO_2)($TPNH_2$), at the same *p*H and temperature is, on the average, 1.3×10^{-4}. The equilibrium constant of reaction Ia can be calculated from these two values, since K_a = (oxalosuccinate)/(α-ketoglutarate)(CO_2) = $K_c/K_b = 0.5 \times 10^{-3}$. Thus, the equilibrium of reaction Ia is so unfavorable for carbon dioxide fixation that by this step alone only about 0.5% of the α-ketoglutarate would be carboxylated.

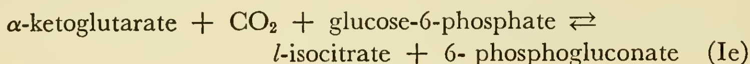
The TPN formed in reaction Ic can be reduced through a coenzyme-linked dismutation as discussed above. This results in the shifting of the equilibrium toward the side of carbon dioxide fixation. Such a shifting has been accomplished with the glucose-6-phosphate dehydrogenase system (see Table I) which catalyzes reaction Id. The combined result of reactions Ic and Id, when a mixture of glucose-6-

GLUCOSE-6-PHOSPHATE DEHYDROGENASE



$$\text{glucose-6-phosphate} + \text{TPN} \rightleftarrows \text{6-phosphogluconate} + \text{TPNH}_2 \quad \text{(Id)}$$

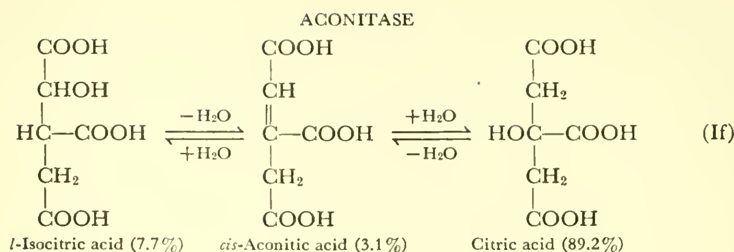
phosphate, α-ketoglutarate, and carbon dioxide is incubated with glucosephosphate dehydrogenase, oxalosuccinic carboxylase, isocitric dehydrogenase, manganese ions, and TPN, is reaction Ie:



$$\alpha\text{-ketoglutarate} + CO_2 + \text{glucose-6-phosphate} \rightleftarrows l\text{-isocitrate} + \text{6- phosphogluconate} \quad \text{(Ie)}$$

The equilibrium constant of reaction Ie has not yet been determined experimentally, but it can be calculated from free energy data. Thus, the free energy change of reaction Id can be estimated from the equation, $-\Delta F \doteq nF\Delta E$, relating free energy change to the potential difference between two reacting oxidation-reduction systems (15). The reacting systems in reaction Id are the glucose-6-phosphate $\rightleftarrows$ phosphogluconate ($E_0' = -0.43$ v. at *p*H 7.0) and the TPN $\rightleftarrows$ $TPNH_2$ system, the potential of which is unknown but can be considered to be near that of the DPN $\rightleftarrows$ $DPNH_2$ system ($E_0' = -0.28$ v. at *p*H 7.0). For a potential difference of $0.43 - 0.28 = 0.15$ v., the ΔF of reaction Id would be -6890 cal., corresponding to an equilibrium constant K_d = (6-phosphogluconate)($TPNH_2$)/(glucose-6-phosphate)(TPN) of the order of 10^5. Hence, the equilibrium constant of reaction Ie: K_e = (*l*-isocitrate)(6-phosphogluconate)/(glucose-6-phosphate)(α-ketoglutarate)(CO_2) $= K_c \times K_d = 1.3 \times 10^{-4} \times 10^5 = 13$.

A further shift of the equilibrium of reaction Ie toward carbon dioxide fixation occurs in the presence of aconitase. This enzyme is widely distributed in animal and plant cells and catalyzes the interconversion between *l*-isocitric, *cis*-aconitic, and citric acids according to reaction If, where the figures in parentheses give the percentage of the



l-Isocitric acid (7.7%) *cis*-Aconitic acid (3.1%) Citric acid (89.2%)

individual components present at equilibrium at 37° C. (18). Under these conditions, over 90% of the *l*-isocitric acid formed in step *3* is converted to *cis*-aconitic and citric acids.

The free energy changes of the various steps of this system are given in Table II. Combination of the four steps gives an over-all balance of about −3000 cal., *i. e.*, the complete system is exergonic by a fairly ample margin.

TABLE II

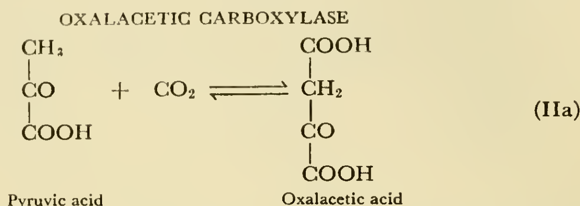
CARBON DIOXIDE FIXATION BY α-KETOGLUTARIC ACID

Step	Reaction	Enzyme	ΔF, cal.	Remarks
(*1*) *Carboxylation*	α-Ketoglutarate + CO_2 ⇄ oxalosuccinate	Oxalosuccinic carboxylase	+4460	Calc. for $T = 295°$ from equil. const.
(*2*) *Reduction*	Oxalosuccinate + $TPNH_2$ ⇄ *l*-isocitrate + TPN	Isocitric dehydrogenase	+ 708	Calc. as above
(*3*) *Reduction of pyridine nucleotide*	Glucose-6-phosphate + TPN ⇄ 6-phosphogluconate + $TPNH_2$	Glucosephosphate dehydrogenase	−6890	Calc. from $-\Delta F = nF\Delta E$
(*4*) *Over-all reaction for first three steps*	α-Ketoglutarate + CO_2 + glucose-6-phosphate ⇄ 6-phosphogluconate + *l*-isocitrate	....	−1722	$\Delta F = \Sigma$ of ΔF of partial reactions
(*5*) *Isomerization*	*l*-Isocitrate ⇄ citrate	Aconitase	*ca.* −1500	Calc. for $T = 310°$ from equil. const.

There is a possibility that the above reaction series might not stop with the formation of citric acid. Some microorganisms have been reported to split citrate to oxalacetate and acetate (2), a reaction that would favor carbon dioxide fixation via the α-ketoglutaric carboxylation system by displacing the equilibrium still further. Since, as we shall see later, acetate can be converted to pyruvate by reductive carboxylation and pyruvate forms carbohydrate in cells, the biological formation of acetate from citrate would be of considerable importance.

Carbon dioxide fixation by pyruvic acid. The primary re-

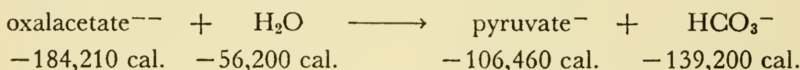
action of this system (5,9,11,12) involves the reversal of the decarboxylation of oxalacetic acid to pyruvic acid (reaction IIa). Reaction IIa



Pyruvic acid Oxalacetic acid

is catalyzed by oxalacetic carboxylase, an enzyme which is found in bacteria and liver and requires magnesium or manganese ions for activity. As in the case of oxalosuccinic carboxylase, the equilibrium of reaction IIa is very far to the left. Reversibility has been demonstrated by allowing the enzyme to act on oxalacetic acid in the presence of isotopic carbon dioxide. By stopping the enzyme action when about half of the oxalacetic acid was decarboxylated, the presence of isotopic carbon in the β-carboxyl group was demonstrated.

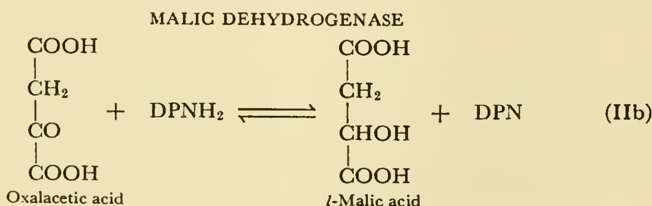
The equilibrium constant of reaction IIa has been calculated from its free energy change, in turn calculated from the free energies of formation (at 38° C.) of the substances involved (5):



oxalacetate^{--}	+ H_2O	$\longrightarrow$	pyruvate$^-$	+ HCO_3^-
$-184{,}210$ cal.	$-56{,}200$ cal.		$-106{,}460$ cal.	$-139{,}200$ cal.

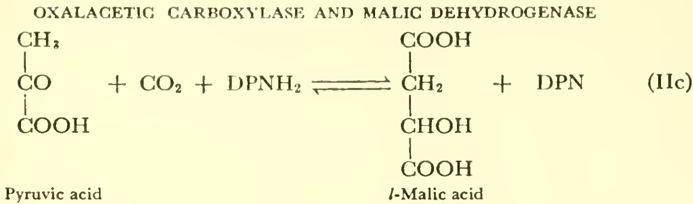
ΔF thus calculated is -5250 cal. for the decarboxylation, and K'_a = (oxalacetate^{--})/(pyruvate$^-$)(HCO_3^-) = 0.2×10^{-3}, a value of the same order of magnitude as that of reaction Ia.

Step *2* occurs when both malic dehydrogenase (see Table I) and reduced diphosphopyridine nucleotide are present, since the oxalacetic acid formed by reaction IIa is then reduced to *l*-malic acid (reaction IIb).

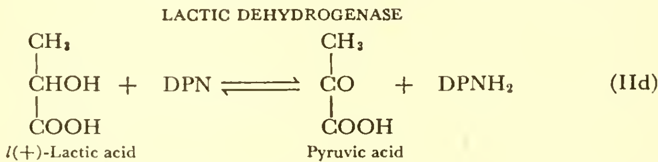


Oxalacetic acid *l*-Malic acid

The combined result of reactions IIa and IIb is reaction IIc:



In step *3*, the DPN formed in reaction IIc can be reduced by lactic acid through a coenzyme-linked dismutation with the lactic dehydrogenase system. Lactic dehydrogenase catalyzes reaction IId.



The combined result of reactions IIc and IId is reaction IIe.

$$\text{pyruvate} + CO_2 + l(+)\text{-lactate} \rightleftarrows l\text{-malate} + \text{pyruvate} \quad \text{(IIe)}$$

The free energy change of this reaction (see Table III), is about +1500

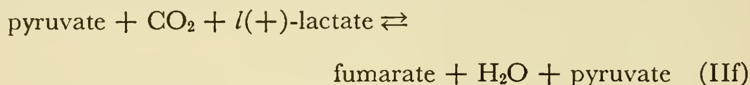
Table III

Carbon Dioxide Fixation by Pyruvic Acid

Step	Reaction	Enzyme	ΔF, cal.	Remarks
(*1*) *Carboxylation*	Pyruvate + $CO_2 \rightleftarrows$ oxalacetate	Oxalacetic carboxylase	+5250	Calc. from free energies of formation
(*2*) *Reduction*	Oxalacetate + $DPNH_2 \rightleftarrows$ *l*-malate + DPN	Malic dehydrogenase	−8300	Calc. from $-\Delta F = nF\Delta E$
(*3*) *Reduction of pyridine nucleotide*	*l*(+)-Lactate + DPN $\rightleftarrows$ pyruvate + $DPNH_2$	Lactic dehydrogenase	+4600	Calc. from $-\Delta F = nF\Delta E$
Over-all reaction (A)	Pyruvate + CO_2 + *l*(+)-lactate $\rightleftarrows$ *l*-malate + pyruvate [CO_2 + *l*(+)-lactate $\rightleftarrows$ *l*-malate]	....	*+1550*	$\Delta F = \Sigma$ of ΔF of partial reactions
(*4*) *Dehydration*	*l*-Malate $\rightleftarrows$ fumarate + H_2O	Fumarase	+ 705	Calc. from equil. const.
(*5*) *Further reduction*	Fumarate + 2 H $\rightleftarrows$ succinate	Hydrogenase; fumaric reductase (?)	−20,450	Calc. from $-\Delta F = nF\Delta E$
Over-all reaction (B)	Pyruvate + CO_2 + *l*(+)-lactate + 2 H $\rightleftarrows$ succinate + pyruvate + H_2O [*l*(+)-lactate + CO_2 + 2 H $\rightleftarrows$ succinate + H_2O]	....	*−18,195*	$\Delta F = \Sigma$ of ΔF of partial reactions

cal. and the calculated equilibrium constant, K'_e = (*l*-malate)/(*l*(+)-lactate)(CO_2), approximately 0.1.

Step *3* could also involve dismutation with any other diphosphopyridine nucleotide dehydrogenase system having an oxidation-reduction potential lower than that of the malic system (see Table I). We have considered the lactic dehydrogenase in this connection because there is evidence that it can participate in carbon dioxide fixation by pyruvic acid. In the presence of the enzyme fumarase, part of the *l*-malic acid formed in reaction IIe would be dehydrated to fumaric acid. Fumarase is widely distributed in plant and animal cells. Reaction IIe would then be replaced by reaction IIf:



$$\text{pyruvate} + CO_2 + l(+)\text{-lactate} \rightleftarrows \text{fumarate} + H_2O + \text{pyruvate} \quad \text{(IIf)}$$

Experimental support for the occurrence of the over-all reaction (IIf) is gained from the observation that fumarate, when added to an enzyme preparation from liver in the presence of pyruvate, DPN, and manganese ions, is converted to lactate and carbon dioxide (5). This indicates that reaction IIf can proceed from right to left. That it also proceeds from left to right is indicated by the presence of isotopic carbon in the carboxyl groups of the residual fumarate when the reaction is carried out in presence of isotopic carbon dioxide (29). The pyruvic oxalacetic system of carbon dioxide fixation has not yet been investigated by the methods used in the study of the ketoglutaric–oxalosuccinic system.

A considerable shift of the equilibrium of reaction IIf in the direction of carbon dioxide fixation can be brought about by reduction of the fumarate to succinate, a reaction that occurs in bacteria and liver tissue. In fermentation of glucose or glycerol by propionic acid bacteria, succinate is found to be one of the end products. Experiments with isotopic carbon dioxide have shown that the carboxyl groups of succinic acid become labeled. The carboxyl groups of malate, fumarate, and succinate, formed by pyruvate fermentation with *Escherichia coli* in the presence of carbon dioxide containing isotopic carbon, also show excess of heavy carbon. This is also the case when pyruvate is incubated with liver preparations. Some bacteria, *e. g.*, *E. coli*, can use molecular hydrogen for fumarate reduction and, in the

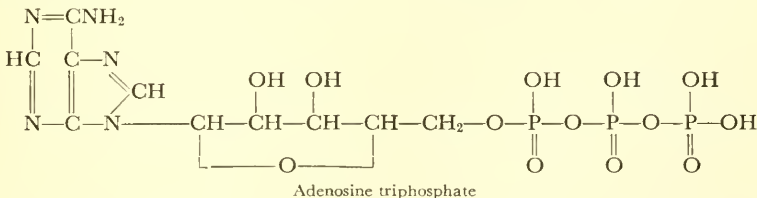
presence of pyruvic acid, carbon dioxide, and hydrogen, all three are utilized to form succinic acid (9,10,28).

The reduction of fumarate to succinate is strongly exergonic and provides ample energy to drive carbon dioxide fixation by this system to completion (see Table III).

Reductive Carboxylation

This type of carbon dioxide fixation has been discussed recently by Lipmann (17) and will only be briefly considered here. The fixation is a consequence of the reversal of the oxidative decarboxylation of α-keto acids catalyzed by specific enzymes. These reactions involve inorganic phosphate, and lead to the formation of an anhydride of the next lower fatty acid and phosphoric acid with liberation of either formic acid, or carbon dioxide and hydrogen; the hydrogen may either appear as molecular hydrogen or reduce a hydrogen acceptor (see reactions IIIa and IIIb). The anhydride bond formed has a high energy content and is generally referred to as an energy-rich phosphate bond. There are other types of energy-rich phosphate bonds of biological importance, such as enol phosphate, guanidine phosphate, and pyrophosphate bonds. The pyrophosphate type of bond has a special significance because, by the action of specific enzymes, it can give rise to any of the other phosphate bonds. Since the reactions involved in these conversions are reversible, it follows not only that any energy-rich phosphate bond can generate pyrophosphate bonds, but also that the various types of bonds can be converted into one another through the intermediate formation of pyrophosphate linkages.

The biologically important pyrophosphate group is the one present in adenosine polyphosphates, that is, adenosine triphosphate (abbreviated ATP), and adenosine diphosphate (abbreviated ADP).

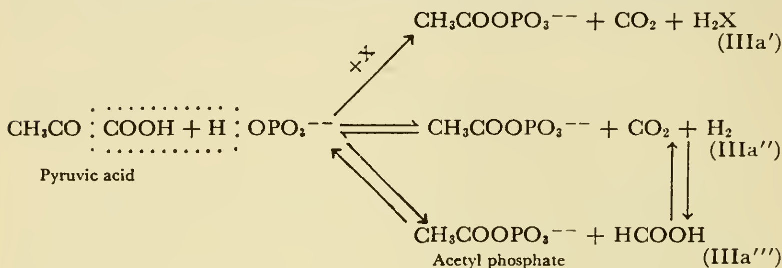


Adenosine triphosphate

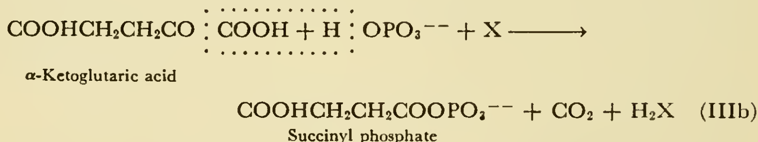
By enzymic hydrolysis, ATP is dephosphorylated to ADP, and this, in turn, to adenosine monophosphate (adenine ribose 5-phosphate or

muscle adenylic acid), with release of inorganic phosphate and of the energy of the bond. By enzymic transphosphorylation, ATP transfers phosphate to carboxyl, enol, and guanidine groups in a reversible manner. Since the energy content of the various types of energy-rich phosphate bonds is nearly the same, in the neighborhood of 12,000 cal., these transphosphorylations involve relatively small changes of free energy (16).

The general type of oxidative decarboxylations can be represented by reaction IIIa (17), where X stands for a hydrogen acceptor.



The oxidative decarboxylation of α-ketoglutaric acid generates phosphate bonds (19) and may be represented by reaction IIIb. All these reactions occur in bacteria; reactions IIIa′ and IIIb also occur in animal tissues.



Two reactions are known at present through whose reversibility reductive carboxylation can take place: (a) the splitting of formic acid to carbon dioxide and hydrogen; and (b) the splitting of pyruvic acid to acetyl phosphate and formic acid. Since acetyl phosphate can be formed enzymically by a reversible reaction between acetic acid and ATP, we have a biological system of reductive carboxylation of acetate to pyruvate.

The pattern of carbon dioxide fixation established for β-carboxylations must be modified here to include two preliminary steps, *viz.*, those of phosphorylation and of carbon dioxide reduction, respectively, and a final step for regenerating the ATP used at the beginning.

A possible series of reactions, modified from Lipmann, is suggested in Table IV. The absolute value of the free energy change of step *6*

TABLE IV

REDUCTIVE CARBOXYLATION

Step	Reaction	Enzyme	ΔF, calc.	Remarks
(*1*) *Phosphorylation*	Acetate + ATP ⇄ acetyl phosphate + ADP	From *Clostridium butylicum*	+3000	Calc. from equil. const.
(*2*) *Carbon dioxide reduction*	$CO_2 + H_2 \rightleftarrows$ HCOOH	Formic hydrogenylase (*E. coli*)	− 200	Calc. from equil. const.
(*3*) *Carboxylation*	Acetyl phosphate + formate ⇄ pyruvate + phosphate	From *E. coli*	+2800	Calc. from equil. const.
(*4*) *Reduction of carboxylation product*	Pyruvate + $DPNH_2$ ⇄ *l*(+)-*lactate* + DPN	Lactic dehydrogenase	−4600	Calc. from $-\Delta F = nF\Delta E$
(*5*) *Reduction of pyridine nucleotide*	3-Phosphoglyceraldehyde + phosphate + DPN ⇄ 1,3-diphosphoglycerate + $DPNH_2$	3-Phosphoglyceraldehyde dehydrogenase	*ca.* + 400	Calc. from equil. const.
(*6*) *Regeneration of ATP*	1,3-Diphosphoglycerate + ADP ⇄ 3-phosphoglycerate + ATP	From yeast	−3000	See text
Over-all reaction	Acetate + CO_2 + H_2 + 3-phosphoglyceraldehyde ⇄ *l*(+)-lactate + 3-phosphoglycerate	....	*ca.* −1600	....

has been considered to be the same as that of the step *1*, on the assumption that the bond energy of the anhydride linkage in 1-phosphoglycerate is the same as that of the acetyl phosphate group. The reaction series postulated in Table IV would lead from acetate to lactate and the over-all reaction would be exergonic. A characteristic feature of reductive carboxylation is that it must be started by input of a large amount of energy from energy-rich phosphate bonds.

Pyruvic acid formed by reductive carboxylation of acetate might also be converted to carbohydrate after reduction to triose instead of being reduced to lactic acid. Such a conversion would require additional phosphate-bond energy which might be amply available in chemosynthesis or photosynthesis.

It is not unlikely that reactions IIIa′ and IIIb might also be reversible. The main difference between these reactions and reactions IIIa″ and IIIa‴ is one of mechanism, *i. e.*, the hydrogen from the keto acid is first transferred to a hydrogen acceptor instead of being released either as molecular hydrogen or in combination with carbon dioxide as formic acid. If such were the case, all of the intermediate reactions involved in the oxidation of foodstuffs (*i. e.*, in respiration) would be reversible.

Carbon Dioxide Fixation in Chemosynthesis and Photosynthesis

It is now known that photosynthesis can be divided into two phases relatively independent of one another: (1) a so-called "dark" reaction which occurs in the absence of light and consists of a reversible fixation of carbon dioxide to form a carboxylic acid; and (2) the photolytic fission of water, yielding hydrogen with liberation of oxygen. Hydrogen produced by photolysis is used to reduce the products formed by carboxylation. In chemosynthesis, hydrogen is obtained by oxidation of inorganic compounds—a process that also supplies energy (6,7,23–25).

It would appear that essentially the same mechanisms that function in carbon dioxide fixation by heterotrophic organisms are operative in both photosynthesis and chemosynthesis, with the difference that, in the case of the heterotrophs, hydrogen and energy are derived by oxidation of organic materials. The widespread occurrence in plants of di- and tricarboxylic acids (malic, citric, isocitric) and of the enzymes that participate in their metabolism (fumarase, malic dehydrogenase, aconitase, isocitric dehydrogenase) lends support to such a view. Phosphorylation processes which, as we have seen, are essential for reductive carboxylations, are connected with carbon dioxide fixation in the sulfur oxidizing autotroph *Thiobacillus thiooxidans*, and photosynthetic organisms may utilize radiant energy for the synthesis of energy-rich phosphate bonds (3,22,26).

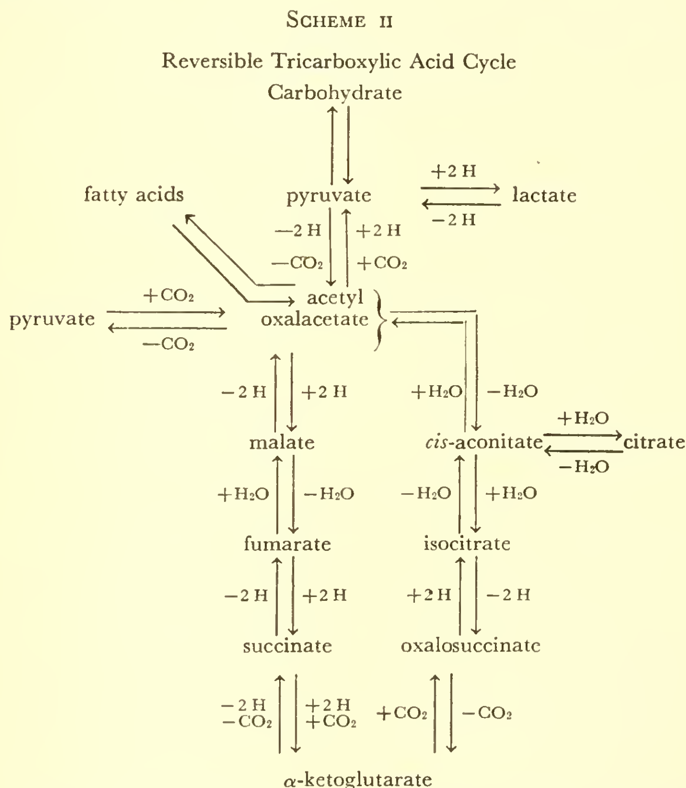
We cannot yet formulate in any detail the course of events in photosynthesis and chemosynthesis, but, using the knowledge gained by the study of the mechanisms of carbon dioxide fixation in heterotrophic organisms, we may attempt to draw a plausible picture of the chemical events. Reversal of the oxidative degradation of foodstuffs, *i. e.*, of respiration, would now seem to be a definite possibility.

We have seen that, on carboxylation and reduction, α-ketoglutaric acid can be converted to citric acid, and have indicated that the latter may be split to acetic and oxalacetic acids. Further, oxalacetic acid can be reduced to succinic acid by way of malic and fumaric acids, and succinic acid could be converted to α-ketoglutaric acid by reductive carboxylation, *i. e.*, by reversal of reaction IIIb. In this way, α-ketoglutaric acid would be regenerated, while the acetic acid

formed earlier can be converted to pyruvic acid by reductive carboxylation, *i. e.*, by reversal of reaction IIIa. We would thus have a cyclic mechanism whereby carbon dioxide and hydrogen entering at various points would emerge as pyruvic acid. The di- and tricarboxylic acids would only act catalytically as carriers of carbon dioxide and hydrogen. This is a reversal of the so-called tricarboxylic acid cycle, which is considered to be an important pathway for the oxidative breakdown of carbohydrate and fat in cells. Scheme II presents this metabolic cycle, incorporating recent findings concerning some of the intermediate reactions (1,4,13,28). For a more detailed discussion, see Lardy and Elvehjem (14).

SCHEME II

Reversible Tricarboxylic Acid Cycle



Another mechanism recently suggested for photo- and chemosyntheses involves a sequence of carboxylations and reductions leading

from a carboxylic acid, through the next higher α-keto acid, to an α-hydroxy acid, which in turn would be carboxylated to the next higher α-keto-β-hydroxy acid, and so on. Such a sequence would be essentially a reversal of a pathway of carbohydrate oxidation by way of phosphohexonic acid, α-ketophosphohexonic acid, phosphopentonic acid, etc., through alternating dehydrogenations and decarboxylations (17).

Obviously, the above schemes are only gross approximations. The main point is that what we know about the mechanisms of carbon dioxide assimilation by heterotrophic organisms strongly suggests that all reactions involved in cellular respiration are essentially reversible. Thus, the processes of both photosynthesis and chemosynthesis may represent reversals of the respiratory process not only from the standpoint of energy but also from the standpoint of the enzymic mechanisms.

References

(1) Buchanan, J. M., Sakami, W., Gurin, S., and Wilson, D. W., *J. Biol Chem.*, **159,** 695 (1945).

(2) Deffner, M., *Ann.*, **536,** 44 (1938).

(3) Emerson, R. L., Stauffer, J. F., and Umbreit, W. W., *Am. J. Botany*, **31,** 107 (1944).

(4) Evans, E. A., Jr., and Slotin, L., *J. Biol. Chem.*, **141,** 439 (1941).

(5) Evans, E. A., Jr., Vennesland, B., and Slotin, L., *J. Biol. Chem.*, **147,** 771 (1943).

(6) Franck, J., and Gaffron, H., in *Advances in Enzymology*, Vol. I. Interscience, New York, 1941, p. 199.

(7) Gaffron, H., *Biol. Rev. Cambridge Phil. Soc.*, **19,** 1 (1944).

(8) Kalckar, H. M., *Chem. Revs.*, **28,** 71 (1941).

(9) Kalnitsky, G., and Werkman, C. H., *Arch. Biochem.*, **4,** 25 (1944).

(10) Kalnitsky, G., Wood, H. G., and Werkman, C. H., *Arch. Biochem.*, **2,** 269 (1943).

(11) Krampitz, L. O., and Werkman, C. H., *Biochem. J.*, **35,** 595 (1945).

(12) Krampitz, L. O., Wood, H. G., and Werkman, C. H., *J. Biol. Chem.*, **147,** 243 (1943).

(13) Krebs, H. A., in *Advances in Enzymology*, Vol. III. Interscience, New York, 1943, p. 191.

(14) Lardy, H. A., and Elvehjem, C. A., *Ann. Rev. Biochem.*, **14,** 1 (1945).

(15) Lewis, G. N., and Randall, M., *Thermodynamics and the Free Energy of Chemical Substances.* McGraw-Hill, New York, 1923.

(16) Lipmann, F., in *Advances in Enzymology*, Vol. I. Interscience, New York. 1941, p. 99.

(17) Lipmann, F., and Tuttle, L. C., *J. Biol. Chem.*, **158,** 505 (1945).

(18) Martius, C., and Leonhardt, H., *Z. physiol. Chem.*, **278,** 208 (1943).

(19) Ochoa, S., *J. Biol. Chem.*, **155,** 87 (1944).

(20) Ochoa, S., *J. Biol. Chem.*, **159,** 243 (1945).

(21) Ochoa, S., and Weisz-Tabori, E., *J. Biol. Chem.*, **159,** 245 (1945).

(22) Ruben, S., *J. Am. Chem. Soc.*, **65,** 279 (1943).

(23) Ruben, S., and Kamen, M. D., *J. Am. Chem. Soc.*, **62,** 3451 (1940).

(24) Van Niel, C. B., in *Advances in Enzymology*, Vol. I. Interscience, New York, 1941, p. 263.

(25) Van Niel, C. B., *Physiol. Revs.*, **23,** 338 (1943).

(26) Vogler, K. G., and Umbreit, W. W., *J. Gen. Physiol.*, **26,** 157 (1942).

(27) Warburg, O., *Ergeb. Enzymforsch.*, **7,** 210 (1938).

(28) Werkman, C. H., and Wood, H. G., in *Advances in Enzymology*, Vol. II. Interscience, New York, 1942, p. 135.

(29) Wood, H. G., Vennesland, B., and Evans, E. A., Jr., *J. Biol. Chem.*, **159,** 153 (1945).

13

HORMONES

B. A. HOUSSAY, DIRECTOR OF THE INSTITUTE OF BIOLOGY AND EXPERIMENTAL MEDICINE, BUENOS AIRES

Definition and Significance

THE HORMONES are specific chemical substances produced by an organ or tissue which, after being discharged into the circulating fluids (*milieu intérieur*), may reach all parts of the organism and in small amounts markedly influence the functions of other organs or systems without themselves contributing important quantities of matter or energy.

This definition differentiates them from other important substances which also reach the circulating fluids, such as: (*a*) *nutritive substances* which supply the tissues with materials and energy as, for example, glucose, amino acids, lipids, etc.; (*b*) *vitamins*, organic chemical regulators contained in the food; (*c*) *chemical mediators of nerve action*, liberated by the nerve endings in the close vicinity of effector or other nerve cells and which exert a localized action; (*d*) *the organizers*, embryonic substances of regional origin which govern the differentiation of a determined organ, even if transplated to another zone or cultivated *in vitro;* (*e*) *the parhormones* (Gley) which, although being excretory substances produced by the metabolic processes of all the tissues, nevertheless perform important regulatory functions, as is the case for carbon dioxide in the regulation of respiration.

The action of what we today call glands of internal secretion is due to the hormones they produce; and the insufficiency of these glands, therefore, is only a matter of hormone deficiency.

In a wider sense, which is not usual, internal secretion would mean any specific cellular elaboration discharged into the internal medium. It is in this sense that Claude Bernard considered glucose, which is produced in the liver and then passes into the blood, as an internal secretion.

As we all know, the glands of internal secretion are referred to as such because they pour their elaborated products into the blood, in contrast to the glands of external secretion which pour their products out of the body or in cavities which communicate with the external medium.

The first demonstration of a hormonal action was the induction of the comb of a capon by testicular graft (Berthold, 1849). The expression "internal secretion" was used for the first time by Claude Bernard (1855) to point out that the liver "shed" sugar in the blood. The concept of internal secretions as we now understand it is due to Brown Sequard who, in 1891, stated in a paper with Arsonval, "Nous admettons que chaque tissue, et plus généralement, chaque cellule de l'organisme sécrète pour son propre compte des produits ou des ferments spéciaux qui sont versés dans le sang et qui viennent influencer par l'intermédiaire de ce liquide toutes les autres cellules rendues ainsi solidaires les unes des autres par un méchanisme autre que le tissue nerveux." The name "hormones" was proposed by Hardy to designate the "chemical messengers" (Bayliss and Starling, 1904) which are secreted in the blood by one organ to stimulate the functions of another.

Hormone etymologically means "I arouse the activity" or "I excite," but we are now aware of the existence of inhibitory hormonal actions. It is useless to classify the hormones as stimulating or inhibitory, inasmuch as the same hormone may often produce both effects, depending upon its concentration or the organ which it affects. It is understandable, therefore, why the designations proposed by Sharpey-Schafer, who gave them the general name of "autacoids" and subdivided them into "hormones" (exciting) and "chalones" (inhibitory), did not gain currency.

Chemical Nature

According to their chemical nature the known hormones can be classified into three groups:

(*1*) *Phenolic Derivatives.*—Both adrenalin, the hormone of the

adrenal medulla, and thyroxine, the hormone of the thyroid, are phenolic substances.

(*2*) *Proteins.*—To the proteins belong the hormones of the hypophysis (two gonadotropins, thyrotropin, adrenocorticotropin, lactogenic hormone and growth-promoting factor from the anterior lobe, and the vasopressor, oxytocic and melanic-expanding principles of the posterior lobe); the hormone of the pancreas (insulin), and the hormone of the parathyroid (parathormone).

(*3*) *Steroids.*—To the steroid family belong the hormones isolated from the ovarium and corpus luteum, testicle, and adrenal cortex, twenty-eight steroids having been isolated so far from the adrenal cortex (Reichstein, 1943). They all have in common a cyclopentanoperhydrophenanthrene ring system.

Attempts to prepare the nonprotein hormones synthetically have followed closely the determination of their chemical constitution. Besides, several synthetic substances of constitution similar to, but not identical with, the natural hormones have been prepared, and their pharmacologic action tested. Many vasoconstrictor and bronchodilator substances have thus been obtained which are more efficient than adrenalin for some therapeutic uses. Several synthetic estrogens (diethylstilbestiol, hexestrol, etc.) have also been obtained which are now largely used instead of natural estrogens to treat ovarian insufficiency. Desoxycorticosterone, a substance which is able to maintain adrenalectomized animals in normal condition, has been prepared by partial synthesis. Probably a wide field for pharmacological investigation of corticoadrenal hormones will soon be opened, since Reichstein (1943) announced a method of preparing steroids containing an atom of oxygen linked to carbon atom 11. Steroid hormones are usually prepared by partial synthesis, starting from stigmasterol, a product of soybeans.

The active substances extracted from an endocrine gland do not always represent the circulating natural hormone. Thus, while thyroxine increases the oxygen consumption if given to the whole animal, it does not have such an action upon isolated tissues. Furthermore, whereas only from 25 to 50% of organically bound iodine is extractable from the thyroid in the form of thyroxine, all the iodine compounds of the thyroid are able to increase the rate of metabolism. These and other reasons make it dubious that thyroxine is a constituent of the thyroid hormone or the hormone itself.

Some of the protein hormones elicit such a small antibody response, as in the case of insulin, that their administration by repeated injections remains effective during scores of years. Other protein hormones become progressively less effective, which makes it necessary to increase the dose each time, as in the case of the parathyroid hormone. There are still other hormones (thyrotropin, gonadotropin, and in different degrees all the anteropituitary hormones) whose action decreases quickly if they are administered daily and which induce the formation of antihormones. Thus, if an animal is treated with daily doses of thyrotropin, it shows, at the beginning, hypertrophy and hyperfunction of the thyroid, but after a few weeks the action disappears and is followed by atrophy and hypofunction of the organ. The serum is then found to contain antithyrotropin, which not only inhibits the thyrotropin action but is also capable of inhibiting the action of the thyroid gland as shown by injecting such serum into another animal.

Antihormones are only produced when protein hormones are administered parenterally. Collip (1934) thought that they were substances of physiological importance and that each hormone should have its corresponding antihormone to balance its effects. But it now seems that antihormones are antibodies or immunity mechanisms (Rowlands) reacting to injection of antigens from another species. It is to be noted that adrenalin, thyroxine, and the steroid hormones do not produce antihormones and are not proteins. For some of the actions of these hormones a certain habit may be produced without any demonstrable antihormones.

The natural protein hormones do not seem to be completely equivalent to those which are extracted in the laboratories, for they do not induce the formation of antihormones. Thus, when a rat is castrated, great quantities of gonadotropins accumulate in its blood, as is readily demonstrated in parabiosis experiments—that is to say, experiments involving sewing two rats together side by side, after opening their bellies laterally, so that their peritoneal cavities, muscles, skin, and blood vessels are fused. In this way, the gonadotropin present in the blood of the castrated rat passes into the circulation of the normal rat, which, as a consequence, shows an intense stimulation of the gonads. This stimulation persists steadily for months, while parabiosis lasts, with no sign of antihormone production, in contrast to the

rapid formation of antihormones and the inversion of effects caused by the injection of gonadotropins isolated from the gland.

The preparation of active protein hormones that will not produce antihormones is one of the outstanding problems endocrinology must solve. Meanwhile, the formation of antihormones imposes an important limitation on prolonged therapeutic application of certain hormones, particularly those of the pituitary and parathyroid. This is one of the causes of the discrepancy existing between, on the one hand, the great functional importance of the hypophysis as demonstrated by experiment and by the study of disease in man, and, on the other, the limited possibilities realized thus far from therapeutic applications.

The chemical mechanisms by which hormones act upon cells are not yet known. It has not yet been proved that they participate in enzyme systems, as is the case for vitamins such as thiamin, niacin, and riboflavin. The hormones are chemical regulators that probably modify some link in the chain of metabolic reactions, a field of study which has remained almost unexplored to date, in spite of its great importance.

Role

Some endocrine organs, phylogenetically, begin as external secretion glands that cast their secretions into the digestive system but afterward lose their excretory function and change into internal secretion glands. The pituitary, thyroid, and pancreatic islets of Langerhans have evolved in some such way. The parathyroid, regulator of the metabolism of calcium and phosphorus, derives from the branchial arches. The ovarium, testicle, and adrenals, which produce the steroid hormones with sexual and metabolic actions, derive from the coelomic epithelium. The adrenal medulla that secretes adrenalin derives from the nervous sympathetic system, and the neurohypophysis from the diencephalon.

The hormones of the vertebrates are better known than those of the invertebrates. In the latter, certain processes have been shown to be regulated by hormones, such as metamorphosis, color (in the case of crustacea), and sexual dimorphism in some instances. The hormones regulate functions that exist before hormones appear and which often persist without them. Thus, all animal and plant cells consume glu-

cose without the intervention of insulin, but in the great majority of the vertebrates insulin is a new regulating mechanism of such importance that, when missing, diabetes results, a disease which is fatal sooner or later depending upon the species. Sexuality exists in many invertebrates without intervention of hormones, but in the vertebrates sexual characters do not reach their full development without the action of ovarian or testicular hormones, the production of which is governed by the pituitary gonadotropins. The functional unity of the organism is assured by nervous and humoral (chemical) mechanisms of correlation, which interrelate the different parts or tissues and regulate their reciprocal activities. Sherrington has pointed out the unifying integrative action of the nervous system; a similar role is played by the humoral factors, among which are the hormones. Both types of mechanism maintain the stability of the *milieu intérieur* (Cl. Bernard) and of the organism as a whole (Cannon's homeostasis) in spite of varying conditions in the external environment and in the organism's activity. Modern studies have thus confirmed, extended, and given a more precise meaning to the old and vague notions about the correlations between the organs ("consensus partium" or "sympathies").

The roles played by the hormones can be classified somewhat conventionally as follows:

(*1*) *Metabolism.*—Some of the hormones regulate the balance of metabolic processes. Their action may be general (stimulation of oxidation processes by thyroxine) or rather specialized (parathyroids and calcium, insulin and carbohydrate). One and the same hormone may modify several metabolic processes, *e. g.*, adrenal steroids act both upon the metabolism of water and salt, and upon the metabolism of carbohydrates.

(*2*) *Morphogenesis.*—The morphogenetic actions are the consequence of the selective role of the hormones in assimilation and growth phenomena. Some endocrine glands such as the pituitary, thyroid, parathyroid, and sexual glands play an important role in growth during a certain stage of development, principally because of their action upon the synthesis of proteins and on the development of bone. In other cases the growth-promoting action is exerted on special organs, as in the case of estrogens which promote the growth of the uterus and mammary gland.

As the result of growth and differentiation, the proper morpho-

logical constitution of each sex and of each individual is attained. The pars distalis of the pituitary gland stimulates the thyroid of tadpoles, and in turn the thyroid secretion promotes the metamorphosis of the larval form. The pars intermedia of the pituitary, and sometimes adrenalin, regulate the color of the skin of amphibians and fishes, by dispersing or concentrating the pigment granules in the chromatophore cells.

(*3*) *Endocrine Interrelation and Balance.*—A close functional relation exists between the endocrine glands. Following Gley, we can, consider "correlation" as the relation of one organ with another; when it is a mutual correlation, we may call it "interrelation." Thus, while the secretion of pancreatic juice caused by secretin is an example of "correlation" between the duodenum and the pancreas, there exists an interrelation between the hypophysis and the gonads, for if the anterohypophysis is responsible for the final development and maintenance of the functional activity of the gonads, at the same time the endocrine secretion of the latter regulates and moderates the gonad-stimulating function of the hypophysis.

Analysis of each separate endocrine gland or of each hormone is clearly artificial. It was initially necessary because in order to study the behaviour of a gland, there were no procedures available other than extirpating the organ, injecting its extracts, and carrying out anatomical and functional studies of clinical cases showing visible alterations of some gland. Afterward the method persisted for didactic reasons; but since no endocrine gland can be considered as if its action were independent of other glands, the method is now being abandoned.

A constellation of endocrine glands exists whose central organ is the hypophysis, the function of each gland being influenced more or less by the function of the others. Because the anterohypophysis contributes to the development and maintenance of the structure and function of various glands, extirpation of the pituitary produces a marked atrophy and hypofunction of the thyroid, gonads, and adrenals, to such an extent that it has been said that the hypophysectomized rat is an endocrinological and metabolic ruin. But, at the same time, the structure and function of the anterohypophysis is governed by hormones secreted by the thyroid, gonads, and adrenal cortex.

Each gland produces very specific hormones which play important roles; nevertheless these hormones constitute simple parts of

complex functional mechanisms. Thus, the regulation of the metabolism of carbohydrates involves the liver, insulin, hypophysial, corticoadrenal, and thyroid hormones, as well as participation of the intestine, kidney, and muscle. Sexual functions depend upon the combined action of the hormones of the hypophysis, ovary, corpus luteum, and placenta (during pregnancy), plus the action of still other glands, such as the adrenal cortex and the thyroid.

As explained further on, in any single function of the organism more than one hormone plays its part, even when one of them exhibits a very specific and preponderant role.

(4) *Sexuality and Reproduction.*—Among the endocrine actions, the sexual functions are especially important. These functions depend on the hormones of the gonads, governed in part by the pituitary gonadotropins, in part by the hormones of the adrenal cortex and, to a certain degree, by the thyroid. They regulate the production of the female (ovule) and male (spermatozoid) germinal cells or the development of the organs which carry them, the development of the impulses and sexual acts that lead to fecundation, the progress of pregnancy and delivery, and the secretion of milk. The sexual hormones are indispensable links for individual sexuality and for the maintenance of the species.

(5) *Mental and Nervous Functions.*—The hormones influence several nervous and muscular activities. We only need to remember the differences shown by the sexes, the mental dullness due to hypothyroidism, the nervousness or mental instability of hyperthyroidism, the tetany of hypoparathyroidism, etc. These actions of the hormones are probably due to their influence on the metabolism of the nervous system, a point that has not received much study.

(6) *Vital Role.*—The extirpation of some endocrine glands causes death. This was attributed to hypothetical poisons which accumulated because they were not neutralized by the endocrine glands (antitoxic role), a theory that has been abandoned because no poison has been demonstrated and because it has been proved that death was due to metabolic disturbances provoked by the lack of the hormones normally produced by the organ in question.

(7) *Resistance.*—The hormones are important factors in building up the resistance of the organism to certain hazards such as high or low temperatures, low or excessive oxygen tension, overexercise, in-

fections, toxins, trauma, various poisons, etc. For example, pituitarectomized or adrenalectomized animals show a low resistance to all of these, and also are very sensitive to hypoglycemia and hypotensor agents or to the circumstances which provoke shock or hypothermy. Many of these responses depend upon metabolic phenomena. Some endocrine factors also influence the production of immunity antibodies or anaphylaxis and resistance to some infections. The harmful agents or circumstances (*e. g.*, cold, fatigue, toxins, etc.) all produce similar reactions in a given organism, but the nature of these reactions varies in different species. The adrenals are largely responsible for these reactions, which pass through several phases: an initial "alarm reaction" followed by temporary compensation, and finally, decompensation. They have been thoroughly studied by Selye.

Abnormal internal secretions. There are internal secretions produced under abnormal conditions. Thus, when the arterial hypertension is produced because the ischemic kidney produces renin, which acts enzymically upon the hypertensinogen of plasma to form hypertensin, a substance which increases the arterial blood pressure. During arterial hypotension, renin is secreted in the blood; but it has not yet been demonstrated with certainty that there is normally a small secretion of renin.

Hormones and cancer. In certain cases, hormones enhance the development of tumors. Their role seems to consist more in promoting the growth of tumors than in initiating the malignant cellular transformation. The extirpation of glands (hypophysis, adrenals, gonads, etc.) or the injection of hormones may hasten or delay the growth of several types of tumors. Thus, testicular castration and injection of estrogenic hormones retard the development of prostatic cancer, whereas testicular hormone accelerates it.

Some hormones promote the development of cancer. Thus, mammary cancer is less frequent in males or in castrated females of a strain of rat showing high incidence of this cancer in the adult female. The injection of estrogens (Lacassagne) stimulates growth of the mammary gland and produces cancer in many of the males of that strain. In these cases, it is debatable whether the estrogens initiate the cancer or merely promote its development. But recent work has shown that estrogenic induction of mammary cancer frequently is obtained in strains of rat in which the cancer occurs spontaneously only in rare

instances (Geschickter and Byrnes). Some estrogenic hormones have curious tumorigenic actions on the guinea pig, while other hormones can prevent this action (Lipschütz).

Benign and malignant tumors of the endocrine glands can produce exaggerated quantities of hormones, as shown in cases of hyperthyroidism, hyperparathyroidism, hyperinsulinism, acromegalia or giantism, and some adrenal, ovarian, and testicular tumors. Complex effects result from the action of adrenal tumors: They may exert either virilizing or feminizing actions, influence metabolic phenomena, alter the body shape, or modify arterial blood pressure. The ovarian tumors also produce varied effects such as feminizing or virilizing actions. Less familiar is the humoral mechanism by which the malignant tumors affect the metabolism of the whole organism.

Specificity

Endocrine glands and hormones have considerable specificity of action. The ovary stimulates the development of the feminine sexual characters, and the testicle that of the male sexual characters, as can be demonstrated by castration, which prevents their development or provokes their regression in both sexes. Conversely, the ovarian graft, or the administration of estrogenic hormones, develops the feminine characters in either castrated or entire animals. In the same manner, testicular grafts or androgenic hormones develop the masculine characters in either male or females, whether castrated or not.

Each hormone produces its characteristic action upon diverse animal species. Thus, the insulin of a given animal produces hypoglycemia in all mammals. It is also a rule that the hormone extracted from an animal is active in all other species of mammals. Thus, insulin extracted from a variety of mammals produces hypoglycemia in the rabbit, and insulin from bovine origin is active on all the vertebrates on which it has been tried. But some rare exceptions are known. Thus, gonadotropins from mammals have no effect on the gonads of the toad, *Bufo arenarum*, the cause of the anomaly being unknown.

Some hormones are elaborated by more than one gland. For example, adrenalin is secreted by the chromaffin tissue but also exists in the cutaneous poison of the toads. The estrogenic hormones are found in the ovary, placenta, and adrenal cortex; and it is remarkable

that one of the most abundant sources for industrial production is the urine of the stallion (in which it decreases after castration). The urine of men and women also contains estrogens and androgens. The androgenic hormones, secreted normally by the testicle, can be secreted by an ovary grafted in the ear (Hill) if the ear is maintained at a low temperature. They are also produced by some adrenal tumors and some ovarian tumors (arrhenoblastoma).

The specificity of response of the reactive organs is not absolute. Androgens can produce some effects on the endometrium, the vaginal epithelium, or the mammary gland. Estrogens can slightly influence the seminal vesicles. The injection of estrogens can provoke masculine erotization in some adult males, either normal or castrated.

The relationship between the chemical constitution and the effect of the hormone is so close that a small modification of its molecule can profoundly change its actions. Thus, ethyl testosterone prepared from the male hormone is very active upon the endometrium. The recent book of Selye shows much of the multiplicity and complexity of the actions of the steroid hormones.

Synergies and Antagonisms

It would be impossible in a short essay to enumerate all the cases in which two hormones either strengthen (synergy) or oppose (antagonism) one another's actions. Estrogens in a certain adequate dose prepare the uterus and mammary glands, and sensitize them to progesterone; nevertheless, in other doses these substances can nullify each other's actions. Estrogenic and androgenic hormones have, in certain cases, antagonistic actions; in others, their actions are independent and do not interfere with each other; and, finally, sometimes they are mutually strengthened.

Regulation of Secretion of Hormones

Even though, for each function, several organs play a part, their participation is regulated so as to maintain a steady balance, as is demonstrated by the constancy of the blood sugar level, of the oxygen consumption, of the blood calcium, etc. These regulations are, therefore, factors in the functional unity of the organism and in the equilibrium of their functions.

REGULATION OF EACH GLAND

Each endocrine gland has its own regulating mechanisms for the secretion of its hormones. As judged by experiments on extirpation and restitution, there must be a basal secretion, generally uninterrupted. In certain instances this has been well demonstrated (adrenalin, insulin). This secretion is submitted to regulating factors, the principal ones being humoral, and in some cases also nervous. Thus the basal amount of insulin secreted by the islets of Langerhans depends on the blood sugar level; and, reciprocally, the blood sugar level depends on the amount of insulin secreted. The basal secretion of insulin increases when glycemia increases; conversely, it decreases when the blood sugar level is lowered. The parathyroid secretion increases when calcemia decreases. In both cases the secretion of either insulin or parathormone tends to restore the altered equilibrium of the internal medium. A similar regulation of the secretion of the thyroid hormone is probable, udging by the constancy of basal metabolism.

The regulation of mechanisms of hormone secretion are very precise and well designed to attain their objective. Thus, one-seventh of the pancreas is adequate to maintain the basal glycemia at the normal level; also, the basal glycemia continues at a normal level in an animal even if four pancreases are grafted by vascular anastomosis. These facts show that a normal secretion of insulin is maintained either with one-seventh of the pancreas or with five pancreases—this because glycemia governs insulin secretion. Only in abnormal cases (diabetes or hyperinsulinism) is regulation of insulin secretion deficient or excessive so that the gland works at a new level. In some cases of hyperplasia, adenomata, or cancers, the endocrine glands have been known to produce hormonal hypersecretion.

Although the reduction in mass of an endocrine organ may not alter its ability to function under basal conditions, it may lead to its insufficiency in cases of emergency. The pancreas of the dog reduced to one-fifth of its mass is enough to maintain normal glycemia, but if grafted to a diabetic dog it does not replace a normal pancreas in correcting the existing hyperglycemia. The resistance of the surgically reduced pancreas is diminished against the action of injurious agents such as extracts of the anterohypophysis and thyroid and, in consequence, diabetes develops readily. When the pancreas is normal, these

injurious agents either do not, as in the case of sugar or thyroid extracts, produce diabetes or else do so, as in the case of anteropituitary extracts, only if much higher doses are used. In most instances the secretion of the hormones is governed by humoral factors, while the nervous factors play only an accessory, dispensable role. Thus, denervation of pancreas, thyroid, adrenal cortex, or gonads does not produce any insufficiency of their endocrine functions. Sometimes it is found that the nervous action makes the secretory regulation somewhat quicker and more precise, as in the case of insulin secretion.

There are cases, however, in which the nervous regulation is important. Thus the splanchnic nerves exert a tonic and emergency action upon adrenalin secretion and their section reduces it to traces. The supraoptic nucleus governs secretion of the antidiuretic hormone of the neurohypophysis and section of the supraoptic neurohypophysial fibers is followed by insipid polyuria. Stimulation by light or desiccation inhibits the melanic-expanding secretion of the pars intermedia of the hypophysis through the medium of the nervous system. After cutting the pituitary stalk, the anteropituitary secretions are produced in sufficient amount so that, generally, the thyroid, adrenal, and gonads are not modified. But there is no increase in the secretion in cases of emergency; and in animals with the pituitary stalk cut ovulation is no longer produced by mating (*e. g.*, doe), embracing (*e. g.*, toad), or vision (*e. g.*, dove) as in normal animals. Intense light can induce the secretion of enough hypophysary gonadotropins to produce hypertrophy of the gonads when they are in the atrophic condition of winter rest. Cold no longer exerts its action on ovarian cycles or upon thyroid or adrenal cortex when the pituitary stalk has been cut in rats. Certain hypothalmic injuries may decrease the secretion of either all or some of the pituitary gonadotropins.

REGULATION OF COMPLEX EQUILIBRIA

The individual regulation of each gland is at the same time submitted to more ample regulations which are often reciprocal. Thus, the blood sugar level is maintained constant in spite of the coexisting secretions which tend to produce either hypoglycemia, such as insulin, or hyperglycemia, such as those of the hypophysis and adrenal glands. There exists, therefore, a functional equilibrium of the secretion of each endocrine gland and, at the same time, a functional equilibrium of all

the secretions of a similar functional constellation (for example, the endocrine secretions which regulate sex or those that govern carbohydrate metabolism, etc.). These regulations are mainly humoral, depending to a great extent on endocrine factors, although the nervous system often has an important share.

The tendency of the endocrine glands to reach a functional equilibrium and then to maintain it without overshooting seems to conform to a sort of general law. Thus, when an ovary is extirpated, the remaining ovary produces the same number of ovules and cycles ("law of follicular constancy" of Lipschütz). A testicular fragment either assures the total development of the sexual characters of a cock, or else it atrophies with no intermediate stable equilibrium being set up (the "all or none" law of Pezard and Gley). The tendency to all or none activity of the gland *in situ*, maintaining its secretion at a constant level, independent of its mass, is not inconsistent with the fact that the pharmacological effects of the hormones vary with the doses, the relationship following the typical S-shaped curve.

The close associations existing among the glands of internal secretion explain why the disturbances affecting one of them are generally reflected in the others. It is rarely, if ever, that experimental or pathological disturbances of an endocrine organ are observed without modifications of the other glands. Thus, the extirpation of the hypophysis leads to atrophy and hypofunction of the adrenal cortex, thyroid, and gonads.

The actions of each hormone can be direct or indirect. The testicle and androgenic hormones, for example, produce hypertrophy of the seminal vesicles and prostate directly. The pituitary gonadotropin (LH or ICSH) produces the same action, but since it induces the secretion of testicular hormones is inactive in the absence of the testicle.

In certain cases the functional interactions of endocrine glands are simultaneous; in other cases they follow each other in sequence, as the proliferative (estrogenic) and secreting (progesteronic) phases of the endometrium during the menstrual cycle.

HYPERSECRETION OF HORMONES

That the secretion of the hormones is not normally at a maximum is borne out by several facts: (*a*) castration very much increases the secretion of anterohypophysary gonadotropins: (*b*) adrenal

ablation very much increases the secretion of pituitary corticotropin; (*c*) in clinical cases of endocrine hyperfunction supernormal effects are observed, similar in many cases to those produced by excessive administration of hormones.

Hyperfunction of the endocrine glands is observed under several circumstances: (*a*) pathologic hyperplasias and tumors (benign, or adenomata, and malign) in which are produced exaggerated secretion of either normal or abnormal hormones; (*b*) injection of either glandular extracts or hormones; (*c*) excitation by a supernumary gland (for example, parabiosis of a castrated animal with a normal one); (*d*) rupture of the endocrine equilibrium.

In some cases either hypo- or hyperfunction takes a certain time to develop. Thus if 95% of the pancreatic tissue of a rat is removed, the residual tissue is still able to keep a normal blood sugar level for two or three months, but later on a progressive diabetes develops. Following certain operations on the ovary (grafting, ligation, partial fragmentation), a hypersecretion of pituitary gonadotropin gradually develops, which, by promoting excessive secretion of ovaric estrogens, produces a marked hypertrophy of the uterus, hyperplasia of the endometrium, etc. In this case a new hypophyso-ovaric equilibrium is established at an abnormal level.

Types of Endocrine Functional Associations

The functional associations belong to various types. (*1*) A gland, such as the anterohypophysis, can develop and maintain the structure and function of one or several other glands, as the thyroid, adrenal cortex, ovary, or testicle. (*2*) One gland can moderate the function of another, *e. g.*, the sexual hormones moderate the gonadotrophic pituitary function.* (*3*) Actions, sometimes antagonistic and sometimes synergetic, can be observed between two glands (or their hormones), as in the case of the ovary (or estrogens) and the corpus luteum (or progesterone). (*4*) Certain hormones increase the sensitivity to others, *e. g.*, estrogens to the effect of progesterone upon the endometrium or upon the mammary gland, and thyroxine to the effect

* In fact, the position is more complex, because estrogens, if given in high doses, may induce an increased secretion of luteinizing, adrenotrophic, and lactogenic hormones of the anterior pituitary, but, in larger doses, suppress them all.

of adrenalin. (*5*) Certain hormones can produce an insufficiency by damaging an organ; thus anteropituitary extracts and, in certain cases, those of the thyroid as well, by a repeated action bring about the disappearance of the β-cells of the islets of Langerhans and produce a pancreatic diabetes, the cause of which could not be deduced by any one not familiar with the previous history.

METHODS OF STUDY

Many methods are required to study the endocrine glands and their hormones.

Morphological. Important data can be obtained by the study of the weight and macro- and microscopic structure of the endocrine organs of different ages and under different conditions, created by extirpation of other organs and injections of hormones. It is also necessary to watch for the initiation and localization of the changes.

Chemical. These include: (*a*) a search for and isolation and purification of the hormones by means of a combination of biological assays and chemical methods; (*b*) a study of the chemical changes and metabolic modifications produced in an animal by the suppression or administration of hormones; (*c*) action of the hormones on tissue slices or on chemical systems *in vitro;* and (*d*) a study of the origin, metabolism, and excretory products of the hormones.

Physiological. These can be subdivided into experimental and clinical methods. The experimental method includes the study of: (*a*) glandular insufficiency and restitution through grafting, implants, or administration of either extracts or hormones; (*b*) hyperfunction induced by the methods already described; (*c*) measurement of the hormones in the organ, in the blood that comes away from it, in that of the general circulation, or even in the urine. It can also be indirectly measured by finding the amount necessary to substitute for the removed organ (substitution method). It is much safer to measure the hormone secreted by an organ than the amount the organ contains, because the amount secreted does not always vary parallel with the amount present in the organ.

The clinical study is very valuable: (*a*) because it furnishes human data and (*b*) because the disease is a spontaneous experiment, the conditions being sometimes more delicate and varied than the experimental methods can secure. Many endocrine functions have

been studied first clinically and afterward experimentally, *e. g.*, Addison's disease, acromegaly, etc.

Quantitative Determination of Hormones. The hormone or its derivatives can be determined in the organ, blood, or urine. The determination can be made by chemical, physical, or biological methods. The low concentration of the hormone rarely allows direct gravimetric estimations; in general, these must be based on colorimetric, spectrophotometric, or chromatographic methods. Biological assays are performed on entire animals, isolated organs, or tissue slices. Animals with insufficiency may be used. In some cases the organ is forced to function overloaded, *e. g.*, a pancreas grafted in diabetic dogs by vascular grafting is tested for the time necessary to correct hyperglycemia.

Hormones are assayed by comparing their effects with those produced by international standards, thus avoiding differences of sensitivity encountered in different races of animals. In order to confirm that a given function depends on an endocrine organ, the following consequences must be etablished: (*1*) the ablation or injury of the gland must produce an insufficiency of that function; (*2*) this deficiency must be compensated for by the graft or implantation of the gland, or by the injection of its extract or its hormone; (*3*) an excess of these substances must induce hyperfunction symptoms opposed to those of hypofunction; (*4*) the anatomoclinical facts must agree with the experimental findings; (*5*) both spontaneous and induced hyperfunction must be improved by treatments which either eliminate the gland or decrease its action.

In order to admit that a given action due to an endocrine gland is produced through the mediation of another gland, it is necessary to show: (*1*) that the removal of the second gland is followed by a condition of hypofunction, just as, or even more, pronounced than that caused by the ablation of the first; (*2*) that the disturbances cannot be corrected either by implantation or injection of extracts or hormones of the first gland when the second one has been removed. For example, (*a*) ovariectomy leads to atrophy of the uterine and vaginal epithelium, just as much or more so than hypophysectomy, and (*b*) the pituitary gonadotropins produce hyperplasia of the uterine and vaginal epithelium only when the ovary is present. These facts lead to the conclusion that the effect of the hypophysis on the uterine and vaginal

epithelium is due to the induced hypersecretion of the ovarial hormones.

To be active, certain hormones require a previous sensitization of the receptor organ by other hormones. Thus, estradiol does not produce hyperplasia of the adrenals in hypophysectomized rats, but this effect is produced if the atrophy of the adrenals is prevented by administration of two daily doses of anterohypophysis (Pinto). Chorionic gonadotropin does not induce ovulation in hypophysectomized rats, but does so if either serum gonadotropin or stilbestrol is previously injected.

Metabolism of the Hormones

At present we live in a period in which the metabolism of the hormones is being energetically studied. The studies include the investigations of origin, transformation, and elimination of the hormones.

The study of origin includes that of their precursors within the organism or in the diet and that of the place and mechanism of elaboration within the endocrine gland.

It is also interesting to know its absorption, its chemical transformations, and the site of these transformations. Thus it is known that the estrogens are destroyed principally in the liver and that thyrotropin is transformed by the thyroid into an inactive compound which can be reactivated at a certain temperature. In some cases the disappearance of the hormone can be quantitatively followed.

The disappearance of the hormone from the blood has been followed in various cases. The unchanged hormone may be eliminated in the urine or, secondarily, in the milk or the bile, but in other cases only the transformation products of the hormone are eliminated through these routes. The urine has the advantage, for purposes of extraction, of being a concentrated ultrafiltrate of the plasma virtually free of protein. Hormones of protein nature may or may not filter through the kidney according to their molecular size. Thus chorionic gonadotropin and thyrotropin are found in the urine, but not so the gonadotropin of the pregnant mare serum. The form in which various steroids with estrogenic, androgenic, or corticoid action are eliminated, as well as the total elimination of the 17-ketosteroids, are being intensively studied. In some cases the metabolism of a particular hormone can be traced by the assay of some transformation product,

e. g., that of progesterone through the urinary elimination of pregnanediol.

Applications

The study of hormones is of interest to medicine and animal husbandry. For medicine it is important to investigate: (*a*) the physiological and nutritive conditions which will assure hormonal equilibrium and which will secure better physical and mental health during rest, exercise, or work; (*b*) the prevention and treatment of the endocrine disturbances and especially the more common disturbances such as diabetes, endemic goiter, endocrine disturbances in women, etc., by genetic, dietetic, and pharmacological methods; (*c*) the prevention and treatment of disturbances due to abnormal internal secretions, *e. g.*, nephrogenic hypertension. The distinction between specialists in diseases of metabolism and specialists in diseases of the endocrine organs is quite artificial. Diabetes, for example, is the most typical endocrine disease. The glands of internal secretion are the regulators of metabolism and it is impossible to study either endocrinology independently of metabolism or metabolism independently of endocrinology.

The study of the hormones is germane to the problem of animal production because it gives a clearer understanding, and thereby wider possibilities, of controlling phenomena such as heat, ovulation, fertilization, pregnancy, number of offspring, breeding without dependence on factors such as lactation, castration, time of the year, etc.

The study and production of hormones has been converted into a problem of national importance. There are enormous commercial interests involved and a great number of technicians devoted to the search, production, and commerce of hormones. This raises the danger of both excessive and inadequate use of the hormones and of exaggerated propaganda.

At the present moment, the main problems of experimental endocrinology may be classified in the following groups: (*1*) isolating pure hormones and studying their actions, either separated or associated, simultaneous or successive; (*2*) establishing the mechanism of action of each hormone to determine whether they are direct or mediated by other organs; (*3*) studying each organ's secretion from the standpoint of the regulating factors and of its relations with the

regulatory mechanisms of other glands and with the wider systemic functions; (*4*) the metabolism of the hormones; (*5*) clinical applications to prevent and cure endocrine diseases; (*6*) applications in animal industry.

This exposition of the hormones for the sake of brevity has been necessarily general in character and, therefore, incomplete. The author had to avoid the double risk of being too elementary or too tedious. Briefness courts the danger of being dogmatic or of not providing the factual basis for each statement. This essay has been written with a view to presenting to scientific laymen, but not to specialists in endocrinology, the current position of the problem of the hormones.

Selected Bibliography

Barker, L. F., Hoskins, R. G., and Mosenthal, H. O., *Endocrinology and Metabolism.* 5 vol., Appleton, New York, 1922.

Bayliss, W. M., and Starling, E. H., "The chemical correlation of the secretory process," *Proc. Roy. Soc. London*, **67**, 310–322 (1940). "Die chemische Koordination der Funktionen des Körpers," *Ergeb. Physiol.*, **5**, 664 (1906).

Berthold, "Geschlechtseigentümlichkeiten," *Wagners Handwört. Physiol.*, **1**, 507 (1842). "Transplantation der Hoden," *Arch. Physiol.*, **1849**, 42.

Biedl, A., *Innere Sekretion.* 2nd ed., 2 vol., 1913; 4th ed., 3 vol., 1919–1922. Urban & Schwarzenberg, Vienna.

Del Castillo, E. B., Reforzo Membrives J., De La Balze, F., and Galli Mainini C., *Endocrinología Clínica.* El Ateneo, Buenos Aires, 1944.

Glandular Physiology and Therapy. Am. Med. Assoc., Chicago, 1935. 2nd ed., 1942.

Gley, E., *Les sécrétions internes.* Baillière, Paris, 1914, 94 pp., 3rd ed., 1925.

Hirsch, M., *Handbuch der inneren Sekretion.* 3 vol., Kabitzsch, Leipzig, 1928–1933.

Lucien, M., Parisot, J., and Richard, G., *Traité d'endocrinologie.* Doin, Paris, 1925 and 1934.

Pende, N., *Endocrinología.* 4th ed., 2 vol., Vallardi, Milan, 1934.

Schaeffer, E. A., "On internal secretions," *Lancet*, **2**, 321, 324 (1895). *The Endocrine Glands*, Longmans, Green, London, 1916.

Starling, E. A., "The chemical correlation of the functions of the body," *Lancet*, June (1905).

Trendelenburg, P., *Die Hormone.* 2 vol., Springer, Berlin, 1929, 1934.

Vincent, S., "Innere Sekretion und Drüsen ohne Ausführungsgang," *Ergeb. Physiol.*, **11**, 218 (1911).

14

FUNDAMENTALS OF OXIDATION AND REDUCTION

LEONOR MICHAELIS, MEMBER EMERITUS, THE ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH

IN ATTEMPTING to define the essential concepts involved in a problem, generally it is found that, because of the flexibility and, often, ambiguity of language, a definition cannot be formulated with perfect clarity. Nature is not so constructed that one can classify all its subject matter within a finite number of distinctly circumscribed terms. A great deal of confusion has arisen, and will henceforth arise again, from the fact that an author may use a given term according to one definition, and then during the discussion consciously or unconsciously forget that definition. Furthermore, it is a frequent fate of a definition that it be based on an assumption which later appears to be either erroneous or, at least, unsuitable. If the term, as commonly happens, is redefined according to the change in the underlying fundamental concepts, the new definition is likely to be in conflict with the older one.

A typical instance of the flexibility of a concept is the term oxidation, and its reverse, reduction. Originally, oxidation meant combination with oxygen; the combination of hydrogen with oxygen to form water is the prototype of oxidation in this sense. According to this definition, the combination of hemoglobin with oxygen to form oxyhemoglobin should be the simplest, purest, and most unambiguous

case of oxidation in organic chemistry. However, according to our present usage of the term, this reaction is so dissimilar from what is now considered to be a typical oxidation that it is no longer classified among the oxidations—it has, in fact, been termed "oxygenation" by Conant. The real oxidation of hemoglobin, according to modern definition, is its conversion to methemoglobin, because, according to the present stage of knowledge and on the basis of our present model of atomic and molecular structures, the oxygen of oxyhemoglobin is attached to hemoglobin without affecting the electronic structure of the iron atom,* whereas in methemoglobin no oxygen is attached to hemoglobin at all—rather, the iron atom of hemoglobin contains one electron more than the iron atom of methemoglobin. The removal of that electron from hemoglobin is now considered as typical for its oxidation. When ferrous chloride reacts with chlorine, ferric chloride is formed. Since ferrous compounds can be converted to ferric compounds by oxygen also, ferric iron has always been considered an oxidation product of ferrous iron. Yet, if this conversion is brought about by chlorine, oxygen plays no part in the "oxidation."

What is common in most processes formerly designated as oxidation, disregarding such an exceptional case as the formation of oxyhemoglobin, can only be stated in terms that would have been quite incomprehensible to the originators of the concept of oxidation. This common property can be defined, in terms of the present state of the atom model, by saying that, after its oxidation, a molecule has been deprived of one electron, or of two electrons; these two cases of oxidation are distinguished by the terms "univalent" and "bivalent." The oxidation of ferrous ion to ferric ion, or of ferrocyanide ion, $[Fe(CN)_6]^{4-}$ to ferricyanide ion, $[Fe(CN)_6]^{3-}$, are examples of univalent oxidation. The oxidation of stannous ion, Sn^{++}, to stannic ion, Sn^{4+}, is a bivalent oxidation. In those cases in which the oxidation, or, in other words, the withdrawal of the electron, is brought about by oxygen, the oxygen is the acceptor of the electron. The fate of oxygen after acceptance of the electron will be discussed on page 220. Other reagents, such as

* This statement involves another difficulty based on facts unknown until recently. Since hemoglobin is paramagnetic, but oxyhemoglobin diamagnetic, as Pauling and Coryell have shown, some change in electronic structure due to the attachment of oxygen must be postulated here too. But it is not of the type one would now call oxidation.

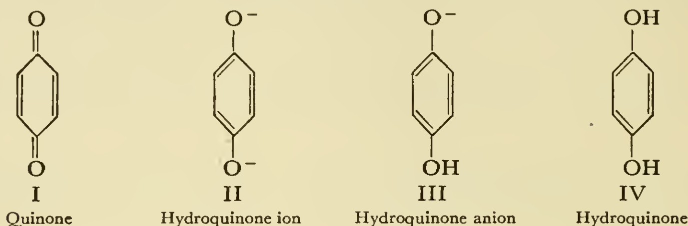
chlorine, can also accept an electron and thus "oxidize" other substances. Oxygen, therefore, is only one among many molecular species which can withdraw an electron from other molecules and so "oxidize" them.

One must remember that the transition of the original definition of oxidation to the modern one has been gradual. Historically, there has been a transition stage inaugurated by Wieland, who developed his concept mainly with respect to oxidation of organic molecules. When alcohol, $CH_3C\begin{smallmatrix} \diagup H \\ -H \\ \diagdown OH \end{smallmatrix}$, is oxidized to aldehyde, $CH_3C\begin{smallmatrix} \diagup H \\ \\ \diagdown\!\!\diagdown O \end{smallmatrix}$, the overall effect is the loss of two hydrogen atoms. According to the original definition of oxidation, one may say that the primary process is the addition of one oxygen atom to alcohol in order to form $CH_3C\begin{smallmatrix} \diagup H \\ -OH \\ \diagdown OH \end{smallmatrix}$ which, by splitting off one molecule of water, becomes $CH_3C\begin{smallmatrix} \diagup H \\ \\ \diagdown\!\!\diagdown O \end{smallmatrix}$. Wieland's suggestion is that the process does not pass through the stage of an addition of oxygen, but that what had been designated as oxidation is in fact a withdrawal of two hydrogen atoms. The oxidizing agent is the acceptor for the hydrogen atoms; for such cases, he replaces the term "oxidation" by "dehydrogenation."

The reconciliation of this idea with the more modern one can be based on the fact that a hydrogen atom consists of a positively charged proton and a negatively charged electron. Then, one may say that the withdrawal of two electrons is essential for oxidation of alcohol. Since the two protons, which should remain, are no longer held by any noticeable force, they are detached also, and become bound to some proton acceptor, such as water, with which they form the hydrogen ion, OH_3^+, as it exists in the presence of water (called also the oxonium ion); or they may be bound to some anion that may be present, such as the acetate ion, CH_3COO^-, with which they form an "acid," CH_3COOH. This hypothesis should not involve the idea that the expulsion of the electron occurs first and the expulsion of the proton thereafter. The sequence is undecided. The principle is, rather, that the withdrawal of each hydrogen atom is the same as the

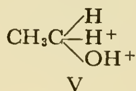
withdrawal of an electron together with a proton; this process is dehydrogenation. The reconciliation with the modern definition is the postulate that only the withdrawal of the electron is characteristic for oxidation, and that the simultaneous detachment of a proton does not belong to the process of oxidation proper.

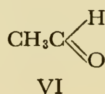
The attachment of a proton is termed a change in the level of "acidic ionization." This will become clear if one first considers an example showing the essential features in a very obvious way. Let us start with quinone (I), the reduction of which, according to modern concepts, consists essentially in the attachment of two electrons. The result is formula II, a doubly negatively charged ion of hydroquinone.



If the reaction occurs in a strongly alkaline solution, the result of oxidation is molecule II. When the reaction occurs in a less alkaline solution, of *p*H about 9 to 10, one proton only is attached to II, and the univalent anion of hydroquinone is formed (III). When the reaction occurs in an acid solution, two protons are attached to II, and (unionized) hydroquinone (IV) is formed. Depending upon the *p*H, either an electron or a full hydrogen atom may be attached to each oxygen. We now agree on the definition that II, III, and IV are on the same level of oxidation-reduction as, but on lower levels of oxidation than, quinone (I). The differences between II, III, and IV are not in their level of oxidation, but in their level of acidic ionization.

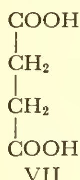
It is easy to transfer these ideas to the case of the oxidation of alcohol. On detaching two electrons only, one obtains structure V, where two protons exist in the molecule. The two protons, however, are held still less firmly than in hydroquinone ion IV, and under all



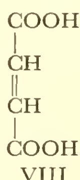


conditions possible, even in extremely acid solution, V changes its acidic ionization level by expelling the two protons and forming acetaldehyde (VI). Here again we do not claim that V is formed first and VI subsequently. It is, however, emphasized that the process of oxidation proper is the detachment of electrons, and that the simultaneous loss of the two protons, although it may be a necessary consequence, does not belong to the process of oxidation itself. Dehydrogenation is thus a special case of oxidation.

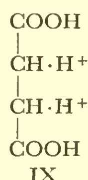
The term dehydrogenation need not be discarded but may be used advantageously for those cases in which the release of an electron involves the simultaneous release of a proton. Whether or not this simultaneous release of the proton occurs depends on the *p*H of the solution. The term dehydrogenation can be reserved for the cases in which even in extremely alkaline solution the release of an electron is accompanied by a release of a proton. The conversion of succinic acid (VII) to fumaric acid (VIII) may be called a dehydrogenation because the intermediate stage (IX) is not capable of permanent existence. Molecular species VIII may be considered as an infinitely strong acid, which even in an acid solution is not capable of holding on to protons. Therefore the enzyme which catalyzes the reaction VII → VIII may



VII



VIII



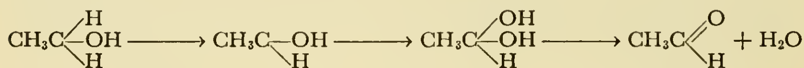
IX

be justly termed a dehydrogenase. When quinone (I) is reduced in an acid solution, we may speak also of dehydrogenation, whereas in the reduction of quinone in a more alkaline solution (IX and VIII) the synonymous use of the terms oxidation and dehydrogenation would involve a more generalized definition of dehydrogenation.

Analogously, if we define oxidation as the loss of an electron, we are compelled to generalize the definition by the amendment that this definition holds whether or not a proton is also released; so whether one uses "oxidation" or "dehydrogenation" is a matter of nomenclature only. In this essay, we shall use "oxidation" and "reduction," and not "dehydrogenation" and "hydrogenation." For

such reactions as the conversion of benzene to cyclohexane, however, it is recognized that one may prefer to use the term hydrogenation instead of reduction. As regards the chain of oxidation reactions as they occur in respiration, it is often said that hydrogen atoms are transferred from one molecular species to another. However, the oxidation of reduced cytochrome to cytochrome involves not the transfer of a whole hydrogen atom, but of one electron only. It is therefore not entirely true that the chain of respiratory processes consists in transferring hydrogen atoms from one molecular species to another, and finally to oxygen, but it is true that this chain consists in transferring *electrons* from one molecular species to another. Sometimes the transfer of the electron is accompanied by a transfer of a proton and sometimes it is not.

It may be added that the loss of a hydrogen atom in dehydrogenation is equivalent to the addition of a hydroxyl group, as far as the level of oxidation-reduction is concerned. When alcohol is oxidized to acetaldehyde, the process may be alternately described as follows. Alcohol detaches one hydrogen atom and accepts a hydroxyl group:



Here, the first step is the detachment of the hydrogen atom, and the second, attachment of the hydroxyl group, while the process formerly was described in terms of loss of two hydrogen atoms. So, *de*hydrogenation is equivalent to "hydroxylation," and both terms can be avoided by describing the level of oxidation proper only in terms of electrons ejected or added, whether or not a proton or a hydroxyl ion is also involved in the process.

Stepwise Oxidation and Reduction of Organic Compounds

If one wants to formulate a simple example of a possible stepwise oxidation of an organic compound, one might propose the series:

each step representing a bivalent oxidation. A univalent oxidation is not imaginable unless one abandons the assumption that carbon is

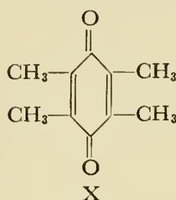
tetravalent, oxygen bivalent, and hydrogen univalent: the univalent oxidation of CH_4 would give, as a step intermediary on the way to CH_3OH, the free radical CH_3 with "tervalent" carbon. To be sure, such free radicals have been recognized for a long time, but those known until recently always have one of two properties: either they are very unstable and have an extremely short lifetime (*e. g.*, CH_3); or they can be produced only from a very restricted number of compounds and only in a solution of a perfectly water-free organic solvent. The famous discovery of triphenylmethyl by Gomberg in 1890 provided a prototype of such free radicals.

It will now be shown that all oxidations of organic molecules, although they are bivalent, proceed in two successive univalent steps, the intermediate state being a free radical, and furthermore, that according to structural conditions these intermediate free radicals may be either just as unstable as CH_3, somewhat more stable, or even perfectly stable compounds.

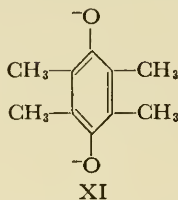
First, however, the concept of stability must be discussed. The criterion of stability may be obtained in two ways. A substance may be said to be unstable if it rapidly undergoes a chemical change when exposed to such ubiquitous reagents as air; pyrogallol, for instance, is unstable in an alkaline solution because it is readily oxidized by oxygen, but is stable in the absence of oxygen. Or, a substance may be said to be unstable if it undergoes a change even in the absence of any foreign substance with which it might react; acetaldehyde, for instance, undergoes the Cannizzaro reaction in an alkaline solution, one molecule of aldehyde being oxidized to form acetic acid, while another is reduced to ethyl alcohol. It will now be shown that the alleged instability of free organic radicals is, in very many cases, essentially due to the latter, bimolecular interaction. It will be shown that the bivalent oxidation in fact occurs in two successive univalent steps as in the example of duroquinone (X).

The methyl groups in the ring of X prevent the secondary, irreversible reactions which are of no interest in this discussion, and which would occur when working with the unmethylated, simple benzoquinone. When duroquinone in an alkaline solution is reduced, as by hydrogen plus palladium, or by any other suitable reducing agent, the faintly yellow solution turns brown first, then colorless. The colorless compound is the corresponding durohydroquinone, which

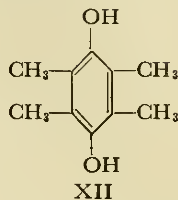
in extremely alkaline solution can be written as formula XI and which



X

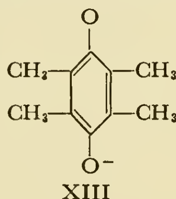


XI

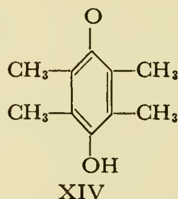


XII

in less alkaline solution would add first one, then another, proton to form XII. This is the customary bivalent reduction. The intermediate, brown substance is the result of a univalent reduction; and it can be shown by various methods, to be described later on (see pages 215 and 217), that it has the same molecular size as XI and differs from it only in so far as one oxygen atom is negatively charged and not the other. This brown substance is a free radical. One might say that one oxygen atom is bivalent and the other univalent, as in XIII. The same molecular species, in a more acid solution, would have formula XIV.



XIII



XIV

Let us consider first the reaction in a strongly alkaline solution. When the solution of the quinone is mixed with increasing amounts of a reducing agent, there will always be, except for the very beginning and the end of the oxidation, a mixture of the quinone, the hydroquinone, and the intermediate free radical which we shall call semiquinone; and an equilibrium is established between these forms. It is important to emphasize that the equilibrium is established instantaneously, and that there is no sluggishness in its formation, as is usually the case in the formation of equilibria with organic compounds (except acidic ionizations). The activation energy of this reaction leading to equilibrium is extremely small. So, in the solution, the intermediate free radical is never present without being in mixture with the quinone

and the hydroquinone. If we designate the hydroquinone (the reduced form of this system) as R, the semioxidized form (the free radical) as S, and the totally oxidized form (the quinone) by T, the equilibrium is established according to the reversible reaction

$$2\,S \rightleftharpoons R + T \qquad (1)$$

and the constant of equilibrium, which may be called the "semiquinone formation constant," is:

$$k = \frac{[S]^2}{[R][T]}$$

Its reciprocal may be called the "dismutation" constant, because reaction (1) is a "dismutation" (or "disproportionation") of the free radical. If k is very small, very little of the radical exists in the state of equilibrium; if k is large, much of the radical can exist, and it can be distinguished by its particular color and by its paramagnetism, which will be discussed presently.

Experience has shown for the case of duroquinone that k is very large in an alkaline solution, but very small in an acid solution. The transition of the behavior from alkaline to acid solution is continuous; and there can be no doubt that in acid solution a small amount of the radical is formed too, even if its concentration is too small to be noticeable by ordinary methods. Thus we are obliged to state that the ionized form of the radical (XIII) is a rather stable compound, and the unionized (XIV) rather unstable, "stability" being judged according to its capability of existing in equilibrium with its parent substances, the quinone and the hydroquinone.

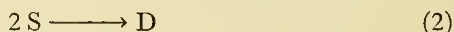
Why is XIII more stable than XIV? In formula XIII, the negative charge has been arbitrarily attributed to the upper oxygen atom, but it can just as well be attributed to the lower one. In fact, the location of the charge is undecided, for it oscillates between the two extreme positions through the chain of the atoms in the ring. Such a condition has been termed resonance. The two limiting structures, one with the charge on top, the other with the charge at the bottom, are indistinguishable molecular species. One speaks about "equivalent" or "symmetrical" resonance, a condition which, according to quantum mechanics, contributes largely to the stability of a molecule. In form XIV, no such equivalent resonance prevails. The

proton attached holds the negative charge rather tightly in place. This molecule does exist but to a much smaller extent than XIII because it is less stable. Free radical formation is a general reaction whether it takes place extensively, as in alkaline solution, or in traces, as in acid solution.

It can be shown by numerous examples that the same behavior is true for organic compounds which form reversible oxidation-reduction systems, such as many dyestuffs, *e. g.*, methylene blue, or phenol-indophenol. No observable amount of any intermediate free radical can be demonstrated for irreversible oxidations, *e. g.*, when alcohol is oxidized to aldehyde. So we are permitted to conclude that the formation of a free radical, in sufficient concentration, in the equilibrium involved, is a prerequisite for the reversibility of an oxidation-reduction system.

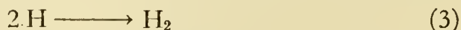
Evidence for Free Radical Nature of the Intermediate Substance

It is appropriate to discuss at this point the evidence for the assertion that the intermediate steps of oxidation are really free radicals of the same molecular size as the parent substance and not dimeric valence-saturated compounds made up in such a way that two radicals are combined to form a bond which abolishes the state of "unsaturation." Such a reaction may be imagined to occur as follows:



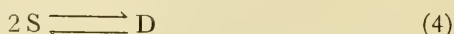
$$2\,S \longrightarrow D \qquad (2)$$

That is, two molecules of the semiquinone radical, S, combine with each other to form a dimeric compound, D, which no longer has the characteristics of a free radical, just as two "radicals" H (hydrogen atoms) combine as follows:



$$2\,H \longrightarrow H_2 \qquad (3)$$

In fact, such reactions do occur, and an equilibrium is established, so that, instead of reaction (2) we should write:



$$2\,S \rightleftharpoons D \qquad (4)$$

Depending upon the conditions, which will not be discussed here, this equilibrium is sometimes greatly in favor of the free radical, sometimes

greatly in its disfavor. The main point is that reaction (2) does not, as a rule, proceed to completion, but that the free radical really does exist, and often even to such an extent that the dimerization is negligibly small. There are two powerful tools to decide whether the intermediate product is S, D, or an equilibrium mixture of the two, *viz.*, potentiometric titration and magnetic measurement.

When a substance such as duroquinone, the example discussed above, is titrated with a reducing agent and the oxidation-reduction potential is plotted against per cent reduction, a curve is obtained the shape of which will depend markedly on whether the intermediate substance is R or D. Straightforward application of the law of mass action yields a complete theory as to the shape of the titration curve. The calculations involved, although not absolutely simple, are of such a nature that only a high-school student might consider them in the realm of "higher" mathematics. Application of the law of mass action has shown in many cases that the intermediate substance is a free radical, almost exclusively under certain conditions and, under other conditions, in equilibrium with its dimer. Furthermore, according to a theory first developed by G. N. Lewis and later amply confirmed by the quantum theory, a free radical, because it always contains an odd number of electrons, must always be paramagnetic, in contrast to the ordinary, valence-saturated organic compounds, which are diamagnetic (provided they do not contain metal atoms such as iron or cobalt). When a solution of duroquinone is slowly reduced in alkaline solution, by glucose, say, evidence can be produced for the gradual appearance of a paramagnetic molecular species and, on further reduction, for its disappearance. The paramagnetism is due to the spin of the odd electron, while in ordinary molecules the electrons always occur in pairs with opposite spin, whereby their paramagnetic effect is quenched.

An elegant method of detecting such free radicals, even of low stability, has been established by G. N. Lewis. In the kind of experiments mentioned above, ejection of an electron is brought about by an oxidizing chemical, which serves as an acceptor of electrons. Lewis effects ejection of the electron by ultraviolet light, the substance being dissolved at the temperature of liquid air, in an organic solvent which has, at this temperature, a rigid, glasslike consistency without crystallizing. In such a rigid medium, no molecular collisions can occur

and free radicals, once they have been established, have no chance to undergo bimolecular reactions by which they may be rapidly eliminated. Radicals can be preserved, in this manner, to an extent which by far exceeds their equilibrium concentration postulated by thermodynamics. Lewis showed that many organic compounds exposed in this way to ultraviolet radiation become colored; in suitable cases he identified, by spectrophotometrical comparison, the colored substance with the free radical obtained previously by chemical oxidation. The peculiar merit of this method is the fact that free radicals can be detected in the cases in which they would be unnoticeable in the state of equilibrium because of their low concentration. On the other hand, since the appearance of color is not necessarily evidence for a free radical, a scrutiny of the phenomenon is necessary for each individual case.

In order to arrive, from these considerations, at the discussion of the final problem we have in mind, the varying degree of inclination of an organic substance toward oxidation (or reduction) must be considered. Here we distinguish two essentially different properties. To state that a substance is easily, or difficultly, oxidized by an oxidizing agent may mean one of two things: that the speed of such an oxidation is great or small (a topic belonging to a discussion of chemical kinetics); or that the final state attainable by the interaction of the oxidizing agent and the oxidizable substrate is a complete, 100% oxidation, an incomplete oxidation, or almost no oxidation at all (a problem of thermodynamics).

We shall start with the problem in thermodynamics, in particular, the case in which both the oxidizing agent and the substance to be oxidized form reversible oxidation-reduction systems. If leucomethylene blue is the substance to be oxidized and potassium ferricyanide the oxidizing agent, when the two are mixed in equivalent proportions practically all of the leuco dye is oxidized to the dye and all of the ferricyanide is reduced to ferrocyanide. But when leucomethylene blue is mixed with an indophenol dye, both the oxidation of leucomethylene blue and the reduction of indophenol will be incomplete, and the final state is a mixture of four substances, the two leuco dyes and the two dyes. Furthermore, if leucomethylene blue is mixed with safranine, no change will occur, at least within practical measurable limits. Safranine does not oxidize leucomethylene blue

to any appreciable extent. On the other hand, leucosafranine readily reduces methylene blue. With this in mind, one can set up a sequence of all reversible oxidation-reduction systems according to their oxidative power. For the four systems mentioned, the sequence is: ferricyanide, indophenol, methylene blue, safranine.

A quantitative expression of the oxidizing power is the "oxidation-reduction potential" of a dye in mixture with its leuco dye. Various substances capable of reversible oxidation and reduction can be arranged in a sequence according to their oxidation-reduction, or "redox," potentials. Each member, when present in the reduced state, can be oxidized by any following member present in its oxidized state; and each member, when present in the oxidized state, can be reduced by any preceding member present in its reduced state.

Let us now discuss the *kinetic* aspect of the problem. If we are dealing only with reversible redox systems, as above, establishment of the equilibrium is always so rapid that the reaction may be considered almost instantaneous. But this is not the case, in general, if an irreversible reaction occurs. In the oxidation of alcohol to acetaldehyde, for example, to oxidize alcohol, a powerful oxidizing agent such as chromic acid must be used; and to reduce acetaldehyde, a powerful reducing agent such as sodium amalgam must be used. Although an excess of oxidative (or reductive) power must be applied in order to make the reaction proceed, the reaction is sluggish. Additional energy is required far beyond the quantity expected on a purely thermodynamic basis because, obviously, an obstacle has to be overcome. Despite the fact that the reaction between alcohol and chromic acid releases energy, energy must first be spent, which is of course eventually released again; and this extra energy is called the activation energy. Although the path of energy is, as a whole, downward, it must first pass over a hill. Generally an activation energy is required for all bimolecular chemical reactions. In reversible reactions the energy of activation is very small, interaction occurring only when the two molecules "collide." The collision is impeded by the fact that molecules of any kind, on approaching each other in the course of thermal motion, will exhibit a mutual repulsion, and only those molecules which happen to have enough kinetic energy to overcome the repulsion will really collide and react with others.

In irreversible oxidations another, more serious, impediment

occurs. We have presented the postulate that all oxidations proceed in a sequence of univalent steps. The first step of oxidation of alcohol would lead to the free radical, in this case, an utterly unstable molecule, in the thermodynamic sense. To generate these radicals, an oxidizing agent of very high potential is required; and even then their concentration remains small, so small that no direct evidence for their existence is available. The free radicals, once generated, will then react by a dismutation:

$$2 \text{ radicals} \rightleftharpoons 1 \text{ alcohol} + 1 \text{ aldehyde}$$

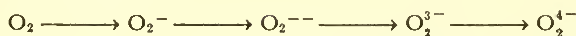
The velocity of the latter reaction depends, among other things, on the concentration of the molecules which are interacting with each other, and therefore on the square of the concentration of the free radicals. If this concentration is very small, it may be the limiting factor for the over-all process of oxidation of alcohol. We may say that the energy of activation for the oxidation of alcohol is essentially the energy necessary for the formation of the free radical. Unless the radical is relatively stable, as in reversible processes, the activation energy is very great. This high energy of activation is the reason why so many organic compounds are "stable." If all thermodynamically possible reactions could proceed unhampered there would be no such thing as organic chemistry.

Inhibition due to high activation energy occurs only when the attainable concentration of the free radical is the limiting factor for the rate of the over-all reaction. It does not matter whether the attainable concentration of the free radical is 1 M or, say, 10^{-8} M. Factors other than the concentration of the free radical, such as the specific constants of reaction velocities, will determine the rate of the reaction. However, if the concentration of the free radical is so small as to be the limiting factor of the over-all process, sluggishness and irreversibility will arise. This consideration fulfills an important requirement for the understanding of reaction rates—it reduces a problem of kinetics to one of thermodynamics.

Oxygen as an Oxidizing Agent

These considerations also explain why oxygen is such a sluggish oxidizing agent despite its very large oxidative power in a thermo-

dynamic sense. If it can be reduced only in successive univalent steps, these steps must be:



$$O_2 \longrightarrow O_2^{-} \longrightarrow O_2^{--} \longrightarrow O_2^{3-} \longrightarrow O_2^{4-}$$

Two of these steps are chemically identifiable. O_2^{--} is, after accepting two protons, O_2H_2, hydrogen peroxide. O_2^{4-} is, after accepting four protons, two molecules of H_2O. However O_2^{-} (or O_2H), and O_2^{3-} (which may be written as O_2H_3, or $OH + H_2O$) are intermediate, utterly unstable steps. Since the reaction must pass through these unstable steps, the activation energy involved in the reduction of O_2 is very high.

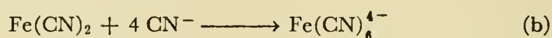
How is this activation energy overcome when oxygen does oxidize a substance? Overcoming the activation energy by working at high temperatures is a usual procedure in the laboratory but is not feasible under physiological conditions. Here the answer is that oxygen reacts with an oxidizable substance very often not only by means of a collision of the molecules but also by the establishment, after collision, of a relatively stable addition compound which can then undergo intramolecular redistribution of electrons.

Certainly no claim is made that this mechanism is always the one involved in activation of oxygen. However, it is *one* of the possible mechanisms and very likely is correct for the particular case to be described in detail. In all probability it is the mechanism by which oxygen is activated in all those cases in which a heavy metal compound, especially of iron or copper, acts as activating catalyst.

It has been shown that, at least in an acid solution, the oxidation of a leuco dye, or of cysteine, and many other substances, by means of free oxygen is accomplished, at least with any appreciable speed, only in the presence of a trace of an iron or copper salt. These metal atoms have two essential properties which render them useful for their catalytic action. First, they readily change their valence; iron may be bivalent or tervalent and copper, univalent or bivalent. Second, these metals are highly inclined to form complex compounds of the Werner type. The nature of such metal complex compounds may be demonstrated as follows. The doubly positively charged ferrous ion, Fe^{++}, can combine, first of all, with two negatively charged univalent ions to form a saltlike compound, for instance, with two cyanide ions:

$$Fe^{++} + 2\,CN^- \longrightarrow Fe(CN)_2 \qquad \text{(a)}$$

However, the combining power of the ferrous ion is not exhausted by this reaction based on opposite charges. In fact, the molecular species $Fe(CN)_2$ has never been shown to be capable of existence. More than two cyanide ions are attached to the iron due to the fact that each CN^- ion has one pair of electrons not used for chemical bonding. Each of such electron pairs can be shared with the iron to fill up its outermost incomplete electron shell to a complete shell, as in a noble gas. In addition to the two CN^- ions of equation (a), four more can be attached, which contribute four negative charges to the complex molecule, which is a ferrocyanide ion:



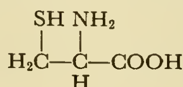
$$Fe(CN)_2 + 4\,CN^- \longrightarrow Fe(CN)_6^{4-} \qquad \text{(b)}$$

The six CN groups are arranged around the central iron atom as the corners of an octahedron. Fe is said to possess "six coordination places" which may be occupied by atoms or atom groups. In analogy, when ferrous ion combines with cysteine, which we may write, briefly, as RSH,* we may imagine that, primarily, a saltlike compound between iron and two molecules of cysteine is formed in such a way that the two hydrogen atoms of the sulfhydryl groups are replaced by an iron atom:

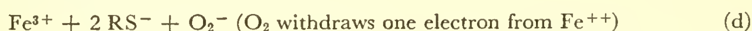
$$Fe^{++} + 2\,RSH \longrightarrow Fe(RS)_2 + 2\,H^{++} \qquad \text{(c)}$$

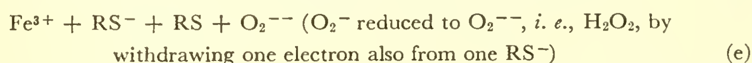
This scheme accounts, so far, for the saturation of two (of the six possible) coordination places. Now atom group R contains another atom with an unused electron pair, *viz.*, the nitrogen atom of the amino group. Thus, the two molecules of RSH will occupy four coordination places. Probably because of the large size of the RSH molecule and steric hindrance involved in it, the two remaining coordination places cannot be occupied by a third molecule of RSH. In fact, no ferrous complex of cysteine can be prepared with more

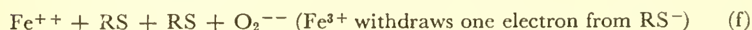
* SH is a sulfhydryl group, and R represents the rest of the molecule, which is, altogether:



than two molecules of cysteine for one iron atom.* The two remaining coordination places may be filled in by other atoms or atom groups of small size having an unused electron pair. One example of such an atom group is carbon monoxide, CO. In fact, the coordination compound, $Fe^{II}(RS)_2(CO)_2$, can be readily prepared by the interaction of Fe^{++}, cysteine, and carbon monoxide in the form of its well crystallizable alkali salt.† Just like CO, O_2 also has (at least) one unused electron pair, and may combine instead of CO, as in the case of hemoglobin which combines either with O_2 or with CO. In hemoglobin, four coordination places of iron are occupied by the four nitrogen atoms of the porphyrin ring, a fifth by protein, and the sixth can combine with O_2 or with CO. So we arrive at hypothetical iron–cysteine–oxygen complexes, analogous to the well-known carbon monoxide complex, $Fe^{II}(RS)_2(CO)_2$. One cannot tell whether one oxygen molecule is attached at one or at two coordination places, or whether even two oxygen molecules can be attached. Suffice it to imagine the complex $Fe^{II}(RS)_2O_2$. This complex is said to be hypothetical because it cannot be prepared, and undergoes a redistribution of electrons, or, in other words, intramolecular oxidation-reduction, which may be symbolized in this way: The original, oxygen-containing complex may be imagined to consist of the following constituents: $Fe^{++} + 2\,RS^- + O_2$. The electron redistribution will occur thus:


$Fe^{3+} + 2\,RS^- + O_2^-$ (O_2 withdraws one electron from Fe^{++}) (d)


$Fe^{3+} + RS^- + RS + O_2^{--}$ (O_2^- reduced to O_2^{--}, *i. e.*, H_2O_2, by withdrawing one electron also from one RS^-) (e)

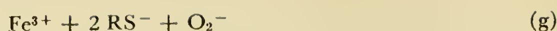

$Fe^{++} + RS + RS + O_2^{--}$ (Fe^{3+} withdraws one electron from RS^-) (f)

An alternative scheme, probably of equal probability, is:

---

* It is interesting to compare cysteine complexes of cobalt with those of iron. For the cobaltous state, no complex with more than two cysteine molecules can be obtained. For the cobaltic state, both a complex with two and another with three molecules of cysteine can be prepared. The cobaltous complex, then, is quite analogous to the ferrous complex. The cobaltic complex stands no comparison, because the ferric–cysteine complex is too unstable, due to the rapid intramolecular rearrangements to be described presently.

† The formula in the first footnote on page 222 shows that atom group R contains a carboxyl group. It is by means of the two carboxyl groups in the complex that alkali salts can be formed.



$$Fe^{3+} + 2\,RS^- + O_2^- \qquad (g)$$

$$Fe^{++} + RS^- + RS + O_2^- \qquad (h)$$

$$Fe^{++} + RS + RS + O_2^{--} \qquad (i)$$

step (f) being identical in both cases. Now, the complex containing the constituents as in (f) is unstable and disintegrates to form three separate molecular species: Fe^{++}; RSSR, cystine; and hydrogen peroxide. The latter may be used to oxidize more cysteine (stoichiometrically, not catalytically), or to oxidize Fe^{++} to Fe^{3+}, and this Fe^{3+} may oxidize (stoichiometrically) more cysteine. Finally, all the iron is again in the ferrous state and the whole cycle is repeated with iron thus acting as a catalyst.

It is an essential prerequisite of this cycle that the change from the ferrous to the ferric state occurs readily and reversibly. In using cobalt instead of iron, the first stages are similar; but, once the cobaltic complex has been established, it shares with all cobalt complexes of the Werner type the property that the cobaltic state cannot be readily reduced to the cobaltous state, even by means of rather strong reducing agents. The final result is, therefore, the formation of the cobaltic complex, stoichiometrically, without starting a catalytic cycle. Copper, but not cobalt, can replace iron as a catalyst.

What is furthermore essential in this process is the fact that each single step in this chain reaction consists of the transfer of a single electron. This assertion is more than a mere hypothesis. Since the change of ferrous to ferric state involves one electron only, the subdivision of the over-all process into one-electron transfers is obvious. It is remarkable that, even for such a simple case of iron catalysis, the whole chain is of such an intricate nature, allowing for different pathways leading to the same final result.

Oxidation Catalysts and Enzymes

The physiologically occurring catalysts (or enzymes) for oxidation or reduction are characterized by their specificity. All oxidation enzymes have been recognized as compounds of reversible redox systems and a specific protein. The same redox system, when attached to different specific proteins, may have a different specificity. The present state of our knowledge is on what may be called a descriptive

level: some of the enzymes can be prepared as pure crystalline entities, and their composition and activity can be examined. Since the first stages in their discovery by Warburg, a vast amount of knowledge has been accumulated which, on the one hand, demonstrates the highly complex nature of the problem and, on the other, shows that certain recognizable features are shared by these enzymes: they are proteins attached to prosthetic groups. The proteins are all specific and not identical with those otherwise occurring in the organism. But the chemical mechanism of the action of these enzymes is not yet worked out, at least to an extent comparable to that given in the simple example of iron catalysis in the oxidation of cysteine.

It is quite natural that enzyme chemists have, thus far, been occupied with the discovery of many kinds of enzymes, the ingenious methods of preparing them, and the measurement of their activity. But at this point we must inquire into the chemical mechanism by which they work; and here only a few speculations can be brought up. One simple suggestion is this: if it is true that the sluggishness of an oxidation process is caused by the instability (in its thermodynamic sense) of the free radical through which the over-all oxidation has to pass, then the function of the enzyme may be that of increasing the stability of the radical, in other words, that of increasing the concentration of the radical which can exist in equilibrium with the reduced and the oxidized state of the substrate. The substrate combines, reversibly, with the enzyme, and the "semiquinone formation constant" of the enzyme–substrate compound may be greater than that of the uncombined compound. We may make another suggestion. Let us suppose that the enzyme can combine not only with the substrate to be oxidized but also with the oxidizing agent. For example, methylene blue can oxidize succinic acid to fumaric acid in the presence of the enzyme called succinodehydrogenase. Suppose this enzyme can combine with both succinic acid and methylene blue. The specific structure of the enzyme brings about a definite spatial orientation and juxtaposition of fumaric acid and methylene blue. When a molecule of one of these two substances collides with a molecule of the other in a solution, the chance of an electron transfer during the short time of collision is nil; but when these two molecules are held close together in appropriate juxtaposition and orientation with respect to each other, they remain in this spatial arrangement for a long time, during which

an electron transfer may occur once in a while. Now, the transfer of a single electron establishes the free radical, and from here on the second step of oxidation takes place readily and spontaneously. In order to account for the fact that the interaction of succinic acid and fumaric acid in the presence of the enzyme and methylene blue is reversible, one must postulate also that the enzyme can combine reversibly not only with succinic acid, but also with fumaric acid; and not only with methylene blue, but also with leucomethylene blue. Such an assumption is not unreasonable and is supported by the observation that, in very many cases, molecular species of a structure similar to that of the specific active substance, inhibit the function of that substance and so compete with it for the enzyme. Out of the numerous examples discovered in recent years, one may recall the antagonism of *p*-aminobenzoic acid and sulfanilamide.

At this stage of the argument, we have reached the realm of speculation. Any further advance depends on the clarification of the structure of the enzymes and especially of the steric structure of the specific proteins. It will be the aim of future work to show that the specific structure of the enzyme forces the substrates attached to the enzyme to stay in such a mutual orientation as to permit an electron transfer, which will not occur with a reasonable probability on a free collision. A similar principle may underlie all specific enzymic reactions, as well as those not concerned with oxidation-reduction. The astounding fact that all enzymes are or contain a protein of specific structure suggests that the attachment of the substrate to the enzyme with its specific protein structure increases the chance of a thermodynamically possible reaction by forcing a spatial orientation of the interacting molecules which has practically no opportunity to occur on spontaneous haphazard collision, and moreover, by holding the interacting molecules in this specific orientation for a length of time very much longer than on the occasion of a haphazard collision by thermal motion. This is the way in which we may imagine the activation energy is overcome, and in which out of a vast number of thermodynamically permissible reactions only a few reactions serviceable for the metabolism are selected.

The road for the exploration of the mechanisms of individual metabolic catalyses will be long. Although it is still far ahead, one is encouraged to believe that the correct road sign has been found.

Selected References

Clark, W. M., *et al.*, "Studies on oxidation-reduction," *U. S. Pub. Health Service, Hyg. Lab. Bull.* No. 151 (1928).

Michaelis, L., "Occurrence and significance of semiquinone radicals," *Ann. N. Y. Acad. Sci.*, **40,** 39 (1940).

Michaelis, L., and Schubert, M. P., "The theory of two-step oxidations involving free radicals," *Chem. Revs.*, **22,** 437 (1938).

Pauling, L., and Coryell, D. C., *Proc. Natl. Acad. Sci. U. S.*, **22,** 210 (1936).

Schubert, M. P., *J. Am. Chem. Soc.*, **53,** 3851 (1931); **54,** 4077 (1932); **55,** 4563 (1933).

Thunberg, T., "Zur Kenntniss des intermediären Stoffwechsels und der dabei wirksamen Enzyme," *Skand. Arch. Physiol.*, **40,** 1 (1920).

Warburg, O., *Über die katalytischen Wirkungen der lebendigen Substanz.* Springer, Berlin, 1928.

Wieland, H., "Über den Mechanismus der Oxydationsvorgänge," *Ergeb. Physiol.*, **20,** 477 (1922).

15

MESOMERIC CONCEPTS IN THE BIOLOGICAL SCIENCES

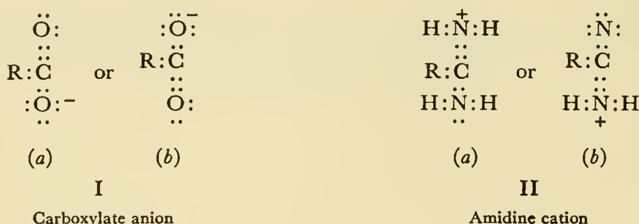
HERMAN M. KALCKAR, MEMBER OF THE RESEARCH STAFF, DIVISION OF NUTRITION AND PHYSIOLOGY, THE PUBLIC HEALTH RESEARCH INSTITUTE OF THE CITY OF NEW YORK, INC.

NO CHEMIST would question that the concept of mesomerism (or resonance), a concept which is actually based on quantum mechanics, has played an overwhelmingly important role in the development of modern physical and organic chemistry. This concept has made it possible to understand and explain the properties of numerous inorganic and organic compounds, and in a number of cases to predict chemical events. There are many indications that this concept will play an equally significant role in the biological sciences, and for that reason it merits the consideration of biologists.

Mesomerism

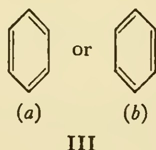
The terms resonance and mesomerism are synonomous. The former is based on quantum mechanical concepts, while the latter, more neutral term merely indicates that a substance exists as a hybrid of at least two more or less symmetric electronic states. The term "mesomerism" is more appropriate in a biological essay in which the symbols and formulas are not those of quantum mechanics, and will therefore be used here.

A group is said to display mesomerism if it can be described by two or more symmetric (or very nearly symmetric) electronic formulas, representing approximately the same potential energy. In that case, the chances that the electrons will occupy one or the other position are equally great. The electrons are therefore moving forth and back between these equivalent positions or, to use a term of physics, oscillating between the positions. The mesomerisms of the carboxylate anion and the amidine cation represent some of the simplest examples of symmetric mesomerism in a molecular group.

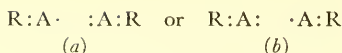


In these two sets of alternative structures, one of the oxygen or nitrogen atoms is surrounded by a complete set of eight electrons, the so-called octet, and the other by only six electrons, the last pair of electrons participating in the double bond. If this pair of electrons were moved up to complete the octet, the opposite oxygen or nitrogen would have to donate one pair of electrons from their octet in order to restore the double bond. Structures *a* and *b* are completely equivalent and indistinguishable.

The well-known benzene mesomerism is usually illustrated by the two structures shown in formula III.



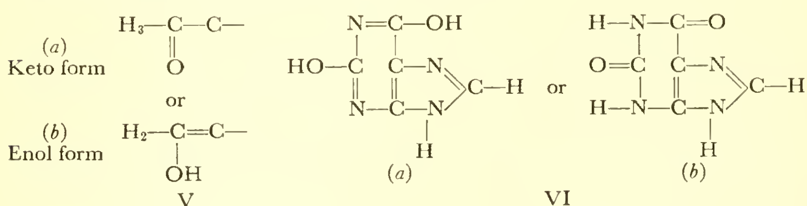
Electronic mesomerism can also exist between two molecules. This so-called intermolecular mesomerism can be illustated by the two symmetric structures of formula IV. In structure *a*, the left A



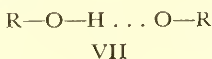
possesses an unpaired electron (odd electron) and the right A, an electron pair; in structure *b*, the situation is reversed. The shift can be represented as an oscillation of one electron between the two symmetrical molecules: R : A. : A : R. The existence of an unpaired electron gives rise to paramagnetism because the neutralized magnetic moments of paired electrons is abolished when the electron lacks its partner. The type of intermolecular mesomerism illustrated in formula IV will be discussed later in this essay.

The frequency of the oscillations which the electrons in mesomeric structures undergo is very high; and it is therefore impossible to consider a mesomeric group as possessing for an appreciable interval of time either structure *a* or *b*. There is, however, a type of mesomerism in which it is possible to distinguish between the two structures. Tautomerism is a classical illustration of this type of mesomerism.

In tautomerism, both an electron and a hydrogen atom participate in the oscillation. Since the hydrogen possesses a significant mass, the oscillation is considerably slower than in the purely electronic mesomerism; and it is therefore possible to distinguish between the two symmetrical states. Two examples of tautomerism are: (1) the enol-keto tautomerism of carbonyl compounds (V); and (2) the lactam-lactim shift of the hydroxy purines (VI).

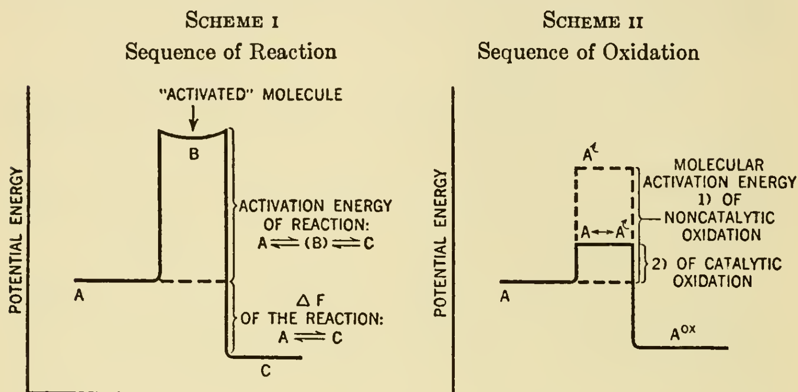


The so-called hydrogen bond results from the attraction of a hydrogen atom attached to one electronegative atom (*e. g.*, fluorine, oxygen, nitrogen) for an unshared electron pair of another electronegative atom (see formula VII):



Biological Significance of Mesomerism

What is the significance and what are the implications of mesomeric phenomena in biological reactions? The great significance of mesomerism lies in the fact that it invariably endows the molecular group with a considerable amount of additional stability. The word stability in this connection is used in its broadest sense. It has been a custom to distinguish between thermodynamic stability and the kind of stability implied in terms such as "willingness" of a group to react



spontaneously. The distinction must be considered artificial. The work of Polanyi, Eyring and Stearn—*cf.* Eyring (3)—indicates strongly that the activation of the substrate, *i. e.*, the problem of how reactive complexes are formed from stable molecular groups, is essentially a thermodynamic one. Scheme I illustrates some of these relationships.

"A" signifies the starting product and "C," the end product of the reaction A → C. ΔF is the change in free energy of the reaction, *i. e.*, the amount of potential energy which is lost when A is converted to C. The free energy or potential energy of the system drops when A is converted to C. However, in order to start the reaction, A must be activated in some way, *i. e.*, the potential barrier which represents the so-called activation energy must be overcome. Both the activation energy and the free energy change (ΔF) are influenced by mesomerism. Since the significance of mesomerism for our understanding of energy coupling in biological systems has been discussed elsewhere (6), this

aspect will be treated rather briefly in the present essay. The influence of mesomerism on ΔF can be summarized as follows: If C has an additional amount of mesomeric stability which A does not possess, the drop in ΔF of the reaction A $\rightarrow$ C is greater than it would have been if C did not possess that extra stability. We have already mentioned the carboxylate ion as a typical representative of a molecular group with extra mesomeric stability. If another molecule is introduced into such a mesomeric group, as by esterification, the symmetry of the group is disturbed and the mesomerism decreases or vanishes, which again implies that the potential energy of the complex is raised. Thus, acetic acid anhydride, CH_3—CO—O—CO—CH_3, in which both carboxyl groups have lost their state of mesomerism, possesses a much higher potential energy than that of the two acetic acids formed by hydrolysis. In terms of scheme I, acetic acid anhydride would correspond to A and the two acetic acid molecules to C.

The living cell contains at least three types of substances in which the mesomerism of two groups is mutually blocked. The first type includes the carboxyl phosphates (acyl phosphates), the second group, the amidine phosphates (phosphocreatine, phosphoarginine), and the third group, the pyrophosphates. The carboxyl phosphates are the primary oxidation products of a reaction in which a carbonyl phosphate complex undergoes enzymic oxidation. It is important to point out that the oxidation is catalyzed by an enzyme specific only for the phosphate complex. Thermodynamically speaking, the oxidation of a carbonyl-water complex to free carboxylate would be greatly favored, and the chances of forming a carboxyl phosphate would be vanishingly small, if the latter reaction were not specifically catalyzed by an enzyme.

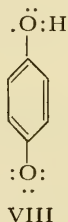
This brings up the question of the nature of enzyme catalysis.

An unusually promising approach toward an understanding of oxidation-reduction catalysis on the basis of mesomeric concepts has been made by Michaelis and his group. Since this topic is discussed in Chapter 14, only certain aspects of the problem will be treated here (8).

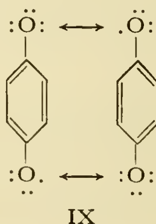
It is now generally recognized that oxidation of organic compounds, involving the removal of a pair of electrons, takes place stepwise. The removal of one electron prior to the other gives rise to the formation of a free radical displaying paramagnetism because of the

presence of an unpaired electron, possessing unneutralized magnetic moment. This free radical has, generally, a very brief existence, since it either accepts an electron again or expels the remaining odd electron. Since the free radical has so little chance of existence, the removal of the first electron is barred, so to speak, by a high potential barrier. In scheme I, the group to be oxidized would be represented by A, the free radical by B, and the final oxidation products by C. The height of the potential barrier would represent the activation energy. The activation energy is the factor which particularly interests us in connection with the concept of catalysis. If the potential barrier is too high, the chance of forming the free radical is practically nil and the rate of the reaction is zero. When the temperature is raised sufficiently, the thermal movements of the molecule become so vigorous that a certain percentage of molecules will slip over the potential barrier. Although thermal movements, of course, are of importance for events in the living cell, physiological temperatures are usually too low to allow most reactions to proceed at measurable rates. In order to bring about a reasonable rate, the living cell has succeeded in lowering the potential barrier by a special device which we call catalysis. It is in this connection that mesomerism may turn out to be of paramount importance, as the model experiments of Michaelis and his associates have so strikingly demonstrated.

The reduction of *p*-benzoquinone to the corresponding hydroquinone goes through a radical (semiquinone). The free radical has very little chance of existence because of its asymmetry (formula VIII). In basic solution, however, the semiquinone will exist as the symmetrical ion oscillating between two symmetrical electronic states. The two equivalent structures interchange the odd electron, similar to what Pauling (9) calls a three-electron bond, either through the benzene ring or by intermolecular bonds (formula IX).



VIII



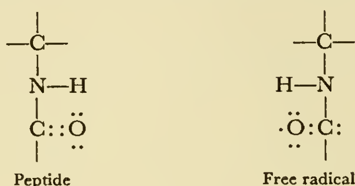
IX

The condition for stabilization of the semiquinone by mesomerism is satisfied only when the two structures are equivalent. This requirement is satisfied when the molecule is dissociated as an anion. The undissociated semiquinone in which the presence of the hydrogen atom eliminates the symmetry of the two structures does not fulfill the required condition. Correspondingly, Michaelis and his group showed that the semiquinone of phenanthrene-3-sulfonate is relatively very stable in alkaline solution in which it exists as the symmetrical anion. The semiquinone of *p*-phenylenediamine, on the other hand, has a fair chance of existence in strongly acid solutions because only then does the symmetrical phenylenediaminium cation exist. In complete accordance with the ideas just developed is the observation that the free radical of paraquinones accumulates only in the alkaline *p*H range, whereas those of the paradiamine compounds accumulate in measurable amounts only at strongly acid reactions. The "catalyzing" effect of hydroxyl ions or hydrogen ions on these two types of oxidation-reductions has actually been explained in terms of mesomerisms.

The intriguing question is whether it is possible to explain enzymic catalysis in terms of the same principles. There are observations which may provide confirmation for such an explanation. Some years ago, Haas (4) found that riboflavin phosphate, when linked to a specific protein, forms a semiquinone when undergoing reduction. This semiquinone is not observed during the reduction of free riboflavin phosphate in neutral solution, but accumulates when the reaction is acid enough to insure complete ionization. In other words, the enzyme is able to stabilize a product at neutral reaction which otherwise would exist only at strongly acid reaction. This observation may be taken as a clear indication that oxidation-reduction enzymes in some way or other are concerned with the formation of mesomeric free radicals.

Before going deeper into the discussion of the nature of enzyme catalysis, it is worth while to introduce a very ingenious theory proposed by Delbrück, dealing with the nature of reproduction phenomena. Delbrück advanced the idea that, in processes like gene reproduction, mesomeric phenomena may play an outstanding role. According to classical concepts, the probability of bringing the same components together in the same sequence as that of the original molecule is infinitely small. Delbrück is probably the first who has expressed the

idea that the reproduction of a chain molecule may be accomplished by the operation of mesomeric forces. According to this hypothesis, the chances for reproduction are reasonable in case the new molecule exists in the transitory state of a free radical (or as numerous free radicals). Each unit of the daughter chain molecule existing as a free radical will attempt to be picked up by the *identical* valence saturated units of the original molecule with which they will form a mesomeric three-electron bond of the type illustrated in formula IX. Delbrück imagines the formation of such a chain of free radicals taking place as the result of a one-step oxidation of an amino aldehyde "polypeptide" chain, as in formula X (1). The peptide represents the fully oxidized group, and the free radical, the partly reduced group.



X

A different quantum mechanical model of autocatalysis of proteins has been discussed in a recent review by Jordan (5). It is hardly necessary to stress that such hypotheses, which are based on quantum mechanical concepts, must be taken only as the very first attempt to explain the type of forces which operate in phenomena like reproduction. As Delbrück (2) expresses it: "The point I want to make is this: Quantum mechanics offers a reason why a two-step oxidation may be catalyzed by a structure which is closely similar to the oxidation product."

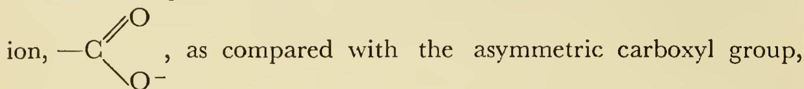
The considerations just outlined ought to encourage us to look at the problem of enzyme catalysis in terms of mesomerism. The idea that a group of specific proteins called enzymes are capable of overcoming potential barriers of 20,000 to 50,000 calories has always been difficult to comprehend. It was mentioned above that the living cell, in order to overcome the potential barrier which prevents the degradation of metabolites, had to evolve the device of catalysis. Suppose the reaction A → C illustrated in scheme I (page 232) were an oxidation (scheme II) involving two electrons and that the intermediate B repre-

sented the first oxidation step, *i. e.*, the product which lost only one electron. In that case, the chance of existence of B, possessing as it does an odd electron, would be infinitely small and the potential barrier, therefore, very high. However, if B were able to form an intermolecular mesomeric bond with A of the type illustrated in formulas IV*a* and *b*, and in IX (that is, a three-electron bond), the chances of getting the reaction started would be much better.

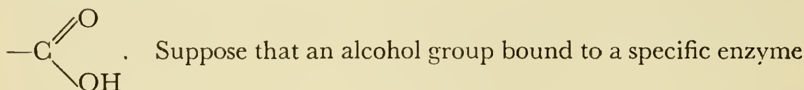
This could be illustrated in the following way: Let A be the product to be oxidized, $\overset{\nearrow}{A}$ the first oxidation product ($^{\nearrow}$ signifies the unneutralized electromagnetic moment of an unpaired electron), and A^{OX} the final product. The broken line signifies the potential barrier (or activation energy) of the noncatalyzed reaction, and the unbroken line, that of the catalyzed reaction in which A forms a mesomeric complex (signified by $A \rightleftarrows \overset{\nearrow}{A}$ or by $A{:} \rightleftarrows .A$) with $\overset{\nearrow}{A}$, both being in close proximity to each other on the specific catalyst. In other words, the state of activation is not confined to one molecule but is "spread" over two or more molecules. Yet only one of the molecules, perhaps that closest to the electron acceptor, donates the second electron and is thus converted to A^{OX}.

It might be useful, in order to make the idea more intelligible, to select as a concrete example the enzymic oxidation of an alcohol to a carbonyl group. The specific catalyst of this reaction could be classified as an alcohol dehydrogenase or as an acetaldehyde reductase (hydrogenase) with equal right, inasmuch as the enzyme is equally specific toward the alcohol and the aldehyde. The first step in the oxidation is the formation of the free radical (see also Chapter 14) which, however, has very little chance of formation if it cannot in some way or other undergo mesomerism with another molecule. Inasmuch as the protein–substrate combination brings the various oxidation products of the substrate in close proximity, one is justified in assuming that a good opportunity for the formation of mesomeric complexes is at hand. The sequence of phenomena in such a dehydrogenation might be described as follows. It is well known that, in oxidations of metabolites such as hydrocarbons, alcohols, and aldehydes, not only electrons but hydrogen too is removed. Most metabolites like hydrocarbons and alcohol are infinitely weak acids and have very little tendency to form hydrogen ions except at extremely alkaline reactions.

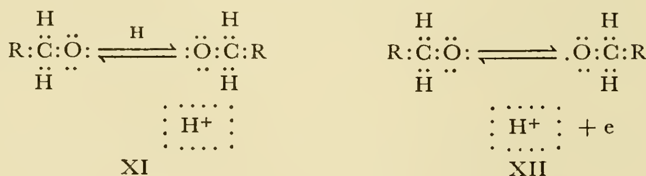
How, then, is the removal of hydrogen ions (which must precede the electron removal) accomplished in biological systems at neutral reactions? In trying to answer this question one may refer to a fact which has been particularly emphasized by Pauling (9), namely, that the high dissociation constant of carboxyl hydroxyl is mainly attributed to the higher mesomeric stability of the symmetric carboxylate



ion, —C(=O)O⁻, as compared with the asymmetric carboxyl group,



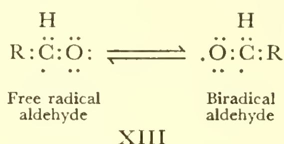
—C(=O)OH. Suppose that an alcohol group bound to a specific enzyme were able to undergo intermolecular mesomerism with two or more other alcohol groups, provided that it existed in the ionized form. This would imply that such an alcohol group showed greater tendency to release hydrogen ions when forming a substrate–enzyme complex with the specific dehydrogenase than when in free solution. If one imagines two alcohol molecules brought together in close vicinity by the specific enzyme protein, one sees immediately that the dissociation of a hydrogen ion from one of the alcohol groups would give rise to the formation of a hydrogen bond (formula XI). Since a hydrogen bond is a structural feature with a certain degree of stability, it means that the addition product would have a fair chance of existence. Now the way is paved for the first oxidation step, *i. e.*, the removal of one electron from one of the alcohol groups. If the other alcohol group simultaneously gives up its hydrogen ion, an ideal opportunity for forming a three-electron bond has been created (formula XII). The free radical, although possessing a certain degree of stability due to the existence of the mesomeric three-electron bond, is nevertheless a labile configuration, which tends to expel a second hydrogen ion and electron.



Waelsch (11) has suggested the possibility that the alcohol group of the substrate may form a mesomeric complex with certain hydroxy

groups (serine, threonine) of the protein catalyst. This suggestion may be in better accordance with kinetic data. Moreover, the suggestion may illustrate the nature of the forces responsible not only for the catalytic action of enzymes but also for their specificity.

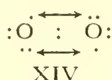
The electron acceptor goes through most of the steps illustrated here, only in the opposite order. The mesomerism between the aldehyde and the free radical is best illustrated as a complex between the free radical and the aldehyde in the biradical (triplet) state in which two electrons in the double bond are separated (formula XIII). In



XIII

the biradical state, the group contains two unpaired electrons which will give rise to the occurrence of paramagnetism. Most double bonds are considered to exist only to a very small degree in the biradical state; yet the occurrence of paramagnetism in molecules containing double bonds is a reminder of the existence of unpaired electrons. In most cases, the biradical state is considered a state of excitation. The energy of this excitation has recently been made accessible to measurement, thanks to the studies of phosphorescence spectra by Lewis and his group. Readers interested in this new and interesting development are referred to the papers of Lewis and his co-workers (7).

There are, however, a few facts concerning biradical double bonds which might be of direct interest for biologists. One of them is that molecules like oxygen and nitrate exist mainly in the biradical state and that the double bond configuration represents the labile state. Correspondingly, these molecules display a high degree of paramagnetism. Pauling (9) describes the stable biradical (or triplet) state of the oxygen molecule as a single bond with two three-electron bonds (formula XIV). It would be interesting to know whether this


XIV

is the form of oxygen which reacts with the coenzymes of the respiration enzymes.

Questioning the value of the speculations and considerations put forward in this short essay might well be justified. The mesomeric description of phenomena like reproduction and catalysis is, admittedly, sheer speculation. Yet, it is a kind of speculation which may prove to be highly constructive, inasmuch as phenomena are introduced which are within the realm of experiments. The studies of Theorell, Pauling, Coryell and Michaelis and his group, on the paramagnetic susceptibility of biologically important substances point out the direction of future research. An adaptation of these methods on a micro scale (10) might very likely pave the way for an experimental attack on problems like catalysis and auto catalysis.

References

(1) Delbrück, M., *Cold Spring Harbor Symposia Quant. Biol.*, **9**, 122 (1941).
(2) Delbrück, M., "Problems of Modern Biology in Relation to Atom Physics." Lectures delivered at Vanderbilt University, 1944.
(3) Eyring, H., *Chem. Revs.*, **17**, 65 (1935).
(4) Haas, E., *Biochem. Z.*, **290**, 291 (1937).
(5) Jordan, P., *Naturwissenschaften*, **32**, 20 (1944).
(6) Kalckar, H. M., *Ann. N. Y. Acad. Sci.*, **45**, 395 (1944).
(7) Lewis, G. N., and Kasha, M., *J. Am. Chem. Soc.*, **66**, 2100 (1944).
(8) Michaelis, L., and Schubert, M. P., *Chem. Revs.*, **22**, 437 (1938).
(9) Pauling, L., *The Nature of the Chemical Bond.* Cornell Univ. Press, Ithaca, 1939.
(10) Theorell, H., *Arkiv Kemi Mineral. Geol.*, **A16**, Art. I, 1 (1942).
(11) Waelsch, H., *private communication.*

VISCOMETRY IN BIOCHEMICAL INVESTIGATIONS

MAX A. LAUFFER, ASSOCIATE RESEARCH PROFESSOR OF PHYSICS, UNIVERSITY OF PITTSBURGH; THE LILLY AWARD IN BIOCHEMISTRY

VISCOSITY has been recognized for many years as an important characteristic of colloidal and biological materials. In spite of this fact, however, in terms of the properties of molecules and macromolecules, the meaning of viscosity remained almost completely obscure until recently. The cause of the sustained popularity of this technique during the past years should probably be attributed to the extreme ease of viscosity measurements, as contrasted with other physical approaches to the study of biological and biochemical systems, rather than to the theoretical significance of the results obtained. The outlook for the future, however, is brighter, for reasonably satisfactory theories have recently become available which describe quantitatively the viscosity of liquids and of macromolecular solutions in terms of molecular size, shape, and thermodynamic properties (22,24). Much progress has also been made in late years from a purely empirical point of view in the correlation of viscosity with the size and structure of molecules, particularly in the field of high-polymer chemistry (11).

In its most general sense, viscosity is a measure of the amount of work which must be expended in order to maintain a certain rate of flow. In the case of some liquids, the rate of flow at a fixed tem-

perature is directly proportional to the force applied per unit area on a plane parallel to the direction of flow. A fluid which flows in this manner obeys Newton's law of flow and is said to be Newtonian. The constant which relates the velocity gradient or the rate of shear in the flowing liquid to the applied shearing force is called the viscosity coefficient or, in less exact usage, simply the viscosity. In many liquids, however, the rate of shear is not directly proportional to the shearing force. Therefore, the ratio of the applied force to the resultant rate of displacement is not a constant but a variable whose magnitude depends upon the applied shearing stress or, viewed from a slightly different but more usual position, upon the resultant rate of shear. Thus, there is no such thing as a viscosity coefficient for such a fluid. However, the ratio of the shearing stress to the resultant rate of shear at a particular velocity gradient can be used as a partial description of the flow characteristics of such fluids. This ratio can be called an apparent viscosity coefficient. A complete description of the flow of a liquid of this sort requires that the way in which the apparent viscosity coefficient varies with the rate of shear be specified. Fluids which flow in this manner are often said to be non-Newtonian.

Until recently, it was fairly generally believed that Newtonian and non-Newtonian fluids were qualitatively different, but newer developments in the theory of viscous flow permit the interpretation that the difference is only quantitative. Eyring and his associates (24) have attempted to understand the flow characteristics of liquids on the basis of concepts analogous to those employed in chemical kinetics. The basic aspects of the Eyring point of view are that, in order for a liquid to flow: (1) there must be regions or holes in the body of a liquid into which molecules can jump; (2) there are potential energy barriers which tend to prevent any particular molecule from jumping into a vacant region near it; and (3) a shearing stress is a mechanical potential which aids molecules jumping in the direction of the stress and hinders those jumping in the reverse direction, thereby resulting in a net displacement in the direction of the stress. Eyring has shown that, based upon these concepts, the resultant rate of shear should be proportional to the hyperbolic sine of the ratio of the work contributed by the shearing force in moving a molecule over the energy barrier to the product of the gas constant per molecule and the absolute temperature. The proportionality constant is a function of the po-

tential energy barrier. The hyperbolic sine of a variable has the interesting property of being practically equal to the variable for values less than about $^1/_2$ and practically equal to the exponential of the variable for all values greater than about 2. Therefore, in a liquid in which the potential energy barrier is low, an observable rate of flow can be obtained by the application of a force small enough to be in the linear portion of the hyperbolic sine function. The rate of shear will thus be proportional to the shearing force. In other words, the liquid will exhibit Newtonian flow. If, on the other hand, the force required to cause an observable rate of shear is somewhat greater than that just mentioned, the rate of shear will not be directly proportional to the shearing force, but may be even an exponential function of the force. Such a liquid obviously would not exhibit Newtonian flow, but the difference between it and a liquid which would exhibit Newtonian flow would be a matter of degree and not of quality.

The application of viscosity to the problems of biology and biochemistry usually involves solutions or dispersions in aqueous solvents. In general, when a solute is dissolved in a solvent, the viscosity of the solution, η, is either greater than or less than that, η_0, of the solvent. The ratio of the viscosity of a solution to that of the solvent under the same conditions, η/η_0, is called the relative viscosity of that solution. When rigid particles of colloidal or macromolecular dimensions are dispersed or dissolved in an aqueous solvent, the viscosity of the resultant solution is greater than that of the solvent; therefore the relative viscosity is always greater than unity for such systems. The relative viscosity of a solution is usually dependent upon the concentration of the solute. Experience has shown that an equation of the sort, $\eta/\eta_0 = 1 + AC + BC^2 + \ldots$, usually can be written to describe the relationship between relative viscosity and concentration, C. A and B are arbitrary constants. For low values of concentration, $\eta/\eta_0 = 1 + AC$ or $\eta/\eta_0 - 1 = AC$. The expression, $\eta/\eta_0 - 1$, is defined as the specific viscosity of the solution. In dilute solutions, this specific viscosity is directly proportional to the concentration. The proportionality constant, A, in the equation written above, was defined by Kraemer (14) as the intrinsic viscosity and is often represented by the symbol, $[\eta]$. In strict mathematical language, the intrinsic viscosity is the limit, as the concentration approaches zero, of the ratio of specific viscosity to concentration. It can be repre-

sented symbolically as follows: $A = [\eta] = [(\eta/\eta_0 - 1)/C]_{C \to 0}$. Obviously, the numerical value of the intrinsic viscosity is dependent upon the manner in which the concentration is expressed. The latter must be specified for the former to have meaning. The intrinsic viscosity is an important constant describing a solute. As will be shown in the paragraphs to follow, it is related to the molecular structure of the solute particles and, in some cases, to the thermodynamic properties of the solution. The principal concern of viscometry is the evaluation of intrinsic viscosities.

All of the constants which have been defined thus far for solutions apply only to those which exhibit Newtonian flow, that is, to solutions for which a true viscosity coefficient can be assigned. Many, if not most, of the solutions of interest to biologists and biochemists do not exhibit Newtonian flow. For such systems, one obtains only apparent viscosity coefficients, which depend upon the rate of shear. Such systems must be described by a series of constants analogous to the relative, specific, and intrinsic viscosities of Newtonian solutions. These can be given the names: apparent relative viscosity, apparent specific viscosity, and apparent intrinsic viscosity. The latter differ from the former only in that each is dependent upon rate of shear. Thus, in a study of a non-Newtonian solution, rate of shear must be held essentially constant, or it must be specified, or the nature of the variation with rate of shear must be described. Apparent intrinsic viscosities are not necessarily any less useful than intrinsic viscosities.

It is the purpose of this discussion to consider the usefulness of viscosity in the study of biological and biochemical systems. In the remaining paragraphs, a brief consideration of the methods of measuring viscosity and a discussion of the meaning of the viscosity of solutions or dispersions of rigid colloidal particles will be presented. Particular attention will be devoted to the examination of direct experimental evidence concerning the correctness of theories used to interpret the various aspects of viscous behavior.

Intrinsic viscosities and apparent intrinsic viscosities are calculated from measurements of the viscosity of the solvent, the viscosity or the apparent viscosity of the solution, and the concentration of the solute. It is entirely outside the scope of this article to discuss the measurement of solute concentration. Suffice it to say that any of the methods of quantitative chemical and biochemical analysis might

be used, depending upon the nature of the solute under study. There are numerous methods of measuring viscosity or apparent viscosity. Perhaps the simplest from a practical point of view is that involving the rate of sedimentation of a sphere in the viscous liquid. According to Stoke's law, the retarding force resisting the translational movement of a large sphere in a viscous liquid is equal to $6\pi r\eta v$, where r is the radius and v the velocity of the sphere. When the sphere is falling under the influence of gravity, the accelerating force, f_a, acting upon it is the mass of the sphere, W_s, minus the mass of the fluid it displaces, W_L, all multiplied by the gravitational constant, G; $f_a = G(W_s - W_L)$. Since the mass of the sphere is its volume, $^4/_3\, \pi r^3$, times its density, d_s, and the mass of the displaced liquid is equal to the volume of the sphere times the density of the liquid, d_L, this equation can be written as $f_a = G\, ^4/_3\, \pi r^3(d_s - d_L)$. The sphere settles at a uniform rate when the accelerating force and the retarding force are equal:

$$6\pi r\eta v = {}^4/_3\, \pi r^3 G(d_s - d_L), \text{ or } \eta = [{}^2/_9\, r^2 G(d_s - d_L)]/v$$

Since it is possible to determine the radius and the density of a large sphere, and the gravitational constant is known, the viscosity of a liquid can be obtained by measuring its density and the velocity of settling of the sphere. It is particularly easy to measure relative viscosity* by this method, for:

$$\eta/\eta_0 = [(d_s - d_L)/(d_s - d_{L0})](v_0/v).$$

The subscript, 0, indicates a quantity describing the solvent. To obtain the relative viscosity of a solution, it is necessary to know only the density of the sphere and to measure the densities of and the velocities of fall of the sphere in both solvent and solution. The velocity of the sphere can be obtained by observing the time required for it to fall a fixed distance. The falling ball method has been used quite satisfactorily, giving results which are in good agreement with those obtained by other methods. This method is not particularly satisfactory for the study of non-Newtonian solutions because the velocity gradient in the liquid surrounding the falling sphere varies

* Practically all viscosity measurements are relative—even those of pure liquids. Much effort has been expended to determine precisely the absolute viscosity of water at 20° C. All other viscosity data are relative to the value assigned to water and are subject to the same error as the absolute viscosity of water.

with the position with respect to the sphere. Thus, the velocity gradient cannot be specified. In addition, there is a tendency for the flow around the sphere to be somewhat turbulent—an effect which cannot be eliminated entirely.

The most commonly used method of determining viscosity is that involving the flow of the liquid through a capillary tube. Poiseuille (20) demonstrated empirically and Hagenbach (11) derived theoretically from Newton's postulate that, when the rate of flow is slow enough to exclude turbulence, the volume of liquid, V, flowing through a tube in time, t, is directly proportional to the pressure difference, P, across the capillary and to the fourth power of the capillary radius, r, and inversely proportional to the length, l, of the capillary and to the viscosity coefficient, η:

$$\frac{V}{t} = \frac{\pi P r^4}{8\eta l} \text{ or } \eta = \frac{\pi r^4 t P}{8 l V}$$

If the dimensions of the capillary are known, absolute viscosity can be measured; but, even if they are not known, relative viscosity can be measured, for $(\eta/\eta_0) = (t/t_0)(V_0/V)(P/P_0)$. The most usual procedure is to observe the time for a definite volume of liquid to flow at a constant pressure difference, in which case $\eta/\eta_0 = t/t_0$, or to observe the time for a definite volume to flow under the influence of gravity with a constant average difference in liquid level. In this case, $\eta/\eta_0 = (t/t_0)(d/d_0)$, where d and d_0 are the densities of the solution and solvent, respectively. These conditions are the ones encountered when using the common Ostwald viscometer. Relative viscosities can be measured with high precision by this technique. The capillary method has the very considerable advantage of requiring easily constructed apparatus. It has one important limitation, however. The rate of flow of a Newtonian liquid at various distances from the wall of the capillary increases parabolically from the wall to the center. The velocity gradient is not constant throughout the cross section of the tube, but decreases linearly from the wall to the center. With non-Newtonian liquids, it is difficult to predict the exact nature of the variation of velocity gradient with distance from the wall, but certainly the gradient varies considerably. For that reason, this type of apparatus is not ideally suited to the study of non-Newtonian systems.

The method most satisfactory from the theoretical point of

view is that of Couette (6). Here, the liquid is placed between two concentric cylinders, one of which is rotated while the other remains stationary. The viscosity is usually determined by measuring the moment due to viscous resistance, M, acting upon the inner cylinder, when the outer one is rotated at a constant velocity. If R_1 is the radius of the inner cylinder, R_2 the radius and Ω the angular velocity of the outer cylinder, and L the depth of liquid of viscosity η between the cylinders:

$$M = 4\pi L \frac{R_1^2 R_2^2}{R_2^2 - R_1^2} \eta\Omega \quad \text{or} \quad \eta = \frac{1}{4\pi L} \frac{R_2^2 - R_1^2}{R_1^2 R_2^2} \frac{M}{\Omega}$$

If the dimensions of the apparatus and the depth of the liquid layer are known exactly, absolute viscosity can be evaluated from the ratios of the moment to the angular velocity. Relative viscosity can be obtained without knowing the dimensions exactly, for $\eta/\eta_0 = (M/M_0)(\Omega_0/\Omega)(L_0/L)$. In practice, the best method is to measure M and M_0 at the same depth of liquid and the same angular velocity. Then $\eta/\eta_0 = M/M_0$. There are numerous possible methods of measuring the moment due to viscous resistance. The most common is to suspend the inner cylinder by a wire of known restoring force, N. In this case, the inner cylinder will turn through an angle θ, equal to M/N, and this angle can be measured. Also, the inner cylinder could be mounted on a mechanism similar to that of a galvanometer, and the moment could thus be estimated in terms of the current through the galvanometer which just prevents the cylinder from turning. Another possibility would be to measure the rate of energy consumption with a wattmeter, for this rate should be proportional to the moment due to viscous resistance.

The velocity gradient at every point within the liquid enclosed between the two cylinders is directly proportional to the angular velocity of the outer cylinder and inversely proportional to the square of the radius at the point in question. In a viscometer in which the distance between the inner and outer cylinders is small compared with the radius of the inner cylinder, the velocity gradient is essentially constant everywhere throughout the enclosed liquid. Even in a viscometer in which the distance between cylinders is one-tenth the radius of the inner cylinder, the velocity gradient will not vary much more than 10% from the mean. This essential constancy of velocity gradient makes the rotating cylinder apparatus the most useful from a theoretical

point of view, for it makes possible a study of the variation of apparent viscosity with velocity gradient in non-Newtonian liquids.

Certain precautions must be observed when any one of the methods just discussed is used. These are adequately described in readily available publications (12). The investigator who desires to use viscometry must become familiar with them. However, since they are not essential to the understanding of the meaning of viscosity data, they will not be discussed here.

Usually the purpose of making viscosity measurements in biological and biochemical studies is to evaluate the intrinsic viscosity, $[(\eta/\eta_0 - 1)/C]_{C \to 0}$. If either specific viscosity, $\eta/\eta_0 - 1$, or relative viscosity, η/η_0, were a linear function of concentration, this would be a simple matter. As pointed out previously, however, specific viscosity is approximately a linear function of concentration only for very low concentrations. Many equations have been proposed to show the relationship between specific viscosity or relative viscosity and concentration. Each one seems to work for some specific system or systems. They have been reviewed by Huggins (13). All can be expanded into series of the form $\eta/\eta_0 - 1 = AC + BC^2 + \ldots.$ The only difference between them is in the relative values of the constants B and A.

The problem of determining intrinsic viscosity reduces itself to the evaluation of the constant A, for it obviously represents the ratio of specific viscosity to concentration as concentration approaches zero. The simplest method of doing this is to plot the observed specific viscosity against concentration and then draw a tangent to the curve at the origin. The slope of the tangent will be A. This practice is subject to the limitation that it uses only the data obtained at extreme dilutions, where the experimental error is bound to be great. For that reason, two other methods for evaluating A are sometimes used. The two simplest expressions relating viscosity and concentration are those of Arrhenius (1) and Bingham (2). The Arrhenius equation is $\ln \eta/\eta_0 = AC$, and the Bingham equation is $\phi/\phi_0 = \eta_0/\eta = 1 - AC$, where ϕ is the fluidity or the reciprocal of the viscosity. Both equations can be expanded into series, such as that shown above. Each can be used to evaluate A, or the intrinsic viscosity. From the Arrhenius equation:

$$\frac{\ln \eta/\eta_0}{C} = A = [\eta]$$

and from the Bingham equation:

$$(1 - \phi/\phi_0)/C = A = [\eta]$$

Thus, in the application of the Arrhenius equation, one plots $\ln \eta/\eta_0$ against C; and in the application of the Bingham equation, one plots $(1 - \phi/\phi_0)$ against C. In both cases, A or $[\eta]$ is equal to the slope of the straight line fitting the data in the region of low to moderate concentrations. These methods employ a wider range of data than the simpler procedure, and thus allow a more precise evaluation of the intrinsic viscosity.

When rigid particles of colloidal dimensions are suspended in a solvent, they increase the viscosity of the system. The first successful theoretical treatment of this effect was made by Einstein (7). His approach was from the point of view of hydrodynamics. When a liquid undergoes plane laminar or simple viscous flow, infinitesimal layers of the liquid glide over one another in the direction of flow, and each layer moves slightly faster than the layer on one side of it and slightly slower than the layer on the other side. In this process, energy is dissipated, resulting in the viscosity of the fluid. If a rigid solid object, large compared with the infinitesimal layers of liquid, is placed in such a flowing system, some of the layers of liquid in which this particle finds itself will move faster than the particle and some will move more slowly. This will tend to cause the particle to rotate, and it will also make it necessary for fluid to flow around the obstruction. The resultant disturbance in the motion of the fluid results in an added dissipation of energy by the system, in other words, in increased viscosity. Einstein derived an equation for the relative viscosity of a suspension of spherical particles as a function of the concentration of the spheres: $\eta/\eta_0 = 1 + 2.5\, C$, where the concentration is expressed as the volume fraction. On theoretical grounds, this equation should be valid only for an infinitely dilute suspension of spherical particles which are very large compared with the size of the solvent particles and very small compared with the dimensions of the viscometer. More recently, Guth (10) has shown that the viscosity of more concentrated solutions of spheres should be given by the equation:

$$\eta/\eta_0 = 1 + 2.5\, C + 14.1\, C^2$$

The Einstein equation has been subjected to a rigorous test by Eirich, Bunzl, and Margaretha (8). These investigators used rotating cylinder, falling ball, and capillary viscometers to measure the viscosity of glass spheres with a radius of 80 microns suspended in a solution of mercuric nitrate in nitric acid, mushroom spores with a radius of 4 microns dispersed in a mixture of olive oil and tetrachloroethane, and yeast cells with an average radius of 2.5 microns dispersed in water. The spherical natures of all of these particles were ascertained by examination with the microscope. The intrinsic viscosities

Table I

Viscosity of Yeast Suspensions

Volume fraction of yeast	$\eta/\eta_0 - 1$	
	Einstein equation	Observed
0.0088	0.022	0.022 0.021
0.0175	0.044	0.042 0.042 0.043
0.0350	0.088	0.082 0.087 0.085 0.083
0.070	0.175	0.195 0.196 0.197
0.140	0.350	0.530 0.524 0.500 0.515

obtained in every case differed from that given by the Einstein equation, 2.5, by less than 10%. This means that the Einstein equation is reliable when it is applied to systems which conform to the conditions for which it was derived. Representative of the data of Eirich and collaborators are the results presented in Table I, where the specific viscosities obtained in a capillary viscometer of a yeast suspension at various concentrations are compared with specific viscosities calculated from the simple form of the Einstein equation. It can be seen that the equation fits the data in very dilute suspensions.

The Einstein equation relates the relative viscosity of a suspension of spheres to the volume concentration. Thus, a viscosity measurement on a dilute suspension of spheres would constitute a means of determining the volume concentration of the suspended matter. There are circumstances under which this information can be of great value. For example, the water content of a pure preparation of particles of influenza virus can be determined by viscometry. Electron micrographs show that the particles of this virus are spheres. Therefore the total volume occupied by the particles (plus enclosed or bound water) can be determined by measuring the viscosity. The amount of solid matter can be determined by chemical analysis. Thus the total volume in suspension associated with unit weight of dry matter can be calculated; and, if the density of the dry matter is known, the percentage by weight or by volume of water in the virus particle in suspension can be calculated.

This problem can be viewed in a slightly different manner. The specific viscosity of the suspension of virus particles would be equal to 2.5 times the true volume fraction of the particles. If these particles were composed of 80% by volume of water, the volume fraction of the solid matter would be only one-fifth of the true volume fraction. Thus the specific viscosity would be equal to 5×2.5 or 12.5 times the volume fraction of solids. Therefore, for such a suspension, the intrinsic viscosity would be equal to 12.5, when concentration is expressed as volume fraction of solid matter. It is not common to find biological materials composed of very much more than 80% by volume of water. Hence, one should not expect intrinsic viscosities of spherical biological materials to have values much greater than 12.5. Yet many biological systems have much higher intrinsic viscosities. For example, tobacco mosaic virus has a minimum intrinsic viscosity of about 39, and values two or more times as great are often observed. Under some circumstances, solutions of tobacco mosaic virus nucleic acid have an intrinsic viscosity of more than 60, and thymus nucleohistone has an intrinsic viscosity of about 82 (16). The interpretation of such high values must be sought in an extension of the theory of Einstein. Today, it is generally believed that the high intrinsic viscosities are caused by extreme departure of the colloidal particles from the spherical shape.

This question was first considered by Jeffery. He extended the treatment of Einstein to include rodlike and platelike ellipsoids of

revolution. Generally speaking, during laminar flow, the long axis of a rigid rodlike or platelike particle will be subjected to a tendency to rotate about the direction in the flow plane perpendicular to the direction of flow. There will be a tendency for the particle to rotate most rapidly when the long axis is perpendicular to the flow plane and most slowly when the long axis is in the flow plane. This will result in an average orientation in the direction of flow, the extent of which will depend upon the degree of anisometry. With particles which are small enough, Brownian motion will oppose the orientation and tend to maintain a random state. The contribution that a particle makes to the viscosity of the solution will depend upon its orientation. When the long axis is perpendicular to the direction of flow the disturbance of the motion of the fluid will be greatest, and when it is parallel the disturbance will be a minimum. Thus, a suspension of randomly oriented ellipsoids will have a higher viscosity than a suspension of ellipsoids oriented more or less parallel to the direction of flow.

Simha (22) has solved the problem for the case of randomly oriented rodlike and platelike ellipsoids of revolution. The intrinsic viscosity depends upon the ratio of the major, b, to the minor, a, axis of the ellipse whose rotation generates the ellipsoid. For elongated ellipsoids, when b/a is considerably greater than 1:

$$\frac{[\eta]}{K} = \frac{(b/a)^2}{15\left[\left(\ln \frac{2b}{a}\right) - \frac{3}{2}\right]} + \frac{(b/a)^2}{5\left[\left(\ln \frac{2b}{a}\right) - \frac{1}{2}\right]} + \frac{14}{15}$$

The factor K is the ratio of the volume occupied by a hydrated particle to its anhydrous volume. For disklike ellipsoids, when b/a is considerably greater than 1:

$$\frac{[\eta]}{K} = \frac{16}{15} \frac{b/a}{\text{arc tan } (b/a)}$$

For the case of elongated ellipsoids oriented parallel to the direction of flow, Eisenschitz (9) has derived the following equation:

$$\frac{[\eta]}{K} = 1.15 \frac{b/a}{\ln \ (2b/a)}$$

By comparing this equation with Simha's, it will be obvious that the contribution of a rodlike particle to the viscosity of its solution is much

less when the particle is oriented parallel to the direction of flow than when it is randomly oriented. The situation with respect to flattened ellipsoids is similar, for Peterlin (19) has shown that the intrinsic viscosity of oriented plates is very little different from that of spheres.

The equations derived by Simha are potentially very useful, for they make it possible to interpret the intrinsic viscosity of a suspension of rigid particles in terms of the ratio of length to thickness of the particles. It is necessary, however, to have some independent means of deciding whether the particles resemble rodlike or platelike ellipsoids of revolution. The Simha equation, particularly the one for rods, is rather cumbersome in the form presented. Mehl, Oncley, and Simha (17) have calculated values of intrinsic viscosity which correspond to various values of the axial ratio, b/a, for both rodlike and disklike particles. These values are reproduced in Table II. Through the use of this table, the determination of axial ratios from viscosity measurements is greatly simplified.

TABLE II

INTRINSIC VISCOSITY OF ELLIPSOIDS OF REVOLUTION (17)

b/a	$[\eta]/K$		b/a	$[\eta]/K$	
	Elongated	Flattened		Elongated	Flattened
1.0	2.50	2.50	20.0	38.6	14.80
1.5	2.63	2.62	25.0	55.2	18.19
2.0	2.91	2.85	30.0	74.5	21.6
3.0	3.68	3.43	40.0	120.8	28.3
4.0	4.66	4.06	50.0	176.5	35.0
5.0	5.81	4.71	60.0	242.0	41.7
6.0	7.10	5.36	80.0	400.0	55.1
8.0	10.10	6.70	100.0	593.0	68.6
10.0	13.63	8.04	150.0	1222.0	102.0
12.0	17.76	9.39	200.0	2051.0	136.0
15.0	24.80	11.42	300.0	4278.0	204.0

The question of paramount importance is whether or not these theoretical equations of Simha are actually valid. Mehl *et al.* attempted to arrive at an answer to the question by comparing the shapes calculated from viscosity data for numerous proteins with those calculated from diffusion and sedimentation data, on the assumption that the protein particles are (*a*) elongated ellipsoids of revolution and (*b*) flattened ellipsoids of revolution. The shapes were calculated from sedimentation and diffusion data by the following method.

First, the molecular weights were calculated using the Svedberg (23) equation, $M = RTS/[D(1 - V\rho)]$, where R is the gas constant, T the absolute temperature, S the sedimentation constant, D the diffusion constant, V the partial specific volume, ρ the density of the medium, and M the molecular weight. From the molecular weights and the partial specific volumes, the particle volumes were computed, and from these the radii, r_0, the particles would have if they were spheres were calculated. From these radii the friction factors, f_0, which the particles would have if they were spheres were calculated by Stoke's law, $f_0 = 6\pi\eta r_0$. The actual friction factors of the various protein particles, f, were calculated from the diffusion constants by using the Einstein-Sutherland (7) equation: $f = RT/ND$, where N is Avogadro's constant. The friction ratios, f/f_0, were then obtained. Perrin (18) and others have shown from hydrodynamic considerations that, for elongated ellipsoids of revolution:

$$f/f_0 = \frac{(b/a)^{2/3}\sqrt{1 - (a/b)^2}}{\ln \dfrac{1 + \sqrt{1 - (a/b)^2}}{a/b}}$$

and for flattened ellipsoids of revolution:

$$f/f_0 = \frac{\sqrt{(b/a)^2 - 1}}{(b/a)^{2/3} \arctan \sqrt{(b/a)^2 - 1}}$$

Thus, with f/f_0 determined from experimental data, b/a values were calculated on the assumption that the particles were (1) elongated ellipsoids of revolution and (2) flattened ellipsoids of revolution. The data considered by Mehl *et al.* are presented in Table III.

An obvious ambiguity exists in the interpretation of these data, for it is possible to assume either a rodlike or a platelike ellipsoid as the model for a protein molecule, unless independent evidence concerning the shape is available. In general, however, it can be seen that the viscosity data interpreted by the Simha equation are in better agreement with the diffusion and sedimentation data as interpreted by the Perrin equation when rodlike ellipsoids of revolution are assumed to represent the shapes.

A far more serious source of ambiguity is derived from the fact that Mehl and co-workers had to assume that the particles were un-

TABLE III

THE SHAPES OF PROTEIN MOLECULES

Protein	f/f_0	[η]	b/a			
			Elongated		Flattened	
			Diff.+sed.	Visc.	Diff.+sed.	Visc.
Egg albumin	1.17	5.7	3.8	5.0	4.0	6.7
Serum albumin	1.25	6.5	5.0	5.6	5.4	7.7
Hemoglobin	1.16	5.3	3.7	4.6	3.9	6.0
Amandin	1.28	7.0	5.4	6.0	6.0	8.5
Octopus hemocyanin	1.38	9.0	7.2	7.3	8.2	11.4
Gliadin	1.60	14.6	10.9	10.5	13.6	21.0
Homarus hemocyanin	1.27	6.4	5.2	5.5	5.8	7.5
Helix pomatia hemocyanin	1.24	6.4	4.8	5.5	5.2	7.5
Serum globulin	1.41	9.0	7.6	7.3	8.9	11.4
Thyroglobulin	1.43	9.9	7.8	7.9	9.2	12.7
Lactoglobulin	1.26	6.0	5.2	5.4	5.7	6.9
Pepsin	1.08	5.2	2.5	4.5	2.6	5.8
Helix hemocyanin, *p*H 8.6	1.89	18.0	16.6	12.0	23.9	26.0

hydrated in order to obtain the shapes from both intrinsic viscosity data and friction ratios. Since both friction ratios and intrinsic viscosities are increased by hydration, both sets of calculations could be in error. Mehl *et al.* tried to assess the effect of hydration. Bull and Cooper (5) discussed the issue in greater detail, and presented a method of estimating the average degree of hydration for proteins from a consideration of essentially the same data. However, since their method involved an assumption not consistent with the Simha and Perrin equations, for the validity of both of which direct experimental evidence has been presented (16), their treatment serves more to emphasize the need for a reliable method of determining hydration than to provide such a method. One of the serious obstacles to the understanding of the physical properties of proteins and other biological materials remains the absence of adequate means for assessing quantitatively the role of hydration.

The question of the validity of Simha's equation was reconsidered in an investigation of the physical properties of tobacco mosaic virus preparations (16). Two preparations of the virus in different states of aggregation were studied in the viscometer and also in the electron microscope. The one preparation gave a relatively low intrinsic viscosity and the electron micrograph showed rod-shaped particles with a unimodal distribution of lengths. The other prepara-

tion had a much higher intrinsic viscosity and a bimodal distribution of lengths, with one maximum at the same particle length as in the case of the first preparation and the other maximum at a greater value of particle length. The ratios of the average length to thickness of the particles of the two preparations were calculated from the intrinsic viscosities by means of the Simha equation for elongated ellipsoids; and the weight averages of the squared axial ratios were computed for the two preparations from the lengths of the particles determined by the electron microscope and from the thickness obtained by x-ray diffraction measurements. The weight averages of the squared axial ratios were calculated because, according to the Simha equation, intrinsic viscosity is approximately a function of the amount of material and of the square of the axial ratio. The results are shown in Table IV.

TABLE IV

AXIAL RATIO OF TOBACCO MOSAIC VIRUS PARTICLES

Virus sample	Intrinsic viscosity	b/a	
		Simha	Electron microscope plus x-ray
A	39	20	23
B	81	31	29

The determination of particle lengths with the electron microscope is a direct measurement subject only to the error of the empirical evaluation of the magnification factor of the microscope. There is evidence to indicate that the tobacco mosaic virus particles in solution are not appreciably hydrated. However, even if they are hydrated to a considerable extent, the error in the intrinsic viscosity would be small for such extremely elongated particles. Therefore, the data shown in Table IV constitute a direct experimental verification of the essential correctness of the Simha equation for the intrinsic viscosity of rodlike ellipsoids of revolution.

The verification of the Simha equation is important because it removes the most serious objection to the use of viscosity data in the determination of particle shapes. It is quite important that one be able to estimate the shape of a biological particle, for, from the shape, the friction ratio can be computed by means of the Perrin equation. The evaluation of the friction ratio is necessary for the interpretation of

the diffusion constant or the sedimentation constant in terms of the molecular weight and dimensions of nonspherical particles. Thus, these quantities can be evaluated through viscosity and diffusion measurements or through viscosity and sedimentation velocity measurements. Viscometry, therefore, assumes dignity as a tool useful in determining the dimensions of suspended particles.

The interpretation of viscosity data is rendered difficult by three complications. Of these, the disturbing question of hydration is the most serious. Until a reliable independent method for estimating the degree of hydration is developed, there will be uncertainty in the evaluation of the shapes of proteins or other biocolloids from viscosity data and also from other physical data. The other two complications are the effect of the charge of the particles upon their intrinsic viscosity and the effect of the velocity gradient in the viscometer upon the magnitude of the apparent viscosity coefficient.

That there should be an effect upon viscosity due to the electric charge of colloidal particles was first predicted from theoretical considerations by von Smolochowski. The equation expressing this relationship has been rederived by Krasny-Ergen (15) and can be presented in the form:

$$[\eta]/K = 1 + \frac{3}{8\pi^2}\frac{\zeta^2\epsilon^2}{\lambda\eta_0 r^2}$$

The symbols ζ and r are the electrokinetic potential and the radius of the colloidal particles, ϵ and η_0 are the dielectric constant and viscosity of the solvent, respectively, and λ is the specific conductance of the solution. Kruyt and Bungenberg de Jong, Bull (4), and Briggs and his associates (3) have all verified the existence of the electroviscous effect. The intrinsic viscosity of colloidal particles with high electrokinetic potentials in solvents of low conductance can be many times greater than the intrinsic viscosity of the same particles measured under conditions under which the charge effects are suppressed. The studies of Bull and of Briggs, however, show that the Krasny-Ergen equation does not afford an adequate quantitative representation of the effect of the electric charge. Thus, the electroviscous effect is reduced to a pure hazard. The results of Bull and of Briggs show that it can be suppressed by adjusting the *p*H to the isoelectric point of the biocolloid or by adding neutral electrolytes such as sodium chloride.

Electrolyte concentrations as low as 0.02 *M* are sufficient to reduce the effect to a relatively small value. An even better, but much more laborious, method of eliminating the effect is to measure the viscosity and the other variables in the Krasny-Ergen equation at various electrolyte concentrations, and then extrapolate the results to zero electrokinetic potential. Briggs *et al.* have shown that this procedure is entirely satisfactory.

The error due to the variation of the apparent viscosity coefficient with the velocity gradient is more difficult to eliminate because it can have more than one cause. One of the possible causes of this error, which will be referred to subsequently as anomalous viscosity, is hydrodynamic. The Simha equations show that the intrinsic viscosity of rodlike and platelike ellipsoids of revolution randomly oriented in the flowing stream should be much greater than the intrinsic viscosity for the same particles completely oriented parallel to the direction of flow, as given by the treatments of Eisenschitz and Peterlin. Thus, the intrinsic viscosity of rodlike or platelike particles will vary with the velocity gradient for purely hydrodynamic reasons, because the orientation of the particles depends upon the velocity gradient. Were this the sole type of anomalous viscosity, the solution of the problem would be simple, for all one would need to do is measure intrinsic viscosity at various velocity gradients and extrapolate the results to zero velocity gradient, where the particles are randomly oriented and where, therefore, the Simha equations are valid. Robinson (21) has shown that the viscosity of tobacco mosaic virus actually does vary with the degree of orientation of the particles as determined by optical means, and that the data can apparently be extrapolated to zero velocity gradient.

However, the anomalous viscosity could be due in part to interparticle forces, which confer upon the solution the properties of a weak gel. This effect will be called structural anomalous viscosity. Structural anomalous viscosity can be understood in terms of Eyring's concepts (24). Eyring has postulated that a solution which shows structural anomalous viscosity is composed of two mechanical systems. The solvent and perhaps some of the dissolved particles constitute a medium which obeys Newton's laws of flow. In addition to this, however, relatively few bonds of high activation energy are assumed to exist between some of the dispersed particles, forming a network

with the properties of an elastic solid. These bonds could be long-range intermolecular forces, ordinary chemical bonds at points of contact between dispersed particles, or forces of some unknown nature. The only important factors are that they be relatively few and that they have very high energies. This condition will result in a structure which will undergo only elastic deformation under low-velocity gradients, but which will flow when sufficiently high shearing forces are applied. The whole solution, consisting of the Newtonian component and the network, will exhibit anomalous viscosity. Eyring has shown that flow curves resembling those actually obtained for many types of systems exhibiting structural viscosity can be constructed on the basis of such assumptions.

The difficulty presented by this situation is obvious. If the anomaly is due solely to hydrodynamic effects, the problem can be solved by extrapolating the viscosity data to zero velocity gradient; but, if the anomaly is due in part to structure, this procedure is defeated because the contribution of the structure—the unwanted factor for our purposes—is at a maximum at zero velocity gradient.

The solution to the difficulty is very simple in theory. Hydrodynamic anomalous viscosity in dilute solutions should be independent of concentration, that is, the ratio of the apparent specific viscosities at any two values of velocity gradient should not vary with the concentration of solute. On the other hand, the structural effect should tend to vanish as the concentration of solute approaches zero. Values of the velocity gradient should exist for which the ratio of the apparent specific viscosities changes with a change in the concentration. Thus, it is theoretically possible to differentiate between structural and hydrodynamic anomalous viscosities and to find experimental conditions under which the latter vanishes. Practically, however, two difficulties remain: first of all, these theoretical considerations have not yet been investigated experimentally; and, second, it might be necessary to work with extremely dilute solutions to realize the conditions postulated. This might not be possible until after the development of much greater precision than has heretofore been required in viscometry.

It has been the purpose of this discussion to present some impression of the simplicity of the viscometry technique, of the meaning of viscosity data in terms of the shape of suspended or dissolved particles,

and of the potential usefulness of viscometry in biological and biochemical investigations. The difficulties which remain as obstacles to the interpretation of viscosity data have been discussed in terms of recent theoretical developments. With the exception of the ambiguity due to hydration, the theories discussed point the way to the resolution of the uncertainties still encountered. It seems not too much to hope that studies in the next few years will establish viscometry as a reliable biochemical or biophysical tool, with dignity and meaning comparable to diffusion and ultracentrifugation.

References

(1) Arrhenius, S., *Z. physik. Chem.*, **1**, 285 (1887).

(2) Bingham, E. C., *Fluidity and Plasticity.* McGraw-Hill, New York, 1922.

(3) Briggs, D. R., Hankinson, C. L., and Hanig, M., *J. Phys. Chem.*, **45**, 866, 943 (1941); **48**, 1 (1944).

(4) Bull, H. B., *Trans. Faraday Soc.*, **36**, 80 (1940).

(5) Bull, H. B., and Cooper, J. A., *Pub. Am. Assoc. Advancement Sci.*, No. 21, 150 (1943).

(6) Couette, M., *Ann. chim. phys.*, **21**, 433 (1890).

(7) Einstein, A., *Ann. Physik*, **17**, 549 (1905); **19**, 289 (1906); **34**, 591 (1911).

(8) Eirich, F., Bunzl, M., and Margaretha, H., *Kolloid-Z.*, **74**, 276 (1936).

(9) Eisenschitz, R., *Z. physik. Chem.*, **A158**, 78 (1931).

(10) Guth, E., *Kolloid-Z.*, **74**, 147 (1936).

(11) Hagenbach, E., *Ann. Physik Chem.*, **109**, 385 (1860).

(12) Hatscheck, E., *Viscosity of Liquids.* Van Nostrand, New York, 1928.

(13) Huggins, M. L., *J. Am. Chem. Soc.*, **64**, 2716 (1942).

(14) Kraemer, E. O., *Ind. Eng. Chem.*, **30**, 1200 (1938).

(15) Krasny-Ergen, W., *Kolloid-Z.*, **74**, 172 (1936).

(16) Lauffer, M. A., *J. Am. Chem. Soc.*, **66**, 1188 (1944); *Chem. Revs.*, **31**, 561 (1942).

(17) Mehl, J. W., Oncley, J. L., and Simha, R., *Science*, **92**, 132 (1940).

(18) Perrin, F., *J. phys. radium*, **7**, 1 (1936).

(19) Peterlin, A., *Kolloid-Z.*, **86**, 230 (1939).

(20) Poiseuille, J. L. M., *Mém. Savants Étrangers*, **9**, 433 (1846).

(21) Robinson, J. R., *Proc. Roy. Soc. London*, **A170**, 519 (1939).

(22) Simha, R., *J. Phys. Chem.*, **44**, 25 (1940).

(23) Svedberg, T., *Chem. Revs.*, **14**, 1 (1934); **20**, 81 (1937).

(24) Tobolsky, A., Powell, R. E., and Eyring, H., in *The Chemistry of Large Molecules.* Interscience, New York, 1943, p. 125.

ISOTOPE TECHNIQUE IN THE STUDY OF INTERMEDIARY METABOLISM

D. RITTENBERG, ASSISTANT PROFESSOR IN BIOCHEMISTRY, COLLEGE OF PHYSICIANS AND SURGEONS, COLUMBIA UNIVERSITY; THE LILLY AWARD IN BIOCHEMISTRY

DAVID SHEMIN, ASSOCIATE IN BIOCHEMISTRY, COLLEGE OF PHYSICIANS AND SURGEONS, COLUMBIA UNIVERSITY

THE STUDY of intermediary metabolism encompasses the reactions undergone by cellular and dietary constituents as well as the reactions involved in the formation of transient intermediates which occur during the degradation of various compounds to products of excretion. In studying these reactions, dietary components were labeled in an attempt to trace their fate after ingestion by the animal. These efforts were but partially successful, since in all cases the labels employed so changed the properties of the substance that the cell treated it differently from its natural analogue. Indeed, in most cases the success attained was actually due to the difference between the metabolism of the foreign substance and that of its natural analogue. When fatty acids are administered to a normal dog the only excretory products produced are carbon dioxide and water; when the ω-phenyl substituted fatty acids are fed (18) there are excreted ω-phenyl substituted acids having a carbon chain with fewer carbon atoms than the original by one or more pairs. From

such data, however, Knoop deduced that the oxidation of fatty acids proceeds by the removal of two carbon atoms at a time; but the possibility still existed that the oxidation of the unnatural acids might proceed by a mechanism entirely different from that of the natural fatty acids. Such techniques were of but restricted use. It was clear that little further advance in the study of the intermediary metabolism could be expected from this direction.

At about the time that Knoop was using the phenyl group to label fatty acids and amino acids, a more powerful technique was in the process of development as a result of the advances being made on the structure of the atomic nucleus. During the 19th century, it was universally believed that the chemical and physical properties of an element were a function of its atomic weight. Mendelyeev, by classifying atoms according to their masses, had been able to predict the properties of elements as yet unknown. In 1906, however, Boltwood (6) discovered a radioactive element, ionium, which was chemically and physically identical with thorium but which had a different atomic weight. Other such examples were found among the radioactive elements and it soon became clear that the properties of an atom were determined, not by its mass, but by its nuclear charge, its atomic number. It was therefore possible to have two or more atoms having the same nuclear charge but differing atomic weights. Such atoms were called isotopes by Soddy (41). In 1910 he stated (40), "Chemical homogeneity is no longer a guarantee that any supposed element is not a mixture of several different atomic weights, or that any atomic weight is not merely a mean number."

With the development of the mass spectrometer by Dempster (11) and Aston (1), it became possible to investigate the isotopic composition of the stable elements; and it is now known that most elements have at least two isotopes, though a few, such as phosphorus and iodine, seem to exist in but one atomic variety. With the exception of the elements from hydrogen to carbon, every element of an even atomic number has three or more stable isotopes, while every stable element with an odd atomic number has two or less stable isotopes.

Our understanding of the laws governing the structure of atomic nuclei is as yet in a primitive stage. We think that atomic nuclei are built up from two elementary particles, the proton, a particle having a mass of one atomic unit and a single positive charge, and the neutron,

an unchanged particle having a mass very nearly the same as that of the proton. A particle containing 6 protons and 6 neutrons will be a carbon nucleus, which has a charge of 6. Using the notation of the nuclear physicists, the particle will be denoted as C_6^{12}, the superscript indicating its mass and the subscript its nuclear charge. If we add one more neutron to the nucleus, we obtain C_6^{13}. This is an isotope of carbon, for we have merely increased the mass and not the nuclear charge. Introduction of a proton into the C_6^{12} nucleus would yield N_7^{13}, an isotope of nitrogen, for both the mass and the nuclear charge have been increased by one unit. The particle C_6^{13} exists in nature, for naturally occurring carbon is a mixture containing one atom of C_6^{13} for every ninety-nine atoms of C_6^{12}. Introduction of one more neutron into the nucleus of C_6^{13} would give rise to C_6^{14}, and subtraction of one neutron from C_6^{12} would form C_6^{11}. Both these particles are carbon atoms and both have been artificially prepared. Neither one exists in nature, since each is radioactive, with a half-life of about 24,000 years and 21 minutes, respectively. Other carbon isotopes may be visualized, such as C_6^{9}, C_6^{15}, or C_6^{16}, but presumably these are radioactive and very unstable, *i. e.*, their half-life times must be very short.

A study of the stable isotopes suggests that a nucleus is stable only if it has approximately the same number of protons and neutrons in its structure. The forces which bind the nucleus seem to be interactions between pairs of protons and neutrons.

In the last ten years, the physicists have prepared artificially radioactive isotopes of practically every element. During the same period, chemists (16, 20, 28, 42–45), notably Urey, concentrated and obtained in more or less pure form the heavy isotopes of hydrogen, oxygen, nitrogen, carbon, and sulfur. At the present time, the biochemist has available for his investigations isotopes, either stable or radioactive, of every element of biological interest. Table 1 lists some of these isotopes with their natural abundance, if stable, and their half-life time, if radioactive. The elements chosen were those of biological interest, and it is fortunate that for every element there is available either a concentrate of a heavy stable isotope or a suitable radioactive isotope.

The choice of the particular type of isotope to be used for any specific problem depends on the isotopes available and the type of problem. In the study of phosphorus metabolism by the isotope

technique, radioactive phosphorus, P^{32}_{15}, must be employed, since only one stable atomic species exists, namely, P^{31}_{15}. The radioactive isotope has a convenient half-life (14 days), and the radiation it emits is easy to measure. Furthermore, the preparation of P^{32}_{15} is relatively simple. The choice for the case of oxygen is also clear; the stable isotope O^{18}_{8} should be used since the radioactive oxygen isotope O^{15}_{8} has a half-life of only 126 seconds, a time too short to permit the carrying out of almost any biological experiment. Similarly, the stable heavy hydrogen, H^{2}_{1}, is preferable to the radioactive isotope, H^{3}_{1}. The reason in this case is not a restricted half-life but the fact that the radiation emitted by H^{3}_{1} is so soft that it is extremely difficult to measure. Also, in this particular case, the measurement of H^{2}_{1} concentrations is simpler than that of the radioactivity.

By this reasoning we conclude that the stable isotopes should be employed for studies of hydrogen, nitrogen, and oxygen, and the radioactive isotopes for sodium, phosphorus, calcium, iron, and iodine. The choice for carbon and sulfur is more difficult. While C^{11}_{6} cannot compete with C^{13}_{6}, isotope C^{14}_{6} can. The actual choice here depends on highly technical points involving the analytical procedures and the relative cost of preparation of the isotopes. It seems likely that, at least for the next five years, C^{13}_{6} rather than C^{14}_{6} will be employed but that thereafter C^{14}_{6} may supplant it. The choice in the case of sulfur is also not easy, since neither form (S^{34}_{16} or S^{35}_{16}) has a definite advantage over the other.

One of the problems which enters into the choice between the stable and the radioactive isotope is the relative ease of the isotope determinations. With the exception of hydrogen, all stable isotopes require the use of a mass spectrometer for their measurement. The apparent complexity and cost of the present type of mass spectrometers have often acted as a deterrent to investigators who have contemplated employing this technique. It is likely that, after the war, the mass spectrometer will be greatly simplified and that its cost from some commercial instrument maker will not exceed that of a good optical spectrograph.

The analysis of organic compounds for deuterium may be carried out by the falling drop procedure (17). The apparatus involved and the procedures used are so simple that no competent investigator need have any hesitation about setting up such equipment. However,

TABLE I

SOME NATURAL AND RADIOACTIVE ISOTOPES[a]

Element	Mass	Relative abundance, atom per cent	Half-life	Element	Mass	Relative abundance, atom per cent	Half-life
H	1	99.99	—	S	31	—	3.2 sec.
	2	0.01	—		32	95.1	—
	3	—	30 yrs.		33	0.7	—
					34	4.2	—
C	10	—	8.8 sec.		35	—	88 days
	11	—	21 min.		36	0.02	—
	12	98.9	—				
	13	1.1	—	Ca	39	—	4.5 min.
	14	—	10^3–10^5 yrs.		40	96.97	—
					41	—	8.5 days
N	13	—	9.9 min.		42	0.64	—
	14	99.63	—		43	0.15	—
	15	0.37	—		44	2.06	—
	16	—	8 sec.		45	—	180 days
					46	0.003	—
O	15	—	126 sec.		48	0.19	—
	16	99.76	—		49	—	2.5 hrs.
	17	0.04	—				
	18	0.20	—	Fe	53	—	8.9 min.
	19	—	31 sec.		54	6.04	—
					55	—	4 yrs.
Na	21	—	23 sec.		56	91.54	—
	22	—	3 yrs.		57	2.11	—
	23	100	—		58	0.28	—
	24	—	14.8 hrs.		59	—	47 days
Mg	23	—	11.6 sec.		124	—	4 days
	24	77.4	—	I	126	—	13 days
	25	11.5	—		127	100	—
	26	11.1	—		128	—	25 min.
	27	—	10.2 min.		130	—	12.6 hrs.
					131	—	8 days
P	29	—	4.6 sec.		>131	—	2.4 hrs.
	30	—	2.6 min.				
	31	100	—				
	32	—	14.3 days				

[a] Data from R. D. Evans in *Medical Physics*. Year Book Publishers, Chicago, 1944.

even in this case the mass spectrometric method is probably superior to that of the falling drop if only for the reason that smaller samples can be analyzed. From 100 to 200 milligrams of organic compound is necessary for analysis by the falling drop technique, but from 3 to 5 milligrams is sufficient for analysis by the mass spectrometer.

The essence of the isotope technique in biochemical research consists in the preparation of a compound in which one or more of the atomic components have an abnormal isotope concentration. In this manner, a substance is obtained which, by suitable isotope analysis, can be detected and estimated quantitatively in the presence of either its normal analogue or of other compounds, and which behaves like its normal analogue in a biochemical system.

Since the electronic configurations of the isotopes of an element are almost identical, isotopic compounds have similar chemical properties. There are, however, some physical properties which are dependent on the molecular mass. Diffusion is one such property, since the rate of this process is inversely proportional to the square root of the mass. Nevertheless, these differences in physical properties are not of importance in cellular metabolism. Molecules of glycine, heavier by one unit of mass because of the introduction of one atom of a heavy isotope of either hydrogen, nitrogen, or carbon, will differ from the normal analogue in their rate of diffusion by about 0.7%. There is direct evidence available that the cell does not distinguish between isotopic isomers. Compounds containing one or more heavy atoms of hydrogen, nitrogen, etc. normally exist in nature. For example, 0.37% of all glycine molecules contain N^{15} atoms instead of N^{14}. Such compounds are therefore not foreign to the cell. If living cells accorded special treatment to isotopic isomers, we should expect the isotope concentrations to vary as we pass from tissue to tissue and from the living into the inorganic world. However, the natural abundances of the heavy isotopes of all elements studied are uniformly the same in organic and in inorganic substances. It is striking that this normal abundance of isotopes even applies to meteorites originating from interstellar space (21).

The labeled compounds prepared must, of course, be of such a nature that the isotope distribution will not be altered by mere exchange. Thus it is impossible to label ammonia in the water system with either deuterium (H_1^2) or tritium (H_1^3), since the hydrogen atoms of ammonia are readily replaced by those of the water. Several other restrictions limit the scope of deuterium (33). When an organic compound is labeled with deuterium, the carbon chain is not being directly labeled, for deuterium may be lost from an organic compound even though no reactions involving the carbon chain take place.

Thus, α,α'-dideuteriosuccinic acid in the living cell may lose all its deuterium by repeated dehydrogenations and hydrogenation (12). The appearance of deuterium in the cellular fluids cannot be regarded as proof that the succinic acid has been oxidized to carbon dioxide and water. If, on feeding a deuterio compound to an animal, one isolates another compound containing a higher concentration of deuterium than exists in the body fluids, it may be taken as positive proof of the utilization of the first compound in the synthesis of the second. The power and elegance of this procedure is illustrated by the finding that the feeding of deuterioacetate to rats leads to the formation of deuteriocholesterol (3). There is no other method at present available which could have disclosed such a complex synthetic reaction. It further illustrates the biological stability of the C—D bond.

While, at first sight, the fact that the C—D bond can be broken in the course of chemical reactions (by dehydrogenations, enolizations, etc.) seems to limit the field of applicability of deuterium, more careful considerations suggest that this is, in actuality, one of the most useful properties of deuterium. The same reactions which will remove deuterium from an organic compound will also introduce it if the medium in which the reaction occurs contains heavy water (D_2O). Injecting D_2O into an animal makes possible the study of a host of reactions and their rates, by determining the velocity of introduction of deuterium into the organic compound. When the body fluids of mice are enriched in D_2O, the isotope concentration in the fatty acids increases and finally reaches a value half that of the body fluids. The isotope enters the fatty acid molecule in the course of the reactions involved in the synthesis of the fatty acids. From such observations (26) it was concluded that half the fatty acids of the mouse was degraded and resynthesized in about seven days.

On the whole, it appears that deuterium is at present the most valuable single isotope available to the biochemist. In the use of heavy nitrogen as a label for nitrogenous compounds, it must also be remembered that reactions exist which may detach the labeled atom and transfer it to another molecule. This is of course especially true in the case of an amino acid. The use of carbon for labeling organic compounds is not subject to such uncertainty because no exchange reactions occur. Further, the conversion of one organic compound to another is, by definition, the utilization of the whole or part of the

carbon skeleton of one compound for the formation of another. Labeling a compound with carbon isotopes is thus the most direct method for demonstrating conversions.

The biochemical isotope technique has been applied in four fields: (1) in detecting the conversion of one compound to another; (2) in studying the mechanism of biochemical reactions; (3) in measuring the rates of reactions; and (4) in determining the amount of a constituent in a mixture. It has the advantage over all other methods that it may be carried out on the intact organism under normal physiological conditions.

The first biological experiment in which the isotope technique was used was worked out by Hevesy in 1923 (15). He traced the transport of lead in plants whose roots were immersed in a medium containing a radioactive isotope of lead. This type of experimentation marked time until Urey, in 1933, discovered and concentrated heavy water. At about the same time, artificial radioactivity was discovered. As a result, there became available to the biochemist concentrates either of heavy isotopes or of radioactive isotopes of all the elements of importance in the study of intermediary metabolism.

Beside the numerous individual problems which in the past ten years have been attacked by this method, the isotope technique has revealed one phenomenon of fundamental significance in the living cell, *viz.*, the dynamic state of tissue constituents. These experiments have shown that practically every component of the animal body is constantly degraded and resynthesized at a rapid rate, and that the former distinctions between structural and metabolic components of the cell or of endogenous and exogenous metabolism are nonexistent. The maintenance of the form and structure of the adult cell or tissue is the result, not of a static state in which no reactions take place, but rather of a dynamic equilibrium in which the synthetic and degradative reactions proceed at equal rates. The existence of this stationary state has been shown by experiments with several isotopes and with every type of tissue, with the possible exception of the tissues of the adult brain. Not only is the form and structure of the organ governed by the rates of synthesis and degradation of its constituents but also its composition. While the autolysis of tissues which becomes apparent on the death of an animal was formerly thought to involve reactions which began after death, we now believe that many of these degrada-

tive reactions occur continuously. During life, the effects of these degradative processes are nullified by synthetic reactions which re-form as much tissue constituents per unit time as are destroyed by the autolytic reactions; at death, the synthetic reactions, which in general are coupled to oxidative reactions, cease and we observe the effects of degradative processes alone. The mechanisms by which the cell regulates these reactions are completely unknown. Since these reaction rates can only be investigated in the intact animal, the only method at present available which can be used is the isotope technique. It can safely be assumed that the determination of these reaction rates will occupy the attention of many investigators in the coming years.

There are several published researches which illustrate the method by which the isotope technique may be applied to the study of the conversion of one compound to another in the intact animal. It was known from feeding experiments with immature rats that, while phenylalanine was an essential component of the diet, tyrosine was not (29). From the chemical similarity of these amino acids, it was suspected that phenylalanine could be oxidized to tyrosine. Furthermore, in tyrosinosis, the feeding of either tyrosine or phenylalanine results in excretion of *p*-hydroxyphenylpyruvic acid (22). It should be recognized, however, that this latter piece of evidence is not as definitive as it appears. The fact that the feeding of a compound to an animal results in the excretion of an excess of another related compound cannot be taken as positive proof that the first compound has been converted to the second, even if, as in some cases, a stoichiometric relationship appears to exist between the amounts of these substances. It is well known that the feeding of mineral oil to rats results in an enhanced excretion of fecal steroids. This is certainly not the result of the conversion of these hydrocarbons to cholesterol but is due to interference by the mineral oil with the reabsorption of cholesterol by the intestinal tract. Further, the addition of glycine to the diet results, in the next 24 hours, in an added excretion of nitrogen in the urine equivalent to that of the added glycine. The conclusion that this extra urinary nitrogen is derived from the dietary glycine would not be correct. We have fed to humans small quantities of glycine labeled with N^{15} (about 10 mg. per kilo weight) and have found that only about 33% of the labeled glycine nitrogen appears in the urine in the next 24 hours. Even after 3 days only about 50% has been

excreted. The other nitrogen which is excreted must have arisen from other sources, probably the tissue proteins (unpublished experiments).

The phenylalanine–tyrosine conversion was studied in rats by feeding them a normal diet to which was added a small amount of labeled phenylalanine (23) prepared by replacing the normal hydrogen atoms of the benzene ring with deuterium atoms. From the tissues of rats which had been fed this labeled phenylalanine, tyrosine was isolated containing a high concentration of deuterium. Here is a direct proof for the conversion. This transformation of phenylalanine to tyrosine took place, not on a diet deficient in tyrosine and calculated to force its synthesis, but on a diet containing an adequate amount. The conversion of phenylalanine to tyrosine is independent of the dietary composition. Such a reaction could have been detected by no previously known experimental technique.

The method just described is of course of the broadest generality, but there are some cases in which its experimental application is difficult. Recourse must then be taken to more indirect methods. The feeding to animals of benzoic acid results in the excretion of hippuric acid, benzoylglycine. In experiments in which glycine labeled with heavy nitrogen was fed to rats, it was found that one-third of the glycine used for conjugation with the benzoic acid was of dietary origin and two-thirds was supplied by the proteins of the tissues (27). The feeding of benzoic acid thus results in the excretion of a sample of glycine of the tissues. Conversions of other compounds to glycine may easily be tested by feeding the labeled test substance, together with benzoic acid. The feeding of these compounds which are converted to glycine *in vivo* will result in the excretion of hippuric acid containing a high concentration of N^{15}. Serine gives such a positive result, whereas leucine, alanine, and ethanolamine give negative results (39). The fact that ethanolamine is not converted to glycine indicates that the mechanism for the conversion of serine to glycine does not involve decarboxylation to ethanolamine and subsequent oxidation of the alcohol, but suggests that the β-carbon atom of serine is split off to yield glycine directly. The proof of the conversion of serine to glycine is not as direct as that of phenylalanine to tyrosine, for only the amino group of serine was labeled and not the carbon chain. It is conceivable that only the amino group was transferred. In control experiments it has been demonstrated that the feeding of

equivalent amounts of ammonia does not give rise to a high concentration of N^{15} in the excreted hippuric acid. This eliminates the possibility that the serine was deaminated and that the ammonia was used to form glycine. The conversion of serine to glycine was confirmed and the mechanism established by labeling the carboxyl group of serine with C^{13}. Serine is converted to glycine by the splitting off of the β-carbon atom (unpublished experiments).

A good example of the use of isotopes is to be found in the experiments which led to establishment of the precursors of creatine. The balance type of experimentation had completely failed to solve this problem, and had not even been able to determine whether creatine was biologically dehydrated to creatinine. Examination of the structure of creatine suggested that glycine, arginine, and methionine might be related to creatine. The feeding of labeled compounds demonstrated that the COOH—CH_2N— group of creatine was derived from glycine (5), the $C\genfrac{}{}{0pt}{}{\diagup NH}{\diagdown NH_2}$ group from the amidine group of arginine (5) and the methyl group from methionine (47).

In the last few years, a beginning has been made in the measurement of the rates of the reactions involved in the synthesis of tissue proteins. Feeding an animal with an amino acid labeled with N^{15} results in the incorporation of this nitrogen into most of the amino acids of the proteins (31), a process which is of course dependent on the rate of the formation of peptide bonds and the rate of transfer of nitrogen from one amino acid to another. Measurement of the rate of increase or decrease of isotope concentration of tissue proteins gives a measure of the rate of formation of the proteins. Such studies reveal a remarkable rate for protein synthesis in rat livers; half of the protein is degraded and resynthesized in six days (36).

Not only are the "normal proteins" of animal tissues in a state of dynamic equilibrium but also the antibodies, as shown with the aid of isotopes in an actively immunized rabbit. The antibody protein is broken down and newly synthesized even when the antibody titer is declining. The rate of decline of the antibody titer is not determined by an uncompensated destruction of the antibody protein but rather is the result of two opposing reactions of which the rate of breakdown is faster than the rate of synthesis. The feeding of isotopic

glycine to an actively immunized rabbit at the stage at which the titer is falling results in the rapid incorporation of the isotopic nitrogen in the antibody even in the phase of decline (32).

The rates of protein synthesis and destruction have been studied in growing tissues, regenerating liver, and tumor tissues. Since it has been shown by the isotope technique that tissues are in a state of flux, it is reasonable to conclude that the synthetic or anabolic rate in a tissue of constant weight must be equal to the degradative or catabolic rate. In a growing tissue, however, the rate must be greater than the degradative rate. This disparity in rates may be the result of a relative increase of the synthetic rate as compared with that in the nongrowing tissue, of a relative decrease of the rate of degradation, or of a combination of the two possibilities. Which of the three possibilities operates and accounts for growth in regenerating liver tissue can be ascertained by the isotope technique. From unpublished data obtained in our laboratory, it appears that, in the regenerating tissue, the synthetic rate is similar to the rate found in the nongrowing liver but the degradative rate is markedly decreased. Growth is the result of an inhibition rather than the initiation or acceleration of reactions. It seems likely that the next few years will see an intensive study of the rate of synthesis of various tissue components under diverse physiological conditions in the intact animal.

An extension of the isotope technique is the isotope dilution method of analysis (25), a technique which has been very useful in various problems, particularly that of accurately determining the amount of different compounds in a mixture. The technique is based on the fact that a compound which contains more than the normal abundance of isotope is inseparable from its normal analogue by the usual laboratory procedures. In this procedure, a small amount, for example, of isotopic glutamic acid is added to a protein hydrolyzate and a representative sample of glutamic acid is isolated from this mixture. From the amount of glutamic acid added (A) and its N^{15} concentration (C_0) and the N^{15} content of the isolated glutamic acid (C), the amount of glutamic acid (B) originally present in the mixture can be calculated from the formula: $B = A[(C_0/C) - 1]$. This technique applied to the determination of the amount of an amino acid in a pure protein has provided the most reliable analytical values at present available (14,38). One of the interesting developments of

the isotope dilution method has been its use in determining whether or not certain substances are involved in animal metabolism. For example, it has been demonstrated that large amounts of acetate are produced daily in the rat though its concentration is never great enough to be measured directly (4).

The role of *d*-amino acid oxidase in the animal organism is obscure, though it appeared possible that its function was to deaminate *d*-amino acids that might arise from a partially or completely symmetrical synthesis of amino acids from α-keto acids (19). The demonstration that *d*-amino acids are not formed in the animal organism would be very difficult with ordinary techniques. If, however, a substance labeled with an isotope is administered to an animal and subsequently isolated either from the tissues or the urine, then the isotope concentration of the recovered compound will indicate whether or not it has been diluted by some of the same nonisotopic substance synthesized by the animal. *d,l*-Glutamic acid and *d,l*-tyrosine labeled with N^{15} were administered to rats and the *d*-amino acids isolated from the urine. Since the isotopic content of the isolated *d*-amino acid was the same as the amino acid administered, it would appear that *d*-amino acids are not formed in the animal organism (35). By a similar procedure Bernhard (2) demonstrated that the long-chain dicarboxylic acids are not intermediates in the metabolism of the fatty acids, as had been suggested by Verkade (46). Bernhard used deuteriodicarboxylic acids, and isolated from the urine of the test animal fatty acids containing the same deuterium concentration as the sample that had been fed.

The finding of N^{15} in the α-amino group of amino acids of an animal fed either labeled ammonia or an amino acid demonstrates that the organism can synthesize these amino acids from the corresponding keto acids. Such evidence throws no light on the question as to whether an amino acid is essential or nonessential from the standpoint of nutrition, since, while the cell may be able to convert a keto acid to the required amino acid, it may not be able to synthesize the appropriate keto acid. This seems to be true for most essential amino acids. On the other hand, if, under these circumstances an amino acid contains no labeled nitrogen, then we can conclude that the animal cannot synthesize it, even though the carbon skeleton is available. Examples of these two cases, in the rat, are leucine (24) and lysine

(31). After the feeding of labeled amino acids, N^{15} can be found in the α-amino group of leucine, which suggests that α-ketoisovaleric acid can be aminated. No N^{15}, however, is found in lysine, indicating either that α-keto-ϵ-aminohexanoic acid is not formed, or that it could not be aminated if formed. In general, it is not possible by the use of isotopic nitrogen to determine whether the carbon skeleton is essential. In certain special cases, such as that of histidine, it becomes feasible. When the amino acid is isolated from rat tissues after the administration of other labeled amino acids, it is found that all the N^{15} is present in the α-amino group (34), none being found in the imidazole ring. This, taken in conjunction with the nutritional replaceability of *l*-histidine by *d*-histidine and by imidazole lactic acid (9,10), forms direct evidence that the organism cannot synthesize the histidine skeleton. It has been claimed that histidine is not essential in man (30); by the feeding of a labeled amino acid to a human it should be possible to settle this and other such problems.

A detailed study of the N^{15} concentrations in the four nitrogen atoms of arginine isolated from the tissue of rats which have received glycine labeled with N^{15} indicates that the α-amino group is not subjected to the reversible deamination-reamination reaction. On degradation of the arginine, it is found that the nitrogen atoms of the α- and δ-carbon atoms contain the same N^{15} concentration (37). The nitrogen on the δ-carbon atom of arginine cannot be involved in a reversible deamination-reamination reaction since it is blocked by the amidine group. Were the α-amino group subject to reversible deamination-reamination, the likelihood of its isotope concentration being the same as that of the nitrogen on the δ-carbon atom would be small. The mechanism for the formation of arginine does not seem to involve the amination or transamination of the corresponding keto acid. The introduction of equal concentrations of N^{15} in the α- and δ-nitrogen atoms of arginine is brought about through a cyclic process involving the formation of ornithine from proline (37). Only in the cases of glutamic acid (7,8,13), aspartic acid (7,8), and alanine (7,8), are the mechanisms for their formation from keto acids known with certainty. It is known, of course, that the keto acids of some essential amino acids may substitute for the amino acid in growth experiments; but the mechanism of the amination is unknown. Therefore, other mechanisms may exist for the formation of amino acids, such as conversion of

one amino acid to another, examples being the conversion of serine to glycine and the conversion of phenyl-alanine to tyrosine.

References

(1) Aston, F. W., *Phil. Mag.*, **39**, 449 (1920).
(2) Bernhard, K., *Helv. Chim. Acta*, **24**, 1412 (1941).
(3) Bloch, K., and Rittenberg, D., *J. Biol. Chem.*, **145**, 625 (1942).
(4) Bloch, K., and Rittenberg, D., *J. Biol. Chem.* (in press).
(5) Bloch, K., and Schoenheimer, R., *J. Biol. Chem.*, **138**, 167 (1941).
(6) Boltwood, B. B., *Am. J. Sci.*, **24**, 370 (1907); **25**, 365 (1908).
(7) Braunshteĭn, A. E., and Kritsman, M. G., *Enzymologia*, **2**, 129 (1937).
(8) Cohen, P. P., *J. Biol. Chem.*, **136**, 565 (1940).
(9) Cox, G. J., and Berg, C. P., *J. Biol. Chem.*, **107**, 497 (1934).
(10) Cox, G. J., and Rose, W. C., *J. Biol. Chem.*, **68**, 781 (1926).
(11) Dempster, A. J., *Phys. Rev.*, **11**, 316 (1918).
(12) Erlenmeyer, H., Schoenauer, W., and Süllmann, H., *Helv. Chim. Acta*, **19**, 1376 (1936).
(13) Euler, H. v., Adler, E., Günther, G., and Das, N. B., *Z. physiol. Chem.*, **254**, 61 (1938).
(14) Foster, G. L., *J. Biol. Chem.*, **159**, 431 (1945).
(15) Hevesy, G., *Biochem. J.*, **17**, 439 (1923).
(16) Huffmann, J. R., and Urey, H. C., *Ind. Eng. Chem.*, **29**, 531 (1937).
(17) Keston, A. S., Rittenberg, D., and Schoenheimer, R., *J. Biol. Chem.*, **122**, 227 (1937).
(18) Knoop, F., *Beitr. chem. Physiol. Path.*, **6**, 150 (1905).
(19) Knoop, F., *Z. physiol. Chem.*, **274**, 291 (1942).
(20) Lewis, G. N., and Cornish, R. E., *J. Am. Chem. Soc.*, **55**, 2616 (1933).
(21) Manian, S. H., Urey, H. C., and Bleakney, W., *J. Am. Chem. Soc.*, **56**, 2601 (1934).
(22) Medes, G., *Biochem. J.*, **26**, 917 (1932).
(23) Moss, A. R., and Schoenheimer, R., *J. Biol. Chem.*, **135**, 415 (1940).
(24) Ratner, S., Schoenheimer, R., and Rittenberg, D., *J. Biol. Chem.*, **134**, 653 (1940).
(25) Rittenberg, D., and Foster, G. L., *J. Biol. Chem.*, **133**, 737 (1940).
(26) Rittenberg, D., and Schoenheimer, R., *J. Biol. Chem.*, **121**, 235 (1937).
(27) Rittenberg, D., and Schoenheimer, R., *J. Biol. Chem.*, **127**, 329 (1939).
(28) Roberts, I., Thode, H. G., and Urey, H. C., *J. Chem. Phys.*, **7**, 137 (1939).
(29) Rose, W. C., *Science*, **86**, 298 (1937).
(30) Rose, W. C., Haines, W. S., Johnson, J. E., and Warner, D. T., *J. Biol. Chem.*, **148**, 457 (1943).

(31) Schoenheimer, R., Ratner, S., and Rittenberg, D., *J. Biol. Chem.*, **130,** 703 (1939).
(32) Schoenheimer, R., Ratner, S., Rittenberg, D., and Heidelberger, M., *J. Biol. Chem.*, **144,** 545 (1942).
(33) Schoenheimer, R., and Rittenberg, D., *Physiol. Revs.*, **20,** 218 (1940).
(34) Schoenheimer, R., Rittenberg, D., and Keston, A. S., *J. Biol. Chem.*, **127,** 385 (1939).
(35) Shemin, D., and Rittenberg, D., *J. Biol. Chem.*, **151,** 507 (1943).
(36) Shemin, D., and Rittenberg, D., *J. Biol. Chem.*, **153,** 401 (1944).
(37) Shemin, D., and Rittenberg, D., *J. Biol. Chem.*, **158,** 71 (1945).
(38) Shemin, D., *J. Biol. Chem.*, **159,** 439 (1945).
(39) Shemin, D., *Federation Proc.*, **4,** 103 (1945).
(40) Soddy, F., *Ann. Rept. Chem. Soc.*, **7,** 285 (1910).
(41) Soddy, F., *Ann. Rept. Chem. Soc.*, **10,** 263 (1913).
(42) Thode, H. G., Gorham, J. E., and Urey, H. C., *J. Chem. Phys.*, **6,** 296 (1938).
(43) Thode, H. G., and Urey, H. C., *J. Chem. Phys.*, **7,** 34 (1939).
(44) Urey, H. C., Brickwedde, H. G., and Murphy, G. M., *Phys. Rev.*, **39,** 164 (1932); **40,** 1 (1932).
(45) Urey, H. C., Fox, M., Huffmann, J. R., and Thode, H. G., *J. Am. Chem. Soc.*, **59,** 1407 (1937).
(46) Verkade, P. E., and van der Lee, J., *Z. physiol. Chem.*, **227,** 213 (1934).
(47) du Vigneaud, V., Chandler, J. P., Cohn, M., and Brown, G. B., *J. Biol. Chem.*, **134,** 787 (1940).

MUCOLYTIC ENZYMES

KARL MEYER, ASSOCIATE PROFESSOR OF BIOCHEMISTRY, DEPARTMENT OF OPHTHALMOLOGY, COLLEGE OF PHYSICIANS AND SURGEONS, COLUMBIA UNIVERSITY; CHEMIST TO THE INSTITUTE OF OPHTHALMOLOGY, PRESBYTERIAN HOSPITAL

MUCOLYTIC ENZYMES may be defined as enzymes which primarily catalyze the depolymerization of highly polymerized mucopolysaccharides.* Mucopolysaccharides occur in many internal and external structures of animals and microorganisms. The biological importance of some of these substances in a few instances is now well recognized, although at present our knowledge of the functions and biological relationship of most of these substances is still very meager. In some instances, either identical or closely related mucopolysaccharides as, for example, hyaluronic acid, occur both in the animal body and in microorganisms. In other instances, mucolytic enzymes, such as lysozyme and hyaluronidase, occur both in microorganisms and in the animal. Apart from their biological interest, mucolytic enzymes are one of the most valuable aids in the isolation and characterization of mucopolysaccharides. Their high specificity is comparable to that of immunological reactions. Their usefulness is somewhat limited by the lack of purity in the available enzyme preparations, a drawback, however, which can be overcome by the judicious application of enzymes from different sources.

* Mucopolysaccharides are hexosamine-containing polysaccharides. For their classification see reference (23).

This essay will deal with some of the problems connected with mucolytic enzymes and their substrates.

Lysozyme

Lysozyme is a bacteriolytic enzyme which effects the lysis of some microorganisms, notably micrococci and sarcinae. Fleming (13) recognized the wide distribution of this agent in nasal secretion, saliva, tears, leucocytes, and egg white. It also occurs in some microorganisms in appreciable concentration, as in some molds (18), and in a strain of *Sarcina* and in a white staphylococcus (*S. muscae*), both of which are susceptible to lysis by lysozyme obtained from egg white (40). The highest concentration of lysozyme is found in egg white, although calculation on the basis of total protein shows that human tears actually have a higher concentration of the enzyme than does egg white.

Lysozyme is a basic protein of a molecular weight of about 18,000, quite stable toward heat and acid reaction, but relatively unstable toward oxidation and alkali (33). It was first obtained in crystalline form by Abraham and Robinson (1). Recently an improved method for its preparation and crystallization was reported (2). The biological activities of amorphous and crystalline lysozyme are identical, as are the activities of lysozyme prepared by our method and that of the California group. This activity is measured by observing the extent of clearing of a suspension of a susceptible test organism, either visually or photoelectrically. The test organism most widely used is *Micrococcus lysodeikticus*, which was isolated from the air by Fleming. This micrococcus is more susceptible than other organisms, although a strain of *Sarcina lutea* has been found by Dr. Rose Feiner (unpublished) which is as susceptible as *M. lysodeikticus*. Most strains of sarcinae, however, have a susceptibility markedly less than that of *M. lysodeikticus*.

Lysis by lysozyme does not take place at an acid *p*H, although the microorganisms have been shown to be killed by the enzyme at such a *p*H. When the *p*H is raised to neutral, visible lysis takes place, but this process is independent of the presence of lysozyme. In fact, the optimum *p*H for the action of lysozyme is in the acid region ($\sim$*p*H 4.5). Some susceptible microorganisms are not lysed at all by lyso-

zyme until, after incubation with the enzyme, the *p*H is raised to about 13, when visible lysis takes place, as in the case of *S. muscae*. There are also instances of organisms which are killed without any visible lysis even at high *p*H values, *e. g.*, *Bacillus megatherium* (35).

Lysozyme, in contrast to most antibiotic agents, acts not only on living but also on killed microorganisms, but in the latter case generally to a lesser degree. For example, *M. lysodeikticus* killed by chilled neutral acetone is lysed almost to the same degree as are living organisms by lysozyme, while organisms killed by acidified acetone, by alcohol, or especially by autoclaving are lysed to a much smaller degree. Visible clearing apparently is a poor, but convenient, method for the estimation of lysozyme.

The phenomenon of bacterial lysis, whether by lysozyme, phage, or other lytic enzymes, obviously is complex. Under the influence of lysozyme the susceptible organisms swell to several times their volume, the Gram stain becomes negative, uptake of oxygen stops, and in the supernatant solution nonprotein nitrogen, inorganic phosphate, and reducing substances appear (31). It seems quite obvious that one enzyme cannot catalyze such diverse reactions.

The mechanism of lysozyme action may be explained in terms of hydrolysis of a substance of mucoid nature which is contained in the bacterial membrane. That this type of hydrolysis takes place was demonstrated by the appearance of reducing substances when lysozyme was incubated with carbohydrate fractions extracted from susceptible organisms following alkaline hydrolysis. The carbohydrate obviously is not a capsular substance which is easily given off into the surrounding media as is the specific soluble substances of pneumococci. The carbohydrate, at least in highly susceptible organisms, occurs in a quantity of about 3 to 5%.

The findings of the writer's laboratory have been confirmed by Epstein and Chain (10) and by Pirie (35). The carbohydrate was said by Epstein and Chain to contain both a ketohexose and an amino sugar, evidence for the former being a positive Seliwanoff reaction in the dialyzates after incubation with lysozyme. The Seliwanoff reaction, however, is not given by the substrate of lysozyme, but by a carbohydrate present in *M. lysodeikticus* with which the preparations of Epstein and Chain must have been contaminated.

As isolated in this laboratory, the substrate of lysozyme is a

mucopolysaccharide of high molecular weight, forming extremely viscous aqueous solutions. It precipitates rabbit antisera against *M. lysodeikticus* in a dilution of 1 : 1,000,000. Crystalline lysozyme in a concentration of 1 to 5 mg. immediately precipitates 10 mg. of the polysaccharide, as does protamine. In a concentration of a few gamma, lysozyme abolishes within a few minutes the viscosity of aqueous solutions of the polysaccharide. On incubation for two hours, the hydrolysis of glucosidic linkages is almost complete. In the presence of 1 mg. of the polysaccharide, the lytic action of lysozyme on living *M. lysodeikticus* is decreased approximately 100-fold, indicating a competition for the enzyme between the added substrate and the substrate contained in the microorganism.

The action of lysozyme on this carbohydrate from *M. lysodeikticus* is highly specific. Organ extracts and enzyme preparations which contain no lysozyme have no depolymerizing effect on the carbohydrate, while lysozyme from sources other than egg white depolymerizes the carbohydrate. From organisms pretreated with lysozyme, no mucopolysaccharide fraction was obtained which liberated reducing groups when brought into contact with lysozyme.

The mucopolysaccharide appears to be firmly bound to the bacterial membrane. The primary mechanism of lysozyme action apparently consists of a depolymerization of this mucopolysaccharide, leading to water imbibition by the organism and to disorganization of the microbial cell. The breakdown of protein and of organic phosphates are effects due to autolytic enzymes, which are apparently responsible for some part of the visible clearing of the bacterial suspensions.

It early became obvious to Fleming and his co-workers that lysozyme was not identical with bacteriophage. But it appears from available data that lysozyme or lysozyme-like enzymes are associated with phage activity. As the result of the work of several investigators—Gratia and Rhodes (14), Twort (41), Wollmann and Wollmann (42)—it seems established that, in some cases, as a result of phage action, a lytic enzyme is released which is capable of dissolving heat-killed and phage-resistant living organisms of the same species as the lysed strain. These lytic agents apparently originate in the bacterial cells and in most instances can be expected to have specificities other than that of lysozyme.

Wollmann and Wollmann (42) found, however, that an enzyme obtained from staphylococci which had been lysed by phage lysed sarcinae, organisms sensitive to lysozyme. Pirie (35) found that lysozyme from egg white released a specific phage from its union with heat-killed cultures of *B. megatherium*, a release which occurred during the same interval of time as hydrolysis of the bacterial polysaccharide by the lysozyme. Heat-killed bacilli incubated with lysozyme could no longer absorb phage.

Anderson, in a recent paper (4), reported the separation of a lytic principle derived by ultraviolet irradiation from a phage of *Escherichia coli*. The phage was purified by low- and high-speed centrifugation. After irradiation with ultraviolet light, the apparently homogeneous and pure phage was split into two components, one of which was a low molecular weight protein found in the supernatant fluid after ultracentrifugation. This protein of low molecular weight caused the lysis of *E. coli* cells killed by ultraviolet irradiation. The killed coli cells also could be lysed by crystalline lysozyme from egg white. The lytic agent obtained from phage, however, could not replace lysozyme in the lysis of *M. lysodeikticus*.

Northrop (34) found that suspensions of a highly purified phage specific for *S. muscae* caused the immediate lysis of suspensions of resting living staphylococci but not of killed organisms. This organism is killed (and lysed after alkalinization) by egg-white lysozyme (40). From this staphylococcus, killed by acetone, a fraction was obtained which not only lysed suspensions of the same species, but also a strain of *Sarcina lutea*.

An explanation of these and other experiments may be attempted along the following lines. Phage virus according to this hypothesis has two components, one a highly specific substance of very high molecular weight, the other a less specific component of lysozyme-like nature which causes the actual lysis of the infected bacterial cells. The lytic agent may be a normal component of the bacterial cell.

What are the functions of lysozyme in the microbial and animal organism? From a teleological standpoint, lysozyme seems to act in the animal organism as a protective enzyme against bacterial invasion. But the most susceptible organisms are harmless saprophytes; pathogenic organisms show little, if any, susceptibility, although some have been reported to be susceptible to egg-white lysozyme.

The protective action of lysozyme, however, is strongly suggested in vitamin A deficiency. Thus Findlay (12) reported a cure of the xerophthalmic condition in vitamin A deficient rats by washing the eyes with human tears. Andersen (3) found a subnormal amount of lysozyme in the tears of human twins suffering from xerophthalmia; on addition of vitamin A to the diet of the twins, the lysozyme increased and the xerophthalmia concurrently improved. A renewed study of nutritional factors in the production and activity of lysozyme seems indicated.

The occurrence of lysozyme and lysozyme-like enzymes in many microbes may be interpreted to mean that these enzymes are involved in some metabolic process connected with the carbohydrate substrates in the bacterial membranes. Since some of these membranes, as in *M. lysodeikticus*, are extremely tough structures, they may serve in facilitating the softening of the membrane in bacterial division. It seems useful also to assume a metabolic role for lysozyme in the animal body. Such a hypothesis would presuppose the occurrence of the substrate of lysozyme in the animal organism and thus far the ocurrence of such a substrate has been reported for egg white only (31). A study of the chemical specificity of lysozyme undoubtedly will be useful in establishing the presence or absence of such substrates in the animal body.

The susceptible organisms become highly resistant to lysozyme action when grown in the presence of sublethal doses of the enzyme. It will be interesting to investigate this adaptation to determine whether the adapted organism fails to produce the substrate of lysozyme at all, or whether the adaptation is due to a chemical modification of the membrane polysaccharide so that it no longer is attacked by lysozyme.

Recently, attention was called to the similarity in occurrence and chemical properties of lysozyme and avidin, the protein which neutralizes the vitamin biotin (22). Avidin and lysozyme are both found in egg white, in the oviduct of birds, and in fish eggs; both are basic proteins, very stable toward acid and unstable toward oxidation and alkaline reaction. They undoubtedly are not identical, either in respect to their chemical or biological properties, the main chemical difference between the two being that of solubility, for lysozyme is a very soluble, and avidin is a highly insoluble, protein, even at acid *p*H values. They differ in their biological activity in that freshly

prepared, pure lysozyme has a negligible avidin activity. It was found that biotin (10 γ) increased the activity of our lysozyme preparation (approximately 1 γ) up to 500-fold. Lysis at such dilutions of lysozyme had never been observed before. However, the activation was shown by only a limited number of batches of *M. lysodeikticus;* and the activating effect of biotin was no longer shown by subcultures of the original batch, while it was still reproducible with old samples. Meanwhile, a failure to activate lysozyme action with biotin was reported (9). It seems probable that we were dealing with a nonpermanent mutant in some of the colonies, which was the cause of an observed autolysis in the affected strain. [Autolytic or "suicide" colonies have been described as a variant of a *Micrococcus tetragenus* by Reimann (38)]. It is hoped that an explanation for the activation will be found in the near future.

The similarity between lysozyme and avidin seems so striking that a relationship between the two deserves further investigation. The formation of avidin, for example, by partial oxidation of lysozyme is being investigated at present.

Other Mucolytic Enzymes

Many bacteriolytic agents have been encountered in animal organs and in microorganisms, some of which may be classified as lysozyme-like enzymes. With increasing knowledge of the mucopolysaccharides, knowledge of specific mucolytic enzymes will undoubtedly increase as well.

Two enzymes have been encountered which may belong in this group. One is an enzyme—thus far found in human saliva, submaxillary glands of animals, and pneumococci—which hydrolyzes one of the two mucopolysaccharides derived from the submaxillary gland. The enzyme is identical neither with hyaluronidase nor with amylase or lysozyme. A search for it in pathogens of the respiratory tract might give some useful information about natural resistance to respiratory infections. The substrate of this enzyme occurs in the glandular extracts as a mucoprotein, composed of about equal parts of a peptide chain and of an acid polysaccharide. The latter contains acetylglucosamine and gluconic acid in equimolar portions (23).

Another mucolytic enzyme group, on which little work has

been done, is that concerned with depolymerization and hydrolysis of the neutral mucopolysaccharide fractions of gastric mucin. One of these substances isolated from pig gastric mucosa is composed of equimolar parts of acetylglucosamine and galactose (21). In these mucopolysaccharide fractions some of the blood group substances occur, substances responsible for the blood group specificities of man and animal. Although the exact chemical composition of the blood group substances is unknown at present, there seems little doubt that the blood group A substance is closely related chemically to the acetylglucosamine–galactose complexes of pig gastric mucosa.

Enzymes which destroy the blood group A activity have been found in human tissues, feces, and a few species of bacteria, among them *Clostridium welchii* (39). From one of these strains Schiff obtained a filtrate which inactivated the A but not the B substance. In the writer's laboratory extracts of one strain of *C. welchii* were prepared which hydrolyzed the polysaccharide of gastric mucosa with the liberation of reducing sugar, while from other strains no such enzymes were obtained (29). The blood group A activity was unaffected by extracts from the nonhydrolyzing strain, while the hydrolysis by the active extracts paralleled the A-inactivating potency (unpublished work of Schiff and Meyer). The wide occurrence of glucosamine–galactose complexes in many bacterial polysaccharides would make a study of these enzymes quite interesting.

Hyaluronidase

Hyaluronidase is the most extensively studied of all mucolytic enzymes. Its main substrate, hyaluronic acid, has been isolated from vitreous humor, umbilical cord (30), synovial fluid (32), skin (26), and some tumors of mesodermal origin, such as fowl leucosis (15,36) and a human mesothelioma (25). One of the most significant findings in this field was its isolation from group A and C hemolytic streptococci by Kendall, Heidelberger, and Dawson (17). Hyaluronidase, an enzyme which depolymerizes and hydrolyzes hyaluronic acid, was first obtained from a type II pneumococcus (28). Although it was at first thought that the enzyme was identical with the autolytic enzyme of pneumococci, this proved not to be the case. Hyaluronidase was found, shortly afterwards, in a group A hemolytic streptococcus, in a

strain of *C. welchii*, and in extracts of spleen and ciliary body (29). Chain and Duthie (6) made the important observation that purified testis extracts, which contained in high concentrations the "spreading factor" of Duran-Reynals (8) and of McClean (19), also contained hyaluronidase; in fact, they proposed that hyaluronidase was identical with the "spreading factor." Based on the findings of Chain and Duthie, an intensive study was made of the hyaluronic acid–hyaluronidase system by a number of investigators.

Hyaluronidase and spreading factor have been found in extracts of leach heads, in a number of snake and other venoms, in skin, in staphylococci, and in many other microorganisms. Aside from some unspecific spreading factors, which may act by oxidatively depolymerizing hyaluronic acid, the identity of hyaluronidase with spreading factor seems well established. The only exception still is the occurrence of a strong spreading reaction in extracts or filtrates of some group A hemolytic streptococci in which no hyaluronidase could be detected (27).

In the spreading reaction, extracts are injected intradermally in rabbits along with a suitable indicator such as certain dyes, hemoglobin, bacterial toxins, or India ink. In the presence of spreading factor, these indicators diffuse over an area of the skin proportional to the concentration of agent, while in the controls the injected indicator remains localized in a small bleb. The reaction apparently is caused by the depolymerization of a hyaluronic acid gel present in the intercellular substance of the dermis. Increased capillary permeability is not caused by hyaluronidase. Obviously the capillary endothelial cement does not contain hyaluronic acid. The spreading reaction can also be demonstrated in the cornea (unpublished work), apparently acting by the depolymerization of the monosulfuric acid ester of hyaluronic acid, which occurs in the substantia propria (24).

Hyaluronic acid, a polymer of acetylglucosamine and glucuronic acid, occurs in different tissues in polydisperse form. Its molecular weight has been estimated from its double refraction of flow as between 200,000 and 400,000 (5), a figure which may be considerably higher when measured on material derived from quite solid gels, as from some malignant cysts or from the nucleus pulposus.

Three methods are available for the quantitative estimation of hyaluronidase and they depend upon: (a) the hydrolysis of the gluco-

sidic linkages in hyaluronic acid, as measured by the increase in reducing power; (b) the decrease in viscosity of hyaluronate-containing solutions; and (c) the decrease of the protein-precipitating power of hyaluronate after enzymic depolymerization. The last method in a modification of that of Kass and Seastone (16) has proved most convenient and accurate.

Group A and C hemolytic streptococci are the only microorganisms in which hyaluronic acid has been demonstrated. It apparently is formed only when these organisms possess a mucoid capsule. Seastone has brought forward good evidence for the parallelism of capsule formation, hyaluronic acid production, and the invasiveness of the organisms. Incubation with hyaluronidase causes the disappearance of the capsule. Furthermore, intraperitoneal injection of hyaluronidase was shown to protect mice against fatal doses of intraperitoneally injected group A and C streptococci (16).

While some strains produce hyaluronic acid, other strains belonging to group A, C, and G produce hyaluronidase. Hyaluronic acid and hyaluronidase production do not seem to occur simultaneously. In a recent study (7) involving 308 strains of group A streptococci, hyaluronidase production was found only in strains belonging to type 4 and 22. The group A strain used mostly in our studies also belonged to type 4.

When hyaluronidase activity was determined chemically with culture filtrates or fractions isolated from streptococci, a peculiar behavior was noted, in that the activity measured viscometrically or by reduction stopped after an initial reaction. In fact in many tests with hyaluronidase-producing strains no enzyme could be demonstrated in the majority of tests. The explanation for the abnormalities may be found in an enzymic destruction of hyaluronidase by streptococcus extracts.

For the concentration and purification of hyaluronidase, bull or ram testis is the most convenient source. Thus, far, the enzyme has never been obtained as a pure protein. Beside inert proteins, testicular preparations contain an enzyme hydrolyzing chondroitin sulfate into disaccharide units containing sulfate (27). The hyaluronidase activity in most preparations runs parallel with that of the enzyme which splits chondroitin sulfate, including samples purified in the final step by electrophoretic separation. The two enzymes are, however, sepa-

rate entities, since the enzyme splitting chondroitin sulfate is absent in pneumococcal hyaluronidase preparations. Furthermore, by treatment with acetone, the former may be destroyed in testicular preparations without affecting the hyaluronidase activity. The concentration of hyaluronidase in testis apparently is far smaller than that found in the leech and probably than that found in some snake venoms.

Hyaluronidase in the testis or, more specifically, in spermatozoa plays an important role in fertilization, as was shown recently (11,20). The cumulus cells surrounding the ovum are embedded in a jelly which apparently contains hyaluronic acid. This jelly is liquefied by hyaluronidase furnished by a sufficiently large number of spermatozoa. Fertilization then can proceed by a single spermatozoon.

In the skin—probably the largest store of hyaluronidase in the body—the enzyme, although generally present in an inactive form (27), may be supposed to regulate the velocity of water and metabolite exchange by decreasing the viscosity of the intercellular matrix.

Clinically, the role of hyaluronic acid and hyaluronidase is yet little explored. Hyaluronic acid concentration appears to be greatly increased in some exudates of the joints. Injection of hyaluronidase into pathological joints lowered the viscosity of the exudates, without permanently improving the pathological condition (37). A disappeaıance of ganglia of the tendon sheaths likewise has been noted on injection of hyaluronidase. Intraperitoneal hyaluronidase injection was used in a case of mesenthelioma of the pleura and peritoneum, to facilitate the removal of a fluid of honeylike consistency: without hyaluronidase injection, paracentesis was incomplete and very difficult; after injection of purified testicular hyaluronidase, a fluid of low viscosity could be completely removed in a short time. The continued injection of large quantities of hyaluronidase apparently had no harmful effects. The continued growth of the malignant tumor finally caused the death of the patient (Stewart and Meyer, unpublished experiments).

Testicular preparations containing hyaluronidase in high concentrations have been shown to lower the high erythrocyte sedimentation rate when added to the blood of patients with rheumatic fever and other diseases. This effect, which has all characteristics of an enzymic reaction, does not seem to be due to a proteolytic action on any of the plasma proteins. At present, it is doubtful, however,

whether the action is due to the hyaluronidase contained in the preparations. The relationship between streptococcus infection and rheumatic fever clearly deserves an intensive study in which hyaluronic acid–hyaluronidase production and hyaluronidase inhibition cannot be overlooked.

In this essay, a few mucopolysaccharides and mucolytic enzymes have been dealt with, and an attempt has been made to show some of the interrelationship of these entities mainly with bacteriological, but also with some physiological and medical, problems. Thus, in lysozyme and its substrate, one is concerned not only with preparative chemistry, but with various problems of bacteriological variation, of immunology, of virus research, and of nutrition. A study of hyaluronidase involves many bacteriological problems, the physiology of tissue permeability and fertilization, and the fields of pathology and medicine. The study of gastric mucin has led to the problem of blood groups and is intimately connected with various tissue and bacterial antigens. The finding in pneumococcal extracts of an enzyme which depolymerizes the acid mucopolysaccharide of submaxillary gland may suggest the presence of similar enzymes in other microorganisms, especially of the upper respiratory tract.

Although other mucopolysaccharides occurring in the animal body and their specific enzymes have not been discussed in this essay, some of these systems are not less important, as, for example, chondroitin sulfuric acid, heparin, and the carbohydrate matrix of amyloid tissue. But because knowledge of the enzyme systems which hydrolyze these substances is too meager, they may only be mentioned at this point. It seems to the author that these and other systems deserve greater attention by the biochemist than they have hitherto received.

Addendum

Recently a new and highly accurate method for the estimation of lysozyme was developed in this laboratory. In this method the depolymerization of the mucopolysaccharide isolated from *M. lysodeikticus* is measured viscometrically. Furthermore, lysozymes, partly in very high concentration, were found in the fresh and dried latex of some species of *Ficus*, in *Euphorbia*, and in papain. The *Ficus* lysozyme is chemically distinct from that of egg white. Its potency is significantly higher than the crystalline lysozyme of egg white.

References

(1) Abraham, E. P., and Robinson, R., *Nature*, **140,** 24 (1937).
(2) Alderton, G., Ward, W. H., and Fevold, H. L., *J. Biol. Chem.*, **157,** 43 (1945).
(3) Andersen, O., *Acta Paediat.*, **14,** 81 (1932).
(4) Anderson, T. F., *J. Cellular Comp. Physiol.*, **25,** 1 (1945).
(5) Blix, G., and Snellman, O., *Nature*, **153,** 587 (1944).
(6) Chain, E., and Duthie, E. S., *Nature*, **144,** 977 (1939); *Brit. J. Exptl. Path.*, **21,** 324 (1940).
(7) Crowley, N., *J. Path. Bact.*, **56,** 27 (1944).
(8) Duran-Reynals, F., *Bact. Revs.*, **6,** 197 (1942).
(9) Elderton, G., Lewis, J. C., and Fevold, H. L., *Science*, **101,** 151 (1945).
(10) Epstein, L. A., and Chain, E., *Brit. J. Expt'. Path.*, **21,** 339 (1940).
(11) Fekete, E., and Duran-Reynals, F., *Proc. Soc. Exptl. Biol. Med.*, **52,** 119 (1943).
(12) Findlay, G. M., *Brit. J. Exptl. Path.*, **6,** 16 (1925).
(13) Fleming, A., *Proc. Roy. Soc. London*, **B93,** 306 (1922). Fleming, A., and Allison, V. D., *Lancet*, **1,** 1303 (1924); *Brit. J. Exptl. Path.*, **8,** 214 (1927).
(14) Gratia, A., and Rhodes, B., *Compt. rend. soc. biol.*, **89,** 1171 (1923); **90,** 640 (1924).
(15) Kabat, E. A., *J. Biol. Chem.*, **130,** 143 (1939).
(16) Kass, E. H., and Seastone, C. V., *J. Exptl. Med.*, **79,** 319 (1944).
(17) Kendall, F. E., Heidelberger, M., and Dawson, M. H., *J. Biol. Chem.*, **118,** 61 (1937).
(18) Kriss, A. E., *Microbiology* (*U.S.S.R.*), **9,** 32 (1940).
(19) McClean, D., *J. Path. Bact.*, **33,** 1045 (1930).
(20) McClean, D., and Rowlands, I. W., *Nature*, **150,** 627 (1942).
(21) Meyer, K., *Cold Spring Harbor Symposia Quant. Biol.*, **6,** 91 (1938).
(22) Meyer, K., *Science*, **99,** 391 (1944).
(23) Meyer, K., "Mucoids and Glycoproteins," in *Advances in Protein Chemistry.* Vol. II, Academic Press, New York, 1945.
(24) Meyer, K., and Chaffee, E., *Am. J. Ophthalmol.*, **23,** 1320 (1940).
(25) Meyer, K., and Chaffee, E., *J. Biol. Chem.*, **133,** 83 (1940).
(26) Meyer, K., and Chaffee, E., *J. Biol. Chem.*, **138,** 491 (1941).
(27) Meyer, K., Chaffee, E., Hobby, G. L., and Dawson, M. H., *J. Exptl. Med.*, **73,** 309 (1941).
(28) Meyer, K., Dubos, R., and Smyth, E. M., *J. Biol. Chem.*, **118,** 71 (1937).
(29) Meyer, K., Hobby, G. L., Chaffee, E., and Dawson, M. H., *J. Exptl. Med.*, **71,** 137 (1940).
(30) Meyer, K., and Palmer, J. W., *J. Biol. Chem.*, **114,** 689 (1936).

(31) Meyer, K., Palmer, J. W., Thompson, R., and Khorazo, D., *J. Biol. Chem.*, **113**, 479 (1936).
(32) Meyer, K., Smyth, E. M., and Dawson, M. H., *J. Biol. Chem.*, **128**, 319 (1939).
(33) Meyer, K., Thompson, R., Palmer, J. W., and Khorazo, D., *J. Biol. Chem.*, **113**, 303 (1936).
(34) Northrop, J. H., *Crystalline Enzymes.* Columbia Univ. Press, New York, 1939.
(35) Pirie, A., *Brit. J. Exptl. Path.*, **21**, 125 (1940).
(36) Pirie, A., *Brit. J. Exptl. Path.*, **23**, 277 (1942).
(37) Ragan, C., and De Lamater, A., *Proc. Soc. Exptl. Biol. Med.*, **50**, 349 (1942).
(38) Reimann, H. A., *J. Bact.*, **31**, 385 (1936).
(39) Schiff, F., *J. Infectious Diseases*, **65**, 127 (1939).
(40) Thompson, R., *Arch. Path.*, **30**, 1096 (1940).
(41) Twort, F. W., *Lancet*, **2**, 642 (1925).
(42) Wollman, E., and Wollman, E., *Compt. rend. soc. biol.*, **110**, 636 (1932); **112**, 164 (1933).

19

SOME ASPECTS OF INTERMEDIARY METABOLISM

KONRAD BLOCH, ASSOCIATE IN BIOCHEMISTRY, COLLEGE OF PHYSICIANS AND SURGEONS, COLUMBIA UNIVERSITY

INTERMEDIARY metabolism is a branch of biochemistry which attempts to trace the pathways of food components in the animal body, with the ultimate aim of describing the sequence of the chemical reactions involved, and of providing a quantitative account of food utilization. Ideally, the study of normal intermediary metabolism should be carried out under conditions which do not of themselves affect the physiological state of the animal. In order to maintain the "normal" state, it is necessary to employ a diet on which the weight of the animal and the composition of the excretion products remain substantially constant. Because the animal cell is unable to distinguish between the various isotopes of an element, substitution of a normal dietary constituent by one which contains an excess of an isotopic element, but which is otherwise the same, does not disturb its physiological state. In order to illustrate the type of information which can be gained with the aid of isotopically labeled substances, it will be useful to consider briefly the scope and the limitations of other methods which have been applied to the study of intermediary metabolic reactions.

Chemical analysis of animal tissues can, of necessity, contribute only limited information about intermediary reactions. It has been

recognized that substances which are suspected of being intermediates of either anabolic or catabolic reactions, but which are present in amounts too small for chemical identification, may acquire quantitative significance in the light of the rate at which they are regenerated. Although their stationary concentration may be small, they arise and are metabolized in large quantities over an extended period of time. In the breakdown of the major food constituents, certain intermediates are known to be produced continuously, *e. g.*, the deamination products of amino acids, ornithine, the phosphorylated trioses, and acetic acid. None of these intermediates has actually been isolated as such from animal tissues. Hence, numerous attempts have been made to arrest metabolism at an intermediate stage by inducing an unphysiological state (diabetes, starvation, etc.) or by administration of substances which contain groupings refractory to attack by animal cells. The latter approach, of which the phenyl substituted amino acids and fatty acids employed by Knoop are the most notable examples, represents an early attempt to tag molecules in order to trace them through the animal body.

A large part of the accepted knowledge of intermediary reactions, particularly of carbohydrates and amino acids, has been gained from studies with surviving tissues, tissue preparations, and isolated enzymes. The *in vitro* techniques deal with reactions which the enzyme systems of cells are able to carry out with an added substrate. The result, although not the primary aim, of this type of experimentation, has been to provide knowledge of the organic chemistry of biological compounds in a "physiological" medium and under the influence of biocatalysts. The use of isolated tissues involves inherent limitations as far as permitting a decision as to whether certain reactions occur in the living organism. The question as to whether and to what extent an *in vitro* reaction is part of the normal metabolism must be decided by an independent method. The fact that, for example, surviving liver and kidney, or enzymes isolated from these tissues, deaminate *d*-amino acids at a rapid rate in no way signifies that the cell normally has to deal with the unnatural amino acids. Evidence that the amino acids of *d*-configuration are not normal constituents of the animal cell is contained in the following experiment (20). When *d,l*-tyrosine or *d,l*-glutamic acid labeled with N^{15} was administered to rats, the unnatural acids excreted in the urine had an

unchanged isotope concentration; if *d*-tyrosine or *d*-glutamic acid had been formed in a normal synthetic process, it should have mixed with the administered isotopic compound, and hence the excreted amino acid should have had a lower isotope concentration.

As a great variety of foreign substances, *e. g.*, the odd-numbered fatty acids, are not only modified but completely metabolized by the animal cell, it must be inferred that the organism is able to handle substances "to which it is ordinarily accustomed" as well as many others which do or do not bear a structural relationship to naturally occurring compounds. Conversely, compounds of established biological activity, such as the lipids, are found to be unaffected by isolated tissues. *In vitro*, the higher fatty acids, cholesterol, and other steroids fail to show any evidence of synthetic or degradative reactions involving carbon-carbon bonds. Limitations of solubility and cell permeability must, to some extent, impair *in vitro* activity toward lipids and may be responsible for their apparent metabolic inertia.

In view of our recent experiences with isotopes, another aspect must be considered for the interpretation of these findings. The studies of Schoenheimer and Rittenberg with isotopic test substances have made it clear that the constancy of tissue composition in the normal adult animal is maintained as a result of balanced synthesis and degradation, and have provided an experimental basis for the concept of dynamic equilibrium of the tissue constituents. It is conceivable that a similar situation would obtain in isolated tissues so that a reaction proceeds without over-all change of the substrate concentration. Evidence that the dynamic equilibrium may obscure the recognition that an *in vitro* reaction is proceeding may be seen in the following findings. The utilization of acetic acid for cholesterol synthesis, first demonstrated in intact animals with the aid of isotopic acetate (4), has recently been observed to occur also in isolated tissue (unpublished results of Bloch, Borek, and Rittenberg). When liver slices were incubated with deuterioacetic acid, deuteriocholesterol was formed at a rate similar to that in the liver of intact animals. As the total amount of cholesterol does not change noticeably under these conditions, it is clear that a quantity equivalent to that synthesized had simultaneously been degraded. The sensitivity of the isotope method is such that syntheses corresponding to as little as 0.1% of the cholesterol present can be detected. Experiments in which liver slices were in-

cubated in the presence of D_2O indicate that the deuteriocholesterol formed amounted to at least 10% of the cholesterol present at the start, an amount of change which, although close to the limits of detection by available analytical methods, nevertheless should be demonstrable.

The available data on the syntheses in which acetic acid is employed indicate that, in the intact animal, fatty acids are regenerated at a faster rate than cholesterol (17). On the other hand, the corresponding results obtained with liver slices would lead to the erroneous conclusion that cholesterol synthesis is by far the more rapid process. The lack of correlation between the rates *in vitro* and *in vivo* must be taken to mean that, in the experiments in which acetic acid was added to liver slices, the conditions chosen happened to be favorable for cholesterol synthesis, but that a component necessary for fatty acid formation was lacking.

The use in tissue slice experiments of enzyme poisons, such as malonate, fluoride, etc., makes possible an accumulation of intermediates whose identification is thereby facilitated. If specific inhibitors for a particular reaction are not available, a suspected intermediate may escape detection because it does not accumulate. The isotope dilution method has recently been applied to establish the occurrence of a reaction which is not detectable by the balance method (14). Incubation of kidney slices with acetic acid does not lead to a demonstrable formation of acetoacetic acid. If acetic acid containing heavy carbon is substituted for ordinary acetic acid and some non-isotopic acetoacetic acid is added, subsequent analysis of the acetoacetate then isolated reveals the presence of heavy carbon in the keto acid. Thus, clearly, acetoacetic acid is synthesized from acetic acid, but its rapid removal by subsequent reactions prevents its accumulation in quantities sufficient for analysis. In experiments of this type, only the isotope concentration and not the absolute amount of acetoacetate is relevant; the rate of acetoacetate synthesis can be calculated from the amount of acetoacetate added initially and from its isotope concentration at the end of the experiment. It is evident that many of the difficulties pertaining to the use of tissue preparations can be circumvented by the use of pure enzymes.

In devising schemes for metabolic interconversions, the biochemist often rests his case on little more than similarities of chemical

structure between various substances. In spite of close chemical resemblances, the biological relationships of creatine to creatinine, of phenylalanine to tyrosine, and of cholesterol to the bile acids and steroid hormones have, until recently, remained hypothetical. The balance type of experimentation with intact animals has, for various reasons, failed to give conclusive answers to these problems. Even if an oversupply in the diet of a suspected precursor leads to an increased production of the correct reaction product, the conversion will remain in doubt until the possibility of an indirect stimulating effect is excluded. Balance experiments may be inconclusive because many biological syntheses appear to proceed at rates which are, within limits, independent of dietary variations. In such cases, the oversupply of a precursor will be without effect on the formation of the reaction product. For instance, the immediate response to addition of excess creatine to the diet is an increased excretion of creatine rather than of creatinine (7). The conversion of creatine to creatinine becomes evident, however, by employing labeled creatine. When small amounts of creatine containing N^{15} are fed to rats, the creatinine excreted contains isotopic nitrogen (6). In this case, it is irrelevant whether or not creatinine is increased in total amount; the presence of isotope is proof of the conversion.

A biochemical conversion may proceed independently of the exogenous supply of the conversion product or of both conversion product and precursor. When an essential dietary constituent is the precursor, its concentration in the diet may regulate the extent of the conversion, but it may be immaterial whether or not the conversion product is supplied. Conversion of phenylalanine into tyrosine and the methyl transfer from methionine in the synthesis of choline proceed normally even when these products of conversion are amply provided in the diet (15,26).

If the precursor can be synthesized by the organism at an adequate rate, its addition to the diet will likewise fail to increase production of the conversion product. Interrelationships established with the aid of labeled test substances (18,21,24), *e. g.*, those existing between the dispensable amino acids, proline, ornithine, and glutamic acid, or between palmitic and stearic acids (19,22), could not be clarified by quantitative methods because these reactions proceed independently of the exogenous supply of both precursor and reaction product.

Recognition of these reactions which are independent of the nutritional state of the animal illustrates the very limitations to which balance experimentation in the intact animal is subject.

In considering the *in vivo* relationships of steroids, it is important to bear in mind that the body possesses a store of cholesterol in every cell. If cholesterol were the parent substance for all other steroids (bile acids and steroid hormones), the precursor would be available in amounts which are large in comparison with the quantities of bile acids and hormones normally produced. Increase of the available cholesterol supply by dietary addition can, therefore, be expected to affect these conversions but little. With labeled cholesterol, the role of this sterol as a precursor of bile acids and of the excretion form (pregnanediol) of at least one steroid hormone (progesterone) could be verified (2,3). For the interpretation of the experiment, the quantities both of precursor administered and of conversion product isolated are immaterial. Appearance of the isotopic label in the reaction product is sufficient proof of the occurrence of the postulated reaction, provided that the uptake of isotope is not due to an unspecific reaction. When deuterium is the isotope employed, the possibility that the reaction product received its isotope by unspecific reactions such as exchange with the body fluids must be excluded. The quantitative importance of the conversion can be evaluated from the relative concentrations of isotope in the circulating precursor and the product formed.

Degradation of Fatty Acids

The apparent inability of *in vitro* systems to synthesize or degrade lipids may well be a consequence of the fact that the metabolism of lipids requires the presence of intact cells. The enzyme systems involved have remained entirely unknown, and the study of intermediary fat metabolism has therefore been carried out primarily with intact animals. Stable isotopes found their first biological application in this field and promise to yield further information particularly when carbon C^{13} becomes readily available. The recent observation that lipid synthesis can be demonstrated to occur in surviving liver if marked test substances are employed may facilitate a more detailed study of reaction mechanisms which are not profitably investigated by experimentation with intact animals.

The view that fatty acids are metabolized by oxidation at their β-carbon atoms with consequent removal of two-carbon units has been held ever since the classical experiments with phenyl substituted fatty acids were carried out by Knoop. The experiments of Dakin with the ammonium salts of fatty acids brought evidence for the chemical susceptibility of the beta positions to mild oxidizing agents in the test tube. The facts that (1) the component fatty acids of animal tissues contain, without exception, an even number of carbon atoms, and (2) that acetoacetic acid is excreted under conditions of deficient or impaired carbohydrate oxidation, have provided additional though indirect support for the principle of beta oxidation. To this theory, the objection has been raised that none of the postulated intermediary keto acids nor any of the fatty acids below C_{10}, including the main breakdown product, acetic acid, has been demonstrated in animal tissues. In view of recent experience, the practical absence of such intermediates is not surprising, since the failure to accumulate is characteristic of the breakdown products of the major food constituents and must be attributed to their rapid rate of removal.

A finding apparently at variance with the beta oxidation theory was the observation that, in isolated tissues, the uptake of oxygen and the formation of ketone bodies are greater with the higher fatty acids than with butyric acid (9), whereas Knoop's theory provided for only one mole of acetoacetate per mole of fatty acid. An attempt to overcome this inconsistency led to the concept of multiple alternate oxidation (8) which visualizes simultaneous oxidation of the fatty acid chain at alternate carbon atoms. In this manner palmitic acid could break down to yield four molecules of acetoacetate. It is noteworthy that, for years, little attention had been paid to the observation of Loeb (12) that ketone bodies were formed on perfusion of liver with acetic acid. If acetoacetate could arise synthetically from smaller units as well as by primary breakdown, then the formation of more than one mole of ketone bodies per molecule of fatty acid would no longer be incompatible with the theory of beta oxidation.

Another fact which could not be fitted into existing theories was the finding that *n*-valeric acid is a ketogenic substance (9,13). Neither Knoop's theory nor that of multiple alternate oxidation could satisfactorily explain this observation unless it was supposed that oxidation of odd-numbered fatty acids could be initiated at the

α-carbon atom. McKay perspicaciously suggested that valeric acid and other odd-numbered fatty acids yield, on oxidation, β-keto acids which then undergo hydrolysis to acetic acid. Condensation of two acetic units would then give rise to acetoacetate.

By employing fatty acids labeled with deuterium and C^{13}, it has been possible to provide more direct support for the principle of beta oxidation and to establish that acetic acid is the source of the "extra" ketone bodies. The first experiment bearing on the question of degradation by elimination of C_2 units was carried out by Schoenheimer and Rittenberg (19) who demonstrated the biological conversion of deuteriostearic acid to deuteriopalmitic acid. It was clear from their data that this transformation was direct and not attributable to utilization of smaller fragments derived from stearic acid. Analytical data for the myristic and lauric acid fractions indicated that the process of C_2 removal did not end at the C_{16} stage. It is noteworthy that the fatty acids with carbon chains longer than C_{10} are constituents of tissue fat but appear to be metabolically inert when added to tissue slices or tissue extracts, whereas the lower acids from C_2 to C_8 are oxidized *in vitro* but cannot be detected in tissue fat. Evidently the rate of oxidative breakdown increases with decreasing chain length, both *in vitro* and *in vivo*. Actually, the lower fatty acids may arise in considerable quantities in the animal body over an extended period of time, but their concentration in the stationary state is insignificant because they are rapidly oxidized further. Evidence supporting this concept has been secured in the case of acetic acid. This acid should be the principal intermediate which is formed by hydrolytic splitting of the β-keto acids in the course of beta oxidation. Although isolation of this compound might be achieved by working up large quantities of animal tissues or urine, its normal concentration in tissues and body fluids is apparently too small to permit ready identification.

An investigation with the aid of deuterioacetic acid of the way in which foreign amines are acetylated revealed that acetic acid is utilized directly in the acetylation reaction (1,5). The excreted acetyl compound contains only a fraction of the isotope present in the dietary acetic acid; dilution of the isotope by acetic acid arising in intermediary metabolism must have occurred in the tissues. If the decrease in isotope content occurring in the acetylation reaction could be ascribed exclusively to the presence of endogenous acetic acid, then the

dilutions would be a measure of the amounts of acetic acid formed by the intact animal. This assumption proved to be valid for the acetylation of the aromatic amines, sulfanilamide and *p*-aminobenzoic acid (unpublished results of Bloch and Rittenberg). The procedure employed represents an application of the isotope dilution technique to *in vivo* systems. Dietary deuterioacetate and nonisotopic endogenous acetate merge, and a representative sample of the mixture becomes fixed in a reaction product which resists attack by the animal cell and is excreted in the urine. From the isotope concentration of the excreted acetyl amino compound it can be estimated that, in 24 hours, the adult rat produces a quantity of acetic acid equal to about one per cent of its body weight. On the other hand, one can calculate from data on fatty acid turnover (17) the amount of acetic acid which should arise in the same period of time on the assumption that one molecule of palmitic acid is broken down to eight molecules of acetate. The value calculated on this basis is in close agreement with that found experimentally by the isotope dilution method. These results suggest that, during oxidation, most, if not all, carbon atoms of the fatty acid pass through the acetic acid stage.

By taking advantage of the fact that, in the acetylation of foreign amines a sample of endogenous acetate is trapped and eliminated in the form of a metabolically inert derivative, it becomes possible to demonstrate acetic acid formation from individual fatty acids. Appearance of deuterioacetyl groups after administration of a labeled fatty acid will be indicative of intermediary acetic acid formation. In accord with the beta oxidation theory, labeled myristic, *n*-valeric, isovaleric, and butyric acids were found to result in the excretion of isotopic acetyl groups, while propionic and isobutyric acids failed to do so (5). Extension of these experiments to other fatty acids which are labeled at specific positions with deuterium as well as isotopic carbon seems desirable.

As some of the intermediates in fatty acid catabolism contain labile hydrogen which is subject to enolization and resultant exchange with the hydrogen of the body fluids, deuterium may be partly lost. Quantitative evaluation of the data on acetic acid formation from the higher fatty acids may, therefore, be difficult and must await the preparation of test substances containing isotopic carbon. On the other hand, the lability of C—H bonds may often be of advantage by

providing information on reaction mechanisms. For instance, butyric acid labeled with deuterium at either the α- or the γ-position will yield deuterioacetyl groups, indicating fission of the molecule into two C_2 fragments. Since α-deuteriobutyrate yields an acetyl group of much lower isotope concentration than does γ-deuteriobutyrate, it can be inferred that the intermediate must be of a nature to permit loss of deuterium at the α-carbon atom. In view of the known *in vitro* lability of the α-hydrogen atoms in β-keto acids, acetoacetic acid appears to be the most probable intermediate in the butyrate–acetyl conversion.

Formation of Ketone Bodies

Theories of fat oxidation have been based, at least in part, on the ketogenic action of certain substances and the ability of others to suppress ketosis. The mechanism of ketone body formation long remained controversial because of the "extra" ketone body formation from even-numbered fatty acids which contain more than four carbon atoms. On administering to fasting rats acetic acid which contained a carboxyl group labeled with heavy carbon, Swendseid *et al.* (25) observed the presence of C^{13} in the excreted ketone bodies. Their data provide unequivocal proof for the early contention of Loeb that acetoacetic acid could be synthesized from smaller molecules. It still remained to be decided whether this process involved a Claisen type of condensation of acetic acid or a coupling with one molecule of pyruvic acid to acetopyruvic acid with subsequent decarboxylation (10). Conclusive evidence pertaining to the mechanism is supplied by recent experiments of Weinhouse *et al.* (27). Octanoic acid labeled with C^{13} in the carboxyl group was incubated with liver slices. Isotope analysis of the resulting acetoacetate revealed the presence of heavy carbon in equal concentrations in the carboxyl and carbonyl groups, a finding which must be ascribed to random condensation of the acetic acid units arising from the oxidation of octanoic acid. The isotope distribution observed can be reconciled neither with the classical view that only the four terminal carbon atoms of a fatty acid are a source of ketone bodies nor with the hypothesis of multiple oxidation at alternate carbon atoms. Consequently, any metabolite capable of forming

acetic acid must be considered a potential source of ketone bodies.* The ketogenic action of *n*-valeric acid, which has also been shown to yield acetic acid in the normal animal (5), can be readily explained as resulting from the condensation of two molecules of acetate.

In liver tissue *in vitro*, the oxidation of fatty acids appears to lead to acetoacetic rather than to acetic acid as an end product. Since such experiments are performed with slices from fasted animals, the accumulation of ketone bodies may well be related to the lack of carbohydrate or carbohydrate intermediates, conditions which, in the intact animal, result in ketosis. It would be unwarranted to conclude that in the normal, well-nourished animal the acetic acid formed by fat breakdown necessarily condenses to acetoacetate prior to its further oxidation. It is equally conceivable that normally fat oxidation takes the course visualized by Knoop and that the acetate–acetoacetate condensation goes into effect only when the catabolism of acetic acid is interfered with.

Fatty Acid Synthesis

It is well recognized that fat can be synthesized from carbohydrate, but information as to the chemical nature of the process is totally lacking. Any proposed mechanism must be in accord with the fact that the component fatty acids of tissue fat without exception have an even number of carbon atoms and comprise all members of the series up to C_{20} and higher. The simplest general scheme for the biosynthesis of all fatty acids would be one involving building units containing two carbon atoms. Proof of C_2 condensation has been established in at least one case by the finding that deuteriopalmitic acid is converted to deuteriostearic acid by the rat (22). Since deuteriostearic acid is also degraded biologically to deuteriopalmitic acid, the reversible removal and addition of C_2 units is clearly a normal event. As to the nature of the fatty acid derivative undergoing the chain elongation, the natural occurrence of fatty acid aldehydes

* In this connection, the question as to whether pyruvic acid can be converted to acetic acid *in vivo* needs further investigation. From our findings with deuterioalanine, which we used as a source of deuteriopyruvic acid, it appears that pyruvate–acetate conversion in the liver constitutes only a minor pathway for pyruvate metabolism.

and the fact that palmitic acid is in biological equilibrium with cetyl alcohol (23), may be of significance.

In testing experimentally the hypothesis of C_2 condensation, the choice among compounds containing two carbon atoms is rather limited. Except in the case of the amino acid glycine, the existence of C_2 compounds in animal tissues has never been substantiated by isolation. Acetic acid suggests itself as a likely intermediate if only for the reason that it arises as a product of fatty acid degradation. When acetic acid containing deuterium in sufficient concentration to induce detectable deuteriocholesterol formation was administered to rats, no significant amounts of isotope appeared in the fatty acids (4). But when the experiment was repeated with an acetate preparation containing very high concentrations of deuterium as well as a carboxyl group labeled with C^{13}, the fatty acids of liver and depot were found to contain both isotopic carbon and deuterium (16). The incorporation of both isotopes proved that the acetate molecule had been utilized as such. However, the ratio of the concentrations of the two isotopes in the fatty acids synthesized differed considerably from that in the fed material, indicating that the conversion involved a loss of carbon-bound hydrogen. Lability of the hydrogen atoms in the methyl group of acetic acid cannot be responsible, since the ratio of the two isotopes remains unchanged when acetic acid is employed in the acetylation of foreign amines (unpublished results of Bloch and Rittenberg). It follows that the intermediates in the fatty acid synthesis are of such a nature as to permit loss of carbon-bound deuterium by exchanging it with the hydrogen of the body fluids. An analogous conclusion was necessary with respect to the intermediates in fatty acid oxidation. The failure to detect fatty acid synthesis from acetate labeled by deuterium alone and in relatively low concentrations, emphasizes the caution which is necessary in the interpretation of negative data. Use of a test compound containing the stable isotope of carbon as well as deuterium not only eliminates the uncertainty arising from hydrogen lability but also permits certain deductions from the change of the $D:C^{13}$ ratio.

The mechanism of condensation leading to the structure of cholesterol must differ in at least some respects from that involved in fatty acid synthesis. Owing to the complexity of the chemical structure of cholesterol the nature of the condensation reactions has re-

mained entirely obscure, apart from the finding that the cholesterol side chain, as well as the nucleus, was derived from acetate. In the case of the fatty acid, a wider variety of degradative procedures are feasible, so that the isotope concentrations at specific positions of the molecule can be ascertained. The acetic acid employed in these experiments contained the carbon isotope only at the carboxyl carbon. If the fatty acids were formed by successive condensation of C_2 units, only the odd-numbered carbon atoms would contain C^{13}. The carboxyl group obtained by thermal degradation of the fatty acid was indeed about twice as rich in isotope as the total molecule. This finding eliminates, as the principal synthetic reaction, the addition of one C_2 unit to either the α- or ω-carbon atom of the preformed fatty acid chain. The position of the isotope in the fatty acid chain can be traced further by oxidative cleavage of the "oleic acid" fraction into pelargonic acid, derived from carbon atoms 10 to 18, and azelaic acid from carbon atoms 1 to 9, of the fatty acid. Both moieties contain deuterium and C^{13}, to about the same extents. It therefore seems probable that the C_2 units are used uniformly for the various parts of the fatty acid chain. If the condensation mechanism involved were of the acetaldehyde-aldol type, in analogy to the *in vitro* synthesis of stearic acid carried out by Kuhn *et al.* (11), a biological reduction of acetic acid to acetaldehyde would be involved, and this is so far without experimental basis. A Claisen condensation with acetic acid or an acetic acid derivative is a more attractive possibility because the simplest reaction of this type, the condensation to acetoacetate, has definitely been shown to occur. The β-keto acids which would have to be postulated as intermediates, should contain readily enolizable hydrogen at the alpha position. The fact that deuterium is lost in the conversion of acetate to fatty acid is in accord with the postulated scheme, which may be viewed as a reversal of the steps responsible for fatty acid degradation. The fact that acetic acid is involved in the synthesis as well as in the breakdown of fatty acids, and the reversibility of two partial steps, *viz.*, the acetate–acetoacetate reaction and the palmitic acid–stearic acid conversion, are in favor of this scheme. The hypothesis is not out of line with general biochemical experience, if one considers the reversibility of individual reactions in the metabolism of proteins and carbohydrates.

The role of acetic acid as a building unit for fat suggests that

one of the pathways by which carbohydrates are converted to fatty acids involves acetic acid. The solution of this problem will necessitate further clarification of the oxidative phases of carbohydrate metabolism, and specifically that of pyruvic acid.

References

(1) Bernhard, K., *Z. physiol. Chem.*, **267**, 91 (1940).
(2) Bloch, K., *J. Biol. Chem.*, **157**, 661 (1945).
(3) Bloch, K., Berg, B., and Rittenberg, D., *J. Biol. Chem.*, **149**, 511 (1943).
(4) Bloch, K., and Rittenberg, D., *J. Biol. Chem.*, **143**, 297 (1942).
(5) Bloch, K., and Rittenberg, D., *J. Biol. Chem.*, **155**, 243 (1944).
(6) Bloch, K., and Schoenheimer, R., *J. Biol. Chem.*, **131**, 111 (1939).
(7) Folin, O., in *Festschrift für O. Hammarsten, Uppsala Univ. Årsskr.*, **1906**, III, 1.
(8) Hurtley, W. H., *Quart. J. Med.*, **9**, 301 (1915).
(9) Jowett, M., and Quastel, J. H., *Biochem. J.*, **29**, 2159 (1935).
(10) Krebs, H. A., and Johnson, W. A., *Biochem. J.*, **31**, 772 (1937).
(11) Kuhn, R., Grundmann, C., and Trischmann, H., *Z. physiol. Chem.*, **248**, 4 (1937).
(12) Loeb, A., *Biochem. Z.*, **47**, 118 (1912).
(13) MacKay, E. M., Wick, A. N., and Barnum, C. P., *J. Biol. Chem.*, **136**, 503 (1940).
(14) Medes, G., Weinhouse, S., and Floyd, N. F., *J. Biol. Chem.*, **157**, 752 (1945).
(15) Moss, A. R., and Schoenheimer, R., *J. Biol. Chem.*, **135**, 415 (1940).
(16) Rittenberg, D., and Bloch, K., *J. Biol. Chem.*, **154**, 311 (1944).
(17) Rittenberg, D., and Schoenheimer, R., *J. Biol. Chem.*, **121**, 235 (1937).
(18) Roloff, M., Ratner, S., and Schoenheimer, R., *J. Biol. Chem.*, **136**, 561 (1940).
(19) Schoenheimer, R., and Rittenberg, D., *J. Biol. Chem.*, **120**, 155 (1937).
(20) Shemin, D., and Rittenberg, D., *J. Biol. Chem.*, **151**, 507 (1943).
(21) Shemin, D., and Rittenberg, D., *J. Biol. Chem.*, **158**, 71 (1945).
(22) Stetten, D., Jr., and Schoenheimer, R., *J. Biol. Chem.*, **133**, 329 (1940).
(23) Stetten, D., Jr., and Schoenheimer, R., *J. Biol. Chem.*, **133**, 347 (1940).
(24) Stetten, M. R., and Schoenheimer, R., *J. Biol. Chem.*, **153**, 113 (1944).
(25) Swendseid, M. E., Barnes, R. H., Hemingway, A., and Nier, A. O., *J. Biol. Chem.*, **142**, 47 (1942).
(26) du Vigneaud, V., Cohn, M., Chandler, J. P., Schenck, J. R., and Simmonds, S., *J. Biol. Chem.*, **140**, 625 (1941).
(27) Weinhouse, S., Medes, G., and Floyd, N. F., *J. Biol. Chem.*, **155**, 143 (1944).

THE STEROID HORMONES

GREGORY PINCUS, VISITING PROFESSOR OF EXPERIMENTAL BIOLOGY, CLARK UNIVERSITY; DIRECTOR OF LABORATORIES, THE WORCESTER FOUNDATION FOR EXPERIMENTAL BIOLOGY

THE MODERN biochemistry of the steroid hormones had its inception in 1929 with the isolation of estrone from human pregnancy urine by Doisy and by Butenandt. The previous findings of Fellner, and of Doisy and Allen, that active ovarian material was probably lipid in nature, and the discovery by Aschheim and Zondek of larger amounts of folliculoid activity in lipid fractions of human pregnancy urine paved the way for the simultaneous chemical identifications of estrone in the United States and Germany. In the fifteen years that followed, the organic chemistry of the steroid hormones has developed explosively. Naturally occurring and synthetic substances having testoid, folliculoid, luteoid, and corticoid effects have been obtained in profusion.*

On the basis of the effects of these compounds, new techniques in clinical medicine have developed. The nature of sex determination, pubertal growth and sexual involution have been delineated experimentally with the aid of the steroid sex hormones. The availability of active corticoid steroid substances has led to a notable elucidation of the hormonal control of electrolyte and water balances,

* The terminology suggested by Selye (10) will be used throughout this chapter. Testoid is synonymous with androgenic, folliculoid with estrogenic or gynecogenic, luteoid with progestational or luteogenic, and corticoid with adrenocorticoid.

as well as of certain phases of carbohydrate metabolism *in vivo*. The remarkable effects of steroid hormones on pituitary activity have led to new conceptions of hormonal balance. The combination of brilliant chemical investigation and the astute application of chemical discoveries to physiological investigation has been so prolific that the journals are crowded with papers in the field; and in fact new journals have been founded specifically for the publication of data in endocrinology.

Nonetheless, on the biochemical side there have been two notable gaps in the development of steroid hormone investigation: (1) the nature of steroid hormone anabolism and catabolism and (2) the role of the steroid hormones in cellular processes.

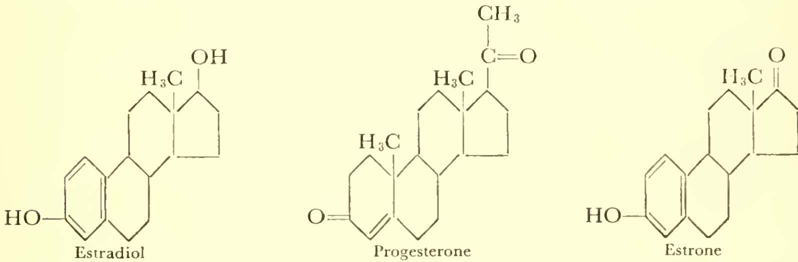
Biosynthesis of Steroid Hormones

Concerning the precursors and synthesis of the steroid hormones, our factual basis is slight. Several authors have indicated possible modes of degradation of cholesterol *in vivo* to active hormonal substance (3,6), and the synthesis of testosterone and progesterone is indeed accomplished commercially in a series of reactions the first step of which is the oxidation of the C_{17} side chain of cholesterol followed by molecular rearrangement. Experimental verification of such hypotheses was practically completely lacking until Bloch's recent demonstration that deuterium-containing pregnanediol is found in the urine of pregnancy after the administration of cholesterol containing heavy hydrogen. The recovery of marked pregnanediol was of the order expected on the assumption that it arose from progesterone, which in turn was synthesized from the cholesterol administered. The nature of this cholesterol degradation and the probable intermediates are not known, although Pearlman's isolation of a series of interesting neutral steroids from bile suggests that a search in bile for the intermediates may be profitable. Also suggestive is the evidence of Long and his collaborators that cholesterol in the adrenal glands declines markedly under conditions involving active secretion of corticosteroid. It must still be demonstrated to what extent cholesterol is the precursor of any or all steroid hormones. The evidence that cholesterol itself is synthesized from simple precursors such as acetic acid renders untenable the numerous theories concerning elementary steroid syn-

thesis, but provides a new approach to the investigation of steroid synthesis.

It is customary to think of three principal glands as the endogenous steroid hormone factories; these are the testes, the ovaries, and the adrenal cortex. There are numerous demonstrations of folliculoid material in animal fluids, especially in liver and bile. Estriol, estrone, and estradiol have been isolated from human placenta and there have been indications of the presence of progesterone in this organ; but the placenta as a synthesizing organ may be considered as a rather special case. In brief, the presence of notable hormonal activity in certain tissues cannot be taken as *prima facie* evidence that the hormones are produced in these tissues. For this reason our concepts even of the sites of steroid hormone synthesis are in a state of flux.

The female in full reproductive vigor leads a double life; she is alternately mate and mother. According to current concepts, two ovarian hormones are primarily responsible for this rhythm of function, estradiol and progesterone. Estradiol, the hormone of the

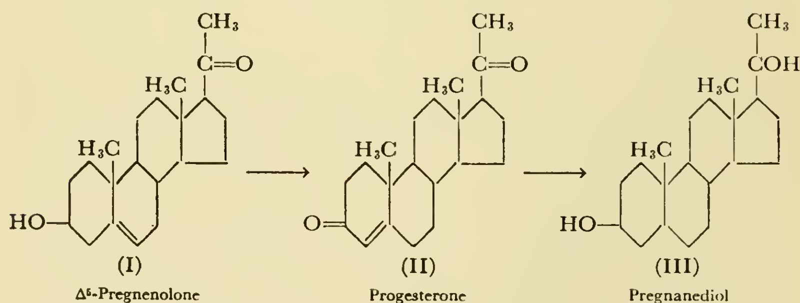


Estradiol Progesterone Estrone

ovarian follicle, is considered responsible for the estrous phenomena that culminate in mating, and progesterone, the hormone of the corpus luteum, is responsible for the maintenance of pregnancy. Doisy and his collaborators, in 1936, isolated and chemically identified estradiol from sow ovarian tissue; and in 1934 progesterone from luteal tissue was chemically identified by four separate groups of investigators. The uniqueness of progesterone as the ovarian luteoid is not open to question, but the evidence is fairly clear that another folliculoid, estrone, also arises from ovarian tissue. Westerfeld and collaborators found it to occur in sow ovarian tissue in an amount nearly equal to that of estradiol, and Dr. J. Schiller and the writer have recently

obtained evidence for its production by human ovaries perfused *in vitro*. Estradiol is oxidized to estrone by various tissues and estrone is reduced to estradiol. Which of the two is formed first in the ovary cannot therefore be determined *a priori*. We need to know the precursor of both to solve the problem.

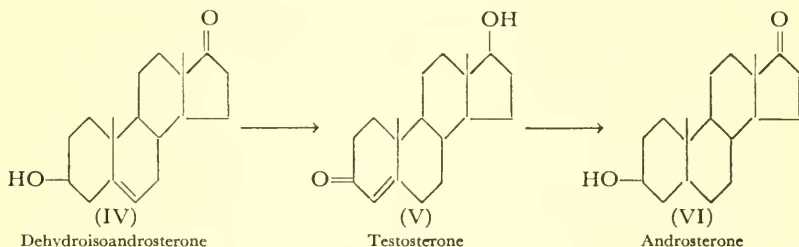
The isolation of estrone (but not estradiol) from adrenal tissue suggests that it may be supplied by the adrenal gland to the ovary for conversion to estradiol. Production of estrone by the adrenal gland must be in minor amount since the excision of the ovaries in adult animals effectively abolishes the estrous cycle. Furthermore, adrenalectomy in animals does not abolish estrous cycles, though irregularities do occur. Finally, Samuels and collaborators have demonstrated that women with Addison's disease (in which the adrenal glands are atrophied or destroyed) excrete normal amounts of folliculoids. It is clear, therefore, that the adrenal glands are not essential for the production of folliculoids. It is interesting to note, however, that very early ovariectomy in mice leads to the eventual establishment of estrous cycles and the excretion of notable amounts of folliculoid. Also, the injection of pituitary adrenocorticotropic hormone into ovariectomized rats stimulates folliculoid production by the adrenals. The most likely conclusion is that the capacity for ovarian hormone synthesis is present in adrenal tissue, but that optimal conditions are found in normal ovarian tissue. This remarkable steroidogenic capacity of the adrenal cortex is also confirmed by the isolation in small amounts of progesterone from adrenal extracts.



The immediate precursor of progesterone may be Δ^5-pregnenolone, for we have recovered pregnanediol from the urine of men receiving pregnenolone by mouth or by injection. The transformation

would appear to be from I → II → III. The conversion of I → II has been accomplished by bacteria. It is interesting that Δ^5-pregnenolone has been obtained as a naturally occurring substance thus far only from testis tissue; yet progesterone is a female hormone *par excellence.* Patient search for pregnenolone in female tissues appears desirable.

Among the naturally occurring 19-carbon steroids, Δ^5-dehydroisoandrosterone (formula IV) corresponds to Δ^5-pregnenolone. We would expect it, therefore, to be the immediate precursor of testosterone (V) (the primary testis hormone). That the transformation of IV → V takes place appears probable on the basis of Danby's data for the perfusion of testis tissue with IV. The resulting high testoid activity in the perfusate might well be attributed to V which is the most active of all testoid steroids. That VI cannot arise directly from IV may be deduced from the fact that Dorfman and Hamilton (2) were unable to recover VI from the urines of eunuchoid men after the administration of dehydroisoandrosterone.



The light thrown on the nature of the immediate precursors of the sex hormones is meager indeed, but in the field of the adrenocorticosteroid hormones the darkness is almost total. We know only that these hormones from adrenal tissue are accompanied by inactive steroid substances which *might* be converted to active hormone by appropriate oxidation and reduction. It would be most instructive to perfuse such substances through adrenal glands to see if they are convertible to active hormone.

While one may thus indicate possible chemical precursors of the known steroid hormones by administering model substances either *in vivo* or *in vitro*, the normal course of metabolism will continue to be a matter of speculation until a more direct attack is devised. The

method of Bloch using deuterized cholesterol is limited by the need for relatively large amounts of the steroid products for mass spectroscopy. Probably it is for this reason that only pregnanediol has been identified as a transformation product of deuterized cholesterol. No similar study has been carried out on other postulated intermediaries. It is to be hoped that the development of micro methods will make possible more extensive utilization of the isotope technique. Similarly, in tracing the synthesis of steroids from postulated triose precursors, the use of isotopes is clearly indicated.

Another approach to anabolic processes, particularly in mammals, involves studies of the metabolism of precursors *in vitro*. The perfusion of active steroidogenic organs should provide useful information. On the basis of data derived from perfusion studies one might attempt the isolation of the enzyme systems concerned with the observed transformations. It is notable, however, that Danby, who obtained increased testoid activity on perfusion of the bull testis with dehydroisoandrosterone, was unable to observe any increase at all on incubating dehydroisoandrosterone with testis mash.

One further note on the problem of steroid hormone synthesis seems pertinent here. Emmens has made a useful distinction between estrogens and proestrogens (folliculoids and profolliculoids in our terminology). In dealing with substances particularly of the stilbene series, he finds certain compounds which on systemic injection are much less active per milligram than when applied intravaginally. These are true estrogens because their action on the susceptible end-organ requires much less material with direct application. Calling the systemic median effective dose S and the median effective local dose L, Emmens finds true estrogens (*e. g.*, estrone, diethylstilbestrol) give in spayed mice an S/L ratio of 50 or more. Proestrogens give an S/L ratio approximating unity (*e. g.*, α-phenylstilbestrol, triphenylethylene). Segaloff has shown that three proestrogens (triphenylethylene, triphenylchloroethylene, and 9,10-di-*n*-propyl-9,10-dihydroxy-1,2,5,6-dibenzanthracene) are definitely potentiated after intrasplenic injection (with the spleen *in situ*). Passage through the hepatic portal system increased their activity some three to six times. Since injection into the isolated spleen does not effect such potentiation, it is probable that the liver converts these substances to more active compounds. The possibility that the liver may play a role in

steroid hormone synthesis should not be overlooked. Estrone perfused through the rat liver or incubated with liver brei is at least in part converted to α-estradiol, that is, into more active substance. Estrone on this basis may be considered a precursor of a highly active folliculoid. Whatever the logical distinction may be, a special role in steroid hormone synthesis may have to be assigned to the liver.

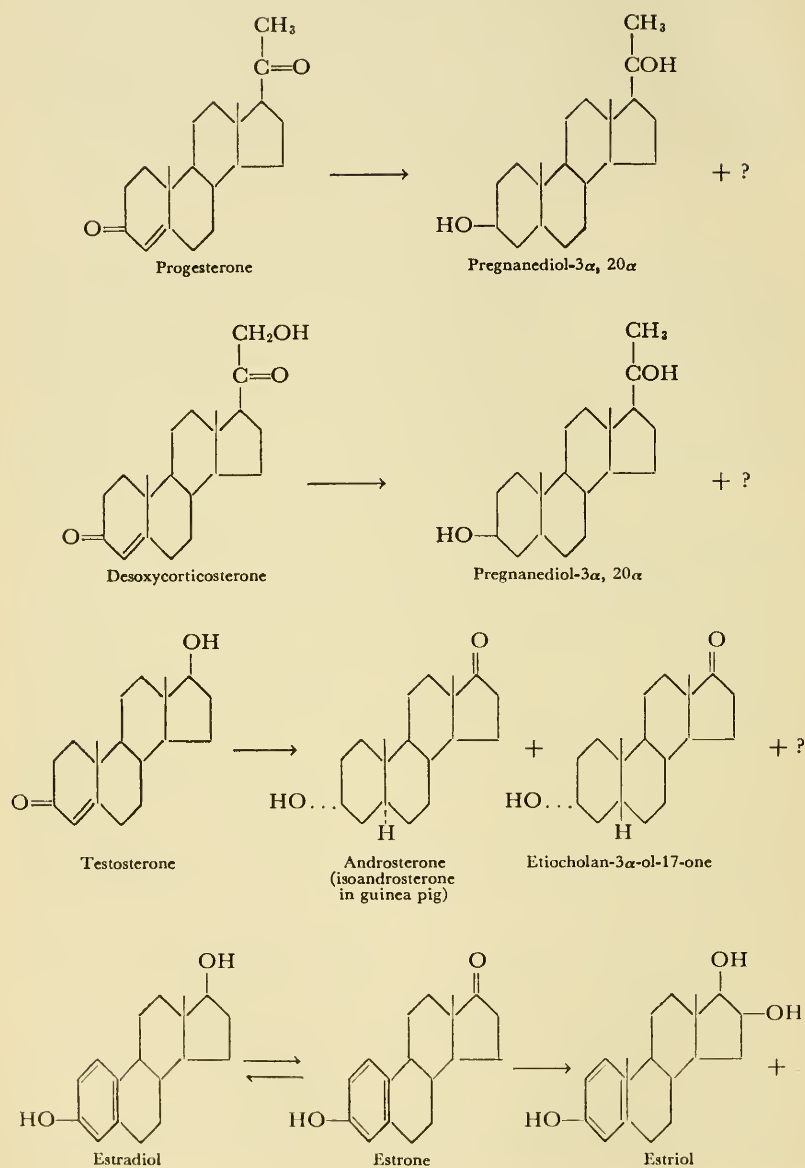
Destruction of Steroid Hormones

The nature of steroid hormone degradation in animals has been more thoroughly investigated than hormone anabolic processes. The available data has been reviewed in some detail (9). With certain exceptions, most of the data have been arrived at by determination of urinary steroid catabolites in normal human and animal urines or similar urinary analyses after the administration of active hormone. In the case of mammalian subjects, certain of the hormonally active steroids probably undergo the series of conversions given in scheme I.

Adopting the convenient classification of the hormonally active steroids on the basis of number of carbon atoms it can be seen that the identified catabolic end product of C_{21} hormones is pregnanediol. It is not established that this is the only end product. Furthermore, pregnanediol is not found in the urine of certain species (*e. g.*, cat and monkey) nor is it found therein after progesterone administration. Finally, quantitative studies indicate that only a portion of administered progesterone and desoxycorticosterone is converted to pregnanediol. It is probable, therefore, that other breakdown products of the C_{21} steroids should be sought. The pregnanolones, which have been found in human pregnancy urine would seem to be likely metabolites, but attempts to identify them in urine after the administration of C_{21} steroids to human and animal subjects have thus far been unsuccessful.

For the C_{19} steroids we have somewhat more detailed information as to possible modes of catabolism. Dorfman and Hamilton have attempted to set up a balance sheet for testosterone administered to eunuchoid men. After making certain assumptions the data indicate that androsterone may be taken as the principal catabolite, accounting for a little over 60% of the administered hormone. In somewhat better controlled experiments with guinea pigs, isoandrosterone (the principal urinary catabolite to be identified in this species)

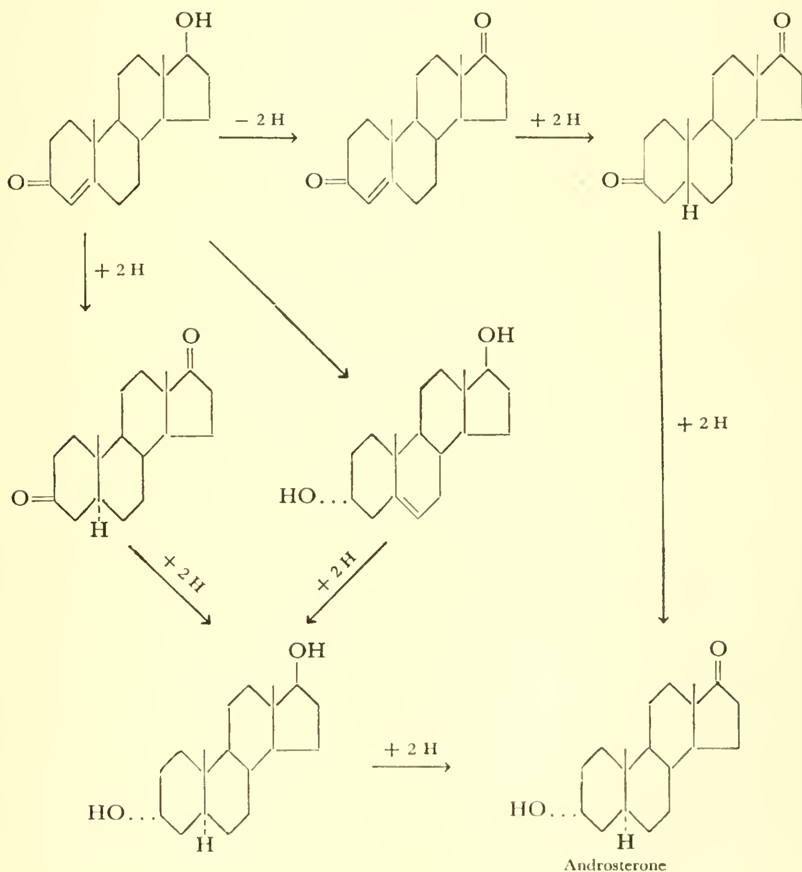
SCHEME I



represents 30% of the administered hormone. From chimpanzees receiving testosterone proprionate orally, Fish and Dorfman obtained in the urine androsterone, androsten-17-one, and etiocholan-3-ol-17-one.

Dorfman and Hamilton (2), on the basis of products isolated from urine following administration of various precursors, have suggested the mechanism of scheme II for the conversion of testosterone to androsterone. It should be noted, however, that this scheme does not exclude other degradation products which may be produced in

Scheme II



smaller yield. The acid test of the scheme for the interconversion of testosterone and androsterone would be met by certain identification of the intermediaries after administration of active hormone. Finally, Dorfman has reported a conversion of exogenous testosterone to etiocholan-3-ol-17-one and to isoandrosterone in man. To account for these reduction products several other mechanisms are indicated (7).

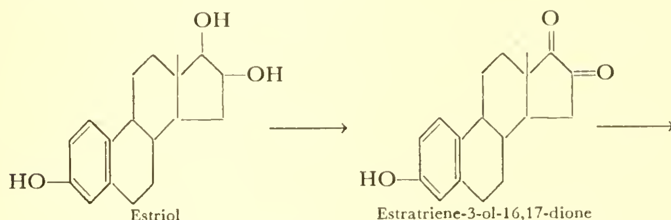
While testosterone has generally been recognized as the true hormone of the testis, certain species may produce other, now unknown, testoids. For example, Ruzicka was unable to obtain testosterone from hog testis tissue, and Hirano has reported the isolation of testalolone and possible related steroids. If we want to ascertain the fate of steroid hormone produced by the testis, a more complete knowledge of just what the testis does produce is obviously required.

The course of conversion of the natural folliculoids (the C_{18} steroid hormones) indicated above appears to be well established. In the rabbit β-estradiol is produced, whereas in man, the monkey, and probably also in the rat and guinea pig α-estradiol is a conversion product of estrone. Estriol is indicated as a metabolite in the various animal species, but only in man has it been isolated as a urinary end product following administration of estrone.

In all species investigated, the recovery of known catabolites from the urine after the administration of estradiol or estrone to the intact animal is low indeed, ranging from 1–2% to at most a possible 15% of the administered hormone. The fate of the balance has been the subject of some investigation. It has been claimed that the missing material undergoes an enterohepatic circulation much like that of the bile acids. This claim requires more evidence, but there is no doubt that a good amount of folliculoid appears in the bile after estradiol and estrone injection. Dr. J. Schiller and the writer have recently shown that the removal of most of the liver of female rats results in a large increase of urinary folliculoid. The injection of estrone into such partially hepatectomized rats leads to the recovery in the urine of folliculoids accounting for about 60% of the administered hormone. From female rats with intact livers, 13%, at most, of administered estrone can be recovered. Zondek has postulated a liver estrinase, for inactivation can be observed *in vitro* on incubating folliculoid hormones with liver mash. Attempts to isolate a phenolase that might be responsible for such inactivation have been unsuccessful.

Since the inactivation of estradiol by rat liver slices appears to depend on adequate supplies of thiamin and riboflavin, an approach to the inactivating systems through enzymes dependent on these vitamins is indicated.

It is fairly well established that the liver is largely responsible for the conversions: estradiol ⇄ estrone → estriol. Liver perfusion not only results in a rapid interconversion of estradiol and estrone, but seems also to hasten the production of estriol. In contrast, estriol is produced in negligible amounts after perfusion of the kidney with estrone. Marrian (8) has suggested that folliculoid breakdown proceeds through estriol to a diketone, thence to other inactive substances. Since the liver is such an active producer of estriol, the postulated diketone should be sought in liver extracts, particularly after the administration of natural folliculoid or following the incubation of the hormones with liver slices or mash.

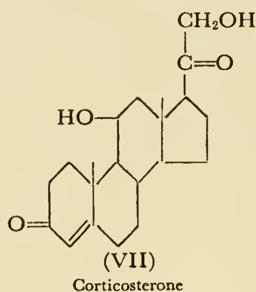


While the catabolic role of the liver has been most intensively examined in studies with the folliculoids, its responsibility for the breakdown of testosterone and progesterone is also established. It is sufficient merely to implant pellets of these hormones into the hepatic portal system (intrasplenic implantation is used principally) to observe the remarkable absence of testoid or luteoid effects. Implantations elsewhere are clearly effective. In contrast, the administration of these steroids to partially hepatectomized animals results in remarkable exaggeration of their typical effects.

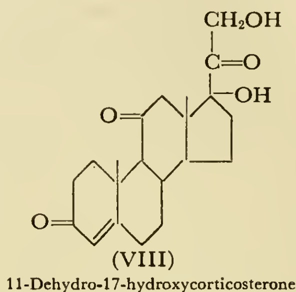
The nature of the products which the liver produces from these hormones is not too well established. Recent work indicates that testosterone is converted to a mixture of 17-ketosteroids one of which, Δ^4-androstene-3,17-dione, has been identified chemically.

In contrast to the information available for the folliculoid, testoid, and luteoid hormones, practically nothing is known about the

fate of the corticosteroids in the organism. We have mentioned that desoxycorticosterone is degraded in part to pregnanediol. No parallel information is available for the other active corticoids such as VII and VIII, largely because these substances are available as pure



(VII)
Corticosterone



(VIII)
11-Dehydro-17-hydroxycorticosterone

chemicals in only very small amounts. (They have been obtainable thus far only from adrenal gland extracts and the yields are very low.) Although we know that they are very rapidly inactivated *in vivo*, their breakdown products are unknown. It is suspected that certain of them may be converted to 17-ketosteroids, since adrenalectomy or atrophy of the adrenals leads to a marked reduction in urinary 17-ketosteroid output; but since other possible precursors of 17-ketosteroids have been isolated from adrenal tissue (*e. g.*, androstenedione) experimental verification is needed.

Eventually we shall have a detailed description of steroid hormone transformation in animals and of the particular enzyme systems involved. If we are to know the extent to which such transformations are tied up with the function of the hormones, we will have to know not only more about these metabolic changes but also the precise role of the steroid hormones in cellular processes.

Role of Steroid Hormones in Cellular Processes

There is an extensive literature containing qualitative and quantitative data on the effects of administering steroid hormones to animals. It is not possible as yet to interpret these effects in terms of the chemistry of the cell. Thus, the remarkable growth of vaginal epithelium induced in spayed animals by folliculoids has been rather intensively studied. That it can be induced by the direct application

of minute amounts of active hormone to the vaginal epithelium has been demonstrated. That these catalytic amounts of hormone cause a remarkable spurt of mitoses in the epithelial cells is also well known; but the means by which the hormones induce this growth have not been elucidated. In the case of the uterus, where parallel growth transformations occur, there has been some indication that the folliculoids cause a vasodilation of the uterine vessels which is followed by an influx of water into the uterine tissues. Similar large water retention occurs in the sexual skin of the pig-tailed monkey. Is this absorption of water accompanied by an intake of growth-promoting substances from the blood? Or is the folliculoid directly a growth promoting factor? Our fragmentary knowledge of the biochemical processes underlying end-organ responses to the steroid sex hormones debars an answer to even such simple questions. In their recent review on the influence of hormones on enzymic reactions, Jensen and Tenenbaum (4) have a section entitled "Sex Hormones," in which they state: "Estrogens and androgens no doubt affect metabolism either directly in the tissues or indirectly by stimulating or inhibiting the rate of secretion of other endocrine organs. *In vitro* studies on the effect of this group of endocrine principles on a given enzyme system have as yet not been reported."

The adrenal cortex steroids which on administration to animals induce gluconeogenesis from protein and suppress the utilization of glucose would seem to offer material for elucidating the role of such steroids in cellular processes. But the few attempts to define the role of these steroids in carbohydrate metabolism have led to apparently conflicting results. Evidence has been presented that adrenal cortex extract inhibits glycogen breakdown and that it also speeds glycogen synthesis in the rat liver. Certainly the adrenalectomized animal has little or no liver glycogen. Glycogen synthesis from pyruvate and *d*-lactate by liver slices is increased on addition of adrenal cortex extract, but there is no increased synthesis from *d,l*-alanine or *d*-glutamate. Kidney slices from adrenalectomized animals show decreased formation of carbohydrate from *d,l*-alanine and no reduction in synthesis from pyruvate. The implication is that adrenocortical hormones affect carbohydrate synthesis in liver and kidney in diametrically opposite ways. What is requisite is not only further data and accurate checking, but a direct attack on the enzyme systems involved.

The tissue slice technique merely allows one to demonstrate an acceleration or inhibition of a set of metabolic events. But unless the enzyme systems involved are isolated and their activity measured directly, the exact participation of the hormones cannot be accurately worked out. They may be coenzymes at some step in the synthesis of glycogen or they may accelerate the synthesis of an enzyme essential for glycogen production.

It has indeed been suggested that the protein breakdown that follows administration of adrenal cortex steroid is due to the increase in liver arginase and not to gluconeogenetic processes. But an increase in kidney arginase occurs in mice receiving steroid substances that promote protein synthesis. Can we explain this apparent discrepancy by saying that liver arginase functions in protein breakdown and kidney arginase in protein synthesis?

This promotion of protein synthesis by various steroid substances (*e. g.*, testosterone, androstanediol, α-estradiol, and others) is a biochemical phenomenon of great interest and importance. Its elucidation in terms of cellular mechanisms is a preliminary to the explanation of the remarkable series of growth phenomena affected by the steroid hormones. Kenyon's review (5) of the data on man clearly defines the effects of testoids and folliculoids on nitrogen balance. The reader is also referred to Albright's engaging Harvey lecture (1) on Cushing's syndrome for an astute and stimulating discussion of the role of steroid hormones in protein metabolism and osteogenesis.

It is notable that the principal protein anabolic steroids are testoids, and one is tempted to ascribe the muscularity of the virile male to testis hormone. The proverbial flaccidity of the eunuch tends to strengthen this notion. But when one examines more closely the sex hormone complement of the male, curious anomalies appear. The male animal produces not only testoid but also folliculoid. Certain stallions excrete more estrone into the urine than do pregnant mares. The bull, according to Marker is a remarkable excretor of pregnanediol, the urinary metabolite of progesterone, whereas little or no pregnanediol is excreted by the steer. Conversely, the female animal produces testoid hormone. Ovaries transplanted to the ear of castrated male rats maintain the reproductive organs of the male, but when they are transplanted to warmer sites (*e. g.*, onto the kidney) their testoid activity does not appear. Can we attribute the difference

between maleness and femaleness to the temperature of the gonads? On the basis of urinary hormone measurements the difference between men and women rests on a difference in the ratio of male to female sex hormone, for the absolute amounts of each excreted varies widely in each sex. It would seem to follow, therefore, that the development of the typical secondary sexual characteristics depends on the balance of certain steroid chemicals in the body. What cellular processes are controlled by this balance? There is no direct neutralization of testoid by folliculoid. Again we find a biochemical mystery.

Conclusion

In this presentation of certain problems in the biochemistry of steroid hormones there has been no attempt at an inclusive review. Essentially we have been engaged on a rather specialized gap-finding mission. If there has seemed to be an emphasis on our ignorance, the purpose thereof is obvious. The frontiersman faces the unknown; but behind him and at his disposal are innumerable known resources.

The known resources and the many frontiers of steroid hormone endocrinology could occupy our attention for many pages. In the decade that has followed the original identification of a few steroid substances as active hormones, investigation has been made of their role in a host of normal and pathological processes. There are data on their involvement in menstruation, pregnancy, lactation, growth, reproduction, embryogeny, puberty, senility, cancer, dermatology, overt endocrinopathies, leukemia, inflammation, diabetes, melancholia, schizophrenia, toxemia—most of the ills and physiological goods of man and beast. These hormones are called upon to regulate the other glands of internal secretion; they depress and stimulate pituitary secretion and thyroid activity. The folliculoids affect not only the organs of reproduction but also bone growth, sugar regulation, skin growth, and carcinogenesis. Progesterone maintains embryo life and pregnancy as well as affecting mammary growth, the salt balance in blood and tissues, and smooth muscle contractility. The testoids are not merely androgenic, they stimulate the synthesis of protein and are often luteoid, acne-producing, and factors in baldness. The adrenal cortex steroids function in sugar, fat, protein, and salt metabolism; they affect emergency reactions to damage, processes in convalescence, and lymphatic function.

From conception to death the steroid hormones regulate, control, arbitrate, and defend. With dexterity and imagination chemists have synthesized compounds that perform and facilitate each of the various effects of the hormones. With the array of activities and compounds available we are predestined to full biochemical explanation of these activities. Unknown only are the rate and manner of the eventual revelation, but therein is the key to the charm and excitement of experimentation.

References

(1) Albright, F., *Harvey Lectures*, **38**, 123 (1943).

(2) Dorfman, R. I., and Hamilton, J. B., *J. Biol. Chem.*, **133**, 753 (1940).

(3) Fieser, L. F., *The Chemistry of Natural Products Related to Phenanthrene.* 2nd ed., Reinhold, New York, 1937.

(4) Jensen, H., and Tenenbaum, L. E., in *Advances in Enzymology*, Vol. IV. Interscience, New York, 1944, p. 259.

(5) Kenyon, A. T., *Biol. Symposia*, **9**, 11 (1942).

(6) Koch, F. C., in *Sex and Internal Secretions.* 2nd ed., Wood, Baltimore, 1939.

(7) Koch, F. C., *Biol. Symposia*, **9**, 41 (1942).

(8) Marrian, G. F., *Harvey Lectures*, **34**, 37 (1938).

(9) Pincus, G., and Pearlman, W. H., *Vitamins and Hormones*, **1**, 293 (1943).

(10) Selye, H., *Rev. can. biol.*, **1**, 577–632 (1942).

PLANT HORMONES AND THE ANALYSIS OF GROWTH

KENNETH V. THIMANN,* ASSOCIATE PROFESSOR OF PLANT PHYSIOLOGY, THE BIOLOGICAL LABORATORIES, HARVARD UNIVERSITY

IT IS probably true that, in the field of biochemistry and physiology, there are at least a dozen workers studying animal material for every one studying plants. It would be expected, therefore, that progress toward an understanding of the chemical mechanisms operative in plants would be achieved more slowly than in the corresponding field of animal biochemistry. Nevertheless, when we survey the plant research of the last twenty years, the results appear surprisingly good. The list of the elements needed for plant nutrition has been more or less completed. The whole field of the relationship between light and flowering has been opened wide, and though the knowledge acquired has not yet been brought to the chemical level, it is of great importance, both for physiology and for agriculture. Somewhat parallel studies of the influence of temperature on flowering are developing. The chemistry of practically all the known plant pigments has been thoroughly worked out; and, while it is true that little is yet known about the formation and interconversion of these pigments, this field is ripe for physiological exploitation. Classical problems such as the interconversion of starch

* At present on leave of absence for service in the Navy Department.

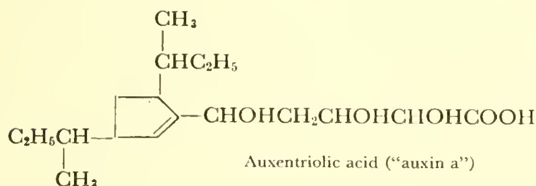
and sugar, and the accumulation of solutes by roots, have been greatly clarified. Even the elusive field of photosynthesis has seen notable progress through the investigation of the effect of intermittent light, and the recent study of the "dark" reaction with radioactive isotopes.

In no field has more striking progress been made than in that of the analysis of growth and its control by hormones. Darwin's work on what was later to become a classical material for study—the oat coleoptile—may be regarded as the first serious attempt to analyze growth. The work was refined and advanced during the fifty years that followed. In the first "shoot" of the dark-grown oat seedling, the extreme tip—a region of about 0.5 mm. in length—was shown (by the combined efforts of a number of workers) to produce a diffusible substance which controls the responses to light and gravity of the whole shoot. Later it became clear that the curvatures toward light, or away from the earth, which constitute these responses, are special cases of normal or symmetrical growth. The growth substance or hormone, diffusing from the tip into the base of the coleoptile, is asymmetrically distributed, with the result that one side grows faster than the other, causing curvature. In the absence of external stimulus, the hormone is symmetrically distributed, and it promotes and controls the normal elongation of the organ.

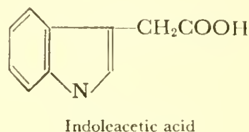
Discovery of the growth-promoting activity in sources far richer and more varied than the tips of seedlings, and the development of quantitative methods for assay by Went and others, soon led to the isolation of several "auxins" active in the test methods. Later it was recognized that a large class of synthetic substances have auxin activity (see below), an activity which is measured by methods that are now standardized. The method most frequently used involves decapitating the oat coleoptile, thus depriving it of its main source of auxin, and applying the test material to one side. The curvature which results is then measured. Alternatively, the auxin may be applied in such a way as to induce straight growth, which is then measured under the microscope. Small sections cut from the coleoptile, pea, or other seedlings can be immersed in solutions of test substances and their growth measured. Many other procedures have been used, all involving the use of auxin-deficient material as the test object.

"Auxin a" (or auxentriolic acid) and "auxin b" (or auxenolonic acid) were isolated by Kögl, Haagen-Smit, and Erxleben from urine

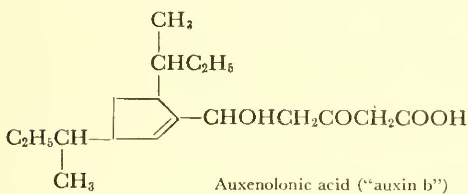
and from corn germ. Later, "heteroauxin" (or indole-3-acetic acid) was isolated by the same workers (3) from urine and from yeast, and by Thimann (10) from cultures of the fungus *Rhizopus suinus*. More recently it has been isolated from wheat germ by Haagen-Smit. The growth-promoting activity of these three substances is quantitatively and qualitatively very similar.



Auxentriolic acid ("auxin a")



Indoleacetic acid



Auxenolonic acid ("auxin b")

In addition to the three known naturally occurring substances, a large number of unsaturated ring-containing acids have growth-promoting activity. These include the propionic, butyric, and valeric derivatives of indole, the acetic derivatives of indene, naphthalene, anthracene, thiophene, and even (weakly) benzene. *cis*-Cinnamic acid is active, but the *trans*-derivative is not. The simple esters and amides, as well as methyl and chloro derivatives of some of these acids, also have activity. Indeneacetic acid is of particular physiological interest because its activity, though high, is localized, and unlike the natural auxins it is not readily transported through plant tissue. α-Naphthaleneacetic acid is also of importance because of its stability, which has led to its wide adoption for certain practical uses in horticulture.

Since the extensive work on the relation between chemical structure and growth-promoting activity has been fully reviewed (7,10), there is no advantage in considering it here. The essential groups are: (a) an unsaturated or an aromatic ring, and (b) a carboxyl or a group readily converted by the plant to a carboxyl. The

carboxyl group must be separated from the ring by at least one carbon (or oxygen) atom. The situation in which a number of chemically different substances exert the same biological activity has occurred also in the field of vitamins, though there, perhaps, the divergence between the natural and synthetic active substances is not so great. For synthetic materials with vitamin activity the term "vitamer" has been proposed (1). On the same basis, synthetic substances with growth-promoting activity in the standard auxin tests could be called "auximers," but at present there seems little need for this term.

It is not proposed to discuss in this essay the chemistry of the auxins but rather to consider the mechanism of their action, with particular reference to the nature of plant growth.

Historically, the first proof that the auxins influence more than one type of growth came from the study of bud inhibition. It is typical of shoot growth that, when one bud is vigorously developing, others below it on the same stem are inhibited from doing so. The apical bud thus "dominates" the lateral buds. Only when it is removed, as in pruning, can the lateral buds develop. After evidence had been brought forward, mainly by Snow, that this inhibition is exerted by means of a substance diffusing out of the growing bud toward the tissues below it, it was finally shown that pure auxin b, or also indoleacetic acid, in quantities comparable to those of the auxin produced by the growing bud, can duplicate the effect. Paradoxically, then, a growth-promoting substance is here acting as a growth inhibitor (6). It should be noted that the movement of auxin in living tissue, as out of a bud into the stem, is invariably in the direction from apex to base. This movement is strictly polar, even taking place against an auxin gradient. Transport occurs in the reverse direction only in dead tissues, or when unphysiologically high auxin concentrations are applied.

Very shortly afterward, two other important cases were brought to light. It was demonstrated, by Went and Bouillenne, that a root-forming hormone diffuses downward out of buds and leaves to stimulate the formation of roots at the base of a cutting. In the purification of this hormone, Went and Thimann found that root-forming activity was concentrated parallel with growth-promoting activity (*i. e.*, auxin), and finally that numerous auxins—synthetic and otherwise—acted as root-forming hormones. At about the same time, Snow showed that,

in young shoots, the cambium is stimulated to divide by a diffusible substance and that pure indoleacetic acid had the same effect. Likewise, the substance in orchid pollen, shown by Fitting to be responsible for the swelling of the ovary on fertilization, was identified with auxin. Finally the elongation of roots, unlike that of shoots, was shown to be inhibited by auxins. Wherever synthetic substances have been tested, their influence on these other growth processes has been to some extent quantitatively comparable with their effectiveness in producing simple growth by elongation (5,10).

Thus, unlike many animal hormones of specialized activity, the functions of auxins are manifold. The stream of auxin coming from the apical bud and the young leaves in it promotes the elongation of the shoot, stimulates the cambium to produce thickening, inhibits development of the lateral buds lower down, and promotes the formation of root initials toward the base. In the flower, the auxin contributed by the pollen, or released from the ovary tissue by fertilization, promotes the swelling of the ovary into a "fruit" and also prevents the fruit stalk from becoming separated (inhibition of abscission).

Morphologically, at least, these functions all appear to be very different. Not only do they comprise effects of opposite sign (namely, growth and inhibition) but even where the effect is positive the types of growth controlled by auxin differ widely. The stimulation of cambium or root initials to divide bears little resemblance to the promotion of simple elongation without cell division, as in the oat coleoptile. It is hard, therefore, to avoid the conclusion that the primary effect of auxin is in all tissues the same—a fundamental reaction whose morphological sequelae depend on the kind and age of the tissue, the availability of other interacting substances, and the external and internal conditions. This concept and the experiments bearing on it will be taken up in more detail below.

The general position described above has been established for some time. During the last few years the field has developed mainly along the following three lines.

(*a*) *Practical Applications.* Originally auxin was most generally used to induce the formation of roots on cuttings. Auxin treatment both accelerates rooting and increases the number of roots formed in cuttings of a wide variety of herbs, shrubs, and trees. It does not, of course, solve all the problems of vegetative propagation.

In particular, the time of year and the age of the plant from which cuttings are taken remain as important variables. A few species also do not respond at all to auxin, their rooting being limited by other factors which are not yet understood.

A second major application depends on the fact that auxin not only inhibits lateral bud development, as described above, but also inhibits the formation of the abscission layer, a special layer of cells at the base of a petiole or fruit stalk, whose walls separate very readily, resulting in the falling to the ground of the leaf or fruit. When the fruit is sprayed with auxin solution, therefore, the formation of this layer is delayed and the fruit remains on the tree for longer. Such spraying has been very successful with apples, since the delay of a week or two helps greatly to solve the long-standing problem of early fruit fall. The most effective auxin for this purpose appears to be α-naphthaleneacetic acid.

The effect of auxin in promoting the swelling of the ovary, mentioned above, has found interesting application in the production of seedless fruit. The auxin is applied to the style of the mature flower, either individually by hand, or by using a liquid spray or even vapor. The ovary begins to swell in a normal-looking way and the process continues, to produce seedless fruit. Tomatoes and holly are the outstanding successes obtained by this procedure, but seedless pears and even watermelons have been produced, and the method may have an important future [see the review by Gustafson (2a)].

Lastly, the treatment of seeds with auxin stimulates root development and may result in a promotion of growth which lasts for many weeks or months. The conditions necessary for this stimulation are, however, not fully understood and the results reported are so conflicting as to cast doubt on the legitimacy of this application. The toxicity of some auxins in high concentrations has recently led to an application as weed killer. An excellent summary of all of this applied work is given by van Overbeek (8).

All these applications have led to a scramble for trademarks, and even patents, and the emergence of special methods of advertising and marketing. Indoleacetic, indolebutyric and naphthaleneacetic acids are now on the market in the form of solutions, pastes, powders, and pills.

(*b*) *Auxin Reserves.* The auxin content of plant tissues is of

great interest. In the early work it was sufficient to place buds, leaves or other auxin-producing organs on such media as agar, and after a given diffusion period to apply the agar to standard oat coleoptiles for assay, a procedure which measures the rate of production. However, since for some purposes it is necessary to know the amount of auxin actually present, attention has recently turned to quantitative extraction of the tissues with organic solvents. This led to the discovery that, while in most tissues a certain small amount of auxin can be obtained by direct extraction, there is commonly from ten to fifty times as much present in an inactive form. This inactive material yields auxin slowly when wetted with water, but the process may be greatly accelerated by hydrolysis with some enzymes, notably chymotrypsin, or with alkali. Whether these inactive forms, or "auxin precursors," are the same in different plants and different types of tissues is not clear as yet. At least one of them appears to be an auxin protein.

The biological significance of these large auxin reserves is worth consideration. It seems probable that growth is commonly controlled, not so much by the rate of transport or even the rate of true formation of auxin, as by its rate of liberation from the stored form. The sudden onset of growth in tree buds in the spring, and its cessation in early summer just when conditions of light, temperature, and nutrition would be expected to favor both auxin production and the formation of organic structural materials, may find its explanation in this way. Many types of infection, as by gall and nodule bacteria and certain fungi, give rise to abnormal growth which corresponds well with the assumption of excessive local auxin formation; in the case of legume root nodules, large quantities of auxin are undoubtedly present. These pathological phenomena may be due to liberation of auxin from the inactive precursors just mentioned. It seems safe to predict that further study of auxin precursors will not only bring to light some interesting enzyme systems, but may lead to explanations of a number of normal and abnormal growth reactions.

(*c*) *Analysis of the Growth Process.* The discovery of an active substance, whether vitamin or hormone, is of importance for its own sake and also because the experimental use of the substance can elucidate the physiology of the processes which it controls. The study of growth in plants has made marked progress through the use of auxin as a tool. Before discussing this use of auxin, however, it

is necessary to consider the distinction between direct and indirect effects on growth. Some of the observed effects of auxin are in part at least indirect. When swellings are produced, or fruits induced to grow, under the stimulus of auxin, there must be a continuous supply of carbohydrates, amino acids, and especially water to the growing zone. The plant, therefore, must have some mechanism whereby materials for growth are accumulated at the place at which they are being used. To take the simplest case, that of the carbohydrates, it could be considered that, as the sugars are converted to polysaccharides and deposited as such in the cell wall, the concentration of soluble sugars decreases and hence more carbohydrates would flow there in consequence of the concentration gradient. However, accumulation in plant tissues does not always follow the gradient. Thus, salts are concentrated in young roots by a factor many times the concentration of the external nutrient solution, and sugars are sometimes accumulated in quite high concentrations in storage organs. Auxin itself, as mentioned on page 324, is freely transported against its gradient. The accumulation of materials in the growing organs, and the transport of materials there, consequently do not necessarily constitute a "gradient" process; and numerous suggestions have been made that auxin exerts its growth-promoting activity through the accumulating or "mobilizing" of food materials. Auxin is envisaged as in some way controlling transport, so that, wherever auxin is, other materials will be collected. This view has been expanded to include mobilizing effects on hypothetical special organ-forming substances, such as factors for root growth, stem growth, and leaf growth (9). The evidence for such organ-forming substances is indirect only; evidence for a mobilization phenomenon as the cause of the growth-promoting activity of auxin is at present by no means convincing.

It remains true, of course, that when auxin produces growth there is an accompanying accumulation of materials. It has been shown, for instance, that application of concentrated auxin paste to decapitated bean plants produces localized swelling which is associated with a large increase in soluble carbohydrates and total dry weight. It has also been shown in ingenious experiments (9) on oat seedlings that a zone of the plant which has been treated with auxin and later cut off is subsequently able to grow in solution better than a control zone. Hence it is not the occurrence of such accumulations, but the

problem of whether they are *causative* of growth or not, that remains unsettled.

To some extent, the need for postulating these indirect mechanisms arises from the difficulty of designing experiments which bear upon the growth process *per se*, and which introduce the minimum of extraneous considerations. In the study of vitamin action, it is usual to employ vitamin-deficient animals, but it is not so easy to obtain with certainty plants which are deficient in auxin. The discussion of auxin reserves in the preceding section brings out the difficulties involved. Green plants in the light, rich in stored auxin, vitamins, and carbohydrates, would appear to provide the worst type of experimental material. But even embryos and dark-grown seedlings, though more nearly ideal, and certainly capable of yielding more quantitative data, have their reservoirs of auxin and food factors. It is evident that isolated *parts* of plants, grown in the dark under controlled conditions and free from storage organs, approach nearest to the rigid requirements of experimental material for the study of auxin action. With this in view many studies have been made of the growth of isolated sections, 3 to 5 mm. long, cut from etiolated oat coleoptiles. Such sections will elongate to a small extent in water, and to a much greater extent in auxin solution, the amount of elongation being proportional to auxin concentration up to a maximum at about 10 mg. per liter. Higher concentrations cause reduction in growth and are toxic. This elongation is achieved at the expense of the stored carbohydrates in the cells. If sugar is added to the solution, the elongation is very greatly increased both in amount and in duration. Nitrogenous substances have not been found to increase growth any further, though some salts, particularly potassium chloride, cause a further increase. Optimal concentrations are: sucrose, 1%; potassium chloride, $M/100$; and indoleacetic acid, 1–5 mg. per liter. These isolated sections provide a simple growing system which is as free from indirect controls as can readily be envisaged. Growth is a function of temperature, of the osmotic gradient, and of the concentrations of auxin, potassium ions, and sugar. It is difficult to see how auxin could cause growth here by mobilizing anything.

Mainly with this material it has been shown that growth is controlled by a specific process of respiration. It had been shown many years earlier by Bonner that cyanide inhibits growth over the

same range of concentrations and to the same degree as it inhibits respiration. Recently, however, certain dehydrogenase inhibitors have been found to inhibit growth at concentrations much lower than those which are effective on respiration. Iodoacetate is particularly effective, inhibiting all growth in coleoptile sections at a concentration of 5×10^{-5} M, a concentration which produces only a very slight decrease in respiration rate (2). This growth inhibition is completely removed by adding malate, fumarate, or to a lesser extent succinate or pyruvate. The four-carbon acids, in fact, accelerate growth somewhat over that of controls, in presence of auxin and sugar. Thus a typical respiratory system appears to exercise control over growth. Also, in the presence of malate, and particularly if the sections are soaked beforehand in solutions of malate, the addition of auxin causes a definite small increase in respiration rate. It is of importance also that the auxin concentrations which are effective in stimulating respiration closely parallel those concentrations at which growth is affected. From the fact that growth can be completely inhibited with only a small decrease in respiration it follows that growth is not controlled by the respiration process as a whole, while the influence of iodoacetate and the four-carbon acids suggests that the dehydrogenation of the four-carbon acids is a controlling process in growth. Evidently this system accounts for only a small part of the total respiration of the coleoptile tissue.

Striking support for these conclusions comes from a study of the effect of auxin on the rate of protoplasmic streaming. The long epidermal cells of the coleoptile are particularly favorable for measurements of streaming rate, which, under controlled external conditions, is nearly constant and reproducible. When auxin solution is applied to the coleoptile, there is an immediate rise in the streaming rate, which, however, if no other materials are present, returns to normal after ten to twenty minutes. The presence of sugar allows the increased rate to be maintained for several hours. Oxygen is necessary for the reaction; and if both oxygen and sugar are available there is a good parallel between the concentrations of auxin which accelerate streaming and those which accelerate growth. Since, also, the increase in rate of streaming occurs before any increase in growth rate can be detected, it is reasonable to conclude that the effect on streaming is one of the first stages in the process, caused by auxin, which

results in increased growth. Further examination was therefore made of the relationship between streaming and respiration.

Iodoacetate, in the concentration which inhibits growth, completely inhibits the acceleration of streaming by auxin, and the inhibition, as with respiration and growth, is removed by the addition of malate. Malate alone has no detectable effect on the streaming rate; but in the presence of auxin it has the following very significant effects. After coleoptiles reach a certain age (about five days at 25° C. in weak red light), they no longer show any acceleration, either of growth or of streaming, when auxin is added. But when malate is added, auxin is able to produce its normal acceleration of streaming in this old material. Similarly, if young sections are given prolonged soaking in water they behave like old coleoptiles, and here too the addition of malate restores the sensitivity of streaming to the influence of auxin (4). Apparently, therefore, aging or prolonged soaking in water depletes the coleoptile of its store of malate or other four-carbon acids.

From these observations it can be concluded that there is a special part of the respiratory system whose activation results in growth. This growth is probably brought about in the first instance through acceleration of the rate of protoplasmic streaming. Finally, the system, involving streaming, a part of the respiration, and growth, is mediated by the four-carbon acids, and auxin apparently acts as an essential coenzyme. It should be added that some unpublished experiments with other materials strongly indicate that this growth-controlling system is present in similar form in materials quite other than the coleoptile.

It will be seen that, while a good deal of information is now available about the mechanism of auxin action, and therefore the mechanism of growth, the interrelationships between the various factors and processes are by no means clear. How, for instance, can a process of respiration be linked to one of growth?

It is important to consider first what constitutes "growth" in its simplest form. To a zoologist growth is commonly thought of in terms of cell multiplication. In the oat coleoptile, however, the number of cells reaches its final value at an early age and thereafter growth is by cell elongation only. Such a situation, with modifications, is common in plant tissues. Furthermore, a coleoptile section

"growing" in pure auxin solution decreases in dry weight as does, for instance, a sea-urchin egg dividing in sea water. Thus the essential minimum of growth, in plant tissues, comprises only an irreversible increase in volume. This in turn means nothing more than the uptake of water, or salt solutions. Now we know, principally from the work of the California group under Hoagland, that the uptake of salts from solution is dependent on respiratory processes, and the uptake of water itself may well be a process of the same type. There is some evidence to support this, for auxin has been shown to induce in potato tissue an increased water uptake, which is dependent on the presence of oxygen. The accumulation and concentration of water is thus seen as a "vital" rather than a purely osmotic process. Respiration in some way provides the free energy for the thermodynamic work involved. The process in some ways parallels that of secretion in the kidney. While a relation between a respiratory process and growth or water uptake can thus be dimly envisaged, the role of the protoplasm, and especially of its streaming, remains completely unassigned. Further analysis may well lead to the correlation of many fundamental, and superficially different, types of process.

Nothing as yet has been said about the part played by the cell wall. It is only the primary wall which can stretch to allow visible growth. Once secondary wall has been deposited, as in the older coleoptile, extension ceases. Yet we have seen that at least one of the reactions leading to growth, that of the acceleration of streaming, can be initiated by auxin in presence of malate in this nongrowing tissue. Part of the growth process may thus be completed, although visible increase in size is mechanically prevented. It may become necessary, then, to define growth in terms of its constituent biochemical processes rather than in terms of its external appearance.

A suggestive parallel to this, though it may have little in common with auxin effects, is furnished by the reaction caused by the "wound hormone," traumatic acid, which is an unsaturated acid present in the brei of many tissues, particularly in bean extract. It causes local growth and swelling, being responsible for the so-called wound reaction. This would appear, superficially, to be a very clearly defined response both anatomically and physiologically, since wound-healing tissue usually consists of masses of rapidly dividing cells. Yet in beans it has been shown that, depending on the variety, the response to wound

hormone may consist of typical rapidly dividing cells or may involve principally enlarged (elongated) and undivided cells. No doubt the underlying response is the same in both cases, but the external visible effect may differ. It does not seem at present that there is any specific hormonal stimulus for cell division *per se*. However, the control of cell division is one of the problems which on further study may yield information about the underlying mechanism of the wound reaction and its relation to other forms of growth.

Lastly, there are a series of problems involving the interaction of auxin with other special substances. Biotin and auxin act co-operatively in the formation of roots by certain types of cuttings. Thiamin and pyridoxine are necessary for the elongation of roots, and auxin inhibits this elongation, but whether the availability of these vitamins is inhibited is not known. In fact, the growth physiology of the root is still largely unexplored, despite a great many ingenious experiments (10). Thus it may be concluded that no more than a beginning has been made in the study of plant growth hormones.

References

(1) Burk, D., and Winsler, R. J., *Vitamins and Hormones*, **2**, 306 (1944).
(2) Commoner, B., and Thimann, K. V., *J. Gen. Physiol.*, **24**, 279 (1941).
(2a) Gustafson, F. G., *Botan. Rev.*, **8**, 599 (1942).
(3) Kögl, F., *Ber.*, **A68**, 16 (1935).
(4) Sweeney, B. M., and Thimann, K. V., *J. Gen. Physiol.*, **25**, 841 (1942).
(5) Thimann, K. V., *Plant Physiol.*, **13**, 437 (1938).
(6) Thimann, K. V., *Biol. Rev. Cambridge Phil. Soc.*, **14**, 314 (1939).
(7) Thimann, K. V., and Bonner, J., *Physiol. Revs.*, **18**, 524 (1938).
(8) van Overbeek, J., *Ann. Rev. Biochem.*, **13**, 631 (1944).
(9) Went, F. W., *Am. J. Botany*, **25**, 44 (1938); *Plant Physiol.*, **13**, 55 (1938).
(10) Went, F. W., and Thimann, K. V., *Phytohormones*. Macmillan, New York, 1937.

22

CHEMICAL MECHANISM OF NERVOUS ACTION

DAVID NACHMANSOHN, RESEARCH ASSOCIATE, DEPARTMENT OF NEUROLOGY, COLLEGE OF PHYSICIANS AND SURGEONS, COLUMBIA UNIVERSITY

In view of the dominating role of the brain and of the central nervous system in the body, the problem of the mechanism of nervous action has always attracted general interest. When, from Galvani's experiments it became clear that nerves generate electricity, the news was received with great enthusiasm, not only throughout the scientific world, but also among all educated people. For more than a century the analysis of the electric changes constituted the only means of studying nervous action. And yet, 150 years later, Herbert S. Gasser, one of the leading electrophysiologists of our time, compared the electric spikes to the ticks of the clock (5). Both are only signs of activity in an underlying mechanism: "It follows then that if spikes are but manifestations of activity in the inherent mechanism of nerve fibers, the story of nerve is by no means told when the spikes are described. We need to know something about the mechanism which produces them—how it is maintained, its capacity for work, and when and how the work is paid for." In spite of all the valuable information obtained by the study of the physical aspect, knowledge of the molecular changes, *i. e.*, of the chemical reactions involved, is necessary for the understanding of the mechanism of nerve activity.

The special function of the nervous system, which is that of carrying messages from one distant point of the body to another, may

be subdivided into three successive phases: a stimulus reaching a neuron must initiate an impulse; the impulse, once initiated, must be propagated along the axon; and, finally, it must be transmitted either to a second neuron or to an effector cell. Early in this century, the idea was evolved that a chemical compound may be connected with the third phase, namely, the transmission of the nervous impulse from the nerve ending to the effector cell. T. R. Elliot suggested, in 1905, that adrenalin may be the transmitter of the impulse from the sympathetic nerve ending to the effector cell. He based this idea on the similarity between the action of adrenalin and the effect of stimulation of sympathetic nerves. In 1921, Otto Loewi discovered that, following vagus stimulation of the frog's heart, a compound appeared in the perfusion fluid which, if transmitted to a second heart, produced an effect similar to that of vagus stimulation. Accepting the basic idea of Elliot, Loewi concluded that the compound (which was later identified with acetylcholine) is actually liberated from the nerve ending and acts as a transmitter of the vagus impulse to the heart cell. Loewi's concept of "neurohumoral" transmission was widely accepted among physiologists.

In 1933, Dale tried to extend this idea of a "chemical mediator" of the nerve impulse to the neuromuscular junction and to the ganglionic synapse. His theory was based essentially on the same type of evidence as previously applied by Loewi in the case of autonomic nerves. In this instance, however, the theory encountered strong opposition. Besides many contradictions and difficulties there were two main objections. The first was the time factor. This factor was of lesser importance in the case of the slowly reacting cells innervated by the autonomic nervous system. But the transmission of nerve impulses across neuromuscular junctions and ganglionic synapses occurs within milliseconds. No evidence was available that the chemical process can occur at the high speed required, and Dale and his associates admitted this difficulty. The second objection was still more fundamental. According to leading neurophysiologists like Sherrington, Fulton, Gasser, and Erlanger, the excitable properties of axon and cell body are basically the same. The electric signs of nervous action therefore did not support the assumption that the transmission of the nerve impulse along the axon differs fundamentally from that across the synapse (6).

The concept of a chemical mediator released at the nerve ending and acting directly on the second neuron thus appeared in many respects to be unsatisfactory.

General Approach

Two features of nervous action are essential to an understanding of the problems and difficulties involved, *viz.*, the high speed of the propagation of the impulse and the infinitely small energy required. In medullated mammalian nerve, the impulse travels at a rate of one hundred meters per second and the energy required per impulse per gram nerve is less than one-tenth of a millionth of a small calorie. The recording of such an event offered many difficulties, even with the use of physical methods. A really adequate electrical recording instrument became available only with the introduction of the cathode-ray oscillograph by Gasser and Erlanger. Still more difficult was the detection of the energy involved. It is not surprising that Helmholtz, who first demonstrated heat production in muscle, failed to demonstrate it in nerves. Even A. V. Hill was unable to detect any heat production in nerve for a long time, and only when he and his associates developed thermoelectric methods of an amazingly high degree of perfection did it become possible to measure amounts of heat of as small an order of magnitude as are produced by nerve activity.

If even physical methods encountered so many obstacles, it is obvious that the study of the chemical reactions connected with an event of this kind must offer serious difficulties. No adequate methods are available for determining directly chemical compounds appearing in such infinitely small amounts and for such short periods. But the development of biochemistry, especially during the last twenty years has shown that, in such cases, much information may be obtained by the study of biocatalysts. Nearly all chemical reactions in the living cell are catalyzed by enzymes. Since Buchner's demonstration, in 1897, that fermentation may occur in cell-free extracts, a great number of enzymes has been isolated, and the chemical reactions of living cells have been studied *in vitro*. While many cell constituents and intermediates, especially those closely connected with cell activities, are extremely unstable and occur in concentrations too low for chemical analysis, a great number of enzymes are relatively stable and, under

appropriate conditions, their activity is suitable for investigation. Moreover, study of cellular constituents as such does not indicate which changes they may undergo in the living cell and at what rate. It is true that study of an isolated enzyme alone does not yet permit direct correlation with its cell function, since there are so many simultaneous reactions in the complex system of the living cell. But, by investigating a great number of reactions and series of reactions, and by studying their connection with events in the intact cell, valuable information of the chemical mechanism of cell function may be obtained. One of the most conspicuous examples of such an analysis is the development of muscle physiology. Because of the pioneer work of A. V. Hill and O. Meyerhof, many physical and chemical changes have been correlated; and our concept of the mechanism of muscular contraction has gone through a real "revolution," according to an expression of A. V. Hill, although we are still far from having a complete picture.

The most promising approach to the chemical reactions involved in nerve activity appears, therefore, to be by way of a study of the enzymes involved. Since Loewi's discovery suggested that the release of acetylcholine may be connected with the transmission of the nerve impulse, it could be expected that a study of the enzymes involved in the formation and hydrolysis of the ester might lead to the elucidation of the precise role of acetylcholine.

But even the best methods of enzyme chemistry available are not adequate for the study of many problems. In many cases, in order to find a satisfactory answer, it was necessary to combine enzyme chemistry with the selection of special cases which offered particularly favorable conditions for the different problems.

Role of Acetylcholine

Investigations based on this line of approach and carried out over a period of nearly ten years have provided evidence for a new concept of the role which acetylcholine may have in the mechanism of nervous action (7,11). According to the new concept, the release and the removal of the ester are considered as an *intracellular* process occurring at points along the neuronal surface and directly connected with the nerve action potential.

The action of acetylcholine may be pictured in the following way: The nerve is, according to the generally accepted "membrane

theory," surrounded by a polarized membrane. The polarized state of the membrane is due to a selective permeability to potassium ions which are many times more concentrated inside the axon than outside. During the passage of the impulse the resistance of the membrane is decreased. From experiments on the giant axon of the squid. Cole and Curtis calculated that the resistance falls from 1000 ohms to about 25 ohms per sq. cm. (3). The permeability of the membrane to all ions is increased and a depolarization occurs. This change in permeability may well be produced by the rapid appearance and removal of acetylcholine (11). The depolarized point becomes negative to the adjacent region and a flow of current results which stimulates the next following point. There again acetylcholine is released and the whole process repeated. The impulse is thus propagated along the axon. At the nerve ending the surface is increased and the resistance therefore decreased. This leads to a greater flow of current which enables the impulse to cross the gap. Whereas, in earlier theories, acetylcholine was considered as a "neurohumoral" or "synaptic" transmitter, *i. e.*, a substance released from the nerve ending and acting directly on a second neuron, in the new concept the transmitting agent is always the electric current, the action potential, but the current is generated by acetylcholine.

The picture is consistent with the idea of the propagation of the nerve impulse as developed by Keith Lucas and Adrian. It becomes unnecessary to assume that the transmission along the axon differs fundamentally from that across the synapse. The assumption of a special mechanism at the synapse different from that in the axon was the chief difficulty which had to be overcome for conciliating the original theory with the conclusions of the electrophysiologists. This appeared necessary for any satisfactory answer to the problem. For, if it is true that physical methods alone are unable to explain the mechanism in a living cell, it is equally true that conclusions based on chemical methods should not be in contradiction to those obtained with physical methods, in view of the much higher sensitivity of the latter.

The facts on which the new concept is based have been recently reviewed and discussed (11). Some essential features are based on studies of choline esterase. A few examples may be given in order to illustrate the new approach.

Choline Esterase

Time Factor. One of the essential results of these enzyme studies is the evidence of the high rate of acetylcholine metabolism in nerve tissue. Significant amounts may be split in milliseconds, that is, a period of time sufficient for the passage of the impulse. Consequently, the potential rate of acetylcholine metabolism is sufficiently high to justify the assumption that it parallels the rate of the electric changes and may therefore be directly connected with the nerve action potential.

The frog sartorius muscle is a special case in which this problem of the time factor has been studied and has received a satisfactory answer. A small fraction of this muscle is free of nerve endings. By determining the concentration of choline esterase in this part of the muscle, in the part containing nerve endings, and in the nerve fibers, it is possible to calculate the concentration of choline esterase at the motor end plates. Since the number of end plates in a frog's sartorius is known, the amount of acetylcholine which may be split during one millisecond can be calculated. This turns out to be 1.6×10^9 molecules of the ester. About $^1/_3$ of the enzyme at the motor end plate is localized inside the nerve ending. On the assumption that one molecule of acetylcholine covers about 20 to 50 sq. Å., the amount which may be hydrolyzed during one millisecond at one end plate would cover a surface of 100 to 250 sq. μ.

Localization at the Neuronal Surface. A high concentration of choline esterase, of similar order of magnitude to that at motor end plates, exists at all synapses, whether central or peripheral. But the difference between axon and synaptic region is only quantitative. Fibers with a very thin myelin sheath, until recently considered as nonmyelinated, like the sympathetic chain of mammals or the abdominal chain of lobster, offer a favorable material for demonstrating this distribution. Hence, the rate of acetylcholine metabolism may be high everywhere in nerves.

Experiments on the superior cervical ganglion of cats indicate that the enzyme might be concentrated at or near the neuronal surface. The giant axon of squid, another special case, was used for testing such an assumption. The axoplasm may be extruded and thus separated from the sheath. Most of the sheath is connective tissue to which are

attached two thin membranes only a few microns thick. The whole enzyme activity is in the sheath. The axoplasm is practically free of choline esterase (1).

Bioelectric phenomena occur at the surface. The localization of the enzyme at the neuronal surface and the high rate of acetylcholine metabolism make possible the assumption that the ester is connected with the electric manifestations of nerve activity. But for the interpretation of the actual role of acetylcholine in the mechanism of nervous action the activity of the enzyme had to be connected with an event in the living cell. Such a correlation has been established in experiments carried out on the third and perhaps most valuable special case, *viz.*, the electric organ of fish.

Parallelism between Enzyme Activity and Voltage of the Nerve Action Potential. The electric discharge in these organs is identical in nature with the nerve action potential of ordinary nerves. The only distinction is the arrangement of the nervous elements, the electric plates, in series. The potential difference developed by a single element is about 0.1 v., which is the same order of magnitude as that found in ordinary nerves. In the species with the most powerful electric organ as yet known, *Electrophorus electricus*, the so-called electric eel, several thousand elements are arranged in series from the head to the caudal end of the organ. Thus, the voltage of a discharge amounts to 400–600 v. on the average; and, in some specimens, more than 800 v. have been observed. In *Torpedo*, another species with a powerful electric organ, the elements are arranged in dorsoventral direction. Since it is a flat fish, the number of plates usually does not surpass 400 to 500; and, consequently, the discharge is only 30–60 v. on the average. In the large *Gymnotorpedo occidentalis* found on the North American east coast, especially in the water surrounding Cape Cod, the number of plates in series and, consequently, the voltage may be more than twice as high.

In 1937, the electric tissue was introduced by the writer as material for the study of the role of acetylcholine in the transmission of the nervous impulse. A high concentration of choline esterase was found in the strong electric organs of *Torpedo* and *E. electricus.* These organs hydrolyze in one hour amounts of acetylcholine equivalent to one to five times their own weight. In the larger specimens, the organs have a weight of several kilograms, so that the amount of acetyl-

choline which may be split in these organs may amount to several kilograms per hour or several milligrams in one-thousandth of a second. The concentration is of the same order of magnitude as was calculated for the motor end plate by Marnay and Nachmansohn (10). It is sufficiently high to make possible the assumption that acetylcholine is directly connected with the discharge, which requires that the compound must appear and disappear in milliseconds. If speculation were to be excluded, the only means of removing this compound so rapidly is enzymic action. The high concentration of a specific enzyme appeared particularly significant in view of the chemical constitution of these organs. They contain 92% water and only 2% protein. The discharge is, in these organs, the main function; there is no question here of a transmission of the nervous impulse to a second unit, since there is none.

In the weak electric organ of the common ray, the concentration of choline esterase is relatively low. If in the three species mentioned, voltage and number of plates per centimeter are compared with the concentration of choline esterase, a close relationship becomes obvious.

A more detailed analysis has been carried out of the electric organ of *E. electricus*. This species is particularly favorable for such studies because the number of plates per centimeter, and consequently the voltage per centimeter, decreases from the head to the caudal end of the organ. The choline esterase activity decreases in the same proportion. If the electric changes are recorded and compared with the chemical values at the same section, a close parallelism is obtained between voltage and enzyme concentration. This is found not only in regard to the variations which occur in the same specimen, but even in absolute amounts for the variations between the individuals (12). In measurements covering a range from 0.5–22 v. per cm., the line correlating the physical and chemical event was found to indicate a direct proportionality.

The voltage developed in the discharge depends upon the electromotive force and the resistance. Two assumptions therefore appear possible regarding the way in which acetylcholine may act. It may produce an electromotive force directly by action on the surface, or it may decrease the resistance by increasing the permeability of the boundary. Resistance and electromotive force are closely related properties. The drop in resistance during the passage of the impulse,

found in the giant axon by Cole and Curtis (3) and in the electric tissue by Cox, Coates, and Brown (4), suggests that the parallelism found between voltage and acetylcholine metabolism may be due essentially to the effect of the ester on permeability in the way outlined above.

Specificity of the Enzyme. In all the experiments on the activity of the enzyme, it was assumed that choline esterase is specific for acetylcholine. In such a case, not only is the conclusion justified that the substrate metabolized is acetylcholine, but the activity of a specific enzyme determined *in vitro* may well be used as an indicator for the rate of the substrate occurring *in vivo*. As pointed out by Schoenheimer and Rittenberg (21), one of the most important general results of the work with isotopes is the conclusion that "enzymes do not lie dormant during life but are continuously active." This, of course, does not imply that all enzymes are working at an optimal rate at every moment. In cells like nerve and muscle a considerable difference must be expected between resting condition and a state of activity. Lactic acid formation for instance, occurs in anaerobic condition even in the resting muscle; but during tetanic stimulation the rate increases several thousand times. Similarly, it cannot be expected that an enzyme directly connected with the events during the passage of the impulse is equally active in resting condition. It appears possible, moreover, and even probable, that enzymes are, to some extent at least, present in excess above the optimum usually required. But all experience in enzyme chemistry appears to indicate that a correlation exists between the concentration of an enzyme in a cell and the rate at which the substrate is metabolized.

It thus appeared imperative to demonstrate the specificity of the enzyme. The ester linkage in acetylcholine shows no peculiar properties. It is therefore to be expected that this ester can be hydrolyzed by other esterases and, on the other hand, that choline esterase can hydrolyze other esters. Specificity in this case would be expected, on the basis of analogy, to be only relative and not absolute. Choline esterase might be expected to split acetylcholine at a higher rate than other esters, while other esterases might be expected to behave differently. By testing a number of substrates a pattern has been established which makes it possible to distinguish choline esterase from other esterases (17).

In the great variety of nerve tissue which has been used as basis for establishing the new concept, the enzyme was found to be an esterase specific for acetylcholine, *viz.*, mammalian brain, lobster nerve, squid fiber containing the giant axon, and the electric tissue. All show a similar pattern, typical for choline esterase. In contrast, the hydrolysis patterns of the esterase of other organs—liver, kidney and pancreas—differ greatly from that of choline esterase. The esterase in these tissues shows several variations; but this could be expected, since the physiological substrate is unknown and probably varies in the different organs. Only in muscle (free of nerve endings) was an enzyme obtained whose properties corresponded with those of choline esterase. It is possible that propagation of an impulse in the muscle fiber has the same mechanism as in the nerve fiber. But the presence of choline esterase alone is not sufficient to permit any conclusion.

Of special interest is the pattern obtained with purified choline esterase. The enzyme extracted from the electric organ of *E. electricus* has been purified to such a degree that one milligram of protein splits three thousand milligrams of acetylcholine per hour. The rates of hydrolysis of different substrates are exactly the same as those obtained with freshly homogenized electric tissue. Thus, the enzyme tested in fresh electric tissue is the same as that which is highly purified and the parallelism established between voltage and enzyme activity becomes particularly significant.

Energy Source

The electric organ also offers suitable material for investigating the chemical reactions supplying the energy for the nerve action potential. Both the electrical and the chemical energy released are in the range of possible measurement, whereas, in ordinary nerves, the methods available are not adequate for quantitative analysis. The organ of *Electrophorus electricus*, for reasons discussed elsewhere, is again particularly favorable for such a study.

Measurements carried out on these fish have revealed some facts about the chemical source of energy for the action potential. The electric energy released externally per gram and impulse was found to be 8×10^{-6} gcal. This is the maximum external energy which may be obtained under the condition that the external resistance is approxi-

mately equal to the internal. The total electric energy is, in this case, about six times as high as the external, or about 48×10^{-6} gcal. per impulse per gram electric tissue (4). Under the same conditions and tested simultaneously, the energy released by the breakdown of phosphocreatine was found to be about 32×10^{-6} gcal. per gram and impulse (average of fifteen experiments). Lactic acid formation released about 17×10^{-6} gcal. per gram and impulse, averaging seven experiments (13). The energy of lactic acid formation is probably used to phosphorylate creatine just as in muscle, where phosphopyruvic acid transfers its phosphate via adenosine triphosphate to creatine ("Parnas reaction"). The figures are consistent with the conclusion that energy-rich phosphate bonds are adequate to account for the energy of the action potential. Hence, if the primary alterations of the surface membrane during the passage of the impulse are due to the release of acetylcholine, the figures suggest that phosphate bonds may yield the energy for the synthesis of acetylcholine.

The amounts of acetylcholine actually released during a discharge are not known. But the amount which may be split by one gram of electric tissue during one discharge—about 5×10^{-6} millimole—may be used as an indication. The amount actually released may be smaller, since the enzyme may be present in excess, but the figures indicate the order of magnitude. The amount of phosphocreatine actually split per gram and impulse is about 3×10^{-6} millimole. Thus, the amounts of acetylcholine and phosphocreatine metabolized seem to be of the same order of magnitude. Since, however, one mole of phosphocreatine yields about 10,000 gcal., while the acetylation of choline requires probably not more than 1500 to 2000 gcal., the fate of the remaining energy has yet to be explained (11). There are, of course, several conceivable processes which could account for this difference. For instance, with the splitting of acetylcholine, a simultaneous change of a protein molecule could occur, due to the acid formed, by which the protein is brought close to its isoelectric point. In such a case, the energy required for bringing the protein molecule back to its original condition would be of an order of magnitude similar to that available. But, so far, there is no experimental evidence for this or any other simultaneous reaction.

One of the essential facts supporting the new concept is, as repeatedly emphasized, the extremely high concentration of choline

esterase at the neuronal surface, making possible a rate of acetylcholine metabolism sufficiently high to parallel the electrical changes. In electric tissue, the rate may be at least 100,000 times, but is probably close to 1,000,000 times, as high as that of respiration. We must distinguish, however, between the possible rate and the absolute amounts metabolized. Acetylcholine is released and hydrolyzed within a very short period. The actual duration of one thousand discharges is about three seconds. The recovery may require one to two hours during which the rate of respiration may be increased. If the absolute amounts of acetylcholine possibly metabolized are compared with those of the phosphorylated compounds actually metabolized and the observed rate of respiration, a satisfactory picture is obtained. Since this whole chain of reactions connected with the nerve action potential is initiated by the release of acetylcholine, it has been called the "acetylcholine cycle."

Choline Acetylase

It appeared essential to test whether or not phosphate bonds are really the energy source of acetylcholine formation as these investigations suggest. Evidence for the correctness of this conclusion would show that the energy of the primary recovery process is really used for the resynthesis of the compound which, by its release, supposedly initiates the nerve impulse. It would, therefore, at the same time, constitute a new support for the assumption that the "excitatory disturbance" (Keith Lucas) which produces a propagated impulse may be indeed the release of the ester. Also, we would have another example of the "reconstruction of the chemical events in living cells," to use an expression of Green (8).

In accordance with the assumption made, a new enzyme, choline acetylase, could be extracted from brain which, in cell-free solution, under strictly anaerobic conditions in the presence of adenosine triphosphate, forms acetylcholine (16).

The enzyme may be extracted from homogenized brain. From one gram of fresh rat or guinea pig brain an enzyme solution may be prepared which forms 120 to 150 μg. of acetylcholine per hour. The presence of eserine and fluoride is necessary to inhibit the action of choline esterase and adenosine triphosphatase, respectively. The enzyme may also be extracted from the powder of acetone-dried brain.

In these extracts, the enzyme is about twice as pure as in those obtained from fresh tissue—one gram of protein may form 3 to 4 mg. of acetylcholine in one hour.

Since acetone inactivates choline esterase, this enzyme is largely or sometimes completely inactivated in the extracts prepared from powder of acetone-dried brain, so that addition of eserine may have either a small effect or practically none on the formation of acetylcholine. It is thus demonstrated that the enzyme mechanism responsible for the formation of the ester is not identical with the hydrolyzing enzyme. But this is not surprising in view of the complexity of the ester synthesis, which probably occurs in several steps involving more than one enzyme. Adenosine triphosphatase is also removed in extracts from acetone-dried brain. No addition of fluoride is therefore required.

The enzyme requires the presence of potassium in high concentration, close to that found in brain. It contains active sulfhydryl groups which may be easily oxidized in air and which are readily inactivated by monoiodoacetic acid or copper in low concentrations. On dialysis, the enzyme rapidly loses its activity. Addition of glutamic acid reactivates it partly. Only the naturally occurring *l*(+)-form is effective (15). With potassium and glutamic acid, 50 to 80% of the original activity may be restored. Further addition of cyanide or replacement of glutamic acid by cysteine may reactivate the enzyme nearly completely. *l*(+)-Alanine also has some effect; other amino acids have either a weak effect or none. Citric acid has an effect nearly as strong as glutamic acid, whereas dicarboxylic acids have practically no effect.

The effect of glutamic acid, although weaker than that of cysteine, appears to be of special interest. Cysteine, like glutathione, may enhance the activity of many enzymes containing sulfhydryl groups which have been oxidized during preparation, whereas the effect of glutamic acid cannot be explained by action on the sulfhydryl groups of the enzyme.

Price, Waelsch, and Putnam (19) have observed a favorable effect of glutamic acid on patients suffering from petit mal attacks. The interest in the effect of glutamic acid on choline acetylase is, by this clinical observation, further increased, since a relation between the clinical effect and the enzyme reactivation is easily conceivable.

The oxidation products of amino acids, *i. e.*, α-keto acids (pyruvic, phenylpyruvic, oxyphenylpyruvic, and α-ketoglutaric acids have been tested) have a strong inhibitory effect on the formation of acetylcholine when present in concentrations of 10^{-3} to 10^{-4} molar. These are close to the concentrations which occur in living cells. Pyruvic acid is known to be a "physiological anticonvulsant" (20). The strong inhibitory effect of α-keto acids on the formation of acetylcholine is therefore obviously of physiological as well as of clinical interest.

Action Potential and Inhibition of Choline Esterase

A new and significant relationship has been recently established between enzyme activity and nerve action potential, this time using the peripheral axon as material (Bullock, Nachmansohn, and Rothenberg). If acetylcholine is the depolarizing agent and if the function of choline esterase is to remove the active ester so that polarization again becomes possible after the passage of the impulse, then inhibition of the enzyme should alter and, in sufficiently high concentration, should abolish the nerve action potential. Experiments carried out on the giant axon and on the fin nerve of squid have shown that eserine, known to be a strong inhibitor of choline esterase, alters and finally abolishes the nerve action potential. When the nerves are put back into sea water, they quickly recover and conductivity reappears. The reversibility of the effect is consistent with the fact that the inhibition of choline esterase is easily reversible *in vitro*. Strychnine, another inhibitor of choline esterase, was also found to alter and, in higher concentrations, to abolish the nerve action potential reversibly.

Prostigmine has *in vitro* the same effect as eserine, but it has no effect on the nerve action potential. Eserine is a tertiary amine and may therefore, if undissociated, penetrate the lipoid membrane. Prostigmine is a quaternary ammonium salt and therefore it cannot penetrate the lipoid membrane. This has been demonstrated experimentally: Eserine was found in the axoplasm of nerves kept in a solution containing this compound, while prostigmine was absent in the axoplasm tested under the same conditions. These observations explain why prostigmine and acetylcholine, both quaternary ammonium salts, act externally only on nerve endings which do not have a myelin

sheath. But such compounds are inactive if applied to the axon. The peculiarity of the synapse in reacting to injected acetylcholine can no longer be referred to a difference in the fundamental physicochemical process underlying the propagation of the nerve impulse, but to the difference in histological structure. This may also be the explanation of Claude Bernard's famous curare experiment, since the active principle of this compound is, according to recent investigations, a quaternary ammonium salt.

Differences between Old and New Approaches

The facts described here may suffice to illustrate the new approach and may be used for analysis of a few general aspects of the problem.

Loewi's discovery that a specific compound is released during nerve activity was important and need not be minimized because of the change in the original interpretation in the light of recent developments. Since the adoption of new methods led to different conclusions, the question appears of interest whether or not it is possible, with our present knowledge, to find an explanation for the observations on which the original conclusions and interpretations were based.

In the older theories, it was assumed that acetylcholine is liberated at the nerve ending and that, having crossed the synapse or motor end plate, it acts directly on the effector cell or the second neuron. Evidence for the role of acetylcholine as such a "neurohumoral" or "synaptic" transmitter was considered as satisfactory if the following three effects could be produced: (1) a stimulating effect by injected acetylcholine; (2) appearance of acetylcholine in the perfusion fluid following nerve stimulation; and (3) the enhancing effect of eserine on nerve stimulation.

The first effect is not necessarily physiological, but may well be pharmacological. Other chemical compounds like nicotine or potassium may have a similar effect. This would be, therefore, no conclusive evidence for the assumption that acetylcholine is the mediator in the original sense. Its appearance in the perfusion fluid following nerve stimulation is a physiological event; this observation of Loewi was significant because it connected the choline ester with nerve activity. The mode of action, however, is not hereby explained. Any compound forming part of an intracellular process may easily

appear outside the cell even if it is rapidly metabolized, because all enzymic reactions follow a logarithmic curve. Therefore, a fraction may persist long enough to escape to the outside. Even if the greatest part of the acetylcholine released during the passage of the impulse is split inside the cell, a small fraction may escape hydrolysis and diffuse to the outside. The appearance of acetylcholine in the perfusion fluid does not, therefore, indicate that the compound acts outside the cell.

Such an assumption would find support if evidence could be provided that the amount leaving the cell is of an order of magnitude similar to that which produces a response. This was assumed by Dale and his associates when they attempted to obtain some quantitative estimates for comparing the two amounts. But in the two cases in which such comparisons were made, a wide gap existed between the amounts appearing in the perfusion fluid and those necessary to produce a single response: The acetylcholine found in the perfusion fluid is about 1/40,000 of the amount required for stimulation in the case of the superior cervical ganglion, and in the case of muscle, only 1/100,000. A difference of such an order of magnitude appears rather puzzling. It becomes even more difficult to explain when one considers that even these amounts were collected only if eserine was present. Choline esterase is present in high concentration outside as well as inside the cell; evidence for this distribution has been obtained in experiments on denervated tissue. According to the new concept, the main function of the choline esterase is to inactivate rapidly the ester inside the cell after its function has been completed, and to restore the membrane resistance thereby before passage of the next impulse. The high concentration of the enzyme outside would have the function of protecting the effector cell against those traces which may escape and which, if allowed to accumulate, may interfere with the normal functioning. The enzyme outside would thus act physiologically as a barrier for the acetylcholine escaping intracellular hydrolysis. This barrier would be nullified, at least partly, by eserine. Even in the presence of this drug, the amount collected is small. It is therefore difficult to imagine that acetylcholine would cross the synapse or end plate in a concentration sufficiently high to produce a stimulus if the barrier were fully active. But the small amounts found are entirely consistent with the interpretation proposed here.

In view of the high affinity of eserine for choline esterase, the enhancing effect of this compound on nerve stimulation, the third kind of evidence which was considered as essential, may well be attributed to the inhibition of choline esterase. But here again no conclusion is possible as to the mode of action of acetylcholine. The enzyme outside the cell will probably be inhibited; but a fraction of the inhibitor may enter the cell. In both cases an enhancing effect of nerve stimulation may result. Such an effect does, therefore, not permit any conclusion as to whether acetylcholine acts physiologically inside or outside the cell. There may be, also, no enhancing effect at all, or it may be difficult to obtain, as in the case of the superior cervical ganglion—a fact which appeared so puzzling to Dale.

We now arrive at a fundamental difference between the implications based on methods of enzyme chemistry and those based on pharmacological tests. All facts observed on choline esterase point unequivocally to the same role of acetylcholine in all nerves. All contain the same specific enzyme, whether they belong to invertebrates or vertebrates, to mammals or to fish. All contain a high concentration of choline esterase whether sympathetic or parasympathetic, autonomic or central nervous system, efferent or afferent. This concentration probably varies because the active surface per unit of tissue varies. Other factors may also be important; but the order of magnitude is always the same and sufficient to permit the assumption that the substrate is metabolized at a rate parallel to that of the electric changes, that is, the passage of the impulse. There are other indications. Wherever tested, a coincidence was found between the time of appearance of the high enzyme concentration and the beginning of nervous function during growth. The possibility of applying findings obtained with the electric organ to the mechanism in brain, as illustrated by the discovery of choline acetylase, is another example. So far, no biochemical fact has been found which is not consistent with the conclusion that acetylcholine has the same function in *all* nerves. It is difficult to conceive that such a specific and well-defined mechanism either should have a variety of functions in different nerves or, as the only other alternative, should be active in some, and inactive—although present—in other nerves.

In striking contrast to the uniformity of the enzymic mechanism, we find a great variety of effects if the pharmacological

method is applied. Only in a relatively few cases are the effects comparable. This variety is due to the fact that all pharmacological action depends on a great number of unknown factors. The limitation of an interpretation of drug effects has been well formulated by Clark (2): "Even in the most favorable cases, where quantitative relations have been established for the action of drugs on cells, there probably remain dozens of unknown variables, and there is usually a considerable range of alternative explanations." The contrast between the effects of eserine and prostigmine on the nerve action potential and, more generally speaking, the great variety of effects obtained with drugs which are strong inhibitors of choline esterase is an excellent illustration of the truth of Clark's statement. The affinity for the enzyme is just one factor since the action depends on many others, like permeability, circulation, concentration, interaction with other cell constituents and enzymes, and so on. Changes in the chemical constitution of the molecule are known to influence profoundly the effect of a drug. Eserine acts mainly on the periphery, and strychnine on the central nervous system. No explanation is as yet available for this difference in preference, but it would be a mistake to conclude that there is a difference in the basic mechanism of a cellular function because a drug acts differently in two cases.

Most elaborate techniques were necessary to demonstrate the three effects on which the evidence for the previous concept was based. But if we keep in mind that the response of the cell to the injection of acetylcholine is not distinguishable from other pharmacological effects, that the enhancing effect of eserine is certainly a pharmacological effect, and that the third observation—the appearance of acetylcholine in the perfusate requires the abolition of the physiological barrier by a drug, then it is apparent why the three phenomena are so difficult to demonstrate and so subject to variations under different conditions. In spite of the fact that all biochemical data indicate an identical role of acetylcholine in all nerves, there is still a great deal of discussion of the question of whether or not this or that synapse is "cholinergic" or whether or not acetylcholine is a "synaptic transmitter" in brain because one or two of the necessary requirements cannot be demonstrated. The answer does not appear difficult if Clark's statement and the limitations of the methods used are kept in mind. If we see how little deviation is found throughout the animal kingdom in the basic

mechanism of muscle function, it is difficult to conceive that a great variety of chemical mechanisms are required for the transmission of the nervous impulse.

There is another aspect which requires some comment. Following the discovery of Loewi that an active compound, the *Vagusstoff*, is released if the vagus is stimulated, it was necessary to identify this compound. Once its identity with acetylcholine had been established and its release following nerve stimulation repeatedly demonstrated, the question arose as to what information could be expected from the determination of the ester itself and of its concentration in different nerves or tissues. Acetylcholine is found in almost all tissues, but the amounts are small, a few micrograms per gram or even a fraction of a microgram per gram. As previously discussed, acetylcholine is, in living tissue, a very unstable ester. We do not know what fraction of the total amount present in the living tissue may be extracted with our crude chemical methods, especially in nerve tissue, where acetylcholine may be so rapidly destroyed. Although the same values will be obtained using the same conditions, the meaning appears obscure. A concentration of 0.2 or 0.4 micrograms of acetylcholine per gram of brain is difficult to interpret, if the same amounts are found in liver, pancreas, or lung. In the spleen of ox and horse are found amounts of 4 to 30 micrograms per gram, although there is no indication at all that the ester has any function in this organ. No acetylcholine has been found in the spleen of other animals. Chemical compounds may occur in small amounts everywhere. Acetylcholine has been found in the potato. The presence or absence of traces of it thus appears less significant and not comparable to the finding of a mechanism such as a powerful and specific enzymic system.

The enzyme systems specifically responsible for the formation. and destruction of acetylcholine are, on the other hand, found mainly in the nerve cell. Most other tissues do not contain them. But, besides this localization, the data obtained on the concentration of choline esterase do make possible the correlation with function. The amounts of acetylcholine which may be split per gram of brain per hour vary usually between 50 and 500 milligrams according to species and center, although they are very constant for each species and each center. This means that the amount of acetylcholine which may be split per gram of brain in one millisecond is of the order of magnitude

of 10^{14} to 10^{15} molecules. On the assumption that one molecule of acetylcholine may cover 20 to 50 sq. Å., 10 to 100 millions of square microns of nerve surface may be covered by the amount which may be metabolized in one gram of brain in one millisecond. Even if half of the enzyme were localized outside the cell, the figures are still impressive and suggestive of a relationship between function and quantity found.

It is of paramount importance for the investigation of a problem to know the inherent possibilities of methods, as well as their limitations. Medicine depends on the use of drugs. Since it cannot wait until the mechanism of their action has been explained, drug effects must be tested on intact cells and intact animals. Such *in vivo* tests are necessary because, owing to the great number of unknown factors in *in vivo* studies, the effect of a drug is unpredictable merely on the basis of its affinity for enzymes *in vitro* and of its possible relation to cell mechanisms. But if we realize this, then, for the same reason, it appears inadvisable to draw any conclusions about the physiological mechanism from drug effects. In any case, it seems difficult to use effects observed with drugs as an argument against conclusions based on biochemical data.

It is equally important to know the limitations of enzyme chemistry. As mentioned above, only if a number of facts or a series of reactions and some relationship with events in the intact cell have been established can conclusions as to the mechanism become possible. In muscle, several physically recorded changes observed during contraction could be correlated with chemical reactions. In regard to the mechanism of nerve activity, the observations on the electric fish have permitted the correlation of biochemical data with occurrences observed on the intact animal, and recent experiments on the giant axon of squid have provided direct evidence for the dependence of the nerve action potential on the normal function of choline esterase. Many more are possible even with presently available methods, and it is reasonable to hope that the continuous improvement of the available methods and the development of new chemical and physical methods will open more opportunities for approaching the great number of unsolved problems. There is not yet evidence for the assumption that all chemical reactions supplying energy have already been determined, since the exact amount of heat produced has not yet been correlated

with the chemical reactions measured. Although the heat production observed in ordinary nerves is of the same order of magnitude as the energy evolved during the discharge of the electric fish, direct correlation would be desirable. Another important problem is the nature of the physicochemical events occurring at the neuronal surface which change the permeability. At present we do not even know in which way acetylcholine is released.

It is needless to enumerate all the problems which still remain unsolved. A great deal of information may still be obtained from further studies which are being carried out on the enzyme systems involved. But the use of more methods will be required until a satisfactory picture is obtained.

Needham (18) stated that "science is the study of the quantitative relationships in the world we live in." Pasteur called Lavoisier the founder of modern chemistry on the ground that he has introduced in chemistry the use of the balance, *i. e.*, the notion of quantity. But since living cells are amazingly complex systems, the interpretation even of quantitative data requires caution. A great teacher, Sir Frederick Gowland-Hopkins (9), has warned us: "All dogmatic teaching about any aspect of the phenomena of life is apt to be checked by the ultimate discovery that the living cell is before all things a heretic." For the ultimate goal, which is that of explaining the mechanism of living cells in terms of physics and chemistry, research based on the combination of a great number of devices will be necessary. Thus far, enzyme chemistry has proved itself repeatedly to be one of the most powerful tools in transforming our approach from pure description of the phenomena of the living cell into scientific analysis. The history of the problem of the mechanism of nervous action adds another illustration of this development.

References

(1) Boell, E. F., and Nachmansohn, D., *Science*, **92,** 513 (1940).

(2) Clark, A. J., *The Mode of Action of Drugs on Cells.* Arnold, London, 1933.

(3) Cole, K. S., and Curtis, H. T., *J. Gen. Physiol.*, **22,** 649 (1939).

(4) Cox, R. T., Coates, C. W., and Brown, U. V., *J. Gen. Physiol.*, **28,** 187 (1945).

(5) Erlanger, J., and Gasser, H. S., *Electric Signs of Nervous Activity.* Univ. Pennsylvania Press, Philadelphia, 1937.
(6) Fulton, J. F., *Physiology of the Nervous System.* Oxford Univ. Press, New York, 1938. 2nd ed. rev., 1943.
(7) Fulton, J. F., and Nachmansohn, D., *Science*, **97**, 569 (1943).
(8) Green, D. E., in *Perspectives in Biochemistry.* Cambridge Univ. Press, London, 1937.
(9) Hopkins, F. G., *Skand. Arch. Physiol.*, **49**, 33 (1926).
(10) Marnay, A., and Nachmansohn, D., *J. Physiol.*, **92**, 37 (1938).
(11) Nachmansohn, D., in Harris, R. S., and Thimann, K. V., *Vitamins and Hormones*, **3**, 337 (1945).
(12) Nachmansohn, D., Cox, R. T., Coates, C. W., and Machado, A. L., *J. Neurophysiol.*, **5**, 499 (1942).
(13) Nachmansohn, D., Cox, R. T., Coates, C. W., and Machado, A. L., *J. Neurophysiol.*, **6**, 383 (1943).
(14) Nachmansohn, D., and John, H. G., *Proc. Soc. Exptl. Biol. Med.*, **57**, 361 (1944); *J. Biol. Chem.*, **158**, 157 (1945).
(15) Nachmansohn, D., John, H. M., and Waelsch, H. T., *Biol. Chem.*, **150**, 485 (1943).
(16) Nachmansohn, D., and Machado, A. L., *J. Neurophysiol.*, **6**, 397 (1943).
(17) Nachmansohn, D., and Rothenberg, M. A., *Science*, **87**, 158 (1944); *J. Biol. Chem.*, **158**, 653 (1945).
(18) Needham, J., *The Sceptical Biologist.* Ghatto & Windus, London, 1929; Norton, New York, 1930.
(19) Price, J. C., Waelsch, H., and Putnam, T. J., *J. Am. Med. Assoc.*, **122**, 1153 (1943).
(20) Putnam, T. J., and Merritt, H. H., *Arch. Neurol. Psychiat.*, **45**, 505 (1941).
(21) Schoenheimer, R., and Rittenberg, D., *Physiol. Revs.*, **20**, 218 (1940).

SOME ASPECTS OF BIOCHEMICAL ANTAGONISM

D. W. WOOLLEY, ASSOCIATE, THE ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH, NEW YORK; THE LILLY AWARD IN BACTERIOLOGY; MEAD JOHNSON AWARD IN NUTRITION

DURING the past five years, a series of observations has been reported which show that compounds having specific antagonistic action to vitamins and to certain other metabolites can be synthesized or obtained from nature. The uses to which these materials may and have been put in the study of biochemistry, as well as their promise in the field of pharmacology, have made it seem desirable to winnow the data thus far obtained and to plant in a prominent place whatever viable seeds may appear.

The agents antagonistic to metabolites may be divided, for purposes of the present discussion, into two groups. The first contains those inhibitory compounds which have structures analogous to the metabolites in question. The second is composed of specific proteins which react with certain metabolites in such a manner as to render them biologically inactive in the systems in which they are studied. The first of these two groups has been examined much more intensively, and at the moment appears to bear more immediate promise than the latter.

The work of Woods with *p*-aminobenzoic acid was the first to cause widespread interest in compounds related structurally to vita-

mins but which were biochemically antagonistic to them. This work was not the first* to demonstrate such antagonism, but it was the first in which the relationship was seen in a light sufficiently bright to cast a shadow ahead. Woods observed that the bacteriostatic action of the sulfonamides was reversed competitively by *p*-aminobenzoic acid, the structural analogue of sulfanilamide. The chemical relationship here was that sulfanilamide was *p*-aminobenzoic acid with a sulfonamide group instead of a carboxyl group. The hypothesis was advanced that the sulfonamides owed their action in inhibiting growth of bacteria to their competition with *p*-aminobenzoic acid in an essential metabolic reaction. This postulate gained some foundation in fact when it was shown that *p*-aminobenzoic acid occurred in yeast and other living forms, and, more especially, that it was an essential growth factor for several species of bacteria.

The discovery of the relationship between *p*-aminobenzoic acid and the sulfonamides prompted the application of similar types of structural change to other vitamins in efforts to produce from these vitamins bacteriostatic compounds. It was soon found that 3-pyridinesulfonic acid and its amide would inhibit the growth of certain bacteria in a manner subject to reversal by nicotinic acid. Likewise, thiopanic acid [pantoyltaurine, *N*-(α,γ-dihydroxy-β,β-dimethylbutyryl)-taurine] acted competitively with pantothenic acid to produce bacteriostasis, and several α-aminosulfonic acids competed with α-aminocarboxylic acids. (The structural formulas of some of these and succeeding compounds will be found on pages 371–373.) Thus it began to appear that the principle of structural analogy was the basis of a general means of producing bacteriostatic compounds.

The next advance was made when it was found that some types of structural analogues of various vitamins would cause the appearance of characteristic signs of vitamin deficiency diseases in animals, and that these signs could be cured or prevented by adequate doses of the vitamin involved. There is an excitement and appeal about a spectacular experiment with animals which is never quite equaled by simi-

* The pioneer observation of Quastel and co-workers on the reversible inhibition of the oxidation of succinate by malonate, and the finding of Woolley *et al.* on the toxic effect of 3-pyridinesulfonic acid in nicotinic-acid-deficient dogs antedated Woods' work; but these two investigations, especially the latter, lacked sufficient appeal and interest to stimulate further searches.

lar demonstrations in lower forms. Perhaps we feel closer kinship with the higher phyla, but more probably we sense the wider range of observation possible in tests with mammals. Be that as it may, Woolley and White reported that the feeding of minute amounts of pyrithiamin to mice caused the appearance of typical signs of thiamin deficiency in these animals.

When mice are fed a thiamin-deficient diet, they do not exhibit most of the characteristic signs which accompany thiamin deprivation in certain other species. The fact that pyrithiamin evoked most of these signs may be of use in the study of manifestations of deficiency in varied species. This idea is strengthened by the observations that antagonistic analogues of some of the other vitamins likewise produce typical signs of deficiency in species in which such signs have not been seen previously.

The disease was prevented or cured by sufficient amounts of thiamin. Pyrithiamin is the analogue of thiamin in which the thiazole ring is replaced by a pyridine ring, or, more specifically, the sulfur atom is replaced by —CH═CH—. It was then shown that glucoascorbic acid, a structural analogue of ascorbic acid, produced a scurvylike disease of rats, mice, and guinea pigs, and that, in guinea pigs, the disease was prevented by adequate amounts of ascorbic acid. Furthermore, signs of riboflavin deficiency were produced in rats by feeding isoriboflavin, and in mice by feeding 2,4-dinitro-7,8-dimethyl-10-ribityl-5,10-dihydrophenazine, the phenazine analogue of riboflavin. Moreover, manifestations of nicotinic acid deficiency were brought about in mice by feeding 3-acetylpyridine, and prevented with nicotinic acid. Shortly before this, it was realized that the signs precipitated by 3,3′-methylenebis-[4-hydroxycoumarin] were those of vitamin K deficiency and were preventable by the vitamin to which the coumarin bears structural analogy. Finally, several results of tocopherol and vitamin K deficiencies were produced in mice by feeding tocopherol quinone.

Interspersed with these observations on animals there have been a long series of findings of similar nature on the reversible inhibition of microbial growth. It is not the purpose of this essay to catalogue the rather impressive number of individual findings in this regard, but rather to attempt to discover what general principles may lie beneath them. Therefore, it will suffice to recount that 3-pyridinesulfonic acid

inhibited growth competitively with nicotinic acid, thiopanic acid with pantothenic acid, pyrithiamin with thiamin, benzimidazole with adenine or guanine, 2,4-diamino-7,8-dimethyl-10-ribityl-5,10-dihydrophenazine and 6,7-dichloro-9-ribitylisoalloxazine with riboflavin, and desthiobiotin with biotin. Many of these examples, along with others, will be brought into the discussion later to illustrate general deductions. In every instance, the inhibitory action of the analogue was negated by increasing the amount of the metabolite in the basal medium. Therefore it could be said that the action of the analogues was in some degree due to the production of deficiencies of the metabolites. In this respect, however, the bacterial examples lacked some of the force of the animal demonstrations since, in the latter cases, it was possible to examine the specific anatomical and physiological changes characteristic of various vitamin deficiencies which were produced by the agents.

One feature of the action of antagonistic structural analogues of metabolites on a wide variety of living things, both animal and microbial, has been the marked species specificity shown by some of these agents. With several of the compounds, inhibitory action was manifest only against the species for which the metabolite concerned was an essential growth factor. Against those species which synthesized their own supply of the metabolite in question, the analogue was ineffective. Thus, thiopanic acid was able to inhibit only those bacteria that required pantothenic acid, and was ineffective in prevention of growth of species not needing this vitamin. A somewhat similar situation obtained with 3-pyridinesulfonic acid in relation to propagation of organisms requiring nicotinic acid. With pyrithiamin, the dependence of inhibition on the type of thiamin requirement was highly developed. Species which demanded the intact vitamin for growth were inhibited by minute amounts of pyrithiamin; those which were content with the pyrimidine portion alone of the vitamin could be prevented from multiplying only by ten times the amount which was effective with the former organisms; and those species which used both pyrimidine and thiazole portions of thiamin needed a further tenfold increase in the concentration of analogue. Those species capable of good growth in the absence of thiamin or of its component parts were quite resistant to any action of pyrithiamin, and were able to thrive in media containing a million times the con-

centration which was effective against thiamin-requiring microorganisms. Another type of manifestation of the correlation of action with requirement for the metabolite preformed in the medium was seen with phenyl pantothenone acting on bacteria and yeasts. This compound is the one obtained by exchanging $-CO-C_6H_5$ for the COOH of pantothenic acid. It inhibited the growth of all microbial species tested irrespective of their pantothenic acid requirements. However, its effect was reversed by pantothenic acid only in the case of those organisms which needed the vitamin. For the other species, pantothenic acid was without effect in reversing the action of the agent. An analogous situation has been recorded in animals, where glucoascorbic acid caused scurvylike conditions in rats and mice, species not requiring dietary ascorbic acid, and in guinea pigs, animals that do demand a dietary source of the vitamin. Ascorbic acid failed to cure the condition in the former species, but was quite active in this respect for the latter.

Thus far, only one instance of the correlation of effectiveness with requirement for the related vitamin is known in the animal world. 3-Pyridinesulfonic acid exhibited no action on mice, a species not needing nicotinic acid in the diet, while it was quite effective in killing nicotinic-acid-deficient dogs, and harmless in equal doses in normal dogs.

In contrast to these correlations, several of the metabolite analogues have been found to be effective against most types of microbial growth regardless of whether or not the related metabolite was a nutritive essential. This was true of sulfanilamide, the analogue of *p*-aminobenzoic acid, and of benzimidazole, the analogue of the purines.

These correlations of action with requirement pose an interesting problem. Why is it that an organism which obtains its vitamins externally is susceptible to the antagonistic analogue while the one which produces its metabolite internally is usually resistant? In other words, why does the ability to synthesize the metabolite make the organism resistant to the deficiency disease caused by the analogue? In only a few instances can this be said to result from formation of the metabolite in amounts sufficient to negate the inhibition. A very small beginning has been made on this problem with pyrithiamin. It has been shown that the types of bacteria which were resistant to pyri-

thiamin—and these were the ones which synthesized thiamin—contained an enzyme system which actively split pyrithiamin to yield the pyrimidine portion, and, presumably, the pyridine fragment of the molecule.

Although it might seem that this pyrithiamin-splitting enzyme is identical with that which synthesizes thiamin, this is probably not so. A pyrithiamin-resistant strain of yeast developed in the laboratory by subculture in the presence of the agent required thiamin as a growth factor. Therefore it is possible to obtain an organism which can split pyrithiamin and yet cannot synthesize thiamin.

In this case, then, resistance in the thiamin-synthesizing forms was correlated with the presence of a mechanism for destroying the antagonist. The possession or lack of such a system for destruction of the inhibitory analogue cannot be the universal answer to the problem posed, because it has been shown that in the case of 3-acetylpyridine, an analogue of nicotinic acid, the resistant organisms were not able to destroy the substance.

Another facet of the problem is the contrast between the action of some metabolite analogues in microbial species and in animals. Thus, thiopanic acid and the sulfonamides caused deficiencies of pantothenic acid and of *p*-aminobenzoic acid in bacteria, but were innocuous in this respect in animals. On the other hand, 3-acetylpyridine caused nicotinic acid deficiency in mice but not in bacteria. Another nicotinic acid analogue, 3-pyridinesulfonic acid, was effective in certain bacteria but not in mice. It may be that these phylal differences represent nothing more than variation in powers of excretion, or it may be that they mirror various degrees of metabolism.

One metabolite may be capable of participation in several reactions, some of which may function in a given type of organism but be absent from the metabolic network of others. If an inhibitor interfered with the action of the metabolite in just one process, it would be expected that it would be effective in only those species which utilized this particular function of the metabolite.

At about the same time that the facts discussed above were being established, a second general type of biochemical antagonism was being brought to light. This latter type dealt with proteins that were antagonistic to vitamins by virtue of the ability of these proteins to

inactivate the metabolites. The first such protein to be studied was ascorbic acid oxidase, an enzyme which is present in several plant tissues and which catalyzed the oxidation, and hence the inactivation, of ascorbic acid. The next example was the curious protein of egg white, variously known as avidin or antibiotin, which combined stoichiometrically with biotin, and thus rendered it ineffective in many biological systems. This case is noteworthy, since it involves a non-enzymic process in which a specific protein is one reactant. Finally, there was the enzyme system in many aquatic animals which actively destroyed thiamin by splitting from it the thiazole portion of the molecule. The suspicion has existed in several laboratories that these proteins have roles to play in the ordered progress of metabolism, and that the avidity with which they attack such vital metabolites may represent only a perversion of their true functions. However, there are no data to indicate their biological significance, and until more is learned about them, they cannot enter further into the discussion of biochemical antagonisms.

With this cursory scanning of the facts in the field of biochemical antagonism, let us look about in an attempt to discover whatever light may be shed on theoretical and on practical problems of biological science. Particularly, let us consider some newly discernible aspects of the study of metabolic mechanisms, and of the relationship of chemical structure to biological activity. It will be apparent as we progress that there are very considerable gaps in our present scanty stock of knowledge, and that in many cases these breaches are of such magnitude that, when they are eventually closed with solid knowledge, the final picture which will emerge may differ rather considerably from the one we are about to examine. With a field as new as this one is, there are sure to be more questions than answers, and the answers available are more in the nature of hypotheses than of facts. Nevertheless, let us integrate present knowledge so that we may better be able to advance in the future.

The first new aspect deals with the use of inhibitory structural analogues of metabolites in the study of biochemical reactions. In this aspect, particularly, hope exceeds realization at the present time, for only local successes have been achieved. The use of specific inhibition of a given reaction has been very serviceable in unraveling the threads of many changes which occur. For example, fluoride

and cyanide have figured prominently in the discovery and differentiation of certain enzymic processes. It has seemed probable that, since the inhibitory structural analogues bring about specifically deficiencies of various metabolites with which they act competitively, these analogues might be highly selective inhibitors of an enzyme system involving the structurally related metabolite. This has proved to be so in the case of *o*-aminobenzylmethylthiazolium chloride, a structural analogue of thiamin. This compound was an effective inhibitor of the enzyme found in certain marine animals which cleaves the thiazole moiety from the vitamin. Likewise, glucoascorbic acid inhibited the action of ascorbic acid in the oxidation of tyrosine by guinea pig liver. The oxidation of tyrosine by liver was apparently a process which involved ascorbic acid, and in this process glucoascorbic acid and its related metabolite behaved competitively. Furthermore, the action of cozymase, a derivative of nicotinic acid, was inhibited by 3-pyridinesulfonic acid. This effect of the analogue, however, was not specific, since several other substances of dubious structural relationship to nicotinic acid gave a similar result. The competitive inhibition by malonate of oxidative systems involving the four-carbon dicarboxylic acids was one of the first instances of biochemical antagonism to be explored, and antedated by several years the observations on sulfanilamide and *p*-aminobenzoic acid. Finally, utilization of pantothenic acid by certain bacteria, a reaction which seemed to be enzymic, was inhibited reversibly by thiopanic acid. Here, the antagonist was of value in differentiating the reaction which involved pantothenic acid from glycolysis. Several inhibitors of glycolysis were also inhibitors of the utilization of pantothenic acid. Thiopanic acid, however, prevented the latter without influencing the former, and thus indicated that the two processes were separable.

The mechanism of action of various drugs is largely unknown. It may be worth while in our gropings for the desired explanations to consider some biological effects produced by benzimidazole, an agent which causes in animals loss of muscular tone and of ability to respond to stimuli. Also, it prevents the growth of many microorganisms. Its growth-inhibiting action is reversed competitively by adenine, which, in the form of adenylic acid and of adenosine triphosphate, is believed to be of considerable importance in muscular contraction. Because of these facts, it would seem that the pharmaco-

logical action of benzimidazole might be an expression of its structural relation to adenine. However, since adenine did not reverse the effect of benzimidazole on animals, more concrete evidence will be necessary before the action of the drug can be explained. Eventually, the *modus operandi* of some drugs may be viewed merely as the result of the production of a deficiency of a structurally related metabolite at the site of action of the drug.* Some drugs now under investigation or in use can be pictured as structural analogues of certain metabolites. Furthermore, the effects produced are not unlike those which might be expected to result from the removal of the analogous metabolite from the field of metabolic action. Nevertheless, it must be remembered that these are speculations and that they demand experimental proof before they can be accepted.

Further conjectures along this line may be applied in consideration of antagonisms known to exist among various pairs of metabolites. For example, it has been shown that androgens and estrogens may occur together in both sexes, and that pharmacologically they exhibit certain antagonisms. Perhaps these are the result of the structural relationship between these hormones.

Although we may think that antagonisms between metabolites and their structural analogues popped into the world with the advent of the nutritionist and the organic chemist, this is probably not so. For example, *Chromobacterium iodinum* has been found to produce a pigment which is bacteriostatic and is a structural relative of some naturally occuring anthraquinones, and more remotely of vitamin K. These anthraquinones, as well as vitamin K, reverse the action of the antibiotic pigment. Similarly, 3,3′-methylenebis-[4-hydroxycoumarin], a structural analogue which elicits some of the signs of vitamin K deficiency, was first discovered in spoiled sweet clover hay.

At this point it is well to contemplate some hypotheses regarding the mechanism of action of the sulfonamides in bacteriostasis, and these same hypotheses as they are applied to inhibitory analogues in general. After the discovery of the reversal of sulfonamide bacteriostasis by *p*-aminobenzoic acid, the hypothesis was advanced that the sulfonamides competed with *p*-aminobenzoic acid for a place on

* It is, of course, not expected that all types of drug action will find explanation in biochemical antagonism. It is merely suggested that some pharmacological agents may owe their biological potency to this phenomenon.

the receptor portion of an enzyme. The enzyme was supposed to require *p*-aminobenzoic acid as a coenzyme, and the introduction of a sulfonamide was supposed to result in a firm combination of the latter with that portion of the enzyme to which *p*-aminobenzoic acid normally would be attached. The resultant foiling of the enzyme was said to be overcome by an increase in the concentration of *p*-aminobenzoic acid which would shift the equilibrium in favor of the *p*-aminobenzoic acid–enzyme combination. Despite the fact that there has been no direct experimental proof of an enzyme system involving *p*-aminobenzoic acid, this attractive hypothesis has continued to flourish with but slight opposition. Its chief buttresses have been the competitive nature of the relationship between *p*-aminobenzoic acid and the sulfonamides, the success realized in application to other metabolites of the same idea of structurally related inhibitors, and the lack of a better hypothesis. A similar explanation has been given for the behavior of other inhibitory analogues. Since some of the metabolites concerned with these antagonists are known to function as coenzymes in well-defined enzyme systems it would seem advisable to put the hypothesis in these instances to experimental test. There are some shreds of evidence that are exceedingly difficult to fit into the original postulate. For example, it has been noted with the sulfonamides, with pyrithiamin, and with benzimidazole that subinhibitory quantities of the compounds actually stimulated the growth of microorganisms. If the coenzyme replacement hypothesis is correct, it is difficult to understand this stimulation. Furthermore, when resting bacterial cells were treated with sulfonamides or with thiopanic acid, no *p*-aminobenzoic acid or pantothenic acid was liberated. This observation, however, does not overthrow the hypothesis because the amounts of metabolite liberated (either *p*-aminobenzoic acid or pantothenic acid) may have been below the detectable amount, or may have been retained inside the cells.

Much has been written about whether the antagonism with *p*-aminobenzoic acid explains the action of the sulfonamides in the production of bacteriostasis and in the cure of certain infectious diseases. Whatever the true explanation of the action may be, the following must not be forgotten. First, *p*-aminobenzoic acid reverses competitively the effect of the sulfonamides, and second, *p*-aminobenzoic acid is a metabolite for microorganisms. In the light of these

facts, it is reasonable to assume that the sulfonamides produce a crippling deficiency of *p*-aminobenzoic acid in the organisms. This view is strengthened when one recalls that structural analogues of certain other vitamins and hormones call forth in animals certain signs associated with specific deficiency diseases. It cannot be said that such antagonisms in any way lend proof to the hypotheses which have been advanced to explain the action. Nevertheless, it is not advisable to allow arguments about mechanism to obscure the well-established facts in the case.

Finally, to end our consideration of the first aspect, let us note examples of some biochemical reactions which have been uncovered by means of microorganisms rendered resistant to the action of an inhibitory analogue. By long-continued culture of a bacterial or fungal species in the presence of gradually increasing concentrations of an inhibitor it is possible to derive a strain which is resistant, or fast, to the agent. By use of a strain of *Endomyces vernalis* made fast to pyrithiamin, it was possible to show that the acquired resistance was correlated with the appearance of a system which cleaved the inhibitor into its component pyrimidine and pyridine portions. This demonstration prompted examination of naturally resistant species, that is, those which synthesized thiamin. These forms were found to possess the pyrithiamin-destroying system. Without the use of the "fast" variant, the recognition of this biochemical reaction would have been more difficult.

The second new aspect to open before us is the possibility of producing new types of drugs by application of the knowledge that pharmacological signs of a predictable nature may be produced by metabolite analogues. The cataloguing of the various specific histological and biochemical manifestations of deficiency of each of the vitamins and hormones is progressing rapidly. It may be desirable to evoke one or more of these signs. If this should be so, then the synthesis and trial of an inhibitory analogue of the metabolite in question would be indicated. Already it has been seen that pharmacological manifestations predictable at least in part can be called forth by the proper metabolite analogue. In this connection, a retrospective glance at the case of 3,3′-methylenebis-[4-hydroxycoumarin], may be profitable. This agent produces signs similar to those seen in vitamin K deficiency, and these manifestations form the basis of

studies carried out with it. The compound was not discovered because of its structural similarity to vitamin K—only after its pharmacological properties were recognized was its structural similarity to vitamin K seen, and its reversal by the vitamin demonstrated. With this glance backward to fortify us, let us look further in this direction.

The study of α-tocopherol quinone represents the next step along the path. Tocopherol deficiency, at least in the first generation of rats or mice, manifests itself in females, to a large degree, by resorption of embryos during gestation. Other signs are generally not seen. Therefore, it was thought that the development of a successful antagonistic analogue of tocopherol would serve as an example of what might be done along lines of application to pharmacology of studies in biochemical antagonisms. In due course it was demonstrated that α-tocopherol quinone in sufficient doses caused hemorrhage and resorption of the embryos of pregnant mice.

Despite the fact that results similar to those of tocopherol deficiency were achieved by administration of tocopherol quinone, tocopherol even in large doses did not reverse these signs. It was of considerable interest to find that vitamin K in very small quantities would prevent the vaginal hemorrhages which immediately preceded resorption. When these hemorrhages were eliminated by vitamin K, the pharmacological effectiveness of the quinone was markedly reduced. It was then realized that tocopherol quinone is a structural analogue of vitamin K as well as of tocopherol. The action of the quinone, however, differed from that of such an antivitamin K as 3,3′-methylenebis-[4-hydroxycoumarin] in its selective action on pregnant animals and in the fact that the coumarin did not produce similar signs.

When similar quantities of α-tocopherol quinone were given to nonpregnant mice, no detectable signs of disease were observed. Although these steps have been made along the path under consideration, the final one has not been taken—the production of a therapeutically useful agent.

One of the features of the action of tocopherol quinone should be discussed since it illustrates a general principle in the use of inhibitory analogues related structurally to metabolites. Almost without exception the effective dose of an antagonist is very large compared with the amount of metabolite involved. In other words, the ratio of inhibitor

to metabolite is generally large.* Therefore, the required dose of an inhibitor depends on two things: the intrinsic activity of the metabolite, and the ratio between metabolite and inhibitor which is necessary for reversal. This latter ratio may vary widely among species for the same inhibitor–metabolite pair. Now, tocopherol is a relatively inactive vitamin as vitamins go. The effective dose in a rat is about three milligrams as compared with a small per cent of this amount for several other vitamins. Hence it was foreseen that very large doses of tocopherol quinone would probably be required to produce the effect; and this proved to be the case.

Antagonistic analogues of hormones may eventually find practical application. It is now believed that certain diseases result from overproduction, or diminished rate of destruction, of various hormones. A possible method of treating such diseases may be by administering an inhibitory analogue of the substance concerned. The analogue would, in effect, remove the excessive amounts of the hormone. A few such inhibitory analogues have already been produced, but the hormones to which they are related have not yet been implicated as causative agents of disease as a result of overproduction in the organism. It has therefore not been possible to test this intriguing postulate.

Because it had its beginnings in investigations of the action of sulfonamides, work on inhibitory compounds related structurally to metabolites has had a strong flavor of bacteriostasis. This has been increased by the fact that a most fruitful method of determining whether a given analogue behaves antagonistically with its related metabolite has been to make observations on the effect on growth of microbial species. Most of the early attention was directed toward the development of chemotherapeutic agents which would be effective against infectious diseases. McIlwain and Hawking were able to show that this method of approach held promise when they demonstrated that thiopanic acid could prevent experimental infection of rats with hemolytic streptococci. However, the amounts of the agent required

* Thus far, only in the cases of benzimidazole and of 3,3′-methylenebis-[4-hydroxycoumarin] has the ratio been 1:1 or less. In these two exceptions the quantity of inhibitor required for an effect was large, but once an adequate level of the agent was established the ratio between inhibitor and metabolite was small.

were so large as to make its clinical application impractical. It is possible that useful agents against infectious diseases may be produced by investigation of inhibitory analogues, but at present this has not been done. Meanwhile, it is well to remember that anti-infection agents form only a part of the chemotherapeutic arsenal, and that magic bullets against noninfectious diseases as well as against infections may eventually be cast in the mold before us.

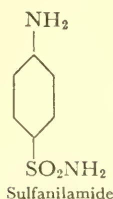
The third new aspect to be considered is an outgrowth of the previous two general aspects. If it is possible to evoke the various types of manifestations discussed, then it is well to examine the classes of structural changes that must be made in a metabolite molecule in order that it may exhibit inhibitory action. At the present time, two general types of structural change which bring about this result can be seen. In addition, there is a third group of structural alterations about which it is not yet possible to make generalizations.

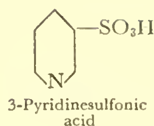
The first general type of structural change involves the replacement of the carboxyl group of acidic metabolites by some other group. The most frequently studied exchange has been that of sulfonic acid or amide for the carboxyl. Examples of this class are: sulfanilamide and its derivatives which are related to *p*-aminobenzoic acid; 3-pyridinesulfonic acid, related to nicotinic acid; the α-aminosulfonic acids, related to the α-aminocarboxylic acids; and thiopanic acid, related to pantothenic acid. With the possible exception of 3-pyridinesulfonic acid, members of this class of inhibitors are not effective in the production of deficiencies in animals. Furthermore, it is from this class that many useful agents against infectious diseases, that is, the sulfonamides and thiopanic acid, have come. Possibly the inactivity toward animals contributes to their chemotherapeutic effectiveness, for it may result in damage to the invading microorganisms without causing undue violence to the host.

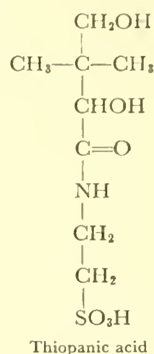
Inhibitory analogues have also been produced by replacement of —COOH by —COR, but the class of compounds so derived has not been studied as intensively as that just examined. Thus substitution of the —COOH of *p*-aminobenzoic acid by —$COCH_3$ (to yield *p*-aminoacetophenone) led to the formation of a bacteriostatic substance whose action was reversible by *p*-aminobenzoic acid. A related ketone protected mice against experimental infections. Similarly the exchange of —COOH in nicotinic acid for —$COCH_3$ to

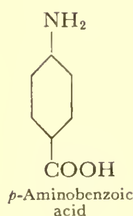
Compounds of General Type I

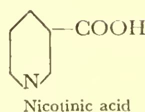
Class A

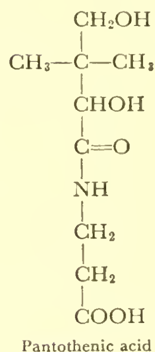

Sulfanilamide


3-Pyridinesulfonic acid

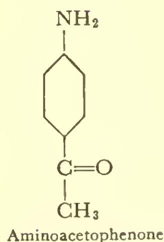

Thiopanic acid

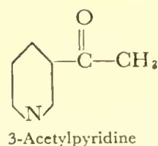

p-Aminobenzoic acid

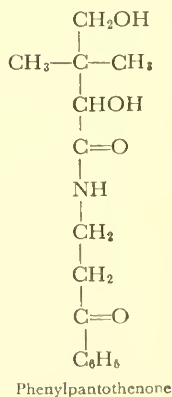

Nicotinic acid


Pantothenic acid

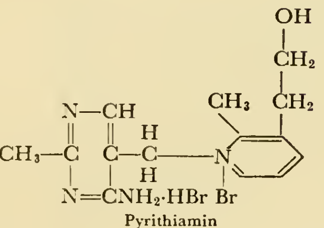
Class B


Aminoacetophenone

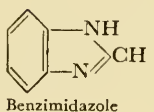

3-Acetylpyridine


Phenylpantothenone

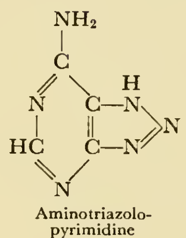
Compounds of General Type II



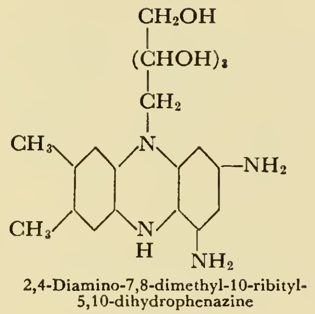
Pyrithiamin



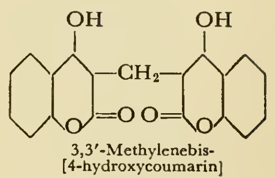
Benzimidazole



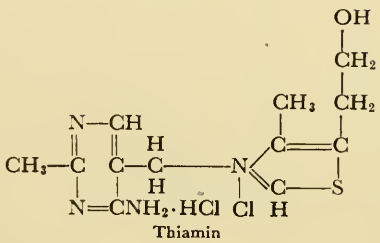
Aminotriazolo-pyrimidine



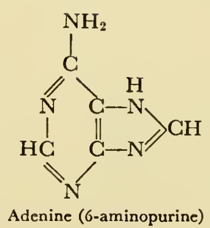
2,4-Diamino-7,8-dimethyl-10-ribityl-5,10-dihydrophenazine



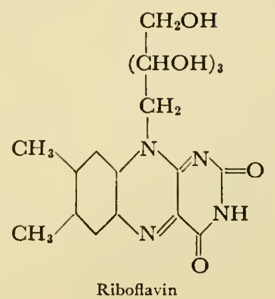
3,3′-Methylenebis-[4-hydroxycoumarin]



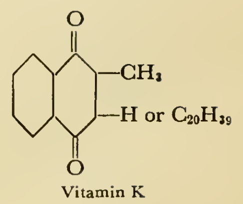
Thiamin



Adenine (6-aminopurine)

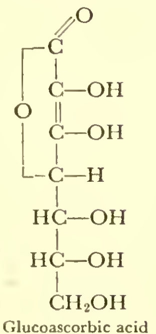


Riboflavin

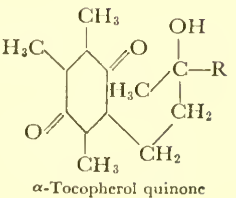


Vitamin K

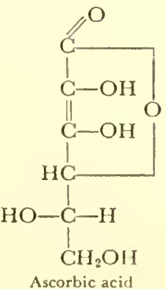
Compounds of General Type III



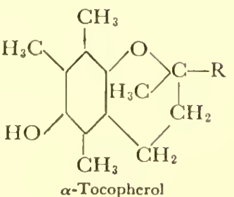
Glucoascorbic acid



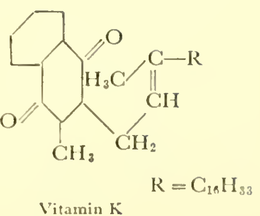
α-Tocopherol quinone



Ascorbic acid



α-Tocopherol



$R = C_{16}H_{33}$

Vitamin K

yield 3-acetylpyridine gave rise to a substance which caused nicotinic acid deficiency in animals but not in microorganisms. Furthermore, substitution of $—CO—C_6H_5$ for the $—COOH$ of pantothenic acid gave rise to phenyl pantothenone which caused pantothenic acid deficiency. The introduction of $—COCH_3$ rather than $—COC_6H_5$ did not yield a compound that produced unequivocal pantothenic acid deficiency, perhaps because a sufficiently negative ketone was not exchanged for the carboxyl. In the cases of nicotinic acid and *p*-aminobenzoic acid, the carboxyl is attached to an aromatic type of nucleus, and therefore the related ketone is more acidic than if this nucleus were aliphatic. On the other hand, in pantothenic acid the carboxyl is aliphatic and may require replacement by an aromatic ketone group.

The second general method for production of inhibitory compounds consists of an exchange of one or more atoms in a ring system. Since so many substances of biological importance are ring compounds the method is rather inviting and has proved to be quite fruitful. The first representative of the series to be studied was pyrithiamin, in which the group $—CH{=}CH—$ replaced the sulfur atom of thiamin. Other examples are: benzimidazole, in which carbon atoms in a benzene ring replace the nitrogen atoms of the structurally related purines; 2,4-diamino-7,8-dimethyl-10-ribityl-5,10-dihydrophenazine, similarly related to riboflavin; triazolopyrimidines, in which a nitrogen atom replaces a carbon atom of the imidazole ring of the purines; benzoureidovaleric acid, in which a benzene ring replaces the thiophane ring of biotin; *o*-aminobenzylmethylthiazolium chloride, in which a benzene nucleus is substituted for the pyrimidine nucleus of thiamin; 3,3′-methylenebis-[4-hydroxycoumarin], derived from vitamin K by exchange of a carbon for an oxygen atom along with changes in the side chains; and iodinin in which two nitrogen atoms have replaced carbon atoms in the ring system of the related anthra- and naphthoquinones. Here, as with the first general type of structural modification, the method has worked apparently in every case in which it has been tried. It will really be remarkable if this record continues to stand in the face of more varied application.

The third, or miscellaneous, group of types of structural change is the scrap heap in which all atypical cases must be collected until a general principle can be seen, and rescued from among them. To

this group belong glucoascorbic acid, α-tocopherol quinone, and desthiobiotin, to mention only a few prominent members.

At the present very early stage of our knowledge, there are many questions demanding answers; and for many of these adequate ones are not yet available. Let us examine two. From the examples now at hand it appears that the exchange of carbon atoms for nitrogen in six-membered rings leads to the production of inhibitory compounds. Witness the effectiveness of conversion of pyrimidine derivatives to benzene compounds. It would be well to know whether the process is unilateral, that is, whether the exchange of nitrogen for carbon would be equally as successful as the reverse process. From the cases of iodinin, and of benzimidazole and triazolopyrimidine (see next paragraph), and from the fact that 6-aminonicotinic acid behaves competitively with *p*-aminobenzoic acid, it would seem that the process actually is not unilateral.

Take now the question of how great the change in structure must be to attain maximal activity in the inhibitory analogue. Is it preferable to make as small a change as possible, and if so, what is the measure of the size of the change which is made? If one examines a case in which progressive change has been made it would seem that a small alteration is better than a larger one, but this cannot be stated with any assurance because of the lack of a standard of magnitude of change. Benzimidazole is derived from the purines, adenine and guanine, by elimination of the nuclear substituents, and by exchange of the two pyrimidine nitrogen atoms for carbon. It is less active in the inhibition of bacterial growth than are the triazolopyrimidines, which differ from the purines only in that the imidazole carbon of the latter has been replaced by a nitrogen atom. Thus, the exchange of one carbon atom for a nitrogen gives a more potent compound than was obtained by trading two nitrogen for two carbon atoms.

Of course, the two exchanges are not strictly comparable because in one case the imidazole ring is altered and in the other the pyrimidine part is transformed. It will be necessary to examine the activity of pyridinoimidazole and its amino and hydroxy derivatives before unequivocal conclusions may be made about the matter. A more unambiguous illustration deals with derivatives of *p*-aminobenzoic acid. Replacement of one carbon atom with nitrogen yields 6-aminonicotinic acid which is antagonistic, while exchange of two carbon for two nitrogen atoms to form 2-amino-5-carboxypyrimidine gives

rise to an analogue with no inhibitory powers but rather with weak *p*-aminobenzoic acid action.

Attendant on this increase in activity is an increase in specificity. Benzimidazole is reversed in its action by either adenine or guanine. By contrast, hydroxyaminotriazolopyrimidine competes only with guanine, and aminotriazolopyrimidine interferes only with adenine. However, this increased specificity may not always be desirable, for benzimidazole has pharmacological powers not possessed by the triazolopyrimidines. Furthermore, the effects of the side chains are of interest. In the case of benzimidazole it was thought that the introduction of an amino group in the proper position would make the compound more analogous to adenine, and hence more potent; but the introduction of this side chain did not materially alter the potency of the analogue. In the case of 3,3′-methylenebis-[4-hydroxycoumarin] it would be of interest to learn the effect of the side chains on biological action, that is, whether the introduction of a phytyl group in position 3 (which would make the coumarin more analogous to vitamin K) would increase its potency or modify its action.

It must not be assumed that it is only necessary to alter the structure of a metabolite indiscriminately in order to achieve an antagonistic agent. Much testing of compounds related structurally to biologically active compounds has indicated that most of these derivatives are inactive under the conditions of test.

On the other hand, there is no unique manner in which the structure must be altered in order to produce antagonistic agents. For example, 2,4-diamino-7,8-dimethyl-10-ribityl-5,10-dihydrophenazine, 6,7-dichloro-9-ribitylisoalloxazine, and 5,6-dimethyl-9-ribitylisoalloxazine, three analogues of riboflavin, all cause riboflavin deficiency in various species even though the type of alteration in structure represented by the three compounds is fundamentally different. Likewise, the structure of *p*-aminobenzoic acid may be altered by changes either in the nucleus or in the nuclear substituents, with resultant production of substances competing with *p*-aminobenzoic acid in bacterial growth. It is of interest to note that only a few of these derivatives are active against infections. Similarly, 3-pyridinesulfonic acid and 3-acetylpyridine represent two inhibitory compounds related to nicotinic acid. Here too, as discussed previously, there are very

marked differences qualitatively in the effects produced by the two agents.

In conclusion it should be said that this essay is intended as a guidepost set beside a new and dim trail to point the direction until the road becomes a much traveled highway. Then a more pretentious marker may be raised and some of the errors of direction and many of the convolutions of the trail may be obliterated. From the work which has been done it seems that a study of antagonisms may be a real current in biochemistry, and not merely an eddy.

Finally, since the present attempt was meant more as an essay and less as a review, no literature has been cited. The reader is referred to references (2) and (7) for documentation of some of the points discussed.

References

(1) Fildes, P., "The mechanism of the anti-bacterial action of mercury," *Brit. J. Exptl. Path.*, **21,** 67 (1940).

(2) McIlwain, H., "Theoretical aspects of bacterial chemotherapy," *Biol. Rev. Cambridge Phil. Soc.*, **19,** 135 (1944).

(3) McIlwain, H., and Hughes, D. W., "Biochemical characterisation of the actions of chemotherapeutic agents. II. A reaction of haemolytic streptococci, involving pantothenate-usage, inhibited by pantoyltaurine, and associated with carbohydrate metabolism," *Biochem. J.*, **38,** 187 (1944).

(4) "Modes of drug action," a symposium in *Trans. Faraday Soc.*, **39,** 319 (1943).

(5) Quastel, J. H., and Woolridge, W. R., "Some properties of the dehydrogenating enzymes of bacteria," *Biochem. J.*, **22,** 689 (1928).

(6) Woods, D. D., "The relation of *p*-aminobenzoic acid to the mechanism of the action of sulphonamide," *Brit. J. Exptl. Path.*, **21,** 74 (1940).

(7) Woolley, D. W., "Some new aspects of the relationship of chemical structure to biological activity," *Science*, **100,** 579 (1944).

(8) Woolley, D. W., "Development of resistance to pyrithiamine in yeast and some observations on its nature," *Proc. Soc. Exptl. Biol. Med.*, **55,** 179 (1944).

(9) Woolley, D. W., "Some biological effects produced by α-tocopherol quinone," *J. Biol. Chem.*, **159,** 59 (1945).

24

CHEMOTHERAPY: APPLIED CYTOCHEMISTRY

ROLLIN D. HOTCHKISS,* ASSOCIATE, THE ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH

THE FUNDAMENTAL procedure for determining the properties of any object is to place that object in a series of new environments and to look for changes produced thereby in it or its surroundings. This procedure is a part of the innate behavior of primitive man—whether he be ancient aborigine or modern infant—and is used, for example, in his examination of a new bauble. The object is picked up, rubbed with the hand to feel its texture, tested perhaps in the mouth, and withal scrutinized minutely in various lights from various angles. A mentally competent primitive soon builds up an enormous "library" of impressions of texture and light-reflecting properties, etc., of the objects of his world, and uses these to guide his activity. The steady accumulation—and, let us be sure to add, the sharing—of these impressions has in time developed large specialized bodies of scientific "knowledge." One of these, more or less arbitrarily defined, may be called cytochemistry. Its basic technique is "looking at" the cells of living matter with such

* On leave of absence, Lieutenant Commander in the Hospital Corps of the U. S. Naval Reserve, Hospital of the Rockefeller Institute.

The views and opinions expressed here are those of the writer, and are not to be construed as reflecting the official views or opinions of the Navy Department.

agents as light or electron beams, electric currents, and especially chemicals or enzymes, and noting changes in the cell or the environment caused by whatever interaction takes place.

Chemotherapy must have had its humble inception in self-medication, the observation of the effect of natural herbs, plants, juices, etc., upon the sick organism. During the course of time this growing branch of medicine has come to make extensive use of the knowledge accumulated by the pharmacologist, and of the materials isolated from nature or synthesized by the chemist. In its present phase, chemotherapeutic medicine is beginning to draw spasmodically upon the pure science of biological chemistry. Despite the commercial significance of chemotherapy, we are hardly justified in speaking of a science of "theoretical chemotherapeutics." We shall entertain here the somewhat limited view that chemotherapeutic medicine is an applied cytochemistry. In so doing, we shall be able to look into the future of what are believed to be some current trends, and to consider some of the challenges that face the workers, and their teachers, in this important field.

Perhaps the bright future of cytochemistry may be the better envisioned if a comparison is made between this science and its more established predecessor, organic chemistry. During the last century, the organic chemist has been systematizing his experiences with the effect of chemical reagents upon various complex molecules. He has learned to detect atom groups which react together as units, to look upon one of these groups as the unifying backbone or nucleus of the molecule, and to locate the position of other groups relative to this one. With the recent confirmation of many of his deductions by physical methods, skepticism as to the general validity of his indirect methods appears to have vanished. In a quite analogous way, the cytochemist is beginning to study the way in which whole molecules and layers of molecules are distributed in the living cell. It may be said to be his task to be aware of the constituents of cells revealed by chemical, physical, metabolic, immunological, pharmacological, or nutritional analysis, and to learn to recognize these constituents, and the evidences of their functioning, in the intact cell. Cytochemistry is indeed the area of common interest which eventually will unite biochemistry with biology in general.

Since most cells are large enough to reveal detail in light, or

electron microscopy, much has been done with stains and fixatives, and we are able to assign generic names like protein, fat, carbohydrate, and nucleoprotein to some cell structures. But enzymes, coenzymes, metabolites, antigens, toxins, and such constituents of molecular size are recognized by functional, rather than visual, attributes. Where any means of identifying one of them in the anatomy of the cell is available, it is likely to be an indirect method subject to the same question and criticism as were the methods of early organic chemical science. A few specific constituents of functional interest, like desoxyribosenucleic acid, ribosenucleic acid, phosphatase, glycogen, cellulose, heparin, etc., can now be recognized by staining methods when they are localized in large amounts in some structure of the cell. Nevertheless, we seldom know both the function, in chemical terms, and the site, of any cell constituent. And even when we do know the chemical potentialities of some unseen constituent, we rarely are able to say when in the life of the cell it is acting and when its activity is latent.

For information about the cell we can look to the stains, fixatives, antiseptics, hypnotics, anesthetics, insecticides, hormones, metabolic poisons, growth factors, etc. In short, we will eventually draw upon every situation in which a substance of known constitution produces observable modifications in cells or cell activities. It is quite natural that cytochemical reagents should all be, in varying degrees, toxic substances. Before they can claim interest as therapeutic agents, however, they must prove to be of less than general toxicity, and capable of administration in a form not too easily hydrolyzed, destroyed, diffused away, precipitated out, excreted, or otherwise removed from the site of activity in the body. The pharmacist and pharmacologist have become so skilled in controlling and modifying these last aspects of drug action that, in a sense, we shall be justified in concentrating our attention upon the first-mentioned factor, cytochemical specificity.

Ideally, chemotherapeutic agents should have maximal effects upon a chosen variety of cells (a parasitic foreign agent, or a diseased, neoplastic, hyperactive, or otherwise abnormal host tissue) and a minimal effect upon other tissue cells of the organism. For this to be possible it is necessary to capitalize upon structural differences between the cells at one of three levels of organization, the morphological, the colloidal, or the chemical level. This is now necessarily

done blindly, and it is possible that it may always have to be done empirically. For, on the one hand, although many visible structural differences have been described between different cells, these can hardly ever be described in terms of function, chemical reactivity or vulnerability. On the other hand, it is safe to say that those few chemical differences which have been detected probably reside in the proteins, polysaccharides, or other macromolecules of the cells. Almost the only agents with specificity great enough to show *qualitatively* different behavior with different cells are the antibody and enzyme proteins or organized agents like the viruses. Less complex chemical agents sometimes behave differently in a *quantitative* sense in staining or penetrating diverse cells, but even here we seldom find the desired situation, *viz.*, that the reactivity of one type of cell far overshadows that of any other type.

Chemical substances of low molecular weight are, however, not devoid of specific affinities and are even capable of "biological" discrimination. Enzymes are enzymes only if there are corresponding substrates, and these latter are in many cases well-defined simple substances. If the enzyme is credited with specificity in its action upon substrates, it is perhaps only fair to ascribe specificity to the substrate when it seeks out an active spot on the enzyme surface and enters there into a transitory enzyme–substrate complex. Statistically speaking, it is the rapidly moving substrate molecule and not the sluggish protein which does most of the hunting. However, we must remember that there is virtually no evidence to indicate how many "false complexes" the substrate enters into, at other sites on the enzyme or other proteins, which do not result in chemical change, and therefore pass unnoticed. We cannot say that a key which fits many locks is specific merely because most of the locks it fits will not work.

With antibody proteins, too, simple substances have a discriminatory ability to inhibit the combination with, and precipitation of, the specific antigens. Indeed large parts of the precipitating capacity of the antibody proteins can be directed toward relatively simple associations of chemical groups contained in the huge antigen particle. Nevertheless, it is just when we turn to the smaller molecules and to the phenomena of "specific inhibition" which they show that we begin to find the broadest overlapping of reactivity shown by distantly related structures. For example, an antibody can be sometimes more

interested in whether a group is on the *ortho*, *meta*, or *para* position in a simple benzene derivative than whether that group is an amino, a methyl, chloro, hydroxy, or nitro group.

As a third illustration of specificity shown by simple substances let us consider crystallization from solution. In general, this process involves a truly remarkably specific selection of particles of a single molecular species, occurring as rapidly as diffusion processes will permit. The molecule of appropriate configuration finds a place where it precisely fits, and kinetic energy of motion becomes transformed into potential energy of position. Symmetrical molecules which can "sit down" in several ways (*e. g.*, para-disubstituted benzene compounds) find this easier to do and crystallize more readily than unsymmetrical molecules; and their crystals are less soluble and higher in melting point. Ions which have affinity for each other are able to react if by reacting they may crystallize out as an insoluble compound.

It may be worth some consideration that the demonstrations of specificity in simple compounds just cited all involve as the specific act the immobilization and orientation of the substance upon a surface, of either a protein molecule or a crystal. We must recall that the energy content of small molecules requires them to vibrate and move about aimlessly in an activity far more rapid and extensive than that of the protein molecules and crystal particles, which move with a sluggish and restricted motion more like the Brownian movement. We may picture the larger particle as furnishing a definite mold or pattern into which the small particle in its bouncing, careening progress makes countless attempts to fit. When, by chance, there is close approach at appropriate orientation to some matching area on the surface, a specific adsorptive act occurs, the duration and outcome of which depend upon intrinsic properties of the molecules. It may well be promptly reversed and thus pass unnoticed in the confusion of molecular traffic, or it may result in enzymic action, or inhibition thereof, in precipitation or its inhibition, etc. We know biochemical events to be replete with interactions of this general nature, and it seems possible that specificity chiefly depends upon the existence of anchoring areas on the surfaces of sluggishly moving large molecules and cells. To learn the characteristics of their specificity we need to know as much as possible about the way they are affected by all kinds

of substances, whether or not these latter are clinically useful drugs. We must expect to find that, in intact cells, where diffusion rates, dielectric constants, dissociation constants, and the like may be very different from what they are in ordinary aqueous solutions, the specificities of biochemical systems also will not be the same as in the test tube. The organizational factors which control the various phases of chemical activity in the cell probably grade continuously from slight "crowding" effects of molecule upon molecule which slightly alter diffusion and reaction rates, through major crowding which in effect makes the environment only semiaqueous, and finally all the way to those special arrangements which amount to actual contiguity, or separation, in space, of reactive elements.

We have attempted to portray as concretely as we are able a general picture of the initial process of drug–protoplasm interaction. It is, more or less, a simple expression in present-day physical terms of views on the mode of drug action which go back at least to 1900. Beginning about that time and throughout the years since, various theories have been put forward, such as those relating narcosis and surface action, associated with the names of Traube, Lillie, and Warburg, among others. Many such hypotheses relating one or another special kind of biological properties to surface behavior have also appeared during this period. Like all old hypotheses, they are old because they have not easily been disproved and discarded, and they have remained hypotheses because they were propounded far in advance of experimental basis. The most inclusive theory is probably that of Ehrlich, who accounted for the phenomena of immunology and pharmacology in terms of combining side chains. Thus, drugs had organotropic or parasitotropic "haptophore" groups able to combine with "chemoreceptor" groups in tissues or parasites; and when so anchored the drugs exerted their effect. To present-day workers this theory has often appeared extravagantly endowed with unsubstantial detail; it is, however, probable that Ehrlich's concepts have remained implicit in the pharmacological thinking of the past few decades, while his terminology has been almost entirely replaced by specific chemical names for reactive groups.

By far the most significant special theory along these lines is one which accounts for the adsorption of the agent, and also, in a somewhat explicit fashion, for the effect which this has upon the cell. In

1940, Fildes postulated that antibacterial substances in general would be found to interfere with essential enzymic reactions of the bacterial cell, either by combining with important substrates ("essential metabolites"), or by competing with them for combination with the enzyme. We shall see in what follows that a number of mechanisms other than "competitive inhibition" can give rise to antibacterial action, but the importance of Fildes' postulation is that it pointed the way clearly for a large amount of experimental work, including the production of many new substances of biological, and probably eventually therapeutic, activity. In Fildes' illustration—toxicity due to mercuric ions—there was no actual demonstration of the identity or even of the thiol nature of the natural metabolite. But an accompanying paper of Woods contained a classic demonstration of competition between the natural *p*-aminobenzoic acid and the artificial drug *p*-aminobenzenesulfonamide and its derivatives. Since that time, others have furnished many more examples of interference with mammalian as well as bacterial physiological processes, caused by substances related chemically to natural metabolites and vitamins.

Since much has been written elsewhere and there are discussions in this volume bearing upon competitive inhibition, let us content ourselves with two or three points relating to our chosen theme. First, let us contrast the rationale applied by Ehrlich and others in synthesizing new agents prior to 1940 with a rationale based upon the Fildes-Woods hypothesis. It was customary previously to begin with a toxic principle, *e. g.*, arsenic, mercury, or a dye, and to build up around it various organic structures and substituent groups in the hope of developing a selectivity toward a particular type of cell. The literature of the last few decades is full of reported attempts to "reduce the toxicity" of such agents by appropriate modifications in structure. On the other hand, the warped metabolite molecules now being synthesized are likely to contain a special arrangement of relatively common chemical groups, which is at the same time the basis for a rather subtle toxicity and for whatever specificity the substance shows. For example, pantoyltaurine, an inhibitory analogue of pantothenic acid, is an amide derivative formed by joining together two natural products. Other synthetic inhibitors are homologues of natural metabolites, which presumably combine with the specific enzymes for the very reason that their structure is not exotic. The spectacular recent

successes in chemotherapy have come through such substances as thiouracil, stilbestrol, sulfonamides, and penicillin, all more or less specific agents which are by no means merely carriers of potent, widely active groups. These particular substances were probably first hit upon somewhat accidentally, but the approach through metabolite inhibition would seem to provide a means of arriving at such combinations rationally.

Second, let us turn to some considerations which may suggest directions to take in the synthesis of future chemotherapeutic agents. One of the main desiderata, we have agreed, is selectivity for a given type of cell. Comparative biochemistry seems to reveal that different cells are very much alike in the way they break down and oxidize substrates; so agents which inhibit catabolic enzymes are likely to have broad activity. On the other hand, since different cells appear to have developed somewhat different proteins and other macromolecules in their protoplasm, it is to be hoped that agents which can inhibit synthetic processes will display greater cell specificity. Furthermore, in designing modifications in chemical structure which will deceive the enzyme proteins, it should be helpful to look at some of the data from immunochemistry. In his book, *The Specificity of Serological Reactions*, Landsteiner describes a large number of modifications in chemical structure which greatly alter the reactivity of an antigen toward its antibody, and in addition, many examples of overlapping specificity, that is, of different chemical groups which to a considerable extent look alike to an antibody protein. These relationships, arrived at by trial and error after innumerable painstaking syntheses, may to some degree decrease the number of errors per trial in the synthesis of pharmaceuticals.

The isolation and characterization a few years ago of two crystalline polypeptides derived from tyrothricin, an antibacterial product extracted from cultures of *Bacillus brevis*, seemed to offer an opportunity of studying some of the factors which are involved in combating bacterial infections. One of these antibacterial polypeptides is basic, the other is neutral. Both appear to have molecular weights of about 2500 and contain enough aromatic and higher aliphatic amino acids to be alcohol-soluble and rather insoluble in water. Both substances markedly depress the surface tension of aqueous solutions. A number of the amino acid residues in each substance

have the usual *l*-configuration but, surprisingly, a considerable fraction of the residues have the "unnatural" and rare *d*-configuration. Whatever their similarities, the basic peptide, tyrocidine, is essentially inert against bacteria *in vivo*, while the neutral peptide, gramicidin, is moderately successful in eliminating certain localized infections. This contrast in activity between apparently similar substances suggested that a study of their mode of action might reveal some significant principles.

It was soon learned that, in the test tube, tyrocidine is in some ways more potent than the other substance; small amounts kill bacteria of many different species, while gramicidin affects only Gram-positive bacteria, principally by inhibiting their growth rather than by killing them. But tyrocidine, like many other potent antiseptics, is too unspecific in its reactions to be effective *in vivo*. It can be demonstrated to precipitate with certain protein anions, and in the presence of protein its action upon bacteria is diminished. It seems probable that tyrocidine, by combining with tissue proteins, becomes at once both harmful to the animal host and almost inert toward the infecting parasite.

Does the milder, more selective, gramicidin also combine with protein, but merely more specifically, so that it combines with, and inhibits, say, some particular enzyme? Neither gramicidin nor tyrocidine has much effect upon ordinary hydrolytic enzymes, but when they are tested upon the respiratory and fermentative systems of whole bacteria the situation is somewhat different. Tyrocidine added in lethal quantities to resting bacteria promptly depresses the respiration to a small fraction of the original rate. Gramicidin seldom shows direct inhibitory action; indeed, when environmental conditions are favorable, this agent commonly stimulates susceptible bacteria to respire at two or three times the normal rate. Yet, with staphylococci, the final end products, acetic acid and carbon dioxide, are produced in the same proportion and eventual yield whether or not gramicidin is present. However, the bacteria are not interested in preparing acetic acid (it may even accumulate sufficiently to hamper their activity sometimes) or carbon dioxide as such: as far as we know today their best reasons for oxidizing glucose are (a) to liberate its energy into useful form and (b) to produce intermediates like pyruvic acid, which instead of being further oxidized may serve as raw mate-

rials for synthesis of protoplasmic constituents in growth. We have some evidence that it is with one of these energy-utilizing functions that gramicidin interferes.

The useful form of energy derived from such processes very commonly appears to be the phosphate ester. An organic substance phosphorylated at the proper place is thereby activated for oxidation or other conversion in much the same way that a wooden match tipped with a phosphorus composition is prepared for a function in which a plain wooden splinter would be inert. Since bacteria, when growing, are performing a great many complicated chemical reactions, it is very important that we are sometimes able to study the first beginnings of synthesis in the phosphorylations and accumulation of intermediates that accompany respiration in washed, nonproliferating bacterial suspensions. In such preparations of staphylococci, normal oxidation of substrates results under certain conditions in the accumulation inside the cells of: (a) phosphate, at the expense of external inorganic phosphate; and (b) an unidentified phosphorylated carbohydrate ester. In the absence of nutrients, these events appear to represent the most that can be accomplished toward growth. But both of these events are virtually abolished by gramicidin in the same small concentrations that suffice to prevent bacterial growth in a more satisfactory environment. In short, gramicidin does not prevent staphylococci from consuming substrate, but appears to prevent them from deriving benefit from the act. It is tempting to conclude that this effect accounts for the growth-inhibitory properties of gramicidin. There is some evidence in favor of this. The prevention by gramicidin of inorganic phosphate uptake, at least, is remarkably clear-cut and requires but small amounts of the agent. Under various conditions with different susceptible and resistant strains, different species, in altered environments, and with and without gramicidin antagonists, this effect parallels the ability to prevent growth.

As yet, it is not possible to suggest which specific enzyme is being inhibited in this effect of gramicidin, although it appears quite likely that the phenomena observed can result from the blocking of a single enzyme. Naturally, therefore, it is not possible to judge whether gramicidin inhibits because it resembles the natural substrate, but it appears improbable that a polypeptide-like substrate would be involved in these reactions. So we may tentatively look upon the action

of gramicidin as an example of noncompetitive inhibition of some enzyme involved in a phosphorylating system. The effect can be studied in cell-free systems, since muscle and kidney extracts are prevented by this same agent from phosphorylating glucose, although the oxidation which normally furnishes energy for the phosphorylation continues unaffected. It is to be expected that future work with cell-free preparations will reveal exactly which of the energy-utilizing systems is being blocked, and at the same time provide a specific inhibitor for use in the study of this important function in the biochemical laboratory.

Tyrocidine is found to contrast sharply with gramicidin in mechanism of action. Bacteria exposed to it respire at such a low rate that syntheses and growth are not possible, if only for the reason that energy is not made available rapidly enough. Yet it is noteworthy that respiration, easily reduced to 5–10% of its normal rate, appears to resist further reduction even when a tenfold excess of agent is applied. This fact suggests that there is no single key enzyme being inhibited by tyrocidine, but either that an organization of enzymes is being rendered more inefficient or that an alternate, inefficient, metabolic pathway is being utilized which is less susceptible. At this point let us recall that tyrocidine, with its basic group and alcohol-soluble side chains, has a pronounced surface activity. Like the natural saponins and bile salts, and the synthetic detergents, wetting agents, etc., it can bring about hemolysis of red blood cells. If it were to produce an analogous injury to bacteria, we should have, as with the hemoglobin of erythrocytes, various soluble cellular metabolites liberated from the cells into the surrounding medium. There is unmistakable evidence that exactly this sort of damage is produced by tyrocidine; and, in exposures at different time, temperature, or concentration, it happens in just those cases in which the bacteria are killed. Appropriate analyses show that five-minute exposure to about one-hundredth their weight of tyrocidine at zero degrees centigrade is sufficient to extract quantitatively from staphylococci the total trichloroacetic-acid-soluble phosphorus and nitrogen compounds present in the cell. Analyses have been made for such typical constituents as inorganic phosphate, adenosine triphosphate, other phosphate esters, total esters, total nitrogen, amino nitrogen, amino acids, and pentoses. Analogous effects have been observed with other microorganisms. So

we conclude that tyrocidine kills bacteria because it brings about a cytolytic injury to their cell membranes.

Some thirty antiseptic substances, including mercury derivatives, halogens, oxidizing agents, formaldehyde, dyes, etc., were able to kill or completely repress growth of staphylococci without releasing cell constituents. The increase in cellular permeability, which is often supposed to occur generally with death, was not noted at all, and certainly cannot, for these cells at least, be confused with the clear-cut, almost instantaneous, physical change that accompanies killing by tyrocidine.

However, the further analysis of the bactericidal effect is of interest to us as cytochemists. It is not hard to explain a drop in the respiratory rate in these dead cells. With the seepage of coenzymes, activating ions, and intermediates out of the cells, there occurs an enormous "dilution" of the multicomponent respiratory systems. It is well known that dilution of a reacting system decreases its rate more or less according to an exponential function of the degree of dilution, in which the exponent is the number of reacting components. Therefore dilution of, say, a three-component system, can be tremendously "inhibitory" in its effect. It is actually found that the residual metabolism is more dependent upon the volume into which this dilution takes place than upon the concentration of cytolyzing agent. Accordingly, it is unnecessary to suppose that enzymes are inactivated by the minute amounts of tyrocidine needed to kill bacteria.

Hydrolytic enzymes such as the peptidases, phosphatases, and nucleosidases, however, are less affected by dilution than are respiratory enzymes, since the hydrolytic systems contain fewer components, and the concentration of one of these, water, is scarcely changed by dilution. In harmony with this, bacteria cytolyzed by tyrocidine begin to undergo enzymic post-mortem changes which result in the liberation of soluble degradation products, and with some species, but not universally, may progress to actual bacteriolysis, *i. e.*, partial or even complete clearing of the suspension. These effects, if not avoided by keeping the temperature low, or by adding enzyme poisons, may obscure the observation of the initial cytolytic injury. It is apparent that tyrocidine does not inactivate these hydrolytic enzymes. Does it activate them, or why do they begin to work after the cytolytic death? Only a few experiments with certain phosphatases have so

far been undertaken but they suggest an answer. Normal washed staphylococci are able to hydrolyze added adenosine triphosphate, and glycerophosphate. Tyrocidine-treated cells, which show a burst of phosphatase activity toward their own small complement of phosphate esters, are not able to operate any faster upon an adequate concentration of added phosphate esters than the normal cells did before their membranes were made permeable. With certain limitations, these experiments indicate that phosphatase and adenosinepyrophosphatase are situated on or near the surface of these bacteria, and cytolytic injury allows substrates from the interior of the cell to leak out into the medium, where they can be hydrolyzed. It will be evident that tyrocidine can possibly afford us a novel means of studying the workings of the cell.

We may stop to inquire whether there is not good reason why phosphatases should be able to act at the periphery of the cell. Like esterases, peptidases, amylases, etc., their function is probably always to break down complex substances since, like these other hydrolytic enzymes, they appear unable to incorporate the energy needed for the corresponding synthesis. Peripherally distributed phosphatases may serve in part as agencies for converting phosphate esters, which appear to be generally nondiffusible, into a form capable of entering the cell.* Those cells able or obligated to use certain complex products of other cells as nutrients must be prepared to degrade these extraneous substrates to an acceptable form. In these cases we may expect other hydrolytic enzymes such as those mentioned to be active, although perhaps not entirely localized, at the cell boundary—more or less in contrast with the enzymes of respiration, fermentation, and synthesis. Many extracellular proteinases and polysaccharidases have been found in the medium surrounding the cells which produce them. And recently there have been interesting indications that the enzymes hydrolyzing trehalose (in yeasts) and lactose (in yeasts and colon bacilli)

* This is not in conflict with the possibility that, after hydrolysis, some substrates may be phosphorylated by the cell itself during absorption, as suggested by Lundsgaard, Verzár and others for the case of glucose absorption from the lumen of the kidney tubules or the intestine. The whole selective mechanism would ensure that foreign, and perhaps inimical, substances would not gain admission.

were active at the surface of cells in those strains capable of fermenting these sugars.

We have waited until now to mention a large group of other substances, which, so far as known, behave in essentially the same way as tyrocidine. It is not surprising that fixatives such as hot water, alcohol, acetone, and trichloroacetic acid also liberate cell constituents. But, in the enormous class of synthetic detergents, wetting, dispersing, and emulsifying agents, etc., there are many bactericidal surface-active agents; and, so far as they have been investigated, these show the same phenomena as tyrocidine when they kill bacteria. Cell solutes are liberated, respiratory activity is depressed, autolytic changes are initiated, whenever the substance and its concentration are adequate to kill. Indeed, tyrocidine deserves no special consideration except for the fact that its behavior has been investigated more fully than that of any of the synthetic products. And the diversity of chemical structures which produce this cytolytic type of killing allows us to conclude that the structural requirements for activity are chiefly physical and are broad and ill-defined. Basic (cationic) agents are generally more effective, especially toward the Gram-negative bacteria, but they may be aliphatic straight-chain, aromatic or heterocyclic quarternary ammonium bases, or simple fatty amines. Various alkyl and aryl sulfates and sulfonates have similar activity, as do some fatty carboxylic acids. When in concentration sufficient to kill, phenols and cresols have the same effect, but this is greatly increased in alkyl-substituted phenols, which have more enhanced surface activity. The small class of nonionizing surface-active agents has apparently not yet produced a representative with notable bactericidal activity.

Somewhat systematic investigations have already been made on the relation of structure to bactericidal activity in some of these compounds, especially the phenols and the aliphatic acids. One of the few conclusions that could be reached was that the most effective substance in each series is usually one having approximately sixteen to eighteen carbon atoms. Branched-carbon-chain compounds appear to be commonly more toxic than those with straight chains. There are almost no evidences of selectivity except that Gram-negative bacteria are mainly vulnerable only to basic substances unless the *p*H is low enough that the cell proteins can combine with anionic detergents. There are other small differences of sensitivity among various species.

Classification of substances as inactive is undoubtedly largely a relative matter, since probably almost any surface-active agent worthy of the name would kill bacteria if used in the concentration of around 1% required for phenol and cresol, which are thought of as substances toxic to bacteria.

The water insolubility of a large portion of the molecule is the driving force that crowds even minute quantities of a surface-active agent out to the boundaries of the aqueous phase. The affinities of this nonpolar group for elements of the bacterial structure determine where upon the cell it will become anchored and to some extent what harm it will do there. The general toxicity of a wide variety of substances in this class suggests that the cellular elements having affinity for hydrocarbon groups do not differ very much from species to species. Evidently, whenever the nonpolar group carries basic or acidic groups along with it the harm it will do may be extensive. If no strongly ionized groups are present, either there is less of the substance adsorbed or its effect is less drastic. Gramicidin, which is a rather insoluble surface-active agent of the nonionizing class, is not notably bactericidal, but does display some specificity and delicacy of bacteriostatic action, as indicated above. Perhaps systematic exploration or syntheses of further compounds in this general physical class of nonionic agents may offer some promise for those anxious to find new medicinals with some capacity for showing specific action upon cells.

The special cases of antibacterial agents just considered have caused us to be concerned with cellular attributes of which we frequently lose sight. Let us attempt to organize our rudimentary information about cell physiology so that we may see whether there are other possible sites vulnerable to chemical action. We consider most cells to be an organized system of enzymes, most of them at present unknown, preserved together with vital coenzymes and metabolites within a structural framework the outer boundary of which displays a selective permeability. As indicated in the table on pages 394 and 395, the chemical tasks of the living cell may be classified as several different, obviously somewhat overlapping, types of function. This tabulation will, one hopes, look very naive indeed in a few years, but it may help us a little now as we try to visualize the problem of exerting chemical control over cells of different types.

Not only is the understanding of the phenomena under "mecha-

Cellular function	Supposed chemical mechanism	Some possible vulnerable sites	Possible examples of deleterious agents
		Degradative enzymes:	
Energy production	Preliminary hydrolysis complex nutrients	Inactivation, destruction	Oxidants, halogens, formaldehyde, acid, heat, ultraviolet light
	Respiratory catabolism of foodstuffs	Inhibition; competitive with substrates	Substrate analogues, malonic acid, carbon monoxide
	Fermentative catabolism of foodstuffs	Inhibition; competitive with coenzymes	Coenzyme analogues, pyrithiamin, pyridinesulfonic acid, sulfonamides?
	Production of metabolic intermediates	Inhibition; noncompetitive	Cyanide, arsenite, barbiturates?, low temperature
	Production of energy carriers; adenosine triphosphate	Inhibition; sulfhydryl groups	Iodoacetate, mercury or arsenic, quinones, maleic acid, clavacin?
		Diversion of coenzymes	Acridine dyes?
		Diversion of enzyme activators	Fluoride, cyanide, cysteine, hydroxyquinoline
	Storage of energy; creatine phosphate, arginine phosphate	Diversion of energy	Arsenate, dinitrophenol?, gramicidin?
	Storage of glycogen, fat reserves	Diversion of intermediates	Sulfite, semicarbazide, hydrazine, cyanide, borate?
		Synthetic enzymes:	
Energy utilization Growth	Synthesis of carbohydrate, protein, fat	Inactivation	As above
		Inhibition	As above; analogues of substrates for synthesis?, phlorhizin
	Synthesis of enzymes	Diversion of prosthetic groups	Coenzyme analogues, hemin analogues, sulfonamides?
	Reduction of oxidized nutrients	Reduced coenzymes	Oxidants, iodine, ferricyanide, oxidizing dyes?
	Phosphokinase systems	Sulfhydryl groups; myokinase	As above; glyceraldehyde
	Elaboration of nucleus, membrane, cell walls, cytoplasmic granules, etc.		
	Repair of cell structures by synthesis	Synthetic enzymes	As above

Cellular function	Supposed chemical mechanism	Some possible vulnerable sites	Possible examples of deleterious agents
Maintenance	Preservation of spatial relationships of cellular elements	Physical structure of cytoplasm	Fixatives, protein precipitants, fat solvents
		Cytoplasmic granules	Acids?
	Semipermeable membrane retaining vital intermediates, catalysts, metabolites	Membrane, damage	Lysozyme, cytolytic enzymes; fat solvents; freezing, grinding
		Membrane, change in semipermeability	Surface-active agents, fatty acids, phenols, tyrocidine
		Increase internal osmotic pressure	Alcohol, acetone, glycols, glycerol, urea
Secretion	Diffusion of nutrients into, and end products out of, cell	Membrane, change in electric charge	Colloidal agents?
		Membrane, occlusion	Antibodies?, protein coagulants
	Production of concentration gradients by enzymic conversion of solutes	Enzymes: inactivation	As above
		inhibition	As above; phlorhizin?, gramicidin?, hormones?
Cell division, budding, sporulation	Rearrangement of already synthesized material	Alteration in physical structure of cytoplasm or nucleus	Colchicine?
	Enzymic changes?	Enzyme inhibition	Penicillin?
Special functions:			
Irritability	Ion, etc., layers at membrane	Surface changes on membrane	Narcotics, urethans?
	Humoral transmission	Cholinergic mechanism	Atropine, curare, quarternary ammonium compounds?
		Choline esterase	Eserine, prostigmine, strychnine?, morphine?, (thiamin?)
		Adrenergic mechanism	Ergotamine; ephedrine
Contractility	Myosin; utilization of adenosine triphosphate energy	Adenosinetriphosphatase Phosphate transfer enzyme?	Sulfhydryl blocking agents as above
Motility	Cilia, flagella	Protoplasmic surface	Surface-active agents

nism" in the table all in the theoretical stage, but we have little more than a crude terminology with which to describe them. Some detail could be given under the mechanism of respiratory and fermentative catabolism, but almost none under any of the anabolic, energy-utilizing mechanisms. The examples of agents given in the last column are, at best, somewhat questionable. Nevertheless, it is hoped that the table will serve to demonstrate that it is almost inexcusable that, when considering the mode of action of new or proposed drugs, we have all almost universally limited ourselves to a consideration of possible effects of the drug upon some aspect of the respiratory or fermentative *degradation* of foodstuffs. As heating engineers trying to see why the house is cold, we have been shortsightedly preoccupied with the flicker of the furnace fire and have never bothered even to feel the pipes and radiators upstairs.

Throughout the above, the viewpoint has been somewhat insistently shifted back and forth from the planning of new chemotherapeutic agents to studying the action of the old ones. It is not pretended that empirical ways of uncovering new agents have been outmoded, but we do maintain that the most successful empiricism will always use, consciously or unconsciously, what reasonable hypotheses are available. And the drugs are among our best cytochemical reagents: they can give us information about the reactivities of living cells, and this information will surely help us to foresee the probable behavior of new drugs. Strictly speaking, there can be no study of relation between pharmacological action and chemical structure and there can be no study of mechanisms of drug action—there can only be the study of drug–protoplasm interaction.

Selected References

LOCALIZATION OF CELLULAR CONSTITUENTS

Caspersson, T., "Studies on the protein metabolism of the cell," *Naturwissenschaften*, **29**, 33 (1941).

Claude, A., Hogeboom, G. H., and Hotchkiss, R. D., "Distribution of enzymes in cytoplasmic constituents," to be published.

Dounce, A. L., "Further studies on isolated cell nuclei of normal rat liver," *J. Biol. Chem.*, **151**, 221 (1943).

Gomori, G., "The distribution of phosphatase in normal organs and tissues," *J. Cellular Comp. Physiol.*, **17**, 71 (1941).

Feulgen, R., and Rossenbeck, H., Microchemical test for nucleic acid of the thymonucleic acid type and the selective staining of cell nuclei in microscopic preparations, *Z. physiol. Chem.*, **135**, 203 (1924).

Lan, T. H., "The *d*-amino acid oxidase, uricase, and choline oxidase in normal rat liver and in nuclei of normal rat liver cells," *J. Biol. Chem.*, **151**, 171 (1943).

Linderstrøm-Lang, K., "Distribution of enzymes in tissues and cells," *Harvey, Lectures*, **34**, 214 (1938–1939).

Myrbäck, K., and Vasseur, E., Lactose fermentation and the localization of enzymes in the yeast cell, *Z. physiol. Chem.*, **277**, 171 (1943).

Wolf, A., Kabat, E. A., and Newman, W., "A study of the distribution of acid phosphatases with special reference to the nervous system," *Am. J. Path.*, **19**, 423 (1943).

ANTIGEN–ANTIBODY REACTIONS

Landsteiner, K., *The Specificity of Serological Reactions*. Harvard Univ. Press, Cambridge, 1945.

Marrack, J. R., *The Chemistry of Antigens and Antibodies*. H. M. Stationery Office, London, 1934.

THEORIES OF DRUG ACTION

Ehrlich, P., *Experimental Researches on Specific Therapeutics*. Lewis, London, 1908.

Lillie, R. S., "Antagonism between salts and anaesthetics," *Am. J. Physiol.*, **31**, 255 (1913).

Traube, J., Theory of narcosis, *Arch. ges. Physiol.*, **153**, 276 (1913).

Warburg, O., Physical chemistry of cell respiration, *Biochem. Z.*, **119**, 134. (1921).

COMPETITIVE INHIBITION IN METABOLISM

Fildes, P., "The mechanism of the antibacterial action of mercury," *Brit. J. Exptl. Path.*, **21**, 67 (1940).

Quastel, J. H., and Wooldridge, W. R., "Some properties of the dehydrogenating enzymes of bacteria," *Biochem. J.*, **22**, 689 (1928).

Woods, D. D., "The relation of *p*-aminobenzoic acid to the mechanism of the action of sulfanilamide," *Brit. J. Exptl. Path.*, **21**, 74 (1940).

GRAMICIDIN AND TRYOCIDINE

Dubos, R. J., and Hotchkiss, R. D., "The production of bactericidal substances by aerobic sporulating bacilli," *J. Exptl. Med.*, **73**, 629 (1941).

Hotchkiss, R. D., "Gramicidin, tyrocidine, and tyrothricin," in *Advances in Enzymology*, Vol. IV. Interscience, New York, 1944, p. 153.

SURFACE-ACTIVE AGENTS

Baker, Z., Harrison, R. W., and Miller, B. F., "The bactericidal action of synthetic detergents," *J. Exptl. Med.*, **74**, 611 (1941).

Hotchkiss, R. D., "The nature of the bactericidal action of surface-active agents," *Ann. N. Y. Acad. Sci.*, **46**, art. 6 (1946).

Kuhn, R., *et al.*, Invert soaps, *Ber.*, **73**, 1080 *et seq.* (1940).

Stanley, W. M., Coleman, G. H., Greer, C. M., Sacks, J., and Adams, R., "Bacteriological action of certain synthetic organic acids toward *Mycobacterium leprae* and other acid-fast bacteria," *J. Pharmacol.*, **45**, 121 (1932).

Suter, C. M., "Relationships between the structure and the bactericidal properties of the phenols," *Chem. Revs.*, **28**, 269 (1941).

25

BIOCHEMICAL ASPECTS OF PHARMACOLOGY

ARNOLD D. WELCH, PROFESSOR OF PHARMACOLOGY, SCHOOL OF MEDICINE, WESTERN RESERVE UNIVERSITY
ERNEST BUEDING, ASSISTANT PROFESSOR OF PHARMACOLOGY, SCHOOL OF MEDICINE, WESTERN RESERVE UNIVERSITY

EARLY MAN, in his almost continuous search for food, encountered a great variety of substances capable of affecting the normal functions of the body. Some of these materials produced effects which ultimately were correlated with their ingestion, a circumstance which led to their use in attempts to alleviate suffering or to combat disease. Undoubtedly the experiences and observations of primitive medicine men added many other substances to the early *materia medica*. Throughout the centuries, testing of the hit-and-miss type has continued, and indeed few substances now exist which have not at some time been used in an attempt to change the course of bodily disorders. Of these, only a small number has proved to have lasting value, although a much larger number of agents of little merit remains to distend the dispensatories of the present day. From the loss of most of these the world would suffer very little.

The search for new drugs continues, although they are usually derived from a different source than formerly, for the synthetic activity of nature has been replaced almost entirely by that of man. Also, in the primary trials of new agents, human beings have now largely been replaced by animals and microorganisms; yet the basic philosophy of trial and error continues essentially unchanged.

Insight into the mechanism of the action of drugs on tissues and organs of necessity awaited the attainment of considerable knowledge in other fields, in particular that of physiology. As the principles governing the function of organs and tissues began to be elucidated and the science of physiology was developed, study of the mechanism of the action of drugs began. Thus, not only has pharmacology developed logically from physiology, but the greater portion of our present knowledge of drugs was acquired through the use of methods devised primarily for physiological studies. The techniques and concepts of physiology were gradually incorporated into the matrix of the newer science, although the emphasis was shifted from a primary interest in function to an exhaustive study of each drug, so that all its multitudinous effects might be uncovered, rather than to a continued study of function and the manner in which it is influenced by *many* chemical substances.

This approach to pharmacology has been to a considerable degree self-sterilizing, since it led in the course of a comparatively few years to relative exhaustion of the stock of drugs capable of exerting sufficient physiological effect to justify extensive study. As a consequence, findings in the field of pharmacology became relatively prosaic, interest in the science to some degree waned, and few young men were attracted to or trained for a scientific career in the field. Where noteworthy advances occurred, they came especially in those fields in which the closest cooperation between the pharmacologist and the synthetic organic chemist was possible; but here too almost all advances, except the most recent, resulted from approaches which were essentially empirical.

In general, pharmacologists must face the indictment of having viewed with too great complacency the limitations of the classical physiological attack upon the basic mechanisms of drug action. The continued emphasis upon the physiological approach led to many improvements in our understanding of the action and proper usage of important therapeutic agents, but, during recent years at least, such methods have rarely opened new fields of research. Indeed, only a few pharmacologists have given attention to the biochemical aspects of the action of drugs or have turned to the investigation of more fundamental systems.

The fact of the matter is that pharmacology, as a separate dis-

cipline of the medical sciences, very nearly ceased to exist, a situation which might not have been long delayed had it not been for the chemotherapeutic advances of the last decade. These, although initiated largely by workers in cognate fields, gave pharmacology a cogent reason for being, and placed renewed emphasis on the fact that the "site" of drug action is not composed of static morphological units. Rather is it to be found in the interrelated chains of chemical reactions which constitute the dynamic equilibrium of the cell.

That the emphasis on the study of the action of drugs has been placed primarily on the responses of organs and tissues, rather than on those of cells, is well exemplified by the investigations of an important pharmacological agent, *viz.*, digitalis. The cardioactive glycosides of the digitalis and related plants have long been known to affect beneficially the function of the decompensated failing heart, an effect now related largely to an increased contractility of the cardiac muscular tissue (5). Yet, few studies have so far been made of the influence of such drugs on the metabolic reactions supplying the energy required for this function of myocardial cells. The validity of such an approach can be examined by investigating the effect of such drugs on various metabolic reactions and isolated enzyme systems in cardiac and in other forms of muscle.

The evidence available today, although in some cases circumstantial, strongly suggests the possibility that the action of many drugs and poisons is mediated through a direct or indirect effect on enzyme systems. This is not a new hypothesis. Myrbäck (23) made a similar suggestion in 1926, and Clark (6) has reviewed the literature of the subject up to 1937. Consideration of various classes of drugs indicates that in only a few cases is it definitely unlikely that enzyme systems participate in the primary action of drugs or poisons, for instance, saline cathartics and diuretics, adsorbents, neutralizers of alkali or of acid, and hemolytic compounds such as the saponins. It is equally true that there are only a few cases in which the action of drugs or poisons has been proved beyond question to involve enzyme systems primarily.

Of perhaps greatest importance, from the standpoint of unquestionable enzymic participation in drug action, is physostigmine (eserine), a potent drug and poison which was shown by Loewi and Navratil (17) in 1926 to inhibit choline esterase specifically. Under

ordinary circumstances the neurohormone, acetylcholine, responsible for the chemical transmission of nerve impulses at many sites, is destroyed promptly, so that the effect of each nerve impulse is very transitory. Partial inhibition of the enzyme which catalyzes the hydrolytic cleavage of acetylcholine permits the accumulation of the neurohormone in higher concentration, a circumstance equivalent in its physiological effects to the injection of acetylcholine at the endings of innumerable cholinergic nerves. Proof that the transitory action of acetylcholine is due to its enzymic destruction and that this enzyme is actually the "site of action" of physostigmine, constituted the first clear-cut demonstration that the primary action of a drug can be based solely on its effect on a single enzyme. It is of considerable practical importance that the rate of hydrolysis of choline esters can be altered by modifying the structure of the choline moiety or of the esterifying acid. Thus, acetyl-β-methylcholine and carbaminoylcholine have the same pharmacological effect as acetylcholine, but they are slowly or not at all hydrolyzed by choline esterase.

Other examples of drugs and poisons, the primary action of which involves enzyme systems, are neostigmine, a compound related to physostigmine, ionizable cyanides, and such vitamins as thiamin, riboflavin, and niacin. However, there is strong presumptive evidence for the involvement of enzyme systems in the action of various other drugs and poisons.

Quastel *et al.* (26) have attempted to explain the mechanism of action of narcotics in terms of their effects on the respiration of brain tissue and on isolated enzyme systems. Various drugs which produce anesthesia (barbiturates, chlorobutanol, and ether) were shown to inhibit reversibly the oxygen consumption of the cortex *in vitro.* Oxidation of glucose, lactate, or pyruvate is inhibited to a larger degree than the oxidation of either succinate or glutamate, while the anerobic utilization of glucose by the brain tissue is inhibited only slightly. In continuation of this work, the effect of chlorobutanol on several enzyme systems involved in the oxidation of glucose was examined by Michaelis and Quastel (22). Since the activity of various dehydrogenases, cozymase, Straub's flavoprotein ("diaphorase"), or cytochrome oxidase is not decreased by chlorobutanol, in concentrations which inhibit glucose oxidation in the brain, it was concluded that the narcotic inhibits an as yet unidentified respiratory enzyme system.

Because of disagreement concerning the concentration of various narcotic agents in brain tissue, at varying depths of anesthesia, evaluation of the observations of Quastel and his co-workers is not a simple matter. Fuhrman and Field (8) concluded that, during anesthesia, the concentration of certain narcotic agents in the brain is less than one-fourth that required for the inhibition of glucose oxidation by brain tissue *in vitro;* however, their conclusions do not appear to have been based on adequate analytical data. Further studies will evidently be required to establish whether those concentrations of narcotic agents which produce anesthesia *in vivo* inhibit to an appropriate degree any of the enzyme systems that are required for the normal functions of the brain. The effect of narcotic agents on the functions of other tissues must also be considered, since it has been shown (22) that certain enzyme systems of skeletal muscle are inhibited by chlorobutanol to the same degree as are comparable systems of brain. However, the functional activity of muscle does not appear to be inhibited by those concentrations of chlorobutanol attained during anesthesia.

Even if it were found that the concentration of narcotic agents required to inhibit certain enzyme systems, or the over-all respiration, *in vitro*, is significantly greater than the concentration attained during anesthesia, a reaction more specifically concerned with the functional integrity of the nervous system may be found that is inhibited by lower concentrations. In favor of the theory that narcosis involves a more general depression of energy-transferring mechanisms in nervous tissue, however, is the wide variety of chemical structures, usually lipid-soluble, which produce a reversible narcotic effect on many types of cells. The importance of further study in this field of investigation is clearly indicated.

Seevers and Shideman (33) have made rather extensive studies of morphine in an attempt to place its action on a biochemical basis. In confirmation of the findings of Quastel and Wheatley (26) it was shown that the drug in high concentration (0.12%) inhibits the oxygen consumption, *in vitro*, of the cerebral tissues of rats, when certain substrates, particularly lactate, are added. In the absence of added substrates no effect was observed. Various mammalian dehydrogenases (lactic, citric, glucose) were shown to be inhibited by similar concentrations of morphine; levels much higher than any which could be attained in the animal body. The results were interpreted very

cautiously by the investigators and serve as a basis for the further studies which are needed.

More recently, Mayer and McCawley (21) have shown that the N-allyl homologue of morphine is even more depressant of oxygen consumption by slices of rat cerebrum than is morphine. Other studies (13,18,35) have shown, however, that this derivative is antagonistic to many of the effects produced by morphine, particularly that of respiratory depression. It would appear that depression of oxidative reactions of brain tissue, *in vitro*, cannot serve to explain the differences between the effects of morphine and the allyl homologue and certainly cannot account for the narcotic and analgesic actions of the parent alkaloid.

A striking observation made during the biochemical studies of Shideman and Seevers (33) is worthy of note. They found that the chronic administration of morphine to rats leads to a sustained acceleration in the oxidative metabolism of skeletal muscle, an effect which is maintained when morphine has been withdrawn, and which is noted when a major portion of the morphine would have been eliminated from the body. The addition of morphine to muscle tissue, *in vitro*, produces an increase in oxygen uptake, an effect observed whether the tissue used is obtained from normal rats or from chronically morphinized animals with an initial rate of oxygen consumption nearly twice that of the muscle tissue of normal rats. Shideman and Seevers do not believe that these observations are to be accounted for by the oxidation of the drug. As was pointed out, the data available do not clearly associate the effects produced *in vitro* with those which occur uniformly in chronically morphinized muscle during withdrawal. The fact that a concentration of azide which is without effect on the respiration of normal muscle reduces to an approximately normal level the oxygen utilization of chronically morphinized muscle should also be mentioned. These interesting observations suggest the need for additional studies of the mechanism of action of this very important drug.

On the basis of studies in their own laboratory (9) and elsewhere (3), Gaddum and Kwiatkowski advanced the hypothesis that the action of ephedrine, a naturally occurring sympathomimetic compound chemically related to epinephrine, can best be accounted for by its inhibition of amine oxidase. This enzyme is one of several which

participate in the inactivation of epinephrine and the closely related or probably identical neurohormone, sympathin, responsible for the chemical transmission of adrenergic impulses. Inhibition of the enzyme leads to the accumulation of epinephrine in much the same way that physostigmine protects acetylcholine from destruction. The stimulant action of amphetamine (benzedrine) on certain functions of the central nervous system was related by Mann and Quastel (20), in a rather complicated theory, to inhibition of the oxidative deamination of compounds such as tyramine, which is derived from tyrosine. It was suggested that the type of mental fatigue relieved by amphetamine is due to the formation, by amine oxidase, of toxic aldehydes from compounds such as tyramine. Central excitation is thus considered to be secondary to an inhibition of amine oxidase. Needless to say, the acceptance of such an hypothesis requires a great deal more evidence than that so far made available.

Amine oxidase is undoubtedly responsible for the oxidative deamination which many of the sympathomimetic amines undergo in the body (2) and is thus responsible, in part at least, for their inactivation. Derivatives of phenethylamine which possess an alkyl group alpha to the nitrogen, such as amphetamine and ephedrine, are not only refractory to the action of amine oxidase, but are capable of inhibiting markedly the action of the enzyme.

The enzymic approach to the study of the mode of action of drugs is by no means limited to those compounds which affect mammalian tissues. The chemotherapy of infectious and parasitic diseases is based on the use of compounds which in suitable dosage have less effect on chemical reactions essential to the well-being of the host than on those reactions requisite to the life or reproduction of invasive organisms. Thus, chemotherapeutic agents vary in their margins of safety from those that are of the same order of toxicity for both host and parasite to those that are essentially devoid of deleterious effects on any mammalian host. Examples of such extremes are actually to be found among drugs which may be used in the treatment of a single disease, like syphilis. Thus, metallic mercury, when used by inunction, is of such toxicity that it is possible only to arrest the progress of the infection, while with highly purified penicillin not only may apparent cures be produced but man appears to be essentially uninjured by any dose so far administered.

In a recent study of the effect of quinacrine (atabrine) on respiratory enzymes, Haas (12) found that the drug inhibits the respiration of *Plasmodium knowlesi* to a greater extent than that of mammalian tissues. Furthermore, at relatively low concentrations the drug inhibits cytochrome reductase and glucose-6-phosphate dehydrogenase (*Zwischenferment*), while it has little or no effect on cytochrome oxidase, cytochrome C, or triphosphopyridine nucleotide. The concentration of atabrine required to inhibit the respiration of malaria parasites *in vitro* is considerably higher than those concentrations of the drug in plasma which exert an antimalarial effect *in vivo*; conceivably, however, the concentration of atabrine in the parasites may be considerably higher than the concentration in the plasma. Quinine was less active than atabrine in inhibiting both cytochrome reductase and the respiration of malaria parasites, a difference which is in agreement with the fact that the antimalarial activity of atabrine is now recognized as definitely greater than that of quinine. In order to test the hypothesis that the antimalarial activity of these two drugs is in part due to an inhibition of the activity of cytochrome reductase, it might be helpful to study the enzyme-inhibiting effect of a group of compounds closely related structurally to atabrine, but among which varying degrees of antimalarial activity are found. The work of Haas would take on added significance if a close parallelism between the relative activities *in vitro* and *in vivo* were found in such a series, and if the discrepancy between the concentration causing enzymic inhibition and that which is therapeutically active were satisfactorily explained. A closer correlation between the effective concentration *in vitro* and *in vivo* was found by Silverman *et al.* (34), who showed that quinine in concentrations only slightly above those therapeutically active *in vivo* inhibits the oxidation of glucose by *Plasmodium gallinaceum*, without affecting glycolysis.

Earlier, but less striking, attempts have been made to correlate the inhibition of enzyme systems with the action of drugs chemotherapeutically effective *in vivo.* In this connection there might be mentioned the action of atoxyl and of quinine on certain lipases, the antifumarase-activity of certain trypanocidal compounds (25), and the inhibition with acriflavine of a cyanide-insensitive hydrogen-transporting system in trypanosomes (30).

Thus, various investigations have been conducted in which the

effect of a drug on isolated systems has been studied in an attempt to explain the physiological action of the compound. In some of these investigations attention has not been directed toward certain factors which, in the opinion of the authors, are of importance. It is suggested that, among others, the following criteria should be considered before the physiological action of a drug is attributed to an effect on an isolated enzyme system.

(1) The concentrations of the drug or poison which produce the effect *in vitro* and which obtain at the anatomical site of action should be closely similar.

(2) If a drug or poison exerts its effect primarily on a specific tissue *in vivo*, either the effect on this tissue *in vitro* should be more pronounced than the effects on other tissues, or the system inhibited must be shown to have more functional significance in this than in other tissues.

(3) Among structurally related compounds there should be close parallelism between the pharmacological activities *in vitro* and *in vivo:* all drugs of the same chemical series which are active *in vivo* must also be active *in vitro*, and those inactive *in vivo* must also be inactive *in vitro*, unless the discrepancy can be accounted for by poor absorption, by inadequate distribution to or penetration of the cells involved, by too rapid excretion, or by metabolic alteration.

These criteria may be of value in avoiding premature explanations of the physiological action of drugs on the basis of effects exerted on isolated systems.

Until very recently, the search for new chemotherapeutic agents has progressed along almost exclusively empirical lines. Such an approach, though tedious, has, nevertheless, yielded many drugs of importance, and will undoubtedly continue to do so in the future. Recently, however, largely on the basis of the theory of the mode of action of sulfonamides, concepts have been developed that promise to be of value in the development of compounds of utility in chemotherapy and in the study of intermediary metabolism.

According to these concepts, which are reviewed in detail elsewhere (19,37,38), interference with biological processes may result from the use of compounds structurally related to, but not utilizable in place of, substances essential for life or reproduction (essential metabolites). Interference of this type can be explained most simply on the basis of a direct competition between the essential metabolite

and its structural analogue for some cellular component for which they both have great affinity; in some cases, however, antagonism involves other factors than simple competitive inhibition. In verification of this concept, there have been developed various structural analogues markedly antagonistic to the biological utilization of niacin, pyridoxine, various amino acids, pantothenic acid, thiamin, riboflavin, ascorbic acid, certain purines, biotin, and vitamin K. Against organisms capable of synthesizing a given essential metabolite, the respective analogues have in general been quite ineffective, the most notable of the exceptions being of course, the sulfonamide group of drugs (analogues of *p*-aminobenzoic acid).

The effect of metabolite-analogues is probably exerted through their affinity for enzymes or for the protein components of enzymes, the substrates, or the prosthetic groups which they resemble structurally. The competitive type of antagonism of the metabolite-analogues is usually comparable to that exhibited by substrates and their analogues competing for isolated enzyme systems, for instance, succinate and malonate in the presence of succinic dehydrogenase. In such cases, the normal and the abnormal compounds antagonize one another competitively and the kinetics of the reaction usually can be explained with mathematical precision. Although none of the analogues of the various essential metabolites so far prepared, with the exception of the sulfonamides, has proved to be a practical chemotherapeutic agent, there is reason to believe that this approach will prove to be fruitful. It would be remarkable indeed if the groping which led to the discovery of the sulfonamides should have yielded the only group of compounds of practical utility in the antagonism of an essential metabolite.

Another biochemical aspect of pharmacology to which more attention should be given is concerned with the metabolic modification of toxic or potentially toxic compounds. This process is generally called "detoxication," a term which is misleading because it implies that these compounds when metabolized are always converted into less toxic substances. While this is generally true for compounds produced in the body (for example, hormones), it is not always the case with foreign substances (drugs, poisons, products of bacterial metabolism in the intestine, such as indole and skatole). In some instances, compounds of equal or even greater toxicity may be formed. It is

suggested, therefore, that the term "detoxication" be reserved only for those metabolic reactions which result in the formation of substances *less* toxic than the compounds from which they are derived, and to apply the term "metabolic alteration" to all changes which foreign compounds undergo in the organism. Metabolic alteration will usually be accomplished by means of oxidation, reduction, hydrolysis, or conjugation.

It is not inconceivable that reactions resulting in detoxication may be favorably influenced and, conversely, that reactions resulting in increased toxicity can be inhibited. In order to investigate these hypotheses, which could have considerable practical significance, biological systems concerned with the metabolic alteration of foreign compounds should be studied.

The transitory action of certain drugs in the body, as a result of their rapid alteration, is often disadvantageous, and in some cases the inhibition of such inactivation-reactions may be useful, as has been shown for the inhibition of choline esterase by physostigmine or neostigmine. With other transiently active compounds an inhibition of inactivation may be difficult to accomplish, because of their susceptibility to several enzyme systems which act on a variety of substrates. Thus, epinephrine is oxidized by the lactic or malic dehydrogenase systems, by cytochrome C (10), by amine oxidase (9), and probably by other enzyme systems.

Richter (27) postulated that epinephrine is inactivated in the body by conjugation with sulfate, a theory based partly on the observation that epinephrine increases the excretion of organic sulfate. However, the quantity of organic sulfate excreted after epinephrine administration is much greater than can be accounted for by conjugation with the drug administered (7). Since epinephrine decreases the formation of glucuronic acid by liver slices (16), it is possible that the hormone stimulates the formation of conjugated sulfates and decreases the production of glucuronates without itself extensively participating in the formation of the excreted organic sulfate.

While many foreign substances are enzymically oxidized in the body, other compounds accept hydrogen as a result of enzymic catalysis. In this manner trinitrotoluene (TNT) is partially reduced in the body by flavoproteins (4). It is interesting to note that one product of the reduction, hydroxylaminodinitrotoluene, is more toxic

Hay... Or Hr

than the parent substance, a metabolic alteration which obviously cannot be termed "detoxication."

As with the other mechanisms of metabolic alteration, conjugation does not necessarily result in a reduction in toxicity. Thus, acetylation of sulfonamides renders these compounds not only therapeutically useless but also more toxic. Conjugations of other foreign compounds with acetic, sulfuric, or glucuronic acids, or with glycine, glutamine, or cysteine, may result, however, in true detoxications. As yet, little is known about the mechanisms and enzymic reactions involved in most conjugations. Conceivably, conjugation-reactions may be enhanced by supplying the organism with the acids which are found in combination with the foreign compounds, or with the precursors of the acids. Acetic acid (15), glycine (11,31), and sulfuric acid (1) are used directly for conjugation, while three-carbon compounds, rather than free glucuronic acid itself, are involved in the formation of glucuronides (16). If a compound is conjugated with more than one acid, one conjugate being more toxic than the other, a favorable result might be attained by increasing the availability of the precursor of the less toxic conjugate. Thus, it might be possible that the potentially harmful acetylation of sulfonamides might be inhibited to some degree if glucuronic acid conjugation (28,32,36) could be favored.

The striking decrease in the toxicity of certain pentavalent arsenicals which results from the simultaneous administration of *p*-aminobenzoic acid is at present an unexplained type of detoxication (29). Although the decrease in the toxicity of such compounds is accomplished without inhibition of trypanocidal action, it has been shown (14,24) that the action of the arsonic acid derivative, atoxyl, on *Escherichia coli* is antagonized by *p*-aminobenzoic acid, in the same manner as are the sulfonamides. Since the mammalian toxicity of the sulfonamides is not antagonized by *p*-aminobenzoic acid, it seems unlikely that the pentavalent arsenicals influence an undetected function of *p*-aminobenzoic acid in animal tissues, unless, by virtue of their different distribution, systems are affected on which the sulfonamides are inert. A study of the effect of *p*-aminobenzoic acid on isolated enzyme systems may throw light on the mechanism of this detoxication.

Discussion of the detoxication of the arsenicals would not be complete without mention of the detoxifying action of sulfhydryl-

containing compounds, particularly glutathione. In fact, extensive studies by Voegtlin, Eagle, and their associates, indicate that the therapeutic effect of the arsenicals in protozoan infections is dependent on a reaction with cytoplasmic sulfhydryl groups.

As understanding of the biochemical mechanisms of energy transfer and of its catalysis is advanced, the fundamental actions and metabolic alterations of more and more of the known drugs and poisons will be elucidated, and unprecedented opportunities will be afforded for the rational development of new therapeutic agents.

References

(1) Bernheim, F., and Bernheim, M. L. C., *J. Pharmacol. Exptl. Ther.*, **78,** 394 (1943).

(2) Beyer, K. H., and Morrison, H., *Ind. Eng. Chem.*, **37,** 143 (1945).

(3) Blaschko, H., Richter, D., and Schlossmann, H., *J. Physiol.*, **90,** 1 (1937); *Biochem. J.*, **31,** 2187 (1937).

(4) Bueding, E., *unpublished research.*

(5) Cattell, M., and Gold, H., *J. Pharmacol. Exptl. Ther.*, **62,** 116 (1938); **71,** 114 (1941).

(6) Clark, A. J., in *Heffter's Handbuch der experimentellen Pharmakologie.* Springer, Berlin, Suppl., Vol. 4, 1937.

(7) Deichmann, W. B., *Proc. Soc. Exptl. Biol. Med.*, **54,** 335 (1943).

(8) Fuhrman, F. A., and Field, J., *J. Pharmacol. Exptl. Ther.*, **77,** 392 (1943).

(9) Gaddum, J. H., and Kwiatkowski, H., *J. Physiol.*, **94,** 87 (1938); **96,** 385 (1939).

(10) Green, D. E., and Richter, D., *Biochem. J.*, **31,** 590 (1937).

(11) Griffith, W. H., *J. Biol. Chem.*, **69,** 197 (1926); **82,** 415 (1929).

(12) Haas, E., *J. Biol. Chem.*, **155,** 321 (1944).

(13) Hart, E. R., and McCawley, E. L., *J. Pharmacol. Exptl. Ther.*, **82,** 339 (1944).

(14) Hirsch, J., *Science*, **96,** 139 (1942).

(15) Klein, J. R., and Harris, J. S., *J. Biol. Chem.*, **124,** 613 (1938).

(16) Lipschitz, W. L., and Bueding, E., *J. Biol. Chem.*, **129,** 333 (1939).

(17) Loewi, O., and Navratil, E., *Arch. ges. Physiol.* (*Pflügers*), **214,** 678 (1926).

(18) McCawley, E. L., *Federation Proc.*, **4,** 129 (1945).

(19) McIlwain, H., *Biol. Rev. Cambridge Phil. Soc.*, **19,** 135 (1944).

(20) Mann, P. J. G., and Quastel, J. H., *Biochem. J.*, **34,** 414 (1940).

(21) Mayer, N., and McCawley, E. L., *Federation Proc.*, **4,** 129 (1945).

(22) Michaelis, M., and Quastel, J. H., *Biochem. J.*, **35,** 518 (1941).
(23) Myrbäck, K., *Z. physiol. Chem.*, **158,** 160 (1926).
(24) Peters, L., *J. Pharmacol. Exptl. Ther.*, **79,** 32 (1943).
(25) Quastel, J. H., *Biochem. J.*, **25,** 898 (1931).
(26) Quastel, J. H., and Wheatley, A. H. M., *Proc. Roy. Soc. London,* **B112,** 60 (1932); *Biochem. J.*, **28,** 1521 (1934); *ibid.*, **32,** 936 (1938).
(27) Richter, D., *J. Physiol.*, **98,** 361 (1940).
(28) Sammons, H. G., Shelswell, J., and Williams, R. T., *Biochem. J.*, **35,** 557 (1941).
(29) Sandground, J. H., *J. Pharmacol.*, **78,** 209 (1943); *Proc. Soc. Exptl. Biol. Med.*, **52,** 188 (1943); *Science*, **97,** 73 (1943).
(30) Scheff, G., and Hassko, A., *Zentr. Bakt. Parasitenk. Orig.*, *I,* **136,** 420 (1936).
(31) Schoenheimer, R., Rittenberg, D., Fox, M., Keston, A. S., and Ratner, S., *J. Am. Chem. Soc.*, **59,** 1786 (1937).
(32) Scudi, J. V., *Science*, **91,** 486 (1940).
(33) Seevers, M. H., and Shideman, F. E., *J. Pharmacol. Exptl. Ther.*, **71,** 373, 383 (1941); **74,** 88 (1942).
(34) Silverman, M., Ceithaml, J., Taliaferro, L. G., and Evans, E. A., Jr., *J. Infectious Diseases*, **75,** 212 (1944).
(35) Unna, K., *J. Pharmacol. Exptl. Ther.*, **79,** 27 (1943).
(36) Weber, C. J., Lalich, J. J., and Major, R. H., *Proc. Soc. Exptl. Biol. Med.*, **53,** 190 (1943).
(37) Welch, A. D., *Physiol. Revs.*, **25,** 687 (1945).
(38) Woolley, D. W., *cf.* page 357.

SOME BIOCHEMICAL PROBLEMS POSED BY A DISEASE OF MUSCLE

CHARLES L. HOAGLAND, MEMBER OF THE ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH, NEW YORK; PHYSICIAN TO THE ROCKEFELLER HOSPITAL

THE MEDICAL biochemist hopes ultimately to provide a basis for the interpretation of disease phenomena which will permit the description of a given syndrome in terms of the nature of chemical and physiological alterations in the cellular mechanisms of the affected organ. To achieve the synthesis of knowledge required for this ambitious task he must rely on information supplied from fields as dichotomous as that of the naturalist, whose efforts may follow only the dictates of his curiosity, and that of the clinician charged with the practical responsibility of alleviating disease. Situated in this no man's land between the ill-defined borders of art and science, the medical biochemist is in some danger of oversimplifying the complex problems posed by his clinical colleagues, and of overextending the concepts of his co-workers in biochemistry and physiology. It follows naturally, therefore, that in cultivating the disputed stretch between the fields of medicine and science he is frequently in jeopardy of litigation from his neighbors on either side.

In no field of biology has intelligent effort been rewarded with more striking or steady advances than in that devoted to a study of the physiology and biochemistry of muscle. To a greater extent than for any other organ, muscle has been the common meeting ground of the biochemist, biophysicist, and physiologist. In the development of

our present knowledge concerning the underlying mechanisms of muscular activity, the resources of practically every field of science have been tapped. For the most part, the quality of the observations made in the study of muscle has been high, due in a measure to the excellent tools available for its study, and in part to the keen minds of the investigators for whom muscle biochemistry and physiology have held a major interest.

Until recently, little interest was exhibited in the application of information dealing with the chemical and physiological processes in normal muscle to an elucidation of aberrant mechanisms at work in the production of muscle disorders. As a consequence, these heterogeneous diseases continue to bear the confusing labels which were given them by the classical pathologists. On occasions, serious efforts have been made to apply the concepts and techniques of biochemistry and physiology toward a solution of the problems presented by these diseases, but the workers have been few, and significant results have been scanty. Nevertheless, on the basis of what is known concerning the normal physiology of muscle, we are perhaps in a better position with respect to this organ than for any other to derive information of an exact and fundamental character which would be helpful in forming a common basis for clarifying the muscle syndromes. Moreover, through a physiological consideration of the affections of muscle, a unique opportunity is open to the biochemist by means of which he may evaluate certain basic information with respect to its usefulness in achieving the goal that he has set for himself, namely, that of defining disease in terms of the nature of specific alterations in cellular mechanisms.

Voluntary muscle constitutes nearly 43% of the total weight of the body and is affected by a variety of diseases in which the primary process appears to be located in the muscle fibers, and by yet another group of diseases in which marked secondary changes in muscle are apparent, but in which the primary process is located in the central or peripheral nervous system. Few syndromes are more confusing than the primary muscular disorders. The confusion is due in part to lack of agreement on the clinical and pathological findings in these affections, and in part to the fact that, although these diseases have been recognized as clinical entities for over fifty years, the pathogenesis has remained obscure. It is becoming increasingly apparent that no

sharp lines can be drawn between many of the primary syndromes of muscle, and that some unknown defect in metabolism is responsible for the biological continuity and sequence of phenomena common to these clinical disorders.

The problem of the secondary affections of muscle resolves itself into a fundamental one having to do with the complex relationship between the integrity of the innervation of muscle and the maintenance of an optimum state of nutrition of muscle cells. This so-called "trophic effect" which cells of the nervous system appear to exert on the metabolism of the contractile cells of muscle is nearly as complex as the problem of muscle metabolism itself, and cannot be sharply divorced from a total consideration of other diseases affecting the muscular system in which no neurological component is recognized.

Although the study of the muscle disorders has not kept pace with advances in the field of biochemistry and physiology of muscle, it has only recently received impetus from several new and important sources of information. The discovery of a disease of dietary origin in animals which resembles progressive muscular dystrophy, but which, unlike the disease in man, is cured with vitamin E, is influencing the study of human affections of muscle. The clinical observation that prostigmine relieves to a variable extent the symptoms of myasthenia gravis has accelerated fundamental studies on the metabolism of acetylcholine in this disease. The discovery of an aberration in potassium metabolism in periodic muscular paralysis had led to some success in a search for effective therapy for this disorder, and to the institution of basic studies on the relation of periodic paralysis to other disturbances in muscle function associated with changes in the concentration of potassium. The presence of a disease in goats, indistinguishable symptomatically from congenital myotonia in man, has greatly facilitated studies on the group of myotonic maladies. In the near future, therefore, increasing knowledge of the physiology of muscle, particularly that pertaining to the physiology and chemistry of contraction, transmission of excitation from nerve to muscle, and the mode of action of specific pharmacological agents on the myoneural junction, should provide many new and effective means for an attack on the problem of the muscle diseases as a whole.

No one syndrome among the diseases of muscle has provided

more challenge to the clinician than progressive muscular dystrophy. This disease, first clearly described by Gowers (6), is characterized by primary degeneration and atrophy of voluntary muscles of the extremities, pelvis, and shoulder girdles. Atrophy is marked in all forms, and hypertrophy or pseudohypertrophy is an early and prominent symptom in one form. The disease has certain well-established familial aspects, although the mode of genetic transmission is not always clear. In the most common type, pseudohypertrophic muscular dystrophy, the onset usually occurs before the fifth or sixth year. The prognosis is grave, and death frequently results within several years after the disease has become clearly manifest. No specific therapy for the syndrome is known and little in the way of supportive therapy has proved useful (16).

Morphological studies on progressive muscular dystrophy have yielded little information, aside from the fact that there is hyaline degeneration and fragmentation of muscle fibrils, and that the muscles show marked increase in fat content. The recent development of a simplified quartz microscope, with the 2537 Å. line of mercury as the light source, has made it possible to obtain ultraviolet photomicrographs of muscle fixed and sectioned by current methods, and in certain instances on surviving specimens without previous fixation (8). Photomicrographs of muscle made by this technique show selective absorption of varying intensity, and reveal more detail than those obtained with visible light on stained material. The technique has particular advantages in the study of muscle when it is desired to correlate morphological and physiological changes. In ordinary light the appearance of unstained muscle is misleading, since the image under these conditions is due to inhomogeneity of tissue and not to the presence of material with selective absorption.

Description of the histological and pathological changes revealed by ultraviolet light is difficult at the moment, because of the inappropriateness of classical nomenclature for description of images produced by specific absorption of light rather than by staining and physical inhomogeneity of tissues. Certain analogies can be drawn, however, between structures photographed in ultraviolet light and those structures which are seen in visible light as a result of the application of tissue stains, on the basis of size and location of the structures within a given cell. It is evident from these studies that ultraviolet

photomicrography to be developed fully will require revision and redefinition of the nomenclature of classical histology in order to include those structures which are due to selective absorption as well as those which are due to the effect of physical inhomogeneity and contrasting stains.

Sections of human muscle photographed in ultraviolet light reveal deeply absorbing transverse zones, spaced at regular intervals, separated by alternate zones of low absorption occurring throughout the length of the muscle fibers. The over-all appearance of the photomicrographs is the familiar one of cross striation which has long been associated with voluntary muscle. Caspersson, using indirect methods, recently concluded that the isotropic, or dark, striae of muscle revealed by polarized light corresponds in position to the strongly absorbing areas seen in ultraviolet light (3). It has been suggested, moreover, that the absorption which is characteristic of these zones is due chiefly to their content of adenylic acid or adenosine triphosphate. A direct comparison of the position of the striae appearing in polarized light with that occupied by striae appearing in ultraviolet light was made recently, in our laboratory, by comparing half of a field of a given muscle section photographed in polarized light with the remaining half photographed in ultraviolet light. When identical magnifications of the images are compared, inspection of the matched photomicrographs leaves no doubt that isotropic zones in muscle revealed in photomicrographs in polarized light correspond exactly to the deeply absorbing striae revealed in the ultraviolet. In sections of muscle taken from cases showing early signs of pseudohypertrophic muscular dystrophy, the transverse zones were very faint, and in some areas almost indiscernible, in marked contrast to the broad compound bands observed in transverse sections of normal muscle. In cross sections, areas of absorption which in normal muscle were punctate and equidistant in character, were revealed as coalescent areas in a state of apparent disorganization.

That the method of simplified ultraviolet photomicrography possesses marked advantages over the classical methods of histology and pathology, which depend on staining and examination of specimens in visible light, has been brought out convincingly in a study of the histopathological changes associated with muscle disease. In this respect muscle tissue has proved to be a happy choice. Because

of its specialized architecture, a fairly satisfactory comparison can be made between structures revealed in ultraviolet light and those which are observed after staining and photomicrography in the visible region of the spectrum. Not only is greater resolution achieved with ultraviolet light photography, but the image which is obtained may provide in addition some idea of the chemical nature of the absorbing areas of the tissue, since it results from the selective absorption of light by proteins and by substances of high absorptive capacity, such as the purines and pyrimidines.

A plethora of micro methods is available for the quantitative determination of tissue constituents, and for almost every type of intermediary compound known to arise in metabolism. However, interpretation of quantitative data secured by analysis of diseased muscle is enormously complicated by lack of information concerning the total mass of muscle cells in these specimens due to their variable content of fat, fibrous connective tissue, collagen, and water. It may be highly desirable to determine quantitatively the various mineral constituents of diseased muscle, and to analyze the material for its content of creatine, adenylic acid, and various organic esters of phosphorus. The results have little significance, however, unless some suitable basis of reference can be found which will permit comparison of the mass of muscle cells in one specimen with the mass of cells in another. Two specimens of muscle removed at biopsy from approximately the same area, in a subject showing incipient progressive muscular dystrophy, may show differences of 50% in fat content, 20% in concentration of collagen and fibrous connective tissue, and 15% in the content of free water and ash. Low values for creatine phosphate, adenosine triphosphate, or other organic constituents in specimens of diseased muscle have little significance unless some factor of correction can be employed which will allow for the differential mass of muscle cells present in the samples. Some success was achieved recently in our laboratory in the use of myosin as a base of reference for certain organic constituents of diseased muscle. Since myosin is the principal protein component of the contractile cells of muscle, its concentration must bear an important quantitative relationship to the total mass of contractile cells in the specimen taken for analysis. Owing to its property of critical solubility in potassium chloride, myosin can be determined quantitatively, and with a fair degree of accuracy, on

specimens of muscle weighing not more than one hundred milligrams.

The problem of adequate bases of reference for constituents of diseased tissue in general is one of the principal barriers to the development of a rational discipline of chemical pathology. It is becoming increasingly evident that practical methods for determination of the residual mass of parenchymal cells of affected organs will have to be devised before full consideration can be given to possible alterations in concentration of enzymes, substrates, metabolites, and metallic catalysts occurring in tissue as a result of affection by disease. Since, in the chemical analysis of diseased organs, we are going to be faced constantly with the need of knowing the net quantities of various tissue constituents, it is safe to assume that increasing attention will be paid in the future to the problem of the relation of the concentration of tissue constituents to differential cellular mass.

Some of the difficulties inherent in the selection of a proper base of reference for constituents of diseased tissue can be avoided by ascertaining the ratio between the concentrations of two substances in a given specimen of material and comparing it with the ratio which is found for these constituents in similar tissues of normal organs. Notable among these is the familiar respiratory quotient, which, since it measures the ratio of the production of carbon dioxide to oxygen consumption, can be considered independently, within certain limits, of the variable extraneous constituents of the specimen. Studies of the respiratory quotient by means of the method of Dickens and Simers (4) of specimens of muscle removed from patients exhibiting various stages of progressive muscular dystrophy did not reveal values differing significantly from those obtained for specimens of normal muscle removed during surgical operation (17). The carbon dioxide and oxygen quotients were extremely low in the diseased tissue, but the ratio of carbon dioxide production to oxygen consumption was normal. It should be pointed out that, with respect to the oxygen consumption, specimens of surviving muscle from patients with progressive muscular dystrophy differ markedly from specimens of muscle taken from animals which have been deprived of vitamin E. Houchin and Matill have shown that muscle tissue removed from animals on diets deficient in vitamin E was characterized by marked increase in oxygen consumption (10), and, moreover, that the oxygen uptake of the muscle speci-

mens was rendered normal by the *in vitro* addition of α-tocopherol phosphate (9).

Notwithstanding the fact that progressive muscular dystrophy has engaged the interest of clinicians since the original description of the syndrome in the middle of the 19th century, almost no attempts were made to study the general metabolism in this disease until the pioneer work of Levene and Kristeller (11) in 1909. These workers showed that the feeding of protein resulted in an excessive excretion of creatine in patients with progressive muscular dystrophy. Subsequently, many studies have revealed that there is marked derangement in metabolism of creatine in this disease, and that endogenous creatine, formed from protein and amino acids, is not retained by the muscles as effectively as in normal subjects. This observation has given rise to the concept that there is, in progressive muscular dystrophy, a diabetic-like state with respect to the ability of the patient to retain either ingested creatine or creatine which is formed endogenously from proteins and amino acids. Whether or not this is a true concept, the recognition of a biochemical aberration in creatine metabolism is perhaps the only truly significant contribution to have been made within the last thirty years toward an understanding of the essential nature of this disease.

Of perhaps even greater significance than an increase in excretion of creatine in progressive muscular dystrophy is a diminished output of creatinine. Because of the specific association of creatinine formation with the integrity of muscle processes, the urinary concentration of this material may, in certain cases, give a more reliable indication of the severity of the disease than the level of urinary creatine (7). Moreover, the relatively great constancy in the excretion of creatinine, even under widely differing conditions of diet and health of the subject, permits the attachment of greater significance to small changes in the urinary concentration of this material than is the case with creatine, which may show wide fluctuations from day to day. The fact that creatinine is not derived from all tissues, but that it arises as a special process of tissue catabolism taking place "largely if not wholly in the muscles," has given rise to the belief that the amount of creatinine excreted in the urine "bears a direct relation to the potential efficiency of the muscles and is a reliable index of the muscular development of an individual" (15). The belief, originally stated by Folin, that

creatinine formation represents a type of endogenous metabolism distinct from the exogenous metabolism of food protein (5) is no longer tenable, as a result of the studies performed by Bloch and Schoenheimer with the aid of isotopic nitrogen (14). The constant daily excretion of creatinine, in contrast to the highly inconstant excretion of other nitrogenous constituents of the urine, however, indicates that the formation of creatinine is an orderly and well-regulated process, the biological significance of which is not entirely apparent. That it arises directly from creatine is abundantly clear. Moreover, there is recent evidence to indicate that it arises in the process of dephosphorylation of creatine phosphate, and that its formation is intimately connected with the phenomenon of muscular activity (2,13).

The marked decrease in the output of creatinine observed consistently in patients with progressive muscular dystrophy has been attributed to a reduction in the total mass and efficiency of the muscular system in this disease. In the normal subject there is a rather remarkable relationship between the level of urinary creatinine and the total mass of the musculature. By taking the muscle mass of the normal subject as 43% of the body weight, a fairly quantitative idea of the mass of functioning muscle in patients with progressive muscular dystrophy can be obtained from a consideration of the value of the urinary creatinine. This value has been found to vary from 35 to 40% in cases of incipient progressive muscular dystrophy to values as low as 17% in patients in whom the disease was advanced. In our laboratory, a marked correlation has been observed consistently between the muscle mass calculated as per cent of body weight and the degree of physical performance of patients with dystrophy, in so far as such performance could be appraised in a quantitative fashion. Calculation of the muscle mass on the basis of creatinine excretion appears to yield important information regarding the extent to which the muscular system is involved in patients with this disease. Moreover, by performing this type of calculation at frequent intervals on a given subject with progressive muscular dystrophy, a fairly quantitative appraisal of the clinical condition of the patient and the rate of progression of the disease at various intervals can be ascertained.

Following the elucidation of the role of methionine in the synthesis of creatine and creatinine by Bloch and Schoenheimer (1) and du Vigneaud (19), speculation arose in the laboratory over the possi-

bility that a deprivation of methyl groups may occur in certain diseases characterized by excessive creatinuria. It has been shown that growth in rats is inhibited by feeding of excess glycocyamine, and that growth is resumed following the administration of choline or methionine (18). Although alternative hypotheses must be entertained, the results of these experiments indicate that, in the presence of an excess of substances which act as methyl acceptors, a loss of methyl groups may occur, and that this deficiency may be prevented by an adequate intake of methylating agents such as methionine and choline. Studies on the total methyl output in patients with progressive muscular dystrophy revealed that the loss in methyl groups, occurring as a result of excessive creatine output, and greater than that seen in normal children, was fully compensated for by a diminution in output of creatinine (7). A surprising agreement was also found between the total quantity of "methyl" excreted by the group of normal subjects and that excreted by patients with progressive muscular dystrophy. These data suggest that there is no absolute increase in the output of total creatine compounds in progressive muscular dystrophy, and that the diseased patient differs from the normal, with respect to creatine output, in the differential partition of creatine and creatinine compounds in the urine. In view of the fact that creatinine has been shown by isotope studies to be derived from creatine, it would appear from a consideration of these results that, in progressive muscular dystrophy, creatinuria does not arise as a result of an increase in the synthesis of creatine, but rather as a result of the incomplete metabolism of this material in the muscle, with a consequent decrease in the amount converted to creatinine.

In view of the intimate association existing between the exothermic decomposition of adenosine triphosphate and the integrity of the contractile processes of normal muscle, no consideration of muscle disease would be complete without some discussion of the metabolism of this compound. Moreover, since the discovery by Lyubimova and Engelhardt of the adenosinetriphosphatase activity of myosin (12), speculation on the nature of myosin in muscle disease would also appear to be in order. The concentration of adenosine triphosphate in the muscles in progressive muscular dystrophy was found to be considerably lower than in specimens of muscle taken under similar conditions and from corresponding areas in normal subjects; and only

traces of the compound were found in the muscles of patients in whom the disease was advanced. However, absolute values for the concentration of adenosine triphosphate in dystrophic muscle have not yet been obtained, because of the factors which have been previously discussed, namely, difficulties inherent in correction of such values for the variable content of fat, connective tissue, water, etc., of the diseased muscle. The difficulties are further enhanced in this instance by the extraordinary lability of adenosine triphosphate. Determination of the adenosinetriphosphatase activity of myosin prepared from specimens of diseased muscle has been more feasible. Although the myosin content of biopsy specimens of muscle from dystrophic patients appears to be relatively low, it was probably not lower than would be expected from a consideration of the diminished concentration of contractile cells in the specimens. The enzymic activity of myosin prepared from the diseased muscles was approximately of the same order of magnitude as that prepared from normal muscle, when tested for its ability to dephosphorylate adenosine triphosphate.

Although attention has been given thus far mainly to a consideration of the chemical and physiological events occurring in the muscles in progressive muscular dystrophy, there has been no adequate evidence presented to show that the primary seat of the disease is actually contained in the musculature. The possibility that the muscular phenomena may be only secondary to an inherent defect arising elsewhere must be kept constantly in mind, lest we become so preoccupied with the organ in which the major symptoms arise that we fail in our over-all search for the site of the initial disorder which sets the muscular syndrome in operation.

Local and systemic changes in metabolism observed thus far in progressive muscular dystrophy have been quantitative rather than qualitative. Unless, and until, genuine qualitative changes in intermediary metabolism in this disease can be demonstrated, the syndrome will continue to escape classification with the rare diseases of inborn metabolism such as alkaptonuria, cystinuria, and others in which such changes are apparent.

Considered on the basis of incidence in population, progressive muscular dystrophy cannot be regarded as a particularly important syndrome, since happily for humanity it is a comparatively rare disease. Viewed from the standpoint of what the successful solution

of this complex disorder would mean to the future development of medical biochemistry, however, few diseases offer more exciting potentialities. It will be recalled that earlier, in a consideration of those tissues which have been of greatest use in a study of the chemical and physical laws of living matter, scarcely a moment's reflection was needed to place muscle first. Indeed, it was indicated that the physiological basis of activity of any organ or organ structure as we regard it today is but largely an extension of concepts first derived for muscle. Is it too much to hope, therefore, that eventually the solution of the clinical syndrome which we recognize as progressive muscular dystrophy may provide many of the new and fundamental concepts we shall require in order ultimately to achieve the description of disease in terms of the nature of altered physiological and chemical mechanisms at work in the cells and tissues of the affected organism?

References

(1) Bloch, K., and Schoenheimer, R., *J. Biol. Chem.*, **138,** 167 (1941).
(2) Borsook, H., and Dubnoff, J. W., *Ann. Rev. Biochem.*, **12,** 183 (1943).
(3) Caspersson, T., and Thorell, B., *Acta Physiol. Scand.*, **4,** 97 (1942).
(4) Dixon, M., *Manometric Methods.* 2nd ed., Cambridge Univ. Press, London, 1943, Chapter 7, p. 94.
(5) Folin, O., *J. Biol. Chem.*, **17,** 469, 475 (1914).
(6) Gowers, W. R., *Pseudohypertrophic Muscular Paralysis.* Churchill, London, 1879.
(7) Hoagland, C. L., Gilder, H., and Shank, R. E., *J. Exptl. Med.*, **81,** 423 (1945).
(8) Hoagland, C. L., Shank, R. E., and Lavin, G. I., *J. Exptl. Med.*, **80,** 9 (1944).
(9) Houchin, O. B., *J. Biol. Chem.*, **146,** 313 (1942).
(10) Houchin, O. B., and Mattill, H. A., *J. Biol. Chem.*, **146,** 301 (1942).
(11) Levene, P. A., and Kristeller, L., *Am. J. Physiol.*, **24,** 45 (1909).
(12) Lyubimova, M. N., and Engelhardt, V. A., *Biokhimiya*, **4,** 716 (1939).
(13) Rosengart, V., *Bull. Med. Coll. Dnepropetrovsk, U.S.S.R.*, **2,** 87 (1940).
(14) Schoenheimer, R., *The Dynamic State of Body Constituents.* Harvard Univ. Press, Cambridge, 1942.
(15) Shaffer, P., *Am. J. Physiol.*, **23,** 1 (1908).
(16) Shank, R. E., Gilder, H., and Hoagland, C. L., *Arch. Neurol. Psychiatry*, **52,** 431 (1944).

(17) Shank, R. E., and Hoagland, C. L., *unpublished data.*

(18) du Vigneaud, V., Chandler, J. P., Cohn, M., and Brown, G. B., *J. Biol. Chem.*, **134**, 787 (1940).

(19) du Vigneaud, V., Cohn, M., Chandler, J. P., Schenck, J. R., and Simmonds, S., *J. Biol. Chem.*, **140**, 625 (1941).

PHYSIOLOGY AND BIOCHEMISTRY

SURGEON CAPTAIN C. H. BEST, C.B.E., F.R.S., DIRECTOR OF MEDICAL RESEARCH, ROYAL CANADIAN NAVY

THE RESPONSIBILITY of preparing an essay under this broad heading, even in normal times, would be very great. Since, under present circumstances, a comprehensive treatment of the subject would be still more difficult, with the consent of the editor I am limiting my discussion to fields in which I have had personal experience.

Insulin and Diabetes

Four of the central problems in this field are: (a) the mechanism of action of insulin; (b) the further study of the insulin molecule and its synthesis; (c) the etiology of diabetes; and (d) the improvement of insulin as a therapeutic agent.

Mechanism of Insulin Action. The broad picture seems bright and clear. The administration of insulin to recently depancreatized dogs or to the uncomplicated case of human diabetes completely restores the organism. If the treatment is carefully continued and no complications occur, the animal or patient may proceed with an essentially normal existence and is, in fact, difficult to distinguish from the nondiabetic. The deranged metabolism of carbohydrates, fats, proteins, and phosphate compounds is corrected, except that the lack of a

physiological mechanism for the increased or decreased rate of liberation of insulin in the severe diabetic permits wider fluctuations from the normal range of values.

The detailed mechanism of the action of insulin is by no means as clear as its specific effect on the diabetic state might suggest. It was soon learned that the formation of the key polysaccharide—glycogen—was stimulated by insulin and that the burning of sugar was increased. The extremely wasteful and very dangerous breakdown of proteins to sugar and of fat to ketone bodies in the liver is checked by insulin. The formation of fat from sugar is accelerated; and there is some evidence that certain of the phosphate compounds of paramount importance in the provision of energy for muscular contraction may be regenerated at a greater rate when insulin is present. Some workers believe that the amount of adenosine triphosphate in the liver is increased by the action of insulin.

A very great deal has been learned in recent years which adds to our understanding of the pathway of carbohydrate storage and breakdown in the body. It would appear that most, if not all, of these steps may be traversed without the action of insulin, but our knowledge of these processes is not complete. A great step forward is represented by the recent demonstration by Cori *et al.* that insulin releases the inhibition of the conversion of glucose to glucose-6-phosphate, which is exerted by anterior pituitary extract.

While in a diabetic animal deposition of glycogen in both liver and muscle is greatly increased by insulin, in the normal animal muscle appears to have a "higher priority" and deposition of glycogen in this tissue, after insulin administration, may be accompanied by a fall in glycogen content of liver.

We have a great deal more to learn concerning the best established effect of insulin, *i. e.*, the promotion of glycogen deposition, and we have even more to learn about the mechanism by which the breakdown of protein and fat is inhibited. It is possible that insulin in some way controls the substrate which is presented to the tissue cells for metabolism.

The intelligent use of such labeling agents as radioactive phosphorus and the stable isotopes of carbon, nitrogen, sulfur, and other elements is certain to illuminate many of the dark passages through which insulin passes in producing its effect on diabetes. We can

confidently expect further advances in this field, which has been widened and cleared even during the present war.

We have no reliable method for the estimation of insulin in small quantities of blood. Perfection of a technique for this purpose would aid, not only in diagnosing the type of diabetes and perhaps in facilitating the treatment, but in throwing a great deal of light on the very real possibility that insulin concentration in the blood is a factor in the control of insulin liberation from the pancreas. We need either a sensitive chemical test or a microbiological procedure that is capable of detecting the small amounts of the antidiabetic hormone which are certainly present in varying quantities in blood. Thus far, very few vigorous attacks have been made on this important problem.

Chemistry of Insulin and the Possibility of Synthesis. Eleven amino acids have now (1945) been isolated from the crystalline insulin protein. Little progress has been made in the identification of any active chemical grouping which might exert an antidiabetic effect in the absence of the whole insulin molecule. The synthesis of this complex protein awaits further development in protein chemistry; there are so many ways possible in which the various amino acids might be linked that few investigators have thought it worth while to give the matter serious consideration at this time. The problem is attractive enough for its own sake, but when we consider that, if the increase in insulin distribution persists at the present rate, all the available pancreases in the world will be utilized for insulin production in the not distant future, the situation acquires a new urgency. For the last fifteen years, the total distribution of insulin in the United States and Canada has doubled every five years—that is, the amount used in 1935 was twice that in 1930, the amount in 1940 twice that of 1935, and the same rate of increase has been maintained up to the present time.

A vigorous search should be made for a source of insulin other than mammalian pancreas. We know that considerable quantities can be obtained from the principal islets of certain fish but this procedure has never been placed upon a commercial basis.

Etiology of Diabetes. There are no data available to tell us what proportion of diabetic cases is produced by a primary lesion of the pancreatic islet cells. There are many well-substantiated statements in the literature which indicate that there is no obvious lesion

of the pancreas in many instances of human diabetes. On the other hand, there is no convincing proof that the islet cells are really normal in these cases; and many investigators feel that the burden of proof that the pancreas is not primarily or secondarily involved rests with those who make this assertion. Careful estimations of the insulin content of pancreas, if it were possible to do this under reasonably standard conditions in autopsies on diabetics, would yield valuable results. It is obvious, however, that the opportunities to make this study in uncomplicated cases will be rare. The administration of insulin or fasting will lower the insulin content of the pancreas in normal animals, a point which must be considered when insulin estimations are made.

In experimental animals, permanent diabetes can be produced by three procedures: by pancreatectomy, by destruction of the islet tissue by the diabetogenic material of the anterior pituitary gland, and by the administration of alloxan. All three methods effect removal or practically complete destruction of the beta cells of the islands of Langerhans. Substances are, however, present in the anterior pituitary gland which aggravate diabetes in the completely depancreatized organism. A part of this effect appears to be due to a deficient carbohydrate utilization in peripheral tissues.

Apart from the possibility of developmental defects in islet structure which would presumably appear in very young children and the well-established arteriosclerotic changes in the pancreas of older individuals, there is little real evidence on which a reasonable theory of the etiology of diabetes may be supported. The association of the diabetic state with overactivity of the anterior pituitary lobe of the pituitary, as in the acromegalic, encourages further study along a path which has already been well outlined. We must develop accurate procedures for the assay of the diabetogenic materials in blood. It has recently been shown that diabetes may be most favorably affected by removal of the exciting cause which, in a few clinical cases, has been found in one of the adrenal glands. Extracts containing the hormones of the cortex of the adrenal may produce diabetes in the partially depancreatized animal, and it is well established that, under certain conditions, removal of the adrenal cortices ameliorates the diabetic state caused by pancreatectomy in much the same way as does removal of the anterior lobe of the pituitary.

The fact that, in the experimental animal and presumably in many diabetic patients, eight-tenths of the insulin-producing capacity of the islet cells must disappear before it is possible to detect the insulin-producing capacity by any procedure except direct examination, presents a challenge to clinicians and experimentalists alike. It is inconceivable that a relatively clear-cut problem of this kind can long resist the onslaught of the vigorous minds which we hope will attack it now that peace is restored. Many physiological and chemical avenues of approach have not as yet been adequately explored.

Improvement of Insulin as a Therapeutic Agent. The problem of giving insulin in a more physiological way presents many difficulties. It will not be easy to make a compound of insulin which is broken down at an increasing rate as the sugar content of the tissues rises. Improvements in protamine zinc insulin are quite possible and should certainly engage the attention of those who have already made great strides in this field. There are indications that a decrease in the amount of protamine with respect to insulin, may serve a useful purpose. Clear solutions of combined insulin which produce a prolonged effect are desirable, but clinical opinion indicates that a completely satisfactory one is not as yet available. Insulin by mouth for mild cases of diabetes is by no means an impossibility, but it will probably always remain a wasteful procedure and in cases in which dosage of insulin must be very accurately regulated, a hazardous one.

It should be possible in the future to make a more definite separation of the diabetic cases into those which have retained a definite capacity for producing insulin and those in whom this function of the islet cells is irretrievably lost. In the former type, it is obviously useless to plan diets with the idea of conserving islet function. In experimental animals, diabetes resulting from partial removal of the pancreas or from administration of the diabetogenic material of the anterior pituitary gland, may now, under certain conditions, be prevented, or in its early stages cured, by the appropriate use of insulin and diets which do not tax the capacity of the remaining islands of Langerhans. These findings have aroused the hope that a similar application may be made in those cases of human diabetes in which a reasonable number of functioning islet cells remains. This application will not be easy until the potential human diabetic can be recognized much earlier than is possible at present and until much more light is shed

on the etiology of the diabetic state in man. We have suggested that it will be important to determine, by actual clinical trial, the extent to which the results of these animal experiments are applicable to the human subject. We have not suggested that insulin should be used prophylactically in children with a diabetic family history until the experimental results have been carefully tested by competent clinicians. We have outlined many of the obvious difficulties which stand in the way of such a clinical research. There may be many more which are not apparent to the experimentalist.

The four problems discussed briefly are a few among the many in this field which demand investigation. We can look forward with confidence to the development of better diets for diabetics; and a lead may have been supplied by the suggestion that certain fats may be used as a source of energy without the production of excessive amounts of the dangerous ketone bodies. The role of some of the new accessory food factors in the intensity of diabetes has recently been investigated and the use of small animals made diabetic by the injection of alloxan will surely be of great value.

It seems assured that there will be a better control of infections which in many cases constitute the main problem in the treatment of diabetics. Penicillin by its gentle and effective action may well introduce a new era of progress in relation to the control of infection in diabetic cases.

Histaminase and Histamine

Some sixteen years ago, we suggested the term "histaminase" for the enzyme system, shown to be present in various tissues, which inactivates histamine. This system has been partially purified as a result of the efforts of various investigators, although a great deal more remains to be done in this field as well as in the delineation of its physiological significance. Histaminase has been useful, in conjunction with other procedures, in the identification of histamine in various tissue extracts and fluids. It is present in liver, lung, kidney, and the intestinal mucosa, but is absent from the gastric mucosa.

A great deal has been learned about variations in histamine content of tissues and about the liberation of histamine in anaphylactic shock, whereas the physiological role of this substance is far from

completely established. A specific chemical procedure for the determination of histamine would be useful in any subsequent researches. The present methods are not specific because they measure the amount of the imidazole nucleus only.

It has been reported relatively recently that histamine may act as a hapten when it is conjugated with certain proteins and that it is possible to produce antibodies which are partially specific for the histamine hapten.

While the histamine–histaminase system is a very useful one for studies in enzyme chemistry, attempts at commercial exploitation of the enzyme are without justification. There is no physiological basis for the claim made by some clinicians that histaminase exerts therapeutic properties.

Heparin

The discovery of heparin at Johns Hopkins University in 1918 stimulated a new series of researches on blood coagulation. The anticoagulant, isolated in crystalline form in Toronto in 1935, prevents the clotting of one hundred thousand times its own weight of blood. The experimental and clinical demonstration that the pure substance prevents thrombosis without the production of any harmful effects opened the way for its clinical use and aroused anew the interest of physiologists and biochemists.

Heparin is one of the mucopolysaccharides. These compounds are of considerable importance in physiology, as they are the chief constituents of connective tissues. The variation in the properties of this group of substances apparently permits the change in characteristics of connected tissues. Heparin is the first substance of this group to be crystallized. Its basic unit consists of two glycuronic acid residues, two glucosamine residues, and five sulfuric acid residues. Its high content of sulfuric acid probably makes it the strongest organic acid found in the body. Since no other mucopolysaccharide has any significant anticoagulant activity, this property must depend on some unique configuration of the molecule. Of considerable interest is the fact that, while heparin in any species is identical from tissue to tissue, on comparison of heparin from different species variations appear. Thus, the relative anticoagulant potencies of the pure heparins from

sheep, pork, beef, and dog tissues are 1:2:5:10. While proteins differ with the species, there are few known examples of species variations occurring with polysaccharides. Chemical studies of heparin would therefore be of interest from several points of view.

The body contains enormous amounts of heparin. A dog weighing ten kilograms contains approximately ten thousand units of heparin, or sufficient to prevent clotting of ten kilograms of blood. This is all contained in the tissues, since none can be detected in normal blood. What is the function of such a large store of heparin? It has been suggested that it provides an emergency mechanism which can prevent intravascular thrombosis by the liberation of heparin into the vessel threatened by thrombosis. Histological studies have demonstrated that heparin is localized in the mast cells located in the walls of blood vessels. Such a position would be suitable for a function of this nature. Heparin is liberated from the mast cells and found in the blood in peptone and anaphylactic shock. What is the initial stimulus to heparin liberation? Is it the appearance of thrombin, of fibrin, or the agglutination of platelets? How much heparin is liberated in twenty-four hours? Is this increased when there is tissue damage, as after operations? Does the heparin content of the mast cells and tissues remain constant, or does it change with age, sex, operations, etc.? Many of these questions could be quickly answered if we had accurate methods for the determination of heparin in blood and tissues. However, the amount in blood is too small for estimation by our present methods, and the only procedure applicable to other tissues is to attempt isolation by the method developed for one particular tissue—beef lung.

On injection, heparin rapidly disappears from the blood. An enzyme, heparinase, has been prepared from rabbit liver. The purification and study of the distribution of this enzyme has not yet been attempted.

Heparin prevents the clotting of blood and the agglutination of platelets, but we still do not know the mechanism of the action. Heparin when added to oxalated plasma and thrombin prevents the clotting of this system. However, if fibrinogen is substituted for the plasma, heparin has no effect. This action of heparin on clotting therefore depends on the presence of some unknown plasma factor. The anticoagulant effect of heparin on blood does not depend on this

antithrombic action of heparin because heparinized blood contains the normal amount of prothrombin. Since, however, the nature of the interaction of prothrombin, thromboplastin, calcium, and other factors is at present unknown, it is not possible to indicate the particular reaction in clotting prevented by heparin.

It can be seen that many questions remain unanswered in the heparin field, the solution of the problems depending upon the development of new methods and experimental approaches.

The Lipotropic Factors

The recognition of the significance of choline as a dietary factor, and the rapid development of our knowledge of other lipotropic agents such as betaine, methionine, and inositol, have opened a new chapter in physiology and have stimulated almost as much experimental work as did the isolation of insulin. The presence of choline or methionine in the diet is essential for the survival of young animals of various species. Death is apparently due to a failure of liver or kidney function. Three of the main problems demanding investigation are: (a) the mechanism of action of lipotropic agents; (b) the physiological significance of the lipotropic factors; and (c) the clinical significance of the lipotropic factors. Choline apparently acts by virtue of its incorporation as an intact unit into the molecule of phosphatidyl choline (lecithin); and, as a result of this step, the rate of regeneration of the choline-containing lipids of liver, kidney, blood, and perhaps other tissue is increased. Betaine acts as a donator of one of its methyl groups which, by methylating ethanolamine, forms choline. Methionine is also a methyl donor; while it may also exert dramatic effects by providing a source of organic sulfur or perhaps by virtue of its intact molecule, its lipotropic action is accomplished by donation of its methyl group to ethanolamine, *i. e.*, it acts as a precursor of choline. Inositol, which has been well established as a lipotropic factor under certain conditions, has been reported, without adequate evidence, to affect particularly the cholesterol fatty liver produced by various procedures. It may stimulate the regeneration of inositol-containing phospholipids but this phospholipid has not, as yet, been isolated from liver, the only tissue on which inositol has been shown to exert a lipotropic action.

Physiological Significance. It remains to be decided whether a supplement of choline is necessary in a diet which contains an adequate amount of choline precursor, methionine. A small part of the lipotropic action of protein may be due to the tyrosine it contains. The lipotropic action of inositol is inhibited in young animals by some component of corn oil, perhaps the essential fatty acids. Work which is in progress should determine the relative importance of these various agents under physiological conditions. The addition of choline to the diets of puppies given an alipotropic ration, prevents failure of liver function and death in this species. Presumably, methionine would have the same effect. Choline or methionine prevents the development of haemorrhagic kidneys which may lead to the death of young rats through failure of kidney function. The extension of this work may throw some light on the etiology of developmental defects in kidneys in man.

Hepatic necrosis in rats is prevented by cystine and hepatic cirrhosis by choline. Methionine, by providing a source of cystine or of organic sulfur, prevents necrosis and, by enabling the body to make choline from ethanolamine, prevents cirrhosis.

Clinical Significance. The provision of diets high in protein may exert a favorable effect on hepatic cirrhosis in man. It is reasonable, in the light of the experimental evidence, to consider that this effect may be due to the methionine contained in protein which acts as a source of choline. A little evidence has been provided that choline is effective in human hepatic cirrhosis.

Evidence has been obtained that the high incidence of human hepatic cirrhosis in certain districts of South Africa may be due to the low "protein" content of the diet. This diet produces hepatic cirrhosis in rats. We can no longer merely speak of *protein* as a dietary constituent without remembering that some of the amino acids, the amounts of which vary in different proteins, have specific physiological functions. At the moment, we are particularly interested in the methionine contained in protein which, as stated above, provides a source of choline.

Various claims have been made for the beneficial effect of choline in preventing liver injuries due to a variety of infectious processes. Methionine is now being studied extensively by clinicians in the hope that its well-established physiological action can be utilized

in the solution of clinical problems. It will require a great deal of time and effort to determine the relative importance of the various possible actions of methionine, that is, as a source of organic sulfur, as a source of a specific amino acid, or as a precursor of choline. It is not reasonable to expect that the lipotropic factors will exert beneficial effects when they are added to diets which already contain an abundance of these agents unless clinical conditions exist in which their absorption or utilization is abnormal.

The metabolism of fat which is, of course, intimately related to that of protein and carbohydrate, can no longer be thought of as the sluggish stream it was widely considered to be relatively few years ago. From some viewpoints it is now easier to study the intermediate metabolism of fats than that of carbohydrate or protein, and we can confidently look forward to great progress in this field during the next few years.

War Medical Research

With the outbreak of war in September, 1939, many productive peacetime medical studies were immediately discontinued. Our enemies had apparently made this move many years previously and thus had, for a time, the great advantage of the results of deliberate and detailed planning for the medical aspects of warfare. Several busy years were required to balance the scales anew.

Our experiences in war medical research in Canada are essentially the same as those of our allies with whom and by whom problems and solutions have been freely shared. The primitive state of our organization for medical research in the Armed Services at the beginning of the war may have introduced an element of novelty into some of the arrangements subsequently made in Canada. In the Air Force and Navy, medical research groups from university departments were set up, in affiliation with the Medical Committees of the National Research Council of Canada, as the nuclei of Service Medical Research Units. The research workers, some placed on active service and some as civilians supported by the National Research Council, continued to use the university buildings and facilities for their Service researches. In the Naval Service, this arrangement has continued with little change up to the present time. In our particular unit, the first re-

searches were on shock, blood substitutes, nutrition, lighting, and the development and application of new tests of vision, hearing, and frequency discrimination. These were followed by attacks on the special and very numerous physiological problems associated with immersion foot, cold and hot weather protective clothing, ventilation, diving, underwater blast, etc.

In spite of the extremely satisfactory progress made under wartime conditions in many countries, a host of Service medical problems remain unsolved at the end of the war. No attempt will be made to list them here. An important matter of concern is that, in our eagerness to return to peacetime medical research, we shall not forget the years when we narrowly missed becoming hopelessly enmeshed in the "coils of slack" that had been allowed to accumulate.

28

X-RAY DIFFRACTION AND THE STUDY OF FIBROUS PROTEINS

I. FANKUCHEN, ASSOCIATE PROFESSOR OF CRYSTAL CHEMISTRY, POLYTECHNIC INSTITUTE OF BROOKLYN

H. MARK, DIRECTOR, POLYMER RESEARCH INSTITUTE, POLYTECHNIC INSTITUTE OF BROOKLYN

THIS ESSAY will review some recent studies with x-rays of the structure of proteins and will discuss a few questions which are still unanswered, but which, presumably, may be successfully attacked with the aid of x-ray or electron diffraction within the not too distant future.

The general purpose in applying x-rays for the elucidation of the structure of matter is to obtain information about the exact position of atoms or ions inside a crystal or molecule. Such information can be obtained from diffraction patterns because the diameters of atoms like carbon, nitrogen, and oxygen are of the same magnitude as the wave length of x-rays as they are produced in commercial tubes. Both are between one and two angstrom units, one angstrom unit being 10^{-8} centimeter. The x-rays diffracted from the three-dimensional atomic lattice of a crystal or a large molecule show particularly high intensities in certain preferred directions, from which one can figure out the distances between the various atoms in the sample.

This line of attack first led to a thorough understanding of the fine structure of most simple inorganic and organic crystals such as

diamond, calcite, quartz, urea, benzene, etc. Later it was possible also to elucidate the molecular structure of more complicated systems such as silicates, dyestuffs, and hormones; and recently the x-ray diffraction method has been successfully applied even to such complicated systems as cellulose, starch, rubber, and protein. Molecular models have been worked out for these high polymeric substances which reflect in a fairly satisfactory way the various chemical and physical properties of the materials.

Proteins of all kinds have for a long time been extensively studied with x-rays. For recent comprehensive articles see particularly references 1, 10, 13, 14, 15, 21 and 30. By 1920 diagrams of several *fiber proteins* had been obtained by Herzog and his collaborators (8, 19); *virus proteins* were first successfully studied in 1936 by Bernal and Fankuchen (7); and *single protein crystals* have been investigated systematically by a number of authors since 1939 (11,12,27). Each of these three types of proteins has challenged the x-ray method in a different way; and the investigations have both originated improvements of the experimental technique and led to refinements of the theoretical interpretation of the diagrams. In this essay we shall describe briefly how fibrous proteins have been studied with x-rays, and will then proceed from the simple qualitative inspection of the diagrams to their more quantitative evaluation.

Typical X-Ray Patterns of Fiber Proteins

Figure 1 shows the x-ray diagram of a bundle of native silk fibers taken with a well-filtered and collimated beam of Cu K radiation. Patterns of this kind can be obtained in about one hour or less with the aid of modern commercial x-ray tubes now being produced in this country by various companies. Since the taking of the diagram itself, until a few years ago a difficult procedure, is now more or less a matter of professional routine, all care and time can be spent on the preparation of the specimens and on the interpretation of the diffraction pattern. It can be seen that the pattern of Figure 1 consists of a number of diffraction spots of moderate sharpness and varying intensities which are arranged symmetrically about the axis of the incident beam.

The primary beam is very intense and would, therefore, produce a large, black spot in the center of the diagram, which might overshadow

part of the pattern. It is therefore customary to stop the primary beam with a small lead plate, the shadow of which can be seen in the center of Figure 1. In addition to these spots, one observes a fairly distinct accumulation of scattered intensity in the immediate neighborhood of the incident beam and a faint, but not negligible, diffuse background upon which the diffraction spots are superimposed.

Theoretical considerations (see 9,24) and extensive experience with x-ray diagrams of many inorganic and organic materials indicate

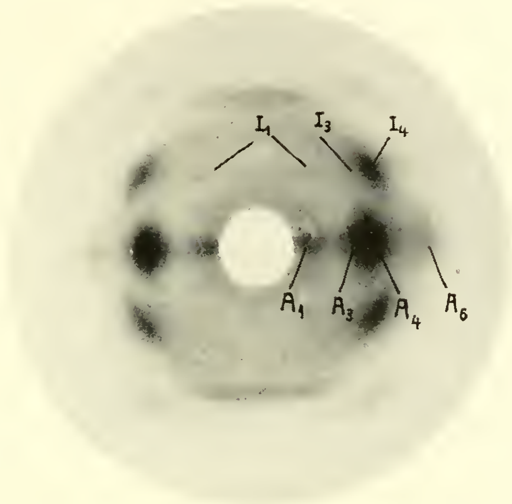


Fig. 1.—X-ray diagram of commercial silk fibroin (silkworm). Beam normal to fiber axis (24).

that, under favorable conditions, conclusions of the following character can be drawn from the analysis of such patterns.

Mere *qualitative inspection* permits the conclusion, on the basis of the *existence* of *distinct* spots, that the sample investigated contains geometrically organized areas inside of which long-chain molecules of the silk protein are localized in a strictly periodic manner, like the individual bristles in a paint brush or the single pencils in a package of pencils. The presence of distinct *spots* instead of rings indicates that these highly ordered domains all have one of their axes parallel (or very nearly parallel) to the common axis of the investigated fibers,

comparable with many paint brushes the handles of which are all parallel. Using a somewhat abbreviated and oversimplified expression, one can also say that the diagram indicates that the sample contains a *multitude* of *oriented crystallites* of silk protein. The fact that each of the spots is larger than the actual geometric projection of the irradiated part of the fiber bundle and that their boundary is somewhat diffuse indicates that the highly organized areas consist of only a limited number of strictly periodically arranged elements, such as atoms, groups of atoms, molecules, etc. In other words, the oriented "crystals" of the fiber protein (or, in our analogy, the bundles of pencils) are *very small.* The presence of the intensive halo around the incident beam —compare Figure 1—shows that there exist in this sample structural heterogeneities the dimensions of which are larger than the wave length of the irradiated light. Cu Kα radiation has a wave length of about 1.5 Å., and the structural heterogeneities as indicated by the central blackening have dimensions of about 100 Å. Finally, the general diffuse background of the pattern is caused by the presence in the sample of a certain amount of disordered material which may, but need not necessarily, be chemically different from the more highly ordered constituents. This would be comparable to a situation in which the bundles of pencils mentioned above are embedded in a mass of randomly arranged pencils pointing in all possible directions.

Using the same simplicity and abbreviation as above, we may summarize by saying that the qualitative inspection of a diagram such as shown in Figure 1 leads to the conclusion that the investigated material consists of very small, highly oriented crystallites, which are embedded in an amorphous matrix in such a way that certain long-range, quasiperiodic *spacings* are somewhat accentuated. Each of these statements corresponds to a special feature of the diagram, as represented in Table I.

It may, perhaps, be appropriate to add a few brief remarks about the justification of the use of "crystallites" and "amorphous areas" in the qualitative interpretation of x-ray diagrams of this kind. With reference to the crystallites, or micelles, we shall consider these to be small volumes of somewhat indefinite size and shape, inside of which the long-chain molecules are arranged according to a fairly regular three-dimensional periodic pattern. There are also, however, in every sample of a high polymer, certain portions which are not crystallized.

These are usually referred to as the *amorphous*, *disordered*, *intermicellar*, or *glassy* fractions. Many investigators, particularly Baker and Fuller (4), Gehman and Field (17), Goldfinger *et al.* (18), and Purves and his collaborators (28) have used a variety of methods to obtain information about these regions in a high-polymer material. It seems that the regions contain the molecular chains in a less perfect arrangement, because, for instance, some irregularity may have prevented them from reaching the proper equilibrium positions corresponding to the crystal-

Table I

Qualitative Inspection and Interpretation of a Typical Fiber X-Ray Diagram

Observed phenomenon in the pattern	Conclusion about the structure
Comparatively distinct, intensive spots of diffracted intensity	Presence of highly organized areas (so-called crystallites)
The spots are not drawn out into long segments or rings	The crystallites in the sample possess a certain orientation
The spots have a certain radial width and their boundary is somewhat diffuse	The crystallites are very small (200 Å. or less)
There is a general diffuse background	The sample contains a certain amount of disordered (amorphous) material
Comparatively intensive halo around the incident beam	Certain long-range quasiperiodicities exist in the samples

lized state. It is believed that one and the same chain can pass through a crystalline area, enter an amorphous portion, go right through it, and enter another crystalline area. This leads to the conclusion that there is no sharp boundary between the crystallized and disordered domains but that the chains of a certain crystalline region somehow became disordered, degenerate into fringes, and, finally, reach a completely disordered arrangement. Hence it may be appropriate, not to make a sharp distinction between chains in a crystallized or amorphous state, but rather to consider *various degrees of disorder*, just as one may observe metallic mixed crystals which have the same composition but a different arrangement of their constituents. Baker, Fuller, and Pape (4) have suggested the existence of a *mesomorphous phase* in such materials as cellulose acetate or nylon; and Taylor (29) has repeatedly empha-

sized that chains with regularly distributed centers of attraction are unlikely to curl up in a completely random way but will lead to structures of an intermediate degree of order.

With these limitations and restrictions in mind, we shall now proceed to a more quantitative discussion and interpretation of fiber diagrams.

Small- and Large-Angle Diagrams

Inspection of Figure 1 has already shown that there are two distinct areas in which scattered intensity can be observed:

(a) In the immediate neighborhood of the incident beam, *i. e.*, within very *small angles* of scattering.

(*b*) All over the photographic film, *i. e.*, at comparatively *large angles* of scattering.

Because of the reciprocal nature of all diffraction phenomena,* *small-angle scattering* permits the study of *large distances* inside the sample, while *large-angle scattering* permits the investigation of short-range order. It has, therefore, become conventional to divide the quantitative study of a fiber diagram into the evaluation of the "large-angle" and of the "small-angle" pattern.

If one measures the position and intensity of all observable points in the large-angle region, one can, first of all, compute from the position of each spot the diffraction angle and, with the aid of Bragg's law, the characteristic spacing which is responsible for each individual spot. In this way one ends up with a table of spacings and relative intensities of all observed diffraction points. In Table II, such a list is given for the diagram of silk protein (Fig. 1, page 441). The "evaluation" of the diagram is now based on the assumption that the spacings of the lattice planes as they appear in Table II are the consequence of the arrangements of all individual atoms in a three-dimensional lattice; and the first goal is to find the fundamental distances and angles of this lattice. Generally, this can be accomplished from x-ray data *alone*

---

* The famous Bragg diffraction law demands that the *product* of the scattering spacing, d, and the sine of the glancing angle, θ, must equal half the wave length of the scattered radiation in order that an *intensive diffracted beam* be produced. Small spacings produce, therefore, large diffraction angles, and large spacings, small angles.

if one has a number of single crystal diagrams, taken under very well-defined geometric conditions, each diagram containing a large number of individual spots. However, if one is willing to combine x-ray data with other evidence, coming from chemical analysis or from physico-chemical investigations, it is possible to advance well-founded structural proposals, provided that the investigated samples exhibit a so-called twofold orientation, such as stretched or rolled thin films of cellulose, rubber, or collagen (5,24). However, if one has only diagrams with

TABLE II

A FEW SPACINGS OF LATTICE PLANES IN THE LATTICE OF NATURAL SILK

Number of reflection in Figure 1	Sine of Bragg's angle, θ	Lattice plane spacing, d, Å.
A_1	0.084	9.2
A_3	0.178	4.3
A_4	0.251	3.1
A_5	0.324	2.4

axial symmetry (so-called fiber diagrams), which contain spots of the number and character as shown in Figure 1, it is, *in general*, not possible to arrive at reliable quantitative conclusions regarding the exact size and shape of the elementary cell of the lattice of the crystalline constituents.

It seems that in some *special* cases, such as the protein in native silk (8), the protein in native wool and hair called β-keratin (3,26) and α-keratin (2,3), and the protein in muscle called collagen (5,6) rather reasonable proposals for the lattice parameters have been suggested by supplementing the x-ray data with chemical and morphological data, and with the aid of a certain amount of intuitive guessing. However, even in these cases one cannot, at present, advance well-founded data as to the *exact location* of the various substituents along the main polypeptide chains in the lattice, although one knows their approximate arrangement. Summarizing, it may be said that the large-angle x-ray study of fiber proteins to date has furnished enough data to arrive at rather probable structures for the most important protein fibers if one combines the x-ray data with all the other physical and chemical knowledge of these materials.

If we now concentrate our attention on the immediate neighborhood of the incident beam in Figure 1, we find there is a distinct blackening of the film in this region. The intensity scattered at *small angles* corresponds to *large distances* inside the investigated sample. If these distances are arranged in some kind of periodic sequence, they produce, according to Bragg's law, diffraction maxima the intensity and sharpness of which depend upon the perfection of the periodic long-range structure. If, therefore, the small-angle scattering consists of a number of (more or less) well-defined lines, one concludes that something like a "superlattice" exists in the crystalline domains of the sample. Figure 2 shows a typical case of such a long-range order in a suspension of tobacco mosaic virus (7). The very elongated rodlike particles of this virus can be perfectly oriented in their aqueous suspension even at such low concentration as 2 or 3%, and assume an order similar to that in a bundle of pencils: their axes are all parallel and they maintain fixed average distances from each other. If such a structure is irradiated with x-rays perpendicular to the axis of orientation, patterns such as those in Figure 2 are obtained. The sharp lines found on the equator of the diagram at small angles indicate that there is a certain periodicity *perpendicular* to the axis of orientation, which roughly corresponds, in the case of the dry virus, to the *average thickness* of the rodlike parallelized virus particles. In the case of tobacco mosaic virus, this thickness was found to be about 150 Å.

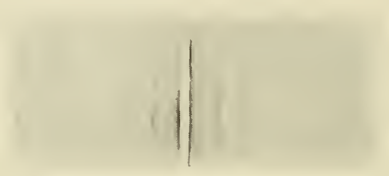


Fig. 2.—X-ray diagram of tobacco mosaic virus. Beam normal to molecular direction.

Comparatively sharp reflections at small angles on the *meridian* of protein fiber x-ray diagrams have also been observed, indicating the existence of a certain long-range periodicity *along* the axis of the parallelized and aligned molecules. Astbury (3) has observed that β-keratin exhibits meridional small-angle reflections, which correspond to a spacing of about 66 Å.; he suggests they are either the third order of a period of 198 Å. or the tenth order of 660 Å. Porcupine quill and feather rachis indicate long periods of 198 and 95 Å., respectively (23). Beef tendon (5) exhibits about thirty orders of a fundamental

period of 640 Å. along the fiber axis; clam muscle fibrils (22) show indications of a long fiber period of 700 Å. In the latter two cases, independent observations with the electron microscope led to similar periodicities in the investigated samples. Bear (6) reports a long fiber period of about 600 Å. for a large number of collagen specimens; and

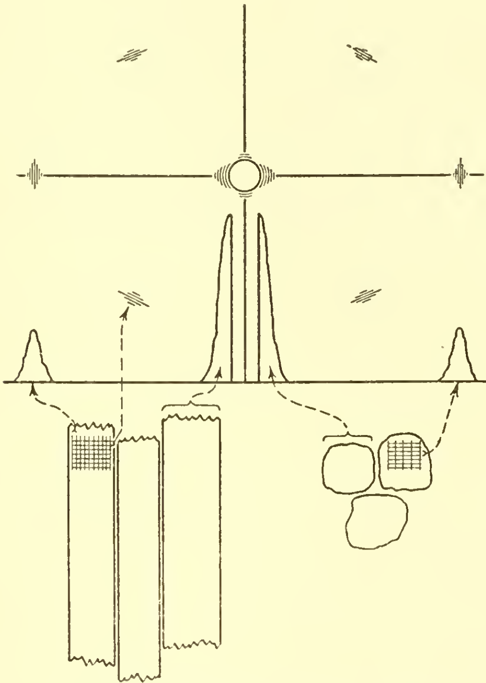


Fig. 3.—Small and large angle scattering of elongated, parallelized crystallites (16).

it is interesting to note that synthetic polyamides also show distinct meridional small-angle reflections, which lead to long fiber periods of about 90 Å. (16,20).

Figure 3 shows diagrammatically the small- and large-angle scattering of a highly stretched and subsequently relaxed nylon filament and indicates the relations between the various reflections of the diagram with the corresponding distances in the sample.

Sometimes the intensity scattered at small angles does not exhibit distinct lines or spots but is of a more continuous character, decreasing gradually with increasing diffraction angle. This points to the fact that there are no repeating, *well-defined* long dimensions within the sample but that these distances either scatter rather widely about a certain average value or are randomly distributed in the sample. In such cases, one must investigate quantitatively the whole intensity distribution around the primary beam before one can draw conclusions about the length of this pseudoperiodicity (16) or more generally say something about the distribution of these long dimensions within the specimen.

Size of Crystallites, Degree of Orientation, and Crystallinity

It has already been mentioned that the sharpness and intensity of the lines or spots of the small-angle diagram depend upon the number of repeating periodicities and upon the accuracy with which they repeat. This is also true for the lines or spots of the large-angle diagram. The more frequently a given lattice plane is repeated in one of the crystalline regions, the more numerous will be the secondary waves which cooperate with a favorable phase difference and, hence, the sharper and more intense is the diffraction spot which they produce. Sharp lines or spots indicate large crystals or crystalline regions; diffuse lines indicate small crystals. In most fiber protein diagrams, the large-angle diffraction spots are fairly broad and indicate that the average width of the bundle-shaped crystallites in a highly oriented protein fiber, such as silk, muscle, or feather, is about 100 Å. or less. Along the axis of orientation the diffraction lines are sharper and, hence, the domains of crystalline order are longer. Stretching and alignment usually increase the size of the crystalline areas, whereas swelling and relaxing decrease it. If the ordered domains become smaller than 20 or 30 Å., the x-ray diagram of the system becomes more and more diffuse and finally approaches that of a liquid.

Next to the approximate average *size* of the crystalline areas, one is interested in the *way in which they are oriented* in a given sample. Completely random arrangement of the scattering elements in a sample produces an x-ray diagram which consists exclusively of *rings*. When some or all of these rings degenerate into segments or spots, the presence

of a certain degree of orientation is indicated. Quantitatively one can measure the length of each individual segment, parallel to the circumference of the ring to which it belongs, and convert this length into the angular opening of the segment. This angular opening describes the degree of orientation in the investigated material: if it is large, the crystallites scatter widely about certain directions and the orientation is poor; if, however, the angles corresponding to all segments in the diagram are small, *i. e.*, the diagram consists of sharp spots rather than smeared out segments, then the orientation of the crystalline domains in the sample is good. In general, orientation of the crystallites is closely connected with important mechanical properties such as strength, rigidity, or toughness.

The last feature of a protein fiber x-ray pattern, which deserves quantitative consideration is the general background of diffusely scattered intensity. This indicates that, in addition to the well-ordered fraction producing the comparatively sharp diffraction lines and spots, a certain amount of material exists which is in a *disordered state* and which cannot produce any characteristic diffraction phenomenon. This *amorphous* or *disordered* matrix in which the crystalline areas are embedded is mainly responsible for the resilience, flexibility, and swelling properties of a protein fiber and hence may represent a very valuable constituent of it. In principle, it would be possible to carry out a measurement of the integrated intensity of this diffuse scattering and by comparing it with the total intensity of the crystalline pattern to arrive at a *quantitative figure* of the ratio of crystalline to amorphous material. In practice, however, it is possible to obtain this ratio only as an estimate rather than as a precise measurement. One can, for example, find out whether a certain treatment of a sample has increased or decreased its degree of crystallinity. In general, procedures such as stretching, rolling, drawing, deswelling, etc., increase the degree of crystallinity, while all steps of relaxation and swelling tend to increase the disorder in the material. Since the ratio between ordered and disordered constituents is of great importance for the mechanical and colloidal properties, even an approximate estimate of it can contribute to an understanding of the behavior of certain systems.

Until recently, comparatively primitive x-ray diffraction techniques were used to study proteins—primitive in the sense that they were not the most advanced available *at the time* of the experiments in

question. More recent papers exhibit an awareness of this fact; and it may be expected that much of the forthcoming work in this field will utilize not only the modern developments of technique but also those in interpretation.

In technique, the new approach starts, properly, with the specimen. While some protein fibers (wool, silk) can be handled with ease, the study of others (like muscles) *must* be accompanied by extreme precautions to prevent any appreciable changes from their natural state. Because the use of normal specimens usually introduces *many* fibers into the x-ray beam, the tendency is to reduce the size both of the specimens and of the x-ray beam. This trend can be carried to quite an extreme degree so that single fibers and beams of 0.05 mm. diameter have been used (16); the results indicate that even much finer beams could be economically employed. When the specimen is a single intact fiber, the x-ray diagram is free from the possibility of disorientation due to lack of parallelism between the individual fibers making up a bundle.

More and more intense sources of x-rays are being made available to the investigator and these are shortening the time required to obtain an x-ray diagram. A more exciting possibility arises from the introduction of Geiger-Muller counters to record the scattered radiation. This innovation permits the almost instantaneous determination of the intensity of scattered radiation and will allow the planning of many experiments which are impossible by photographic methods. Studies of structure as a function of time (a contracting muscle, perhaps) could be made now that a continuous recorder of scattered radiation is available. It will also be possible to study specimens which deteriorate rapidly.

Efficient monochromatization of the x-ray beams will also be useful in many cases. Biological specimens often give poor x-ray diagrams regardless of technique; and the additional background due to incomplete monochromation may be enough to obscure important details in the x-ray diagram.

The interpretation of the data is also being steadily improved. Increased use of reciprocal-lattice concepts and the more widespread realization of the shortcomings of ordinary fiber diagrams (*i. e.*, diagrams from specimens which possess orientation with reference to one direction only) lead the experimenter to attempt to prepare

specimens of higher orientation. As we have seen, much more *certain* conclusions can be drawn from such studies.

Fourier methods have been widely used in the study of single crystals (both protein and nonprotein). There seems to be no good reason why these methods should not yield useful results when applied to fibrous protein. Certainly, many recent experiments have yielded data suitable for such studies.

When one considers in retrospect the considerable effort that has been expended in the study of protein structures with x-rays, perhaps the conclusion is inevitable that the return has been somewhat meager; but at the same time there seems to be considerable justification for an optimistic view of what is soon to come. The methods and techniques of x-ray diffraction which are *now* available should tell us much that we want to know about the structure of proteins.

References

(1) Astbury, W. T., *Cold Spring Harbor Symposia Quant. Biol.*, **2,** 15 (1934.)
(2) Astbury, W. T., *Ann. Rev. Biochem.*, **8,** 113 (1939).
(3) Astbury, W. T., and Sisson, W. A., *Proc. Roy. Soc. London*, **A150,** 533 (1935).
(4) Baker, W. O., Fuller, C. S., and Pape, N. R., *J. Am. Chem. Soc.*, **64,** 776 (1942).
(5) Bear, R. S., *J. Am. Chem. Soc.*, **64,** 727 (1942).
(6) Bear, R. S., *J. Am. Chem. Soc.*, **66,** 1297 (1944).
(7) Bernal, J. D., and Fankuchen, I., *J. Gen. Physiol.*, **25,** 111 (1941).
(8) Brill, R., *Ann.*, **434,** 204 (1923).
(9) Buerger, M. J., *X-Ray Crystallography.* Wiley, New York, 1942.
(10) Bull, H. B., in *Advances in Colloid Science*, Vol. I. Interscience, New York, 1941, p. 1.
(11) Crowfoot, D., *Chem. Revs.*, **28,** 215 (1941).
(12) Crowfoot, D., and Riley, D., *Nature*, **144,** 1011 (1939).
(13) Fankuchen, I., *Cold Spring Harbor Symposia Quant. Biol.*, **9,** 198 (1941).
(14) Fankuchen, I., *Ann. N. Y. Acad. Sci.*, **41,** 157 (1941).
(15) Fankuchen, I., in *Advances in Protein Chemistry.* Vol. II, Academic Press, New York, 1945.
(16) Fankuchen, I., and Mark, H., *J. Applied Phys.*, **15,** 364 (1944).
(17) Gehman, S. D., and Field, J. E., *J. Applied Phys.*, **10,** 564 (1939).
(18) Goldfinger, G., Siggia, S., and Mark, H., *Ind. Eng. Chem.*, **35,** 1083 (1943).

(19) Herzog, R. O., and Jancke, W., *Ber.*, **53,** 2162 (1920).
(20) Hess, K., and Kiessig, H., *Naturwissenschaften*, **31,** 171 (1943).
(21) Huggins, M. L., *Chem. Revs.*, **32,** 195 (1943).
(22) Jakus, M. A., Hall, C. E., and Schmitt, F. O., *J. Am. Chem. Soc.*, **66,** 313 (1944).
(23) McArthur, I., *Nature*, **152,** 38 (1943).
(24) Meyer, K. H., *Natural and Synthetic High Polymers.* Interscience, New York, 1942.
(25) Nickerson, R. F., *Ind. Eng. Chem.*, **32,** 1454 (1940); **33,** 85, 1022 (1941).
(26) Pauling, L., and Niemann, C., *J. Am. Chem. Soc.*, **61,** 1860 (1939).
(27) Perutz, M. F., *Nature,* **149,** 491 (1942); **150,** 324 (1942).
(28) Purves, C. B., Asaaf, A. G., and Haas, R. H., *J. Am. Chem. Soc.*, **66,** 59, 66 (1944).
(29) Taylor, H. S., at several lectures, particularly during the discussion at the meeting of the American Chemical Society in Memphis, Tennessee, April, 1942.
(30) Wrinch, D., *Trans. Faraday Soc.*, **43,** 1368 (1937).

IMMUNOCHEMISTRY

MICHAEL HEIDELBERGER, PROFESSOR OF BIOCHEMISTRY, COLLEGE OF PHYSICIANS AND SURGEONS, COLUMBIA UNIVERSITY; CHEMIST TO THE PRESBYTERIAN HOSPITAL, NEW YORK

THIS WAS the title, perhaps whimsically given, of a series of lectures delivered in 1904 by Svante Arrhenius, a great Swedish physical chemist, at the University of California, and published in book form (1a) three years later—so hectic was the pace in immunology. True, Ehrlich (5) had insisted on the chemical nature of immune processes and immune reactions, but his views were overshadowed by the widely accepted and facile concepts of Bordet (2a), whose colossal contributions to immunology gave him vast influence. He first classified immune reactions as essentially physical, then, when this position became untenable, maintained that immune reactions were "colloidal." The persuasive, amorphous terminology of the early colloid chemistry, descriptive of everything but accounting for little, found its way into books and papers on bacteriology and immunology, where, be it sadly whispered, it may often still be seen. Doubtless, many of those who once found it useful and comfortable still think of it with nostalgia.

But Arrhenius, with his Danish pupil, Madsen, proceeded to show how, with certain assumptions, the laws of classical chemistry could be applied to typical immune reactions, and so the term *immunochemistry* came into serious use and was seriously taken, even though the proposed analogy of antigen–antibody reactions to the union of weak acids and weak bases was quickly shown to be invalid.

Looking back from the present admittedly none too elevated observation post, in which one must still duck occasional snipers' shots, one cannot fail to be astonished that so many sound observations and so many keen deductions were made in those early days. Next to nothing was known of the chemical nature of antigens and antibodies, the protagonists in the drama of immunity, while complement, the powerful mediator of cell lysis, flitted in and out like the ghost in *Hamlet*, between material existence and a mere "colloidal state" of something else. Worse still, measurement of quantities of any of these elusive unknowns could only be carried out in relative terms, such as by the volume which just would kill or not kill an animal, or by the dilution at which an inflamed skin area or a haze in a test tube just faded out. Immunology, and the early immunochemistry as well, staggered under the dictatorship of dilution and the tyranny of titer.

Relief, happily, was just around the corner. Discovery of the specific polysaccharides (2,8) in Avery's laboratory gave concrete direction to the relation between chemical constitution and bacterial specificity. Extension of chemical concepts to protein antigens was facilitated by the imaginative and painstaking studies under Landsteiner's (29) direction, in which aromatic radicals, optically active acids, and amino acids were coupled to proteins through the diazo reaction and their effects on specificity studied. As for antibodies, prolongation of the discussion as to whether these were serum globulins or unknown substances adsorbed on globulins seemed futile after Felton's simple method for the concentration of pneumococcus antibodies led him to the demonstration that zinc and aluminum salts of these globulins were completely precipitable by the homologous type-specific polysaccharide of pneumococcus (5a).

Though the chemistry of antigens and antibodies began to be better understood, the fetters of the old, relative analytical methods remained unbroken until Kendall and the writer (11), putting their faith in the rigorous criteria of analytical chemistry, devised quantitative micro methods for the accurate estimation of many antigens and antibodies in absolute terms, that is, units of weight rather than of titer. With Sia (19) they demonstrated the parallel between mouse protection and the amount of antibody nitrogen precipitated from anti-pneumococcus type I serum by the specific polysaccharide of type I pneumococcus. This was in 1930, but only in 1938 did the "brass

hats" permit a substitution of antibody nitrogen values for mouse protection, and then only in rabbit antisera to certain of the other pneumococcus types. Millions of mice were finally saved, however, from succumbing to protection tests by the advent of sulfa drugs and penicillin, which made antipneumococcus sera practically obsolete. Although large volumes of antipneumococcus horse sera, at least, have enriched the effluent from certain production centers of biologicals in order to release storage space, infants and children are still being cured of influenzal meningitis by known amounts of antibody from rabbit antisera to *H. influenzae* type b (1), for a drug to cure this dread disease has not yet been found,* nor is the mouse protection test free from difficulties in both its application and interpretation.

Combining proportions by weight in the precipitin reaction were studied, and with their aid a quantitative theory of this reaction was derived from the law of mass action, leading to a simple linear relation applicable to many immune systems (12). The data obtained were most easily explicable on the basis of the union of multivalent antigen and multivalent antibody, giving concrete and quantitative mathematical expression to the less definite "lattice" or "framework" theory of Marrack (30).

The assumptions which were made admittedly involved oversimplification, but this seemed pardonable in a pioneer theory which explained much that had appeared mysterious and provided simple, usable equations that could be readily understood. There was time enough for taking care of the complications later, as several friends and colleagues have done. Pauling and his group (33), brushing aside the original assumptions as "arbitrary and unlikely," have made a different application of the laws of chemical equilibrium but arrived at the same approximate equations (were their faces red!). By a statistical approach Kendall (28) has paralleled and outstripped the original theory, once more obtaining equations of the same form for the precipitin reaction and succeeding in explaining quantitatively even the limited range of toxin–antitoxin flocculation. One would like to say more and know more about Hershey's (23) painstaking and rigorous, but painfully complicated, analysis. Boyd (3) has been a consistent Rightist in his attitude toward the theory of mutual multivalence, holding to the old notion that antigen–antibody combination

* Streptomycin shows promise (H. E. Alexander, *Science*, in press).

is a static affair and aggregation a nonspecific aftereffect. Insisting that the evidence for the multivalence of antibody, at least, is inadequate, he has produced no satisfying alternatives and his publications show evidence that underground movements are weakening his hold.

Study with Kabat (9) of a system of bacterial agglutination showed this reaction to be subject to laws similar to those underlying the precipitin reaction, and for the first time placed the important function of salts in a perspective consistent with modern protein chemistry. But a theory that does not lead to useful and exciting predictions is like a sterile "Man O' War." In its explanation of salt effects observed in a study with Teorell (14), the quantitative theory developed with the new analytical methods led to a prediction as to how analytically pure antibody globulin might be obtained (13,14); and this was realized in two laboratories, (5b,10). Since, with Pedersen and Kabat (17,24) it was shown, with the help of Svedberg and his ultracentrifuge and Tiselius (35) and his electrophoresis apparatus, that purified antibodies have the properties of serum globulins, the protein nature of antibodies can no longer be doubted.

The technically important toxin–antitoxin reaction has been worked out analytically by Pappenheimer and Robinson (32), so that it is now possible, aided by a single pair of nitrogen estimations, to measure the unitage both of an unknown diphtheria toxin and an unknown antitoxin. To Ehrlich this would have seemed a prodigious feat, but the "brass" of immunology has failed to legitimize it.

Quantitative immunochemical methods are, however, being applied in many other directions. They were used, with Mayer, for a study of reversibility and velocity in the precipitin reaction (20,31). Through further modifications, with MacPherson, they have been extended to the measurement of micrograms of antibody in human sera (7,16). The new methods have also supplied an absolute measure of complement, furnishing a weight unit independent of the lytic function of this strange and little understood complex (6,15,18,22). With this measure, the reacting proportions of complement, hemolysin, and red cells could be calculated, as also those of antigen, antibody, and complement involved in the process of complement fixation. Oddly enough, complement could be fitted handily into the quantitative theory of the precipitin reaction, which had been elaborated in blissful neglect of what might have been an embarrassing apparition.

For the first time, complement "belonged" and this Flying Dutchman of immunology was at rest.

Stepping out with these quantitative methods, immunochemistry has also contributed to the interpretation of isotope studies in protein metabolism. Compounds of heavy nitrogen were found by Schoenheimer and his group to slip in and out of body proteins and their constituent amino acids with unexpected speed and thoroughness. In collaboration with these biochemists it was found possible to work not only with labeled nitrogen but with labeled protein—antibody—as well. In a rabbit immunized with type III pneumococci, antibody, quantitatively separable from the other serum globulins, was found to take up dietary nitrogen at about the same rate as these other proteins (21,34). All of this seemed eminently proper until another test was made adopting Treffers' idea of injecting a similarly immunized rabbit with a different antibody, type I pneumococcus anticarbohydrate, from another rabbit. The type III-immunized rabbit was fed heavy nitrogen. Here, then, were labeled dietary nitrogen and two labeled indicator proteins, one being produced (and destroyed) by the animal, the other merely being metabolized. Each indicator protein in the rabbit's serum, drawn at intervals, was quantitatively and independently precipitable by the corresponding specific polysaccharide and so separable from the other serum proteins. What happened was that, while the type III antibody quickly took up heavy nitrogen as before, the passively injected type I antibody, which was merely circulating in, but was not being elaborated by, the test rabbit, remained free from heavy nitrogen. This was taken to indicate that heavy nitrogen fails to enter serum protein molecules which were complete at the start of heavy nitrogen feeding, but becomes part of the structure only of those proteins which are being constantly synthesized. As a by-product, quantitative studies of the rate of disappearance of heavy nitrogen from antibody indicated a half-life (with respect to synthesis *vs.* destruction) of about two weeks, much the same as that of other serum proteins in this animal.

Incidentally—and this was overlooked at the time—the absence of N^{15} in the passively injected type I antibody, after precipitation with its homologous type-specific polysaccharide and washing under standard conditions, affords as rigorous a test as has ever been made of the specificity of immune precipitation and the validity of the assumption under-

lying the estimation of antibody in absolute terms; namely, that antibody, and antibody only, is quantitatively measured. Although the serum of the type III-immunized rabbit contained up to 1 atom per cent N^{15} excess, the type I antibody isolated from this medium containing other antibodies, globulins, and albumins rich in heavy nitrogen retained none of the N^{15} within the error of the mass-spectrographic method, about ± 0.03 atom per cent N^{15} excess.

The immunochemist has learned quantitative relationships with even so complicated a test tube as the guinea pig, for Kabat and Landow (26,27) have shown that 0.2 mg. of antibody suffices to sensitize the animal so that fatal anaphylactic shock ensues with 1 mg. of antigen such as egg albumin or 0.1 mg. of a type-specific polysaccharide of pneumococcus. Either of these shock doses is far in the region of antigen excess in the *in vitro* precipitin reaction, suggesting at once the first simple explanation of how much smaller quantities of antigen may desensitize the animal. Because small quantities of antigen combine with relatively much larger amounts of antibody, after the injection of "desensitizing" small doses of antigen relatively little antibody remains for the final destructive effect of an injection of excess antigen.

All of which must not be taken to indicate that quantitative immunochemistry, like a parachute flare, has had its episodic illuminating burst and is fading out. It has shown the way to objectives of theoretical and practical interest, and many more applications are apparent. One quantitative immunochemist has at last irritated our native virologists into an appreciation of its potentialities in their vast field of endeavor (25). Quantitative immunochemical methods are also indispensable guides in the study of bacterial antigens and the fractionation of many other natural carbohydrate and protein mixtures. A recent application in the little studied field of lipoproteins was the report of Cohen and Chargaff (4) on thromboplastin. Suspensions of bovine thromboplastin removed measurable amounts of antibody from specific antisera in rabbits, and, remarkably enough, the specific precipitates showed greater thromboplastic activity than the antigen contained in them. Displacement of the phosphatide in the antigen by means of heparin did not destroy the power to precipitate with antiserum.

Moreover, if proteins or proteinlike structures are ever synthe-

sized, with or without the mediation of enzymes, immunochemistry will afford incisive aids to their study. Protozoan and other parasitic diseases are intruding upon us with a new insistence born of the war and increased speed of travel, and it is confidently to be expected that immunochemistry will become an essential factor in the understanding and conquest of these scourges. The broadening horizons of immunochemistry will steadily augment its opportunities for fundamental contributions in the struggle against infectious disease.

References

(1) Alexander, H. E., Heidelberger, M., and Leidy, G., *Yale J. Biol Med.*, **16,** 425 (1944), and earlier papers.

(1a) Arrhenius, S., *Immunochemistry.* Macmillan, New York, 1907.

(2) Avery, O. T., and Heidelberger, M., *J. Exptl. Med.*, **38,** 81 (1923); **42,** 367 (1925).

(2a) Bordet, J., *Traité de l'Immunité.* Masson, Paris, 1920, 1939.

(3) Summarized in Boyd, W. C., *Fundamentals of Immunology.* Interscience, New York, 1943.

(4) Cohen, S. S., and Chargaff, E., *J. Biol. Chem.*, **136,** 243 (1940).

(5) Summarized in Ehrlich, P., *Studies on Immunity.* Wiley, New York, 1906.

(5a) Felton, L. D., *J. Immunol.*, **22,** 453 (1932).

(5b) Goodner, K., and Horsfall, F. L., Jr., *J. Exptl. Med.*, **66,** 437 (1937).

(6) Heidelberger, M., *Science*, **92,** 534 (1940); *J. Exptl. Med.*, **73,** 681 (1941).

(7) Heidelberger, M., and Anderson, D. G., *J. Clin. Investigation*, **23,** 607 (1944).

(8) Heidelberger, M., and Avery, O. T., *J. Exptl. Med.*, **38,** 73 (1923); **40,** 301 (1924).

(9) Heidelberger, M., and Kabat, E. A., *J. Exptl. Med.*, **60,** 643 (1934); **63,** 737 (1936); **65,** 885 (1937).

(10) Heidelberger, M., and Kabat, E. A., *J. Exptl. Med.*, **67,** 181 (1938).

(11) Heidelberger, M., and Kendall, F. E., *J. Exptl. Med.*, **50,** 809 (1929); **61,** 559 (1935).

(12) Heidelberger, M., and Kendall, F. E., *J. Exptl. Med.*, **61,** 563 (1935); **62,** 467 (1935); **65,** 647 (1937).

(13) Heidelberger, M., and Kendall, F. E., *J. Exptl. Med.*, **64,** 161 (1936).

(14) Heidelberger, M., Kendall, F. E., and Teorell, T., *J. Exptl. Med.*, **63,** 819 (1936).

(15) Heidelberger, M., and Mayer, M., *J. Exptl. Med.*, **75,** 285 (1942).

(16) Heidelberger, M., and MacPherson, C. F. C., *Science*, **97,** 405 (1943).

(17) Heidelberger, M., and Pedersen, K. O., *J. Exptl. Med.*, **65**, 393 (1937).
(18) Heidelberger, M., Rocha e Silva, M., and Mayer, M., *J. Exptl. Med.* **74,** 359 (1941).
(19) Heidelberger, M., Sia, R. H. P., and Kendall, F. E., *J. Exptl. Med.*, **52,** 477 (1930).
(20) Heidelberger, M., Treffers, H. P., and Mayer, M., *J. Exptl. Med.*, **71,** 271 (1940).
(21) Heidelberger, M., Treffers, H. P., Schoenheimer, R., Ratner, S., and Rittenberg, D., *J. Biol. Chem.*, **144,** 555 (1942).
(22) Heidelberger, M., Weil, A. J., and Treffers, H. P., *J. Exptl. Med.*, **73,** 695 (1941).
(23) Hershey, A. D., *J. Immunol.*, **48,** 381 (1944), and earlier publications.
(24) Kabat, E. A., *J. Exptl. Med.*, **69,** 103 (1939).
(25) Kabat, E. A., *J. Immunol.*, **47,** 513 (1943).
(26) Kabat, E. A., and Boldt, M. H., *J. Immunol.*, **48,** 181 (1944).
(27) Kabat, E. A., and Landow, H., *J. Immunol.*, **44,** 69 (1942).
(28) Kendall, F. E., *Ann. N. Y. Acad. Sci.*, **43,** 85 (1942).
(29) Summarized in Landsteiner, K., *The Specificity of Serological Reactions.* Harvard Univ. Press, Cambridge, 1945.
(30) Summarized in Marrack, J. R., *Chemistry of Antigens and Antibodies.* H. M. Stationery Office, London, 1934; 2nd ed., 1938.
(31) Mayer, M., and Heidelberger, M., *J. Biol. Chem.*, **143,** 567 (1942).
(32) Pappenheimer, A. M., Jr., and Robinson, E. S., *J. Immunol.*, **32,** 291 (1937).
(33) Pauling, L., Campbell, D. H., and Pressman, D., *Physiol. Revs.*, **23,** 203 (1943).
(34) Schoenheimer, R., Ratner, S., Rittenberg, D., and Heidelberger, M., *J. Biol. Chem.*, **144,** 545 (1942).
(35) Tiselius, A., and Kabat, E. A., *J. Exptl. Med.*, **69,** 119 (1939).

LIBRARY

30

SOCIAL ASPECTS OF NUTRITION

W. H. SEBRELL, MEDICAL DIRECTOR, U. S. PUBLIC HEALTH SERVICE; CHIEF, DIVISION OF PHYSIOLOGY, NATIONAL INSTITUTE OF HEALTH

THE PHENOMENAL scientific progress in biochemistry in the past few decades has opened up new vistas in the field of nutrition. The most important aspect of this progress has been our greatly increased knowledge of the chemistry of the human organism. Although we still have much to learn about the chemistry of our vital processes, enough is now known to show that, by controlling his nutritional environment, man can make far-reaching contributions to his health and welfare.

Our enemies in this war have demonstrated how manipulation of the food supply can be used to weaken nations by creating conditions which cause malnutrition, ill health, and increasing mortality. But our knowledge of nutrition also can be used to strengthen the people of the world. If this knowledge is intelligently applied, progress toward the ideal of adequate nutrition for everyone can become one of the main roads leading to a greater degree of health and prosperity than the world has ever known. This, if attained, can be a major contribution to a permanent international peace. Since scientists possess this vital knowledge, the question is: how can scientists best play their parts in transforming the ideal into reality?

Most scientists feel they have completed their work when the results of their experiments have been published in a technical journal

and presumably made available to all. In fact, of course, this is not true, for their results are actually made available only to a select few who are capable of interpreting them correctly. The job of making the results available to everyone in terms of their practical application still remains to be done.

In our world of today, the biochemist has become so much a part of medicine, and his findings are so important in the diagnosis and treatment of disease, that his work can no longer be regarded as finished when his new discoveries are published. It is a part of his job and responsibility to assist in making his results of real benefit to people. This does not mean that every biochemist should become a crusader and public lecturer, but rather that some qualified investigators should take part in cooperative local, State, and national efforts to improve nutrition, adding their knowledge to that of the physician, agriculturist, food distributor, social worker, and representatives of government in the attempt to produce, distribute, and utilize our food supply in the best possible manner.

Thus pellagra, a deficiency disease, has been known for about twenty-five years to be due to a deficient diet. The foods which are high in pellagra-preventive value are known and widely available in the United States. Niacin was found to be the specific essential vitamin, and has been available for seven years. Yet in the United States in 1943 at least 1303 people died of pellagra, and a much larger number suffered from its effects. Why did all this suffering and death occur? The reasons were either that the necessary foods were not available to the victims or that they were ignorant of the information existing about this disease. Poverty and ignorance are the two great factors involved in malnutrition of all kinds.

Would we not have been more successful in eradicating this disease had all the scientists who knew the facts taken a more active part in correcting the underlying causes? Although willing to argue over technicalities, and the validity of observations, research workers seem to be slow to see that it also should be one of their obligations to devote at least some of their energy to devising practical measures for the prevention of disease, and in imparting their knowledge to those who can use it in the field. If we are determined to make democracy work, scientists must make their special knowledge widely available.

In the less highly organized society which existed during the

early development of this country, there was an opportunity to obtain food by hunting, fishing, and farming. Today, it is worse than useless; it is cruel to tell a malnourished man that he needs better food when his lack of food is due to his inability to buy it, or when the food is not available in his locality. A few years ago we were destroying so-called food surpluses although many thousands of people in this country had deficiency diseases. The trouble was not overproduction in terms of adequate food for everyone but a failure in distribution. Much food had to be destroyed because there was no market for it, the producer was threatened with economic ruin, and there was no mechanism by which the food could be made available to those who needed it for better nutrition.

This anomaly of want in the midst of plenty was partly due to the fact that scientists were working independently on narrow aspects of the problem and there was no organization for putting all the knowledge together into a unified plan. The existing knowledge had not been made available either to the public or to those responsible for handling the food supply. This defect has been remedied to some extent by action taken as a result of the National Nutrition Conference in 1941, and by the establishment of the Food and Nutrition Board of the National Research Council, which has played a useful part during the war and which should continue to render valuable service to the nation, by making scientific knowledge more readily available, by enabling scientists to see nutritional problems as a whole, and by giving them an opportunity to participate in the application of their knowledge to human welfare. Had we, as a nation, recognized and met our obligation to make adequate nutrition available to everyone who needed it, the so-called surpluses never would have existed. No nation has ever produced enough of the right kinds of food to meet the needs of adequate nutrition for its entire population.

It is only within the past few years that research in nutrition and biochemistry has brought our knowledge to the point at which we could with some assurance formulate a dietary we know is adequate for health. Such a dietary has been set forth in the recommended dietary allowances of the National Research Council (4). With this basic and essential information, the scientist is now ready to ask society if it is willing to accept as a principle of democracy that everyone is entitled to an *opportunity* to secure a diet adequate for health. This

responsibility was accepted, not only by the representatives of the United States, but also by those of forty-three other nations at the United Nations Conference on Food and Agriculture. This is a great step forward, but it is still a long way from transforming the ideal into actuality. It is important for this country and the world that the objectives of the United Nations as regards nutrition should be recognized and put into practice in every state and community in this country. The implications are tremendous. A great expansion in agriculture along selected food lines would be necessary to meet this need. In 1940, Cavin, Stiebeling, and Farioletti (1) estimated that, if the average consumption of protective foods could be raised to the level of families whose food intake was rated as "good" from the standpoint of nutrition, the *increase* in national consumption of milk would be 20%; butter, 15%; eggs, 35%; tomatoes and citrus fruit, 70%; and leafy green and yellow vegetables, 100%. That much of our dietary inadequacy is due to lack of purchasing power is clearly indicated by extensive data showing that it is the members of the lowest income families who have the greatest deficiencies. It is also indicated in the changes occurring in our food consumption during the war. Stiebeling (5) estimates that, with the increased power to buy due to high employment and better wages, at least one family out of every seven which had a poor diet in 1936 was able to obtain a fair or good diet in 1941.

What increased purchasing power will do is also shown by the fact that rationing became necessary during the war *although the amounts of food available for civilians were greater than ever before.* This has been true of all foods except sugar and syrups, coffee, tea and cocoa, and fats and oils. The increases in 1943 as compared with prewar levels were about as follows: milk and milk products 17%; meat 6%; eggs 16%; potatoes and sweet potatoes 9%; pulses 22%; tomatoes and citrus fruit 17%; leafy green and yellow vegetables 9%; and other vegetables 5%. It must be remembered that these figures represent classes of foods and that there were decreases in certain selected, popular food items such as beef and canned fruits. In terms of nutrients, the increases are even more amazing: animal protein, 9%; vegetable protein, 3%; calcium, 15%; iron, 14%; vitamin A, 8%; ascorbic acid, 7%; thiamin, 37%; riboflavin, 18%; and niacin, 14% (2). Full employment and higher than usual wages made many

good foods so scarce that they had to be rationed. That our agriculture can produce the foods we need if there is an effective demand was shown by the fact that, in addition to the greater supplies available for civilians, huge amounts were allotted to the armed forces and Lend-Lease.

Having accepted our responsibility as the nation which sponsored the Hot Springs Conference, the basic problem now is how the food adequate for health can be made available to our low-income families. That it can be done has been shown by war experience. We must leave the method of doing it to our legislators. Once it is done, we can expect the greatest era of national well-being and agricultural prosperity we have ever known.

The attack on the problem is already under way. During the days of the food "surpluses," one outlet for food was the school lunch room. At that time, this use of the food did not conflict with the usual trade channels and much food was disposed of in this manner. Now, the school lunch program is just beginning to stand on its own feet in the sense that the need for such a program is recognized: it is not being carried out as a means of disposing of food surpluses. Many people are at last seeing the fallacy of providing expensive buildings, equipment, instruction, and books to a child with a mind and body retarded by malnutrition. Yet, in many places, the idea still persists that it is no concern of the educational system to make an adequate meal available at the school. In many schools the child does not even have the opportunity to buy an adequate meal, although the modern school is often a long distance from the child's home; there are no feeding facilities whatever. In other schools the lunch room is operated by volunteers or with poorly paid and poorly trained personnel with neither the necessary knowledge nor funds to supply the needed food. Every medical appraisal of groups of school children reveals large numbers with evidence of malnutrition. Children are found arriving at school without breakfast, while many do not have the funds to purchase an adequate lunch. Those who carry lunches from home frequently bring inadequate food for lack of knowledge of nutrition in the home and because of the difficulties of packing and carrying an adequate, appetizing lunch. An obvious step would be a method of supplying an opportunity to obtain an adequate lunch to all school children throughout the country.

In the past, the application of our knowledge of nutrition to industry has not received any more attention than we have given to nutrition in our schools. The right of a workman to have an opportunity to obtain an adequate meal at his work received little recognition until, under the stimulus of the necessity for the utmost production in war industries, attention has finally been paid to malnutrition as a factor affecting a man's ability to stay on the job and to do a good job under pressure. In-plant eating facilities vary from none at all (and little or no lunch interval) to the finest type of adequately supervised feeding arrangements with ample time to eat. The benefits both to the management and the worker have been so great that within the past two years a large number of industrial establishments have made great improvements in this respect even under wartime difficulties. It has been estimated that before "Pearl Harbor" less than 20% of our industrial workers had access to any type of in-plant food service. By January, 1944, about 6,500,000 of the 22,000,000 people engaged in war industries were receiving meals through industrial feeding facilities (3). Here again we see a movement toward the goal of an opportunity for an adequate diet for everyone, but with the goal still a long way off. From the scientists' point of view, the mere supplying of food is not the objective, although many think the goal has been achieved when this is attained. The fundamental fact which always should be in the forefront is that the objective is to supply *nutritionally adequate* food. The mere installation of feeding facilities in schools and industries does not furnish the opportunity the scientist has in mind. The installation must also include supervision by adequately trained dietitians and nutritionists so that the knowledge of nutrition is applied in making every meal furnish a maximum of nutritive value in appetizing form.

Although these two programs attack important parts of our national problem, there are many other groups for which no wide general provision has yet been made in this country, such as pregnant and lactating women, the preschool child, and the housewife. Recent research has demonstrated the very great importance of adequate nutrition in the prenatal period to the welfare of both mother and infant. The complications of pregnancy and delivery are less and the infant is healthier, grows better, and is more likely to survive if prenatal nutrition has been adequate. With our future national welfare

at stake, is not every infant entitled to the *opportunity* to obtain a diet adequate for its growth and normal development? No national program with such an aim exists in this country today. In some countries adequate infant feeding formulas are prepared, bottled, and distributed under public supervision to infants in low-income families. We are woefully behind in our application of the knowledge of nutrition our scientists have supplied. We have hardly begun to create the *opportunity* for our people to obtain adequate nutrition.

The other great cause of malnutrition in the United States, in addition to the general dearth of opportunity to obtain nutritionally adequate food, is the lack of knowledge of the need for, and what constitutes, good nutrition, as reflected in bad food habits, poor food preparation, and general ignorance and indifference about the types of food necessary for good nutrition. The practical aspects of our scientific knowledge of nutrition should be conveyed to everyone in some manner other than through advertising or the educational material supplied for the purpose of promoting the use of some particular food or medicine. Nutrition education should be a part of the curriculum of every school. Clear, understandable, correct, and unbiased information should be available to every family. A start has been made in this field also, but it is only a start; although it has been going on for many years, it still has reached and affected only selected groups of people. There is a wealth of literature available to those who know how to obtain it and with enough education to read and understand it. Agricultural agencies and health departments have been especially active in this field. Under the war stimulus thousands of women have attended Red Cross courses in nutrition; nutrition committees and other organizations interested in nutrition have reached many more thousands with educational material.

However, much more is needed; for literature is only part of the answer. Demonstrations, cooking schools, exhibits, radio and newspapers, moving pictures, and every other educational medium should be used to the fullest extent. Lunch rooms in schools and industrial establishments should be used, not only for feeding, but also for practical education in good nutrition. Studies have demonstrated that educational efforts here pay immediate dividends in influencing food selection. The weakest point in nutrition education is that it has not been given at the point of application, namely, the grocery

store, the restaurant, and the lunch room. The educational material has not been tied in with the food seller and distributor. Until nutrition education is applied at the point of sale of food and the available food is related to the educational material, we cannot expect to obtain maximum educational results. Furthermore, the group to whom education would mean the most, namely, the children, has been largely neglected. Nutrition education in our secondary and high schools still consists very largely of teaching the way to prepare appetizing food. Little attention is given to nutritive value, to nutritive losses by various methods of cooking, or to the necessity for an adequate diet. The chance to establish good food habits at an early age in the elementary schools is ignored. No attempt is made to correct bad food habits established at home, and a golden opportunity is lost to establish good lifetime food habits. Class room lectures and cooking classes should be closely correlated with the foods served in the school lunch room. A start has been made in this field also—and an increasing number of schools under the stimulus of the Office of Education are making progress in this field, but again we still have a long way to go.

Recently, great advances were made in our knowledge of the effects of different methods of food preparation and handling on the nutritive value of food as it is eaten. It is now well known that some methods of preparation result in practically complete loss of some nutrients, while other methods of preparing the same food will conserve much of its nutritive value. The effect of this knowledge of food preparation on the population at large has not been material because it has not been applied: we still have mashed potatoes, overcooked vegetables, and the use of too much water in cooking; cooked foods frequently are held on warming tables for hours before consumption. These are all well known as methods of handling food which result in a high loss of nutritive value; much educational effort has been spent on attempts to correct these practices. In addition, most of us eat certain foods only because we like them and not because they are good for us. Too many of our cooks do not know how to prepare food that is both appetizing and nutritious.

A new factor of great public interest has been introduced into the nutrition picture by the discovery of the chemical structure of several vitamins and the development of methods of synthesis on a commercial basis. Here science has furnished us with a new, powerful, and valuable

weapon against malnutrition if we use it properly. Pure vitamins are invaluable in the treatment of deficiency diseases, but their widespread use by the public requires technical guidance if we are to avoid economic waste and a false sense of security. It is obvious that the use of multivitamin and mineral pills can be no substitute for an adequate diet. Since we still have not identified all of the dietary essentials, a suitably varied diet covers these defects in our knowledge. Even if we knew how to put all of the dietary essentials into one capsule, attempts at mass use of such a preparation would fail because of lack of public acceptance. Our aim should be to teach people how to meet their nutritional requirements with food. The widespread use of vitamin and mineral supplements seems clearly indicated in situations in which, for any reason, an adequate food supply is not obtained. In such cases, a survey of the diet available and clinical examination for evidence of deficiency should be the basis for dietary supplementation which then may be handled individually or collectively as the situation indicates. One of our great nutritional advances has been the introduction of enriched white bread and flour. Here, a cheap, widely used food has been employed as a vehicle for additional supplies of vitamins and minerals which are not being obtained in sufficient amounts by that part of the population using bread and flour in the greatest amount.

However, the addition of vitamins to foods is not something to be done haphazardly. Carefully selected, widely used, suitable foods supplemented by substances for which there has been shown to be a widespread need are indicated. Other examples of the beneficial and successful use of this principle are the addition of iodine to salt, of vitamin D to milk, and of vitamin A to oleomargarine. The fortification with vitamins and minerals of a large variety of foods of limited use would be both wasteful and unnecessary. It might defeat the very purpose of the practice by increasing the price of such foods, thus tending to place them beyond the reach of those who need them most.

An important trend in nutrition research has been the formation of the Nutrition Foundation by the food industry in order to support fundamental research in the nutrition field. Grants for research on nutrition are made to universities and scientific laboratories through an independent scientific advisory committee. This enables the food industries to contribute to the advance of our knowledge and to obtain and make the best use of the most recent advances in nutrition research.

The organization of the Nutrition Foundation represents the far-sighted viewpoint of these industries, which realize that the closer we approach dietary adequacy for everyone the sounder the basis for their business, a viewpoint shown also in a change in much of the food advertising, which is now frequently directed at the general field of improved nutrition. The leaders in this field are not only promoting the best interests of their business but also contributing to the welfare of the country. The scientist can do much to direct this great public educational force along sound lines. If the results of research in nutrition are to be of public use and benefit, the scientist must come out of his laboratory and assist in their application. He cannot, like a famous mathematician, hope that his research will be useless so that he can work without distraction.

In the field of medical science, the biochemist has given the physician new and valuable methods in the diagnosis of nutritional failure; but, because of difficulties in the interpretation of clinical signs of early deficiencies, a controversy continues as to the extent and importance of deficiency diseases in this country. We still need better methods of diagnosis to enable us to recognize the earliest manifestations of nutritional deficiencies.

There has been a rapidly growing recognition that adequate nutrition is of great importance in convalescence from any disease, and that nutritional deficiencies frequently occur as complications of many unrelated diseases, especially where there has been marked loss of appetite, nausea and vomiting, diarrhea, disturbed metabolism, etc. With the increased awareness of the importance of adequate nutrition in medicine and the growing complexity of the subject, we now see medical and public health schools teaching nutrition as a part of clinical and preventive medicine, as well as of biochemistry and physiology. There is reason to hope that, in a few years, our need for physicians well grounded in the subject will be met.

The lack of statistics on the distribution and prevalence of the deficiency diseases makes it exceedingly difficult for the public health officer to attack the problem effectively from the viewpoint of preventive medicine, although here we find an active interest being taken and a start being made. Many health officers now see that the prevention of malnutrition is part of their obligation to provide an opportunity to attain the best degree of health. This aspect of nutrition

could not develop until our knowledge had advanced to the point at which the scientist could show the health officer that malnutrition is a widespread health problem and that human dietary needs are sufficiently well known for the health officer to apply adequate remedies. In spite of the present lack of knowledge about the extent of the problem, the health officer is laying the foundation for more extensive work in the field. A start has been made in almost every State, although there is still no clear definition of the best method of approach or of the way in which the health department can work to the best advantage. The problem is made much more difficult for the health officer by the fact that he recognizes that an adequate solution, involving as it does education, agriculture, and food distribution, means that he must cooperate with many agencies not concerned in most public health problems. One of his most important duties is to determine what deficiency diseases exist in his community, how many people are affected, and where these people are located. He also can carry some of the educational work, and his voice in pointing out the importance of the problem can do much to secure action by other interested agencies. In order to bring about the necessary coordination, nutrition committees have been set up in every State, bringing together for the first time the various agencies interested in different phases of nutrition. These committees are wartime committees, but they should continue in peacetime to form a unified approach to all aspects of nutrition without which we cannot obtain a permanent correction of our difficulties.

While the ultimate goal of an opportunity to obtain an adequate diet for everyone is still far in the future, just the recognition of its importance and the realization that it may be obtainable and that it should be a part of any plan for permanent international peace is a long step toward success. Unfortunately, only a few recognize these factors today. The next great forward step must be a general recognition that the future development of the nation to a maximum of health and prosperity depends on adequate nutrition for everyone. Scientists can do much to make their work of the greatest use to mankind by assisting in bringing this about.

This war has shown that great advances in nutrition and health are within our reach. Gains during the war were made possible by some factors on which we could not count in peacetime, but there

is one factor which it will be our opportunity to continue to make available now that the war is over—the participation of scientists in the task of improving the nutrition of people. They must play a larger role than in the past in applying the results of their work. Research work in nutrition has, in recent years, achieved an enviable record. It will earn the gratitude of mankind if a greater effort is made to give human beings the benefits of its results. The nutrition scientist can play an even more important part in the future than he has in the past if he will, on occasion, come out of his laboratory and mingle with the people in the market place. He will go back to his essential work refreshed by a greater appreciation of the way in which his results contribute to the improvement of human welfare and by his contact with the every-day problems of life.

References

(1) Cavin, J. P., Stiebeling, H. K., and Farioletti, M., "Agricultural Surpluses and Nutritional Deficits," in *1940 Yearbook of Agriculture, Farmers in a Changing World.* U. S. Dept. Agr., Washington, 1940.

(2) *Food Consumption Levels in the United States, Canada, and the United Kingdom.* Report of Joint Committee of the Combined Food Board, U. S. Dept. Agr., April, 1944.

(3) Goodhart, R., "Protecting the Health of the Industrial Worker: Nutrition," in *New Steps in Public Health.* Milbank Memorial Fund, New York, 1945.

(4) "Recommended Dietary Allowances," *Natl. Research Council Reprint Circ. Series*, No. 115 (1943).

(5) Stiebeling, H. K., "Adequacy of American Diets," in *Handbook of Nutrition.* Am. Med. Assoc., Chicago, 1943.

ORGANIZATION AND SUPPORT OF SCIENCE IN THE UNITED STATES

L. C. DUNN, PROFESSOR OF ZOOLOGY, FACULTY OF PURE SCIENCE, COLUMBIA UNIVERSITY

THE WAR and the sudden need to improvise means for supporting and directing war research have brought into high relief an important fact which has been dimly recognized for many years: there has been in the United States no orderly means for the continuous support of fundamental scientific research, and no policy or method for the deliberate utilization of science by our society. Science has been a hardy plant which grew where and how it could, thriving in the comfortable greenhouse of a research institute, or turning ample fertilizer into real fruit in an industrial laboratory, or in the more usual case struggling for sustenance in the thin soil of colleges and universities, occasionally enriched by temporary growth stimulants from a foundation or private donor. Except in the case of certain industrial developments and in a few government departments, the support of science in the United States has not been the result of decision but of chance, operating in a milieu which contained good scientists and a good deal of fluid wealth.

The most blunt and truthful statement we can make about the reason for the lack of continuity and of public policy regarding science

is that, as Americans, we did not want either continuous support or direction or planned application of science. The detailed causes of this attitude trace in part to reasoned premises and in part to prejudice; and from these there has resulted a confusion of thought which the war has now revealed.

The contradictions come out most clearly in the views of scientists concerning the support of science after the war. Most of them hope for release from the capricious and precarious methods by which fundamental research was chiefly supported before the war, namely, by periodic begging from donors, such as foundations who chose the researches to be supported. Scientists generally hope for a more orderly and stable means of support than this, yet most of them would not turn to the Federal Government as the source of more continuous support. They profess to fear infringements on their freedom more when support comes from their government than when it comes from private sources.

There is no sense in dodging or belittling the dilemma in which this places science. On the one hand, the war agencies which have guided and financed a large segment of scientific research propose to withdraw from this function. If they do, the public investment in scientific research will drop to a third or a quarter of its present level. At the same time, the principal sums in the hands of the great foundations are declining and science must adjust itself to diminishing support from this and other private sources, and possibly to the extinction of this sort of financial aid within another generation. There will eventually remain as sources of support chiefly industry and business, through their research laboratories and foundations, and the government, through its own scientific agencies or through new channels yet to be created.

Most scientists who do not like "domination of science by government" like "domination of science by industry" even less; and many have already objected to the influence which the foundations wield because of their control of the fluid funds with which to supplement the fixed investments of universities and research institutes in men and permanent plant. It has often seemed that this small tail of free funds has wagged the larger dog of solid investment.

Moreover, scientific research depends upon trained men and women as much as upon material facilities, and we have as yet made

no provision for assuring a steady flow of young scientists into research. For advanced training we have relied upon the existing scholarships and fellowships of the universities, which are so meager that most young scientists can devote only a portion of their time to learning, the rest being needed for earning a living; and upon the advanced fellowships supplied by foundations, private philanthropy, and industry. The same considerations of approaching exhaustion of private funds apply to the training of persons as to the provision of research funds.

The facts that must be faced are, then, that the present means of support of science are running out and, whether we like it or not, changes in the sources and form of support will occur; and that a chief desideratum for scientists will be to keep science under the new conditions as free as possible to develop according to its own inner needs and according to its function in society.

In the following pages I propose to discuss, first, what the function of science is that entitles it to support; second, what determines the attitudes of scientists toward forms of support; third, what general public policy toward science would represent the best interests of science and scientists; and, fourth, how this policy could be implemented in practical ways.

At the bottom of every consideration of science in its public aspects must lie the question: "What is science for?" When this question is squarely and thoughtfully faced, scientists will agree that science exists for man and not for itself alone. As a means of understanding the material world, it leads toward the improvement and control of the environment in which human society must always operate. Eventually, its results and the methods of thought which it develops accrue to the public good, not merely by increasing the physical well-being of the people through technological applications, but also by extending the domain of reason and by increasing our understanding and appreciation of nature. In discussing the material means which have to be provided for scientific research, it is often forgotten that the great and lasting changes wrought by science are in men's minds, and that, in the end, science is to be supported for the same reason that education is to be supported. The products of science are primarily increase and diffusion of knowledge and increase in the number of trained minds, and secondarily increase of technical facilities and production of goods. Like other knowledge, scientific under-

standing is one of the "rights" to which all citizens should have equal access. Its support, like that of education generally, is thus to be shared, as most essential activities are in our society, by the State and by "public spirit" as it acts through foundations, private citizens, and industry. At the material level, science in the modern world has become a public necessity without which technical advances and social developments determined by them cannot occur in an orderly way. It has become so "affected with the public interest" that its support must be a matter of public concern. The scientist has thus become in some sense and in spite of himself a public servant.

The many scientists serving their country during the war *as scientists* are less likely now than formerly to forget their public function; but in the past a failure to recognize this led scientists as a class to have too little confidence in seeking support for scientific work. They were not sure that science was worthy of public support, because oftentimes science was not what the world needed, but only what they enjoyed doing. They did not generally think about a public policy for science because they were not clear about the public function of science. Can we really expect (they would say) the public to support this kind of work? Or as a small boy said to a scientist after a visit to his research laboratory, "Uncle, do they really pay you for doing this?"

When questions about the organization and support of science were raised, however, other reasons were generally given for either opposing the formulation of policy or avoiding the question altogether. These reasons took different forms, but in general had their roots in our tradition of individualism. Since scientists have usually been strong individualists, the traditional public objections to schemes for the support and direction of science have been strengthened and rationalized by the scientists themselves. They said: "Organization kills initiative," "Planning interferes with free enterprise," or "Continuously assured support removes the need for periodic justification of each research on its own merits." "Support implies direction, and he who pays the fiddler will call the tune; and only scientists can know what tunes can or should be played."

These are valid and weighty objections and they must be squarely met by any general proposal for the maintenance or direction of science. It is nevertheless true that these are not the primary or

real reasons for opposing the formulation of a public policy or even specifically for opposing the support of science from public funds, since the same scientists who use them against government support approve the use of organization, planning, continuous support, and central direction when these are employed, *as a matter of policy*, in the great industrial laboratories. In fact, many scientists point with pride to the splendid results which industrial laboratories have achieved under the very conditions which they allege would impede and stifle scientific research done at the expense of government. Moreover, public support and direction appear to have been quite acceptable in the great program of agricultural research which has been in operation since 1887 through the United States Department of Agriculture and the State Agricultural Experiment Stations. These facts are not cited to minimize the difficulties involved in planned continuous support and direction of research. They do show clearly, however, that the objections are generally not to support and direction as such but to these only when the authority which wields them is the Federal Government. As the attitude toward agricultural research shows, the objection does not apply with similar force to the State governments. Many scientists have expressed the fear that central and especially federal support of scientific research would put an end to "scientific freedom" and lead to "regimentation." In most cases, it is the threat to scientific individualism or "free enterprise in science" that is the real cause of fear. Since such changes in modern society as the decline of individualism are not due to deliberate acts of governments but result from the social and economic and technical developments of our age, they call, not for fear, but for a greater effort to understand them.

I believe that most scientists have come to realize the nature of such objections to discussing general policies for the public support of science. The central position that "pure science," especially physics, came to occupy in war research revealed facts about science in the modern world which simply could not be evaded or overlooked. Even the need of "coordination," the blackest of the beasts which threaten the research scientist, became evident as soon as the war imposed pressing requirements which an unplanned, uncoordinated science could not meet. The knowledge that our enemies had succeeded in so organizing their research and development programs that they had

"got the jump" on us in numerous ways persuaded even reluctant individualists that coordination was absolutely necessary.

The war emergency also revealed the lack of balance which obtains when science is directed by chance. Many fundamental problems, upon which other inquiries depended, had not been touched and efforts had suddenly to be made to straighten the front. If this was borne in upon those scientists who participated in war research, it became even clearer to those who through lack of organization were left out. There are now many biologists who would sacrifice their cherished individualism for the sake of being identified with a great national effort. They realize that the neglect, the omission almost, of biology and biologists from the hastily improvised war agencies was bad not only for biology and for other sciences, such as the medical and agricultural sciences which depend upon biology, but for the nation. Their state of mind is not improved by the reflection that, by and large, the fault was their own.

Still other changes in the attitudes of scientists are due to the growing realization that research workers need to recognize the connection between their own special work and the general scientific structure in which it will find its place and its function. It is difficult for the research worker to envisage this larger field without inquiring too about the still wider frame of society in which science operates. Many more scientists than formerly now believe not only that this social awareness of the men who do the work of science is needed to make a social being and a citizen of the scientist, but that this is essential in the national interest. Those who so believe will want to face the questions involved in the public support of science.

By these paths we come to the problem itself: what public policy toward science would encourage the best growth of science and its use for the welfare of the people? The aims of policy must be to reconcile two basic requirements, about which there is probably general agreement.

(1) Science and scientists must be free to grow and change in ways determined in part by the discoveries of science itself. This is the way in which science has progressed in the past—and the autonomy of small groups and the feeling of freedom of the individual to follow the new idea wherever it may lead are goods which must be preserved. This freedom must be accepted and guarded as a matter of principle;

and provisions for freedom of publication and the prevention of arbitrary censorship must be a part of the basic policy.

(2) The forms of support and organization of science must be determined by social needs and purposes and are therefore matters of concern not only to scientists but to government and to the ultimate beneficiaries of science, that is, the people, as consumers and workers. Those who most directly need and use the results of scientific research in education, industry, agriculture, medicine, and public health have a special interest in the development of science, and means must be provided by which this influence can be exercised. The two primary conditions should therefore be: (a) a central organization by which the conduct of science is made responsive to public requirements and needs; and (b) the representative character of the directing agency or agencies, insuring democratic methods in administration.

These two requirements of autonomy, on the one hand, and subservience to social needs, on the other, have seemed antithetic to some, but I do not believe this need be the case. There is much evidence of the vitality and progressiveness of science in other countries where it is largely under public control. The extreme example of public control is in the Soviet Union, where the direction of scientific research is centralized in the Academy of Sciences, through which the support of the state flows to all of the research agencies. Other European countries occupy positions intermediate between this maximum and the minimum reached in the United States, where almost alone among modern nations science has retained a predominantly private character. Even here, the wartime activities of the Office of Scientific Research and Development and the Committee for Medical Research show that no essential incompatibility exists between research and public control; while the long peacetime history of United States Government scientific departments and especially of the Department of Agriculture illustrate the feasibility of accomplishing at once a scientific and a social purpose.

Much experience in the United States and in other countries indicates that, to obtain the maximum results from a given effort in scientific research, the interests of the research workers themselves must be consulted, but that these are not fundamentally different from those of the community around them. Scientists traditionally are primarily devoted to their work, often sacrificing other interests to it and exclud-

ing other interests which tend to interfere with it. Yet, as the war shows, they will voluntarily and gladly place this devotion and their technical ability and intelligence at the service of an objective which is clearly defined and compelling. On the other hand, directing agencies, public or private, do not grudge to the scientist a greater measure of freedom than to other workers, provided they are assured of his adherence to the principles of service and to the general purpose which they consider essential, and that this freedom actually produces the results expected from it. Freedom within a general plan is a practical ideal at which to aim, as the comparative freedom of local political units within the general frame of Federal Union of the United States shows.

Voluntary cooperation of scientists with public agencies in the planning and execution of research would seem to provide the soundest base. The greater tendency toward teamwork and pooling of ideas by groups of scientists, the distribution of responsibility and credit for scientific work among the whole staff of a laboratory, the greater diffusion among younger scientists of the sense of social responsibility, and the resulting tendency for social incentives to supplement more purely personal motives—these facts all indicate that it is reasonable to expect that scientists can and will participate in formulating the plans they will execute. This leads to the kind of self-government to which democratic administration tends, and which industry has found valuable as an incentive.

A further question that policy must meet is the ultimate disposition of the new knowledge which accrues from science. In the large segment of scientific research under private control, it is generally agreed that the ownership of valuable processes arising from research is to be vested, not in the individual scientist, but in the laboratory or the industry which has financed the research. Patents therefore generally become the property of the corporation by which the scientist is employed.

The question of ownership has already arisen concerning values accruing from war research, and it must enter inevitably into all plans for the future support of science.

The clearest basis for policy in this regard is that research done for a social or public purpose must be brought as quickly as possible to serve this purpose. If it is carried out for the public and at public

expense, it should belong to the public; and there is no more direct way of making it public property than by publishing it as soon as the facts are clear. Publication would preclude patenting and, with certain precautions to be discussed below, would prevent the results of public science from becoming private property. But, by the same token, the results of private science would remain private, subject to patent or other ownership rights and restrictions.

A division of this sort already exists. Most agricultural research in the United States is done at public expense and results are freely published and can be consulted and used by anyone. The greatest change in American agriculture in the present century, the introduction of crossbred or hybrid corn, resulted chiefly from cooperative research between the United States Department of Agriculture and the State Agricultural Experiment Stations. The results were quickly utilized by private seed companies, none of which was able to obtain a patent or found a monopoly on it. Crossbred corn therefore came very quickly into general use and its benefits were soon spread over all agricultural communities.

Side by side with this development, it was possible for private individuals and corporations to produce and patent new varieties of other plants, such as roses, which could be propagated asexually. The ownership of new rose varieties is thus (in general) private; but the new method of corn breeding belongs to the public.

The question of property rights need then be faced only when new values are created by publicly supported research; and the basic policy stated above—that is, free publication of the results of public research—need not interfere with existing arrangements under which private research operates. As a matter of fact, the more fundamental the research in the sense that the more general the truth that arises from it, the less will property questions arise. It is hard to find a patentable value in the general theory of relativity, or in the periodic system of the elements, or in the theory of the gene. It is the fate and the function of such ideas to become common property, and no man-made rules should be allowed to interfere with their free circulation. It is usually only the specific applications of general ideas which become subject to property restriction; and public policy can only aim at preventing such restriction from interfering with the advance of science or with the spread of the benefits to the people.

It is time now to deal briefly and in bare outline with the last question: how can these ideas and hopes about the support of science be brought into practical operation?

It seems evident that there must be an agency having as its chief concern the preservation, advancement, and diffusion of scientific knowledge. There are, in the United States, dozens of organizations having this aim in limited spheres, but that not one of them fulfilled the required functions in the national interest became evident when, in the war emergency, a wholly new and temporary agency, the Office of Scientific Research and Development, had to be created. The importance of the work assigned to this office, and the power and facilities which accompanied the responsibility, pointed not only to the need but to the method of meeting the need for a central agency of government concerned with science.

It is probable that nothing less than the creation of a cabinet department of science under a Secretary of Science can permanently meet the need. It ought to be connected directly with the central executive body of the government, because only in such a position can it be made aware of the basic problems which face the nation, and only through the political power which attaches to cabinet rank can it gain the means and facilities with which to support the study of both immediate and long-term problems.

The structure of such a department may well be different from that of other government departments because, in addition to policy making and administrative functions, it would have to serve as a coordinating agency for many existing scientific agencies, both public and private. To name only two groups of interests, it would have to be closely connected with the universities and research institutes, and with industry, since in each of these institutions needs for new knowledge are likely first to become apparent, and from each flows scientific and technical information which can be put to use in national defense and development.

At the heart of such a department could well be a board or council of scientific research which could act at once as a granting agency, allocating funds for specific researches, and as a board of strategy, seeking out neglected areas, mobilizing disparate facts and distant persons, and shifting its forces from time to time to explore new avenues of research. If it fulfilled its best purpose, it could not be

content to sit and sift, but would itself have to search and ponder in a more active way. Its basis of operation as a granting agency might well be patterned upon the Office of Scientific Research and Development in that it might receive applications for research funds from universities, research laboratories, other government agencies, or even individuals, and might enter into contracts with those it judged as offering the best prospects for needed scientific advance. Like O. S. R. D., it might find no need to become an operating agency with plants and facilities of its own, although it should have some freedom to use those methods best calculated to promote the best research.

Much would depend upon the composition of this board. It should consist of working scientists who can judge the merits of various research proposals and policies, and of representatives of those for whose benefit the research is done and who in the end pay the bills, that is, the public as represented by labor, consumers, and industry, small or large. Perhaps a proportion of eight scientists and four public representatives would express both the purposes and responsibilities of the board; and some of the scientists should be drawn from, or be primarily interested in the scientific work of, the government departments.

Since there should be no disposition on the part of such a board to displace any existing research agencies, but rather to supplement and aid them, its most important function might well turn out to be, especially in its initial operations, that of coordinating and facilitating research generally. It would undoubtedly avoid competition with industrial research, and direct its first attention to "unprofitable" fields such as exploration looking toward new natural resources, housing, public health, etc. It would probably be concerned with such public services as the provision of adequate means of publication, of bibliographic and library services, of abstracts and translations of foreign scientific literature, and similar functions.

Either this board or another one in the Department of Science would of necessity concern itself with one of the basic questions in all scientific research: how to insure an adequate supply of trained scientists for research, for education, for industry, and for public service. Its operation in this respect could well be patterned upon the fellowship boards of the National Research Council, which at present administers limited and temporary funds supplied from private sources.

Two main criticisms of the proposal outlined above may be

anticipated. One is that research cannot be free under a central direction, but will wither and die. Scientists, it is said, will not submit to regimentation, nor can new ideas, the life blood of science, be created by subsidy. The other criticism is that the needs are already met by such existing agencies as the National Academy of Sciences and the National Research Council.

The first criticism is certainly a cogent one when central *control* is proposed, but it applies with less force to a board which judges applications initiated by working scientists as individuals or groups, especially when many of the judges are themselves working scientists who know how delicate a plant original research is and how necessary is the atmosphere of freedom to its growth.

Much will depend upon the degree to which members of the board realize that any organization of this sort exists primarily to provide a material body for the mind of science. There are scientists and others who know this and who apply to organizations proposed for science two essential criteria: Does it provide the mind with adequate and proper facilities? Does it leave the mind free to strike out in new directions? Men who ask these questions are the ones whose sense of public duty would bring them into the service of such a board, just as it brought such men into the direction of war research.

In regard to the second criticism, it must be pointed out that in the war emergency neither the National Research Council nor the National Academy of Sciences proved to have the character needed for an agency to guide and administer the organization and support of science. Neither is an operating agency; and, as constituted at present, neither could provide the initiative and the administrative services which are required. The relative isolation in which they have functioned has removed them from that close connection with problems of public policy so essential for an agency to have which is to be responsive to public needs. They have the confidence of scientists and close connection with academic research and with the scientific societies and organizations and are thus well prepared to serve an important advisory function. The National Academy of Sciences, as a council of elder statesmen, could well be called upon to pass upon the qualifications of scientists proposed for membership in the Board of Scientific Research. The Academy would be less able to maintain sufficiently close relations with consumers, with labor, and with industry, and it

would be less competent to advise on questions bearing on the social relations of science in these fields.

The Board might conduct its relations with the scientific societies through the National Research Council, which could then be incorporated into the Department of Science and carry out other important functions, such as maintaining a permanent roster of scientific personnel.

It is of course possible that the Academy and the present National Research Council might be so changed as to assume the functions it is proposed to assign to the Board. The changes would be so fundamental as to constitute conversion of these older organizations into a new department of the government; and it is probable that the traditions of both institutions would make such conversion a slow and difficult process, for, in spite of their "national" character, neither has felt itself to be a truly public agency.

The foregoing brief sketch has given merely in outline form, without any attempt to develop detailed procedures for giving them effect, some of the general ideas and principles which might underlie a policy for the public support of fundamental research. Since it was written (January, 1945) several events of cardinal importance have occurred. One of these was the publication in July, 1945, of "Science, the Endless Frontier," a report to the President of the United States by Vannevar Bush, Director of the Office of Scientific Research and Development. This is at once a report on the state of scientific research in the United States and a set of recommendations concerning its future support, in answer to specific questions posed by President Roosevelt in 1944. Since it embodies the result of extensive study by committees of competent and experienced scientists serving as advisers to Dr. Bush and is documented with data on current scientific research activity, it is destined to constitute one of the bases on which public policy for science will rest. Another foundation had been provided by the discussions and hearings centering around a bill submitted to the Senate in 1943 by Senator Harley S. Kilgore of West Virginia. These two major influences, one provided primarily by scientists, the other by legislators, came together in the autumn of 1945 during the joint public hearings on five bills for the public support of science. The message to Congress of President Truman on September 6, 1945, definitely committed his administration to the support of fundamental research in

both the natural and the social sciences, including the training of young scientists and the dissemination of scientific information. This fusion of the efforts of statesmen and scholars is a fact of the first importance and transcends in its significance for the future the particular provisions for support which are now under discussion. There is no doubt that the atomic bomb greatly accelerated the speed with which scientists and legislators were approaching each other and the swift realization of the significance of science on the part of the public created the atmosphere in which progress could be made.

As a result of these recent events I think it can be said that the public, through its representatives, has now acknowledged that the support of science is a public responsibility, to be assumed in a broad and permanent form through a new agency of government to be known as a National Science Foundation. This foundation will resemble the Board of Scientific Research described in earlier paragraphs of this article, but will have even broader powers since it will include a Division of Social Sciences, and provisions for the training of young scientists and for international collaboration in science which go beyond those previously suggested.

It appears that there is now substantial agreement among scientists concerning the major purposes and functions of this Foundation, although there is still some reluctance to grant that the administration of the Foundation should be primarily responsible to the government rather than to scientists.

However, it has now become so clear that the aggressive advancement of fundamental scientific knowledge is a primary condition for the maintenance of democratic government and for the attainment of a good society that we may confidently expect methods to evolve in practice by which public control may be exercised without unduly limiting the freedom which science needs. The attainment of this goal will be hastened if scientists, acknowledging their public responsibility, will build the means, through a guild or federation with a social purpose, by which their influence on public policy can be brought to bear.