Proteins

BIOLOGY
COLLOQUIUM
1956

OREGON STATE CHAPTER OF PHI KAPPA PHI
OREGON STATE COLLEGE ⋆ CORVALLIS ⋆ 1956

Proteins

EDITED BY
TSOO E. KING
Associate Professor of Chemistry
OREGON STATE COLLEGE

This book is not copyrighted. Proper credit should be given the Biology Colloquium and to the authors if parts of the book are quoted.

TABLE OF CONTENTS

iii

FOREWORD

The Biology Colloquium is conducted in a spirit of informal discussion and provides opportunity for participation from the floor. The colloquiums are sponsored by the Oregon State Chapter of Phi Kappa Phi with the collaboration of Sigma Xi, Phi Sigma, and Omicron Nu. Sigma Xi assumes special responsibility for the colloquium luncheon. Phi Sigma, Phi Kappa Phi, Omicron Nu, and Phi Lambda Upsilon provide afternoon teas. The College Library arranges special displays of the writings of colloquium leaders and notable works on the colloquium theme.

Grateful acknowledgment is made of the cooperation and interest of the several faculties of Oregon State College concerned with biology, of those biologists contributing to the program, of Chancellor John B. Richards, President A. L. Strand, and other executives of Oregon State College.

The first Biology Colloquium was held March 4, 1939, with Dr. Charles Atwood Kofoid of the University of California as leader, on the theme "Recent Advances in Biological Science." Leaders and themes of succeeding colloquiums have been: 1940, Dr. Homer LeRoy Shantz, Chief of the Division of Wildlife Management of the United States Forest Service, theme "Ecology"; 1941, Dr. Cornelis Bernardus van Niel, Professor of Microbiology, Hopkins Marine Station, Stanford University, in collaboration with Dr. Henrik Dam, Biochemical Institute, University of Copenhagen, theme "Growth and Metabolism"; 1942, Dr. William Brodbeck Herms, Professor of Parasitology and Head of the Division of Entomology and Parasitology, University of California, theme "The Biologist in a World at War"; 1943, Dr. August Leroy Strand, Biologist and President of Oregon State College, theme "Contributions of Biological Sciences to Victory"; 1944, Dr. George Wells Beadle, Geneticist and Professor of Biology, Stanford University, theme "Genetics and the Integration of Biological Sciences"; 1945, Colloquium omitted because of wartime travel restrictions; 1946, Dr. Robert C. Miller, Director of the California Academy of Sciences, theme "Aquatic Biology"; 1947, Dr. Ernst Antevs, Research Associate, Carnegie Institute of Washington, theme "Biogeography"; 1948, Dr. Robert R. Williams, Williams-Waterman Foundation, theme "Nutrition"; 1949, Dr. Eugene M. K. Geiling, Head of the Department of Pharmacology, University of Chicago, theme "Radioisotopes in Biology"; 1950, Dr. Wendell M. Stanley, in charge of Virus Laboratory, University of California, theme "Viruses"; 1951, Dr. Curt Stern, Professor of Zoology, University of California, theme "Effects of Atomic Radiations on Living Organisms"; 1952, Dr. Stanley A. Cain, Conservationist, University of Michigan, theme, "Conservation"; 1953, Dr. Wayne W. Umbreit, Head of the Department of Enzyme Chemistry, Merck Institute for Therapeutic Research, "Antibiotics"; 1954, Dr. Daniel Mazia, Professor of Zoology, University of California, theme "Cellular Biology"; 1955, Dr. Ernst Mayr, Curator, Museum of Comparative Zoology, Harvard University, theme "Biological Systematics."

Committees

COLLOQUIUM COMMITTEE:

Tsoo E. King (chairman), Vernon H. Cheldelin, G. E. Crossen, Ernst Dornfeld, P. R. Elliker, D. M. Goode, H. P. Hansen, H. Krueger, L. F. Remmert, H. W. Schultz, W. H. Slabaugh, W. P. Stephen, C. A. Storvick, J. G. Collins, Sally Hornecker, Nicholas J. Kirkmire, Paul Kitos, Patricia Morton, Marilyn Sherburne, Janet Smetts, Douglas Van Fleet, Felicia Wirtz.

PHI KAPPA PHI EXECUTIVE COMMITTEE:

William H. Carlson (president), J. R. Dilworth (president elect), H. P. Hansen (past president), Kenneth Munford (journal correspondent), Henry Ten Pas (secretary-treasurer), Bill Brumfield (vice president), Nancy Allworth (assistant secretary).

SIGMA XI EXECUTIVE COMMITTEE:

Hugo Krueger (president), Allen B. Scott (vice president), David Nicodemus (secretary), Arvid T. Lonseth (treasurer).

PHI SIGMA EXECUTIVE COMMITTEE:

Denzel Ferguson (president), Theodore Dyrness (vice president), Jack Foster (treasurer), Felicia Wirtz (corresponding secretary), Sam Dunn (recording secretary), Bob Hughes (editor), Ivan Pratt (faculty adviser).

OMICRON NU EXECUTIVE COMMITTEE:

Marilyn Sherburne (president), Priscilla A. Horning (vice president), Janet G. Smetts (secretary), Sally A. Hornecker (treasurer), Dorothy J. Eddy (bulletin board chairman), Diane Griswold (editor).

PHI LAMBDA UPSILON EXECUTIVE COMMITTEE:

Paul Kitos (president), Don Reed (vice president), Will Johnson (treasurer), Jerzy Bialy (recording secretary), Vance Gritton (corresponding and alumni secretary).

Colloquium Speakers

Otto K. Behrens, Ph.D., Director of Biochemical Division, The Lilly Research Laboratories, Eli Lilly & Co., Indianapolis.

Henry Borsook, Ph.D., M.D., (Colloquium Leader), Professor of Biochemistry, California Institute of Technology, Pasadena.

Frederick H. Carpenter, Ph.D., Associate Professor of Biochemistry and Virus Laboratory, University of California, Berkeley.

Walter B. Dandliker, Ph.D., Associate Professor of Biochemistry, University of Washington, Seattle.

Edmund H. Fisher, Ph.D., Associate Professor of Biochemistry, University of Washington, Seattle.

Heinz Fraenkel-Conrat, M.D., Ph.D., Research Biochemist, University of California, Berkeley.

Rollin D. Hotchkiss, Ph.D., Member, Rockefeller Institute for Medical Research, New York.

Choh Hao Li, Ph.D., Professor of Biochemistry and Experimental Endocrinology, University of California, Berkeley.

A. L. Strand, Ph.D., President, Oregon State College, Corvallis.

Presiding Officers

William H. Carlson, M.A., Librarian, Oregon State College; President, Oregon State Chapter of Phi Kappa Phi.

Vernon H. Cheldelin, Ph.D., Professor of Chemistry and Director of Science Research Institute, Oregon State College.

Tsoo E. King, Ph.D., Associate Professor of Chemistry, Oregon State College.

Hugo M. Krueger, Ph.D., Professor of Physiology, Oregon State College; President, Oregon State Chapter of Sigma Xi.

Howard S. Mason, Ph.D., Associate Professor of Biochemistry, University of Oregon Medical School.

Francis J. Reithel, Ph.D., Associate Professor and Acting Chairman of Chemistry Department, University of Oregon.

Seventeenth Annual Biology Colloquium

Theme: PROTEINS

Colloquium Leader

HENRY BORSOOK, Ph.D., M.D.

Professor of Biochemistry,
California Institute of Technology, Pasadena

Opening of the Colloquium

DR. T. E. KING: Welcome to the Seventeenth Biology Colloquium. This is the first time the colloquium has been extended over two days.

Two years ago we had a colloquium on cellular biology. If one wants to break down the cellular unit to the subcellular or molecular level, one will immediately find a common denominator in cells from all tissues of all phyla and even in viruses which do not have cellular structure. This common denominator is a class of compounds known as proteins.

Proteins are not only building blocks for any organism but also possess catalytic actions. These facts have been recognized for a number of years, and research on proteins has been vigorously pursued for over a century. Unfortunately, the systematic information on this subject is still incomplete. It is a pity to see that continuous attempts have been desperately made to achieve understanding about the functions of cells and mechanisms of biological processes in spite of only fragmentary knowledge of proteins available.

For this reason the topic of proteins is selected for this year's colloquium. The program is planned on the basis that studying proteins is a preface and foundation for pursuing life sciences. A general discussion of the role of proteins in biology will be given by the colloquium leader.

The debate about the homogeneity of proteins is still going on. Proteins, however, can be considered as chemical entities and follow the fundamental laws of chemistry and physics.* By this reasoning, one would wonder whether those actions due to proteins can also be profitably pursued with these fundamental laws. In other words, can we apply the fundamental laws of physics and chemistry to study life phenomena? Three speakers will present aspects of this approach. The present knowledge on physicochemical properties of proteins is relatively more than that on physicochemical interpretations of biological action. Dr. Fischer will discuss, from the standpoint of chemistry, one of the actions of proteins—enzymic reactions.

Another remarkable property of proteins is the actions of some key hormones. Minute quantities of hormones can control many vital biological processes. Dr. Behrens and Dr. Li will review the chemistry and biology of glucagon and anterior pituitary hormones.

Now if we go a step further into cellular constituents we will find that another class of compounds, nucleic acids, occur in cells and viruses. Even the simplest or smallest virus contains nucleic acid. In fact some viruses are *pure* nucleoproteins. Nucleic acids do not occur coincidently with proteins. There is intimate interrelationship between them. The importance of nucleic acids in genetics has been rapidly explored in the last few years. Dr. Hotchkiss will give us a review along this line. Related to the genetic role of nucleic acids, Dr. Fraenkel-Conrat's talk will be on reconstitution of tobacco mosaic viruses with proper nucleic acids and proteins. The question about how proteins are formed in organisms is certainly the one many people would wish to ask. Here nucleic acids come out again. Dr. Borsook will discuss biosynthesis of proteins and its relationship with nucleic acids.

I want sincerely to thank all speakers who have come here to spend a weekend with us for

---

* These fundamental laws include thermodynamics of irreversible processes or steady state. I, just like you, have been aware of the complexities of life phenomena. But I don't believe it is profitable and would object to think of having one set of laws for the "non-living" world and another for biological reactions. However, in studying the latter, one has to consider *organization*. It is a well-known fact that enzymes in a cell are not like enzymes in a bag. There is a pattern of organization. This very fact of organization is one of the characteristics of living organisms and may be one of the determining factors of the biological evolution.

the Seventeenth Biology Colloquium. Since the colloquium is conducted in a very informal way I am sure many spirited discussions will come from the speakers and the floor alike.

In starting this Seventeenth Biology Colloquium, I should like to pay tribute to Professor Delmer M. Goode, curriculum consultant and editor of publications. As president of the Oregon State Chapter of Phi Kappa Phi in 1939 he farsightedly initiated the annual academic event of the Biology Colloquium. Since then, his constant and untiring contribution and enthusiasm have helped to make the annual colloquiums successful.

The official welcome to the leaders and participants in this colloquium will be given by Dr. A. L. Strand, himself a biologist.

PRESIDENT STRAND: Dr. King, Dr. Borsook and other Visiting Leaders, and Members of the Colloquium: This week on the campus, when we have the Condon lecturer as well as the speakers of the Colloquium here, is a stimulating time. However, this particular year we've overdone things with two other big attractions, the Parade of Progress and the Model United Nations.

I understand a new science called Group Dynamics has come into being. I wish some Clausius among the group dynamicists would show us how to be in several groups at the same time.

Some years ago on this campus, in speaking of the interrelatedness of all things, Professor Childs used a text from the New Testament: "Now the coat was without seam, woven from the top throughout," or substantially "the garment was of one piece." A good text, and a broad concept to be sure, and one, as applied to Science, which has been tested thoroughly during the last hundred years—that is, since the idea of the "continuity of forces" was first stated.

The exciting things in science are really not the sort of side-show spectacles, rubber exploding out of pop bottles, etc., as we have going on now on another part of the campus. The exciting things are the discoveries that relate phenomena, that relate sometimes fields that seemed to be far removed from each other.

As Dr. King said, when it comes to this, proteins are the great continuum. They live up to their name which I understand means "first" or foremost.

Last year the subject of the colloquium was "Systematics." That would seem to be a distant field from biochemistry. At least when I was learning the names of the phyla of the animal kingdom some forty years ago, and about orders, families, genera, and species, neither I nor my teachers, I'm sure, imagined any connection with proteins. Yet, within a few years, a young man in the department was using precipitin reactions to show the relationship between the genera of Noctuid moths.

Year before last the subject was "The Cell, or Cellular Biology." Suffice it to say that back in some far distant time, when life was first formed, some protein molecules got together to make the first cell.

And so it goes with most of the subjects chosen for the colloquium in the past. Proteins are one mechanism, evidently, by which the garment of life processes becomes of one piece.

It is probably redundant to point out one difficulty in modern science. The garment is of one piece all right, but in our examination of it we get down to such fine threads that the shape and character of the whole becomes a bit obscure. Some scientists dwell on such small areas or particles of the garment that nobody but their own subspecies confederates can understand the special terms they have generated to describe what they have seen.

I am not one who holds to the proposition that complex things can be made simple if you are clever enough. That is not so, and I suppose it is impossible to name anything more complex than proteins. Nevertheless, it is the old ideal of the colloquium to sometimes step back from the micrometric position and take in more of the fabric, to point out relationships, to say why certain things are important, to turn the garment around, or inside out, even at the risk of offending fellow specialists.

Accordingly, I was happy to see that our distinguished leader has chosen a broad synthesis, "The Role of Proteins in Biology." The other day Dr. Beadle called proteins exciting. That is an apt description, and I think we have some exciting things about them coming up.

We're glad to have you here. I wish, too, to express our thanks to the visiting scientists who have made the program possible.

The Role of Proteins in Biology

HENRY BORSOOK

It is commonplace among historians that every generation needs to rewrite history; it is the same with us. Proteins are a familiar topic, and it is good every once in a while to take up the familiar facts, look at them again, turn them around, and see whether they still appear as we thought. This I propose to do this morning. I shall have very few new facts to present to you.

It would take a very long time to read merely a list of all the different processes in which proteins participate. I use the word "participate" because the title suggests the particular aspects of proteins I shall talk to you about this morning. It is their role, in other words, what they do. It would be silly for me to take up one protein after another and say what I can about each. Rather I shall talk about certain general aspects of what proteins do and how they behave. I shall begin by going back a little and say that, as far as the role of proteins in biology is concerned, the 19th century ended in 1935. Until then we were dominated by the 19th century mechanical way of looking at things. I say "mechanical" advisedly rather than mechanistic. The whole of the 19th century, a very great century for science, was a mechanical one. Its science as well as its commerce was dominated by the steam engine. The discovery of the steam engine led to the development of a science which taught men how to make better steam engines. They learned the limits of their size and efficiency. This was the science of thermodynamics. Thermodynamics dominated all of science, including biology, of the 19th century. It was common for biologists, I suppose in order to appear respectable to their more impressive colleagues in physics and chemistry, to refer to living things as machines. And so the animal, man, was referred to or described as a machine. Our structure was likened to that of a steam engine. The cast iron or steel structure of such an engine changes very little, as it burns fuel, except for a slight frictional wear and tear. Our bodies were thought to be the same; the food we ate was the fuel. Only a small fraction of the daily food was used to replace the frictional wear and tear losses. As I say, this was the picture that dominated physiology and biochemistry, as far as proteins were concerned, up to 1935. At this time the mechanical picture was seriously and explicitly challenged. An alternative interpretation of the evidence was put forward that there is no useful distinction to be made between the machine and the fuel, that all the substances in the body are changing, they are all being broken down and rebuilt continuously and rapidly. For example, in a man in nitrogen balance about $\frac{1}{2}$ of the nitrogen in the urine in a 24-hour period comes from the food and the other $\frac{1}{2}$ from the breakdown of his body proteins, and so, since the man was in balance, a moiety corresponding to the latter half was resynthesized. I remember talking about this at that time with a very distinguished physicist. "I can't see what you're making such a fuss about," he said. "What is the difference? The whole difference is between 3 grams and 8 or 10 grams of nitrogen a day in the urine. Why all this fuss?" The difference was between a machine, a lifeless thing, and a biological view of the organism. In order to try to impress him, I said, "Do you realize that about every seven days half of your liver is new?" That set him back. The half life of liver proteins in animals is about seven days. I say it is a much more biological view for these reasons. Once we become accustomed to the concept of the cell, of the proteins in the cell as in a state of continous flux, it becomes easier, even without knowing the details, to envisage how a cell or an organism adapts itself to the great vicissitudes to which it is exposed in the course of its life history, to great changes in diet, activity, reproduction, growth, and disease. It was hard to see how this happened with the old steam engine concept.

It was hard to see how antibodies and adaptive enzymes could be formed as needed. From the new view it is easy, and logical—"natural" we say; and now, of course, we take it all for granted. Antibody formation, adaptive enzyme formation, these are great wonders; and to have won to a point of view which says "why yes, of course, this must happen," is something for those of my generation to wonder at still. I remember when I first considered this question, many years ago, I was struck with a phenomenon that must be very familiar to you who live in this part of the world: when the salmon go up to the spawning grounds, they may travel as far as 750 miles.

3

During this whole time they do not eat. They convert muscle proteins into gonadal protein, and the process is so regulated that, throughout this long journey during which a fish is being transformed to little more than a skinny carrier of gonads, there is no change in the free amino acid concentration of its blood and tissues. It was facts like this that led one to challenge the old mechanical concept.

The biological point of view is all very well, but we nevertheless must come face to face with hard chemical facts. We have to look at proteins as chemicals and the biological point of view must grow out of such facts. Let us do what a biologist always does when confronted with a new field of investigation: he tries to make a classification. A good classification I tell my students is an idea; it is not merely a list. We may have many classifications according to how many different ideas we may have. What kind of a classification can we make of proteins? There are so many proteins that it might take a whole hour to read their names. We might say there are skin proteins, parenchymatous proteins, storage proteins as in seeds, the proteins of the tissue fluids, proteins concerned with respiration, contractile proteins, and so on. A chemist would turn his nose up at such a classification because it does not tell us anything about what proteins are chemically; and he would be right.

We might turn around by classifiying proteins from a functional point of view. Can we make any kind of useful classification by considering stable proteins, and labile proteins? This is the chemical way of looking at function. Let us look at stable proteins; it will at least illustrate some of the different kinds of proteins. Last year, some of my students raised the question, "How labile is hair?" So we did an experiment. We fed a rat a C^{14} labeled amino acid; we painted the hair of the rat before we started and fed the labeled amino acid for a few weeks, then we cut the hair off to see which part of the hair, if any part, was labile. We could tell that by the presence on the labeled amino acid in it. This was very easy. One slid the hair along its length from its root to tip under a Geiger counter, to measure the C^{14} if present. The only part of the hair that was labile was the part immediately emerging from the hair follicle; the hair protein that had emerged prior to the experiment was inert. It had no metabolism, it had no turnover, it had nothing in it alive. It is not merely that hair or horn or nails are in the skin

as against proteins in solution that determines whether a protein is labile or not. Hemoglobin is a most active protein functionally. But in the mature erythrocyte it has no turnover. It is stable until the red cell is destroyed. Proteins on the skin as a whole are rather stable. The skin gets worn off and so is continually renewed, but once the proteins are in the skin they have a very slow turnover. Muscle appears to be composed in the main of two classes of protein: those which have a slow turnover and those which are much more labile. The most labile proteins are in viscera, liver, intestines, and in the circulating proteins of the plasma. If one looks a little closer one sees that function does not provide a useful distinction between labile and nonlabile proteins. Enzymes are the machinery of the cells. Their turnover in some cells is very fast. The digestive enzymes of course are continually being made and secreted. Their turnover is very fast, but this is not unexpected. It was not predictable that the turnover of enzymes would be fast in the liver; they are among the most labile proteins in liver. Antibodies circulating in the blood stream are as labile as the serum proteins, which are continually being made and continually being broken down. As far as we can see now, whether a protein is labile or stable depends largely on where it is situated; in other words, it depends on what other proteins are present, what happens to it.

This is a nice biological concept, but it is the very devil for a chemist to have to deal with. Biologists deal in large units; cells until recently were the smallest of such units. Virus particles are smaller; they have their biology, as it were. For the chemist the cell is a universe; the units which make up a cell for the chemist are at the molecular level. He would like to have a classification of proteins at the molecular level, as he has of other organic molecules. We have no such classification of proteins at present.

Nevertheless let us look at proteins as a chemist does. As you know, we can write proteins as:
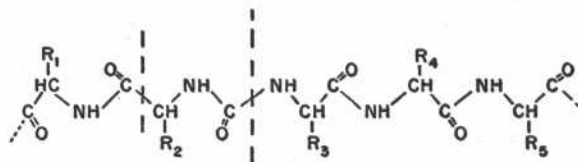
The segment between the dotted vertical lines is an amino acid unit: an acid radical, a basic radical, and a side chain, which is the part that

sticks out. Most proteins have many amino acids. There are twenty odd different R's. The amino acids are linked as shown through the conjugation of their acid and basic groups into peptide bonds. There are two amino acids, cysteine and methionine, that contain sulfur. Cysteine contributes to another type of linkage that Dr. Fraenkel-Conrat spoke of yesterday, the S-S linkage. The sulfur of one cysteine residue is linked with that of another, very often on another chain. Such a linkage between peptide chains gives a very firm structure. It is like a beam across a roof. This is one way, then, in which a chain is so linked. Phosphorous may in some cases contribute to make a cross linkage. Dr. Schweet, in my laboratory, has found evidence of still another kind of linkage, the ε-amino group of lysine to the γ-COOH of glutamic acid. Such a linkage makes for bending and twisting of the whole molecule.

The carboxyl and amino groups of amino acids are really only a means to an end. The important chemical feature, as far as biological action of proteins is concerned, is not really the peptide bonds, but the structure that is formed. The R groups, twenty odd, have a great variety of shapes and sizes. They may be 1 hydrogen, a CH_3 group, a straight chain of 4 or 5 carbons, they may be branched, they may carry a plus charge or a negative charge, they may be an aromatic ring, and so on. The long peptides of proteins are not straight chains, but spiral coils, and sticking out on all sides are different shapes and sizes of the side chains. Because of their difference in shape and size the R groups do different things. It is from the topography, the three-dimensional topography, from which the really important biological properties of proteins come. All the important and interesting biochemical properties depend on this topography. The biological specificity comes from the topography. I shall return to this topic shortly.

Before that I want to say something about the structure of proteins. We may refer to the peptide links as the primary structure; the bonds holding peptides together, such as the SS bonds, as the secondary structure. The peptide bonds are in spiral coils and there are loose bonds between the whorls of the spiral; these bonds may be called the tertiary structure. Then larger complexes are held together loosely, in what may be called a quaternary structure. It is the last two, the tertiary and quaternary structures, which are so important for the biolo-

gical properties of proteins. The tertiary and quaternary structures depend on hydrogen bonds largely; they are not very strong bonds, but they are strong enough to determine the shape of the protein molecule, they determine the coiling, the way coils stick to each other, the way they bend and how the large molecules stick to each other. Now all of this: the kind of coiling, where the SS bridges go, where the phosphorous bridges go, the epsilon linkages of the lysine, all depend on and in turn constitute the topography of the protein. The topography of the protein is a 3-dimensional, relatively very large area. The side chains that constitute the topography are fairly reactive. There are positive charges, minus charges, hydroxyl groups, aromatic linkages, and so on, scattered all over this surface. It is not a surface, it is an up and down affair, and so, as you might expect, a lot of things can and do happen on it.

One of the quite wonderful chemical properties of proteins is that they are so large. Their largeness, their amino acid pattern, and the details of their topography are all inherited. How this comes about we do not understand at all. If one set out to duplicate exactly all the topographical features of a village, to do it again and again, we would find great difficulty even to draw it exactly the same again and again and again. Yet, in two egg whites, whether the egg was made and laid now or 50 years ago, whether it was in this country or in China, as far as the chemists can tell, and as far as the immunologists can tell, the protein in those egg whites all have the identical topography. When the chemist isolates and purifies a protein he is creating an artifact. Chemistry is artifact; biochemistry is reality. The chemist makes things simple because all he can handle are simple things. In biology, that is in living cells, there are no pure proteins, they are all conjugated proteins. They are all combined with something and with each other. We can begin only with the properties of pure crystalline proteins, and then it is our task to see what happens when proteins are conjugated, in other words, when they enter into the constitution of the mess of a cell, the mess, that is, from the point of view of a chemist.

I shall turn now to the topic that was discussed at some length yesterday afternoon by Dr. Fraenkel-Conrat and by Dr. Fisher, to some of the consequences of the large size of protein molecules. Proteins, as molecules go, are pretty big. I try to think of some analogy that might

illustrate their relative size. Let us compare in this respect a common chemical, sodium chloride, and a common protein, hemoglobin. Their molecular weights are roughly as 1 to 1,000 times— that is, sodium chloride is 58.5 and hemoglobin is about 67,000. Let us say that roughly the specific gravities are the same, so that hemoglobin has about 1,000 times the volume of sodium chloride; but volume is a little misleading here. As we saw, proteins are made up of long chains of amino acids; human hemoglobin is really 4 proteins stuck together, so in order not to be unfair to sodium chloride, we divide the number of amino acids in hemoglobin, 583, by 4 and say there are 146 amino acids in a hemoglobin unit, and that 4 of these units make a molecule. If we take sodium chloride as our unit of size, then the hemoglobin molecule is 73 times as long and about 10 times as high. Sodium chloride is to hemoglobin as one small building to a fairly large village. This, I think, is a fair comparison. It is a fair comparison as to the different kinds of things that can go on in a village, the organized association of many structures, as compared with the many fewer activities possible in a single building. In the picture Dr. Dandliker drew of a protein molecule there were channels within the molecule. These spaces permit traffic through a protein molecule, traffic of water and of dissolved substances. Compared to sodium chloride, or to any one constituent amino acid, a protein is really like a community.

Dr. Fraenkel-Conrat pointed out that some of the reactive radicals in native proteins are tucked away within the cells so that other things cannot get to them. This property also is inconceivable in an ordinary small molecule. This is one of the consequences of the large sizes (and the consequent coiling) of protein molecules. The osmotic effects of proteins in tissues depend on their larger size; they cannot get through the cell walls. At least we say that proteins stay in the blood stream or they stay within the cells because they are big molecules and they cannot get through pores in cell walls which smaller molecules can get through. I suspect this is wrong, and that it is more a matter of solubility whether proteins can get through the cell wall or not. Leucocytes can get through cell walls; why not a protein? We know nothing of the organization and function of cell membranes. On the proteins' extensive topographical surfaces charges are scattered; these are neutralized at the approximate neutrality of the tissues, and this is respon-

sible for most of the familiar osmotic effects. Another well known phenomenon, contraction, whether of a muscle or of a flagellum, obviously depends on the large size of the protein molecules that make up the contracting fibers; they must be long to be able to fold up. But something much more subtle and important to the life of all cells is a consequence of the large size of the protein molecule. Figure 1 is, I think, not a false picture
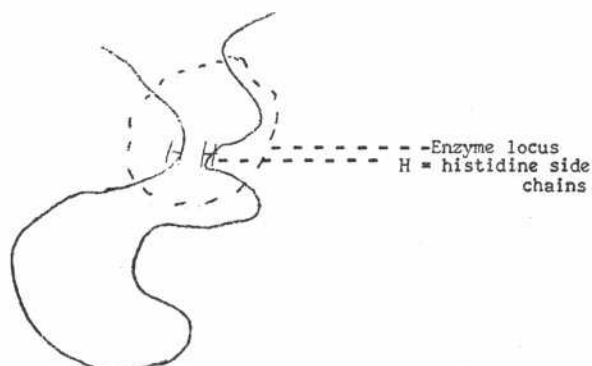

Figure 1. Diagram of the locus of an enzyme: fumarase.

of the nature of an enzyme locus—in this case, fumarase—even though it is as simple as a child's drawing. The enzyme locus consists essentially of two histidine side chains in the right relation and distance to each other. On either side of the locus the configuration derived from the other side chains fits fumaric acid and malic acid so that the enzyme can insert or remove water from these two acids. In order to bring these two histidine radicals just the right distance apart, or together, whichever way you may want to put it, the molecule had to be able to be folded on itself and such folding is possible only if the molecule is big. The same applies to the area around the two histidine side chains. As Dr. Fisher pointed out yesterday, there is considerable evidence that, whenever we have an enzyme reaction which undergoes a large change in activity in the neighborhood of neutrality, histidine side chains are involved in the enzyme locus. We may say this because, of all the amino acids, only the side chain of histidine undergoes large ionization changes in the neighborhood of pH of 7.

One of the most important properties of hemoglobin, as you all know, is its ability to combine with oxygen, and when it is combined with oxygen it is 20 times stronger as an acid than when it has lost its oxygen. It is this property of

hemoglobin that makes hemoglobin the most important pH buffer in the blood. The radical in hemoglobin which undergoes the change in its strength as an acid is that of histidine. I do not want to make out that histidine is the only amino acid involved in enzyme action, or in hemoglobin; that is obviously nonsense. The point I do wish to make is that we can begin to understand something as deepseated, as subtle, and wonderful as an enzyme locus, in terms of an amino acid side chain, and as a consequence of the fact that it is built into a molecule which is big and can fold; the folding is a consequence of the bigness. Certain proteins can become antibodies. Corresponding to the antigen, they are folded into a shape that is specific for the antigen. Again it is the large size that confers on it the subtlety of being folded in different ways, and each different kind of folding makes it biologically, though probably not chemically, a different protein.

With such simple ideas one can begin to make pictures, that on the whole are reasonably satisfactory, of these profound and subtle functions. We do not need to invoke anything mysterious or strange, only the well-known properties of the side chains of the amino acid and certain simple consequences of the folding of the protein molecule. A change in the folding of the protein molecule may be what happens in disease. A bacterial toxin, the diptheria toxin, spills out into the blood stream: the heart becomes weak, the nervous system is affected. How? It may be that the proteins are denatured by the toxin—i.e., the folding of the proteins is changed. This is a speculation, of course. A phenomenon that has been known for a very long time is that of the denaturation of proteins. In other words, what happens when you boil an egg. It is now fairly sure that denaturation consists in a change from the specific form of folding to a nonspecific, random folding. Denaturation is reversible if the initial unfolding and subsequent refolding have not gone too far. Then the chances are pretty good that the protein will fold back to its original form. It is easy to see that if the initial unfolding has gone far the chances are small of its folding back again exactly as it was before. The probability is very small that, once a protein is really unfolded or uncoiled, it will ever fold back again the way it was, and we say this is irreversible denaturation. This picture of denaturation also depicts how radicals like SH groups may be so buried within the protein in the native protein

that they are inaccessible to reagents, and that denaturation makes them accessible. Whenever I get a chance, I ask pathologists why they do not begin to study the denaturation of proteins by various toxins in order to get a more deepseated understanding of pathology. As far as I know very little attention is being given to this. The tools, the methods, the ideas, the mathematics are available now for the study of denaturation.

The main point that I wish to make in this connection, however, is that to understand some of the most important biological properties of proteins we need neither invoke nor invent any special or unique chemical properties of proteins. All we need to take into account are the well-known chemical properties of the radicals of the amino acids, whether they are plus or minus charges or hydroxyl groups, or imidazole groups, and so on, and the consequences of the size and folding of proteins on the relations of the side chains and the topography. Even in the case of viruses we need invoke nothing basically new. When I was an undergraduate there were attempts to create a new science, called Colloid Chemistry. The professors of that science asserted that the laws governing the behavior of big molecules such as proteins were different from the ordinary laws of chemistry and physics, which were for small molecules. We have abandoned this position even when we enter the realm of the most fundamental biological properties.

I now come to a more general, a more biological part of my topic. I shall discuss, in a very summary manner, the concept of protoplasm. It is a funny word; it is a bad word, really. Its meaning has continually been changed, and I think I can best introduce it by reading the comment to you of one of the wisest and most profound of the biologists of a generation or two ago, the great cytologist, E. B. Wilson. Asked whether we may in any measure advance our understanding of cell activity by the assumption of an ultramicroscopical organization or metastructure of protoplasm as distinguished from its chemical and molecular constitution, he said:

Up to the present time [this was in 1926] no single theory of protoplasmic structure has commanded general acceptance, and it is more than doubtful whether any universal formula for this structure can be given. We are driven by a hundred reasons to conclude that protoplasm has an organization that is perfectly definite, but it is one that finds visible expression in a protean variety of structures, and we are not in a position to regard any of

tion. It is not the macromolecular organization that E. B. Wilson referred to. He wanted something bigger. Do we need anything bigger?

DISCUSSION

DR. ROLLIN D. HOTCHKISS: This has been a very pleasant lecture, and not content with merely beating my hands together, I want to make a few sound waves of another kind which tend to show how much I appreciate it. Dr. Borsook has given us a picture of some of the things we ought to know about protein in order to know how it acts, and this is very important because it touches on the question of how it comes about that these proteins are synthesized and put together by cells over and over again in the same way, and always do precisely the same kind of job. This, Dr. Borsook has so well indicated, the kind of thing that we have to know about proteins, is not merely in what order these units are put together, these amino acids which are such convenient units for chemists to find. We tend to think that, because these are the convenient breaking points of the molecule for chemists, these building blocks tell the whole story. Probably those amino acids are often very convenient to the cell too, but to understand the enzyme fumarase you have to get the concept of di-histidine-ness—for example, you have to know something about how two histidines are folded into its surface. You have to know this kind of thing about a protein, and it may be this kind of thing that matters most when it acts, and when it is synthesized. The picture that he has given us of a village is also useful in this connection. We can analyze villages into the dairy store and the meat market, the dry goods store and so on, but we have other things that make up a town, and it may be more important to know that the laundromat is next to the library, and while you are having your washing done you can read a book, or drop into the church for a bit of meditation. It is this kind of concept that I think Dr. Borsook has put in the picture so well, and we must remember it as we follow the attempts biophysicists and biochemists are making these days to explain how proteins are organized.

DR. HEINZ FRAENKEL-CONRAT: I'd like to comment on one particular idea that Dr. Borsook

expressed. It is a very interesting one that I think should be stressed: the idea that disease is a protein denaturation phenomenon. There is some work going on right in his own back yard which really fits in with this concept very well, if one is willing to accept evidence from diseases of bacteria. For certain bacteriophages can attack bacteria only in the presence of a cofactor which is tryptophane, and there is good evidence that that tryptophane locally denatures the cell wall of the bacteria. All rules of denaturation seem to hold for this reaction, so that it seems that trytophane is a denaturing agent for locally loosening up protein structure so that the phage can get into the cell.

DR. VERNON CHELDELIN: I wonder in that connection, Dr. Borsook, would you want to view degeneration and old age along the same lines?

DR. BORSOOK: Well, this comes to mind at once, but one shies away from the simple explanation. The simple pictures we make, that seem so logical, nearly always have turned out to be wrong. But I will say, what brought this to my mind was what happens to a child with diphtheria. What made the child sick so long, even when there was no obstruction to breathing, was the diphtheria toxin poured into the blood stream. Just what does this diphtheria toxin do? Why does the heart become weak? The heart muscle is not chewed away. My guess is that in disease processes such as this we are dealing with protein denaturation of the heart muscle or nerve cell, and so on. As for what happens in old age, let me tell of a man who came to see me a week or two ago. He was really sophisticated and experienced in the chemistry of plastics. He knew all about cross linkages, how to make them, and how to measure them, and so on. He said, "Old age is that we gradually accumulate too many cross linkages in our proteins." My answer to him was, "May be." This is the only answer I can offer to Dr. Cheldelin's question.

COMMENT: I would like to mention a disease known as "hemoglobinalysis" which we have observed in our department at the medical school. This disease is due to hemoglobin A. I think it is a very good example of a disease being directly referred to the denaturation of the protein in the body.

these as a universally diagnostic of living substance. As far as *visible* structure is concerned, no satisfactory distinction, practical or logical, can be drawn between a "primary" or "fundamental" structure, and a secondary one. The fundamental structure of protoplasm lies beyond the present limits of microscopical vision. [It does not lie beyond the limits of our vision with the electron microscope.] Probably the only element of protoplasm that will be admitted by all cytologists to be omnipresent is the "homogenous" hyaloplasm, which offers to the eye no visible structure.

Just a word about the history of the word "protoplasm." It was invented by Purkinje about 1840, and he used it (he was an embryologist) to denote the formative material, as he called it, of the animal embryo. Even this is a little vague now. About a decade later Von Mohl used the same word for the contents of plant cells. The word got kicked around and was never very clear until Strasburger in 1882 tried to clarify the usage of "protoplasm" by designating the material of the nucleus as the "kayroplasm," and the material which is outside the nucleus but inside the cell as the "cytoplasm," and kayroplasm and cytoplasm together as "protoplasm." In 1919 the cytologist Harper said:

> The structure of protoplasm is the structure of a cell. The search for some ultramicroscopic structure of living substances as such, and more deepseated than cell structure, has so far proved as vain as the older attempt to demonstrate the existence of a vital force.

Harper is obviously wrong. There are the nucleus, the genes within them, in the cytoplasm there are mitochondria, microsomes. These are organized particles of great biological importance. There is a structure. What have the biochemist and the protein chemist to say about protoplasm? They might well ask the cytologist, "What do you demand of this macromolecular structure? What does the organized structure have to do? What phenomena need to be accounted for that a simple solution of protein molecules will not account for?" This is the fundamental question. It is a fairly old one. You remember the long and bitter controversy between Pasteur and Liebig about fermentation. Pasteur held that fermentation was a property only of the living cell. Liebig held that fermentation is a result of the vibration of certain molecules within the cell and it is the transmission of this vibration which causes the glucose molecule to fall apart into CO_2 and alcohol. Liebig was right, of course, and Pasteur was wrong. Fermentation is an enzymatic affair and, as Buchner

was the first to show in 1896, can be made to go on, apart from the intact cell. Is the cell, or a particle, any more than a bag of proteins in solution, of RNA and DNA? From the present evidence we must say "yes." A smashed cell cannot reproduce. Immature red cells make hemoglobin very fast. Disrupt these cells by the gentlest means, even without disruption merely disturb their internal structure by changing the internal osmotic pressure, and protein synthesis stops. Dr. Fisher pointed out that in the cytoplasmic particles the rates of oxidations are very much faster than can be attained by any concentration in solution of oxidative enzymes, substrates, and coenzymes, and so on. Oxidative phosphorylation stops when one destroys the structure of the cytoplasmic particles. But one may say, "Yes, but all the structure does is to keep certain enzymes close together. Once you smash it up, the enzymes, etc., float around, and taking into account microscopical and molecular distances, they are too far apart, and this is the reason that things are slower, and certain phenomena like reproduction and oxidative phosphorylation are no longer possible." There is something in this. For example, let us take one of the well-known biochemical cycles, the Krebs cycle. Start anywhere you like, let us say with pyruvic acid and oxalacetic acid. There follows a sequence of changes, to active acetate, through coenzyme A on to citric acid, isocitric acid, oxalosuccinic, ketoglutarate, and so on. We do not need any organization for this sequence. All one needs to do is put the enzymes with their coenzymes into a beaker and it will go. Acetate and oxalacetic will go only to citric and citric to isocitric and isocitric to oxalacetic, etc. The whole ordered sequence of events is not dependent at all upon organization of the catalysts. And so I shall turn around and ask a biologist, "Suppose we look forward to a time when we shall know how proteins and nucleic acids are held together in an orderly way. This will then tell us the organization of a nucleus, the organization of a microsome or mitochondrion, the organization of a cell membrane. Will this be enough for you? Do you need any more organization (I come back to the bad word "protoplasm") than this?" With the methods, the ideas, the knowledge that we have, I think we can look forward to a time when we shall have adequate working pictures of protein molecules in all their foldings, of nucleic acids and all their foldings, of the forces between these. This is an organiza-

Size, Shape, and Activity Coefficients of Proteins in Solution

WALTER B. DANDLIKER*

The methods available for determining the size and shape of protein molecules in solution may be conveniently divided into two classes: the equilibrium and the kinetic. The first class includes osmotic pressure, sedimentation equilibrium, and light scattering and can be treated by the methods of classical thermodynamics. In the limit of low concentrations the chief result of the equilibrium methods is usually the molecular weight while data at higher concentrations indicate the deviation from dilute solution laws and thus make possible the determination of protein activities.

In the kinetic methods the protein particle is subjected to an applied force resulting in movement of the particle and some change in the properties of the fluid. This group of methods includes sedimentation velocity, diffusion, viscosity, and double refraction of flow. To interpret the results of these experiments requires some hydrodynamic theory relating particle size and shape to observable macroscopic quantities.

During the last few years advances in this field have been made along several lines including new or improved theoretical treatment, more exact application of existing theories, and greatly improved instrumentation. It is mainly with some of these developments that this review concerns itself.

The hydrodynamic properties of a rigid body depend upon both the shape and the volume of the body. For protein molecules, the practice has been to represent the shape by an ellipsoid of revolution and to relate shape to macroscopic properties by the theories of Perrin (frictional ratio) and Simha (viscosity increment). Classically, the "volume of the protein molecule in solution" was assumed to consist of two parts: the volume of the anhydrous molecule and the volume of "bound" water.

Numerically this was computed from $\bar{v}$, the partial specific volume of the protein, M, the molecular weight, w, the mass of water bound per g.

of protein, ρ, the density of water and N_o, Avogadro's number. The volume,

$$\frac{M\bar{v}}{N_o} + \frac{Mw}{N_o\rho}$$

was taken to be both the domain of the protein molecule and the hydrodynamically effective volume. Scheraga and Mandelkern (1) pointed out that it is not reasonable to obtain the effective volume as an excess over $M\bar{v}/N_o$ since $\bar{v}$ must reflect interactions between solute and solvent, making it impossible to identify $M\bar{v}/N_o$ with the anhydrous protein. In fact, $\bar{v}$ may be positive, negative, or zero (as for example, $MgSO_4$ in H_2O). In addition, the identification of this volume with the hydrodynamically effective volume ignores the possible flow of solvent through the domain of the protein molecule, electrostriction, and the deformation of the domain by hydrodynamic forces.

The procedure proposed (1) is simply to find the dimensions of the rigid ellipsoid of revolution, which reproduces all of the hydrodynamic properties of the protein in solution. That such an ellipsoid exists may not be intuitively obvious but is shown by the treatment that follows. The relation between this ellipsoid and the actual configuration in solution cannot be determined by present methods.

In the limit of low concentration the intrinsic viscosity, $[\eta]$, can be written as

$$[\eta] = N_o V_e v/100M \qquad (1)$$

where V_e is the hydrodynamically effective volume, M is the anhydrous molecular weight, and v is a shape factor dependent on only the axial ratio $p = b/a$ of the ellipsoid. The semiaxis of revolution is a, while the remaining two axes are equal to b. Numerical values of the function $v = f(p)$ have been tabulated (2). The frictional ratio $f/f_o = 1/F$ is likewise a function of p only and may be combined with Stokes law to give

$$f = (162\pi^2)^{\frac{1}{3}} (V_e)^{\frac{1}{3}} (\eta/F) \qquad (2)$$

where η is the solvent viscosity. From (1) and (2) it is evident that both $[\eta]$ and f depend on both p and V_e so that a simultaneous solution of these two equations should give both p and V_e. Combining

* The author acknowledges support from Grant No. H-2217 of the National Heart Institute.

equations (1) and (2) with the Svedberg equation to eliminate f gives a new function β depending on p only and defined by:

$$\beta \equiv \frac{N_0 S [\eta]^{\frac{1}{3}} \eta}{M^{\frac{2}{3}} (1-\bar{v}\rho)} = \gamma F v^{\frac{1}{3}} \qquad (3)$$

where $\gamma = N_0^{\frac{1}{3}}/(16200\pi^2)^{\frac{1}{3}}$

A similar relation for β is obtained by combining equations 1 and 2 with the relation $D = kt/f$ resulting in

$$\beta \equiv \frac{D [\eta]^{\frac{1}{3}} M^{\frac{1}{3}} \eta}{kt} = \gamma F v^{\frac{1}{3}} \qquad (4)$$

It is thus evident that the measurement of $[\eta]$, M, and s, or $[\eta]$, M, and D, as well as η, $\bar{v}$, and ρ serves to determine both p, and V_e. In the past the role of V_e has been somewhat overlooked and for example high values of $[\eta]$ and f have usually been associated with high asymmetry, whereas from equations 1 or 2 it may be seen that these could result from changes in V_e just as well.

Since there is no ambiguity in solving (3) or (4) for p and since the functions involved (Perrin's equations and Simha's equation) are all continuous, we can always find the dimensions of the hydrodynamically equivalent ellipsoid so long as the value of p is 2.12 or greater. This minimum value is found (1) for a sphere. It should be pointed out that β does not vary rapidly with p so that very accurate data are needed to fix p with certainty. The dependence of β on p is especially small for oblate ellipsoids ranging from only 2.12 to 2.15 for a change in p from 1 to 300. A value of p greater than 2.15 rules out an oblate ellipsoid as the equivalent model.

Scheraga and Mandelkern give an analagous treatment for rotary diffusion. In this case the measurement of M, $[\eta]$, and θ, the rotary diffusion constant is sufficient to determine p and V_e.

An interesting application of these equations was made with viscosity and diffusion data on horse serum albumin in urea as shown in table 1 (1).

The increase in $[\eta]$ and the decrease in D at high urea concentrations were originally ascribed to an unfolding process resulting in a very asymmetric particle. The Scheraga and Mandelkern treatment indicates instead that the asymmetry changes little or none but that the molecule swells and V_e increases.

SEDIMENTATION VELOCITY

The development and general availability of the Spinco ultracentrifuge has greatly increased the interest and use of both sedimentation velocity and sedimentation equilibrium. Several improvements in the use of ultracentrifuge data have stemmed from the work of Archibald (4) who carried out the intergration of the differential equation of the ultracentrifuge

$$\frac{\partial c}{\partial t} = \frac{1}{x} \left(\frac{\partial}{\partial x} \left\{ \left[D \frac{\partial c}{\partial x} - \omega^2 sxc \right] x \right\} \right) \qquad (5)$$

and indicated how numerical computations could be made to yield the solute concentration as a function of position, time, centrifuge speed, and the sedimentation and diffusion constants of the solute molecule, so that if the solute transport can be satisfactorily measured, then the values of s and D can be found. The interpretation of solute transport is very simple if the time of centrifugation is not too long and the sedimentation constant not too large.

Under these circumstances there will be a region in the centrifuge cell where the solute concentration, c, depends only on the time, t, and not upon x, the distance from the center of rotation. This is equivalent to saying that the diffusion process occurring in the region of the boundary or in the region of the meniscus (if the boundary does not separate) has not yet disturbed the concentration in some region of cell well away from the boundary. One can then neglect diffusion in computing the rate of solute transport in such a region and in so doing greatly simplify the mathematical problem.

This approach was first applied by Gutfreund and Ogston (5) in their studies on the oxidation products of insulin. It was recognized by these

Table 1. CALCULATION OF DIMENSIONS OF EFFECTIVE ELLIPSOIDS IN UREA SOLUTIONS OF HORSE SERUM ALBUMIN USING DATA OF NEURATH AND SAUM (3) AT 25°

Urea concn. M	$[\eta]$	$D \times 10^7$	$\beta \times 10^{-6}$	r(Å)
0	0.049	6.85	2.23	a=82; b=16.4
0.5	0.050	6.20	2.04	38
1.5	0.056	6.08	2.07	39
3.0	0.065	5.69	2.04	41
4.5	0.123	4.45	1.98	51
6.0	0.147	4.27	2.01	54
6.66	0.170	4.15	2.05	56

workers that measurements of this sort would serve to give the sedimentation constant of rapidly diffusing materials even when it is not possible to separate the boundary from the meniscus. Baldwin (6) made some additional refinements of the treatment which avoids carrying out a double integration of the experimental data. The method is simply to apply the conservation of mass to two regions above and below a fixed value of x—, i.e., to compute the concentration in a region on the basis of the mass of material being transported into or out of the region by the process of sedimentation. Baldwin's (6) form of the result is:

$$c_o x_o{}^2 = x_o{}^2 \left(c_o e^{-2\omega^2 s t} - \int_{x_o}^{X} \frac{dc}{dx} dx \right.$$

$$\left. + \int_{x_o}^{X} x^2 \frac{dc}{dx} dx \right) \qquad (6)$$

where c_o is the initial concentration of solute at time, $t = o$, x_o the position of the meniscus, and X the fixed position lying in a region where c, the concentration, varies only with time (sector shape and inhomogeneity of field). By graphical or numerical integration of the refractive index gradient curves one can determine the sedimentation constant, s, knowing, of course, the angular velocity, ω and the initial concentration, c_o.

Klainer and Kegeles (7) combined Baldwin's calculation with the suggestion of Archibald (8) that the equation of sedimentation equilibrium can be applied to the region at the meniscus and at the bottom of the cell even during the transient state. Archibald showed that the function $(1/cx)\,dc/dx$ is a constant (independent of time at x_o and x_b) equal to $M(1-\bar{v}\rho)\ \omega^2/RT = \omega^2 S/D$. Thus, if the function $(1/cx)dc/dx$ can be extrapolated to x_o or x_b then the molecular weight can be obtained directly with only short times of centrifugation (up to a few hours). The value of dc/dx is given directly from the refractive index gradient curve (by dividing by the refractive index increment) and x_o is readily measured. The problem is to evaluate c from the refractive index gradient data. Klainer and Kegeles (7) expressed the concentration at x_o as the concentration at the arbitrary surface X minus the difference in concentration between X and x_o:

$$c_{x_o} = c_o e^{-2\omega^2 s t} - \int_{x_o}^{X} \frac{dc}{dx} dx \qquad (7)$$

The relation $c_x = c_o e^{-2\omega^2 s t}$ is readily obtained by considering the combined dilution effects due to sector shape and inhomogeneity of field. Now the integral appearing in equation 7 is the same as one in equation 6 so that the same calculations can be used to determine M by the Archibald method and s by the Gutfreund-Ogston-Baldwin method. The calculations for x_b are quite similar to 6 and 7, the forms being:

$$c_o x_b{}^2 = x_b{}^2 \left(c_o e^{-2\omega^2 s t} + \right.$$

$$\left. \int_{X}^{x_b} \frac{dc}{dx} dx \right) - \int_{X}^{x_b} x^2 \frac{dc}{dx} dx \qquad (8)$$

$$c_x = c_o e^{-2\omega^2 s t} + \int_{X}^{x_b} \frac{dc}{dx} dx \qquad (9)$$

Further use of Archibald's method (4) of integrating the ultracentrifuge differential equation has recently been made by Yphantis and Waugh (9, 10). These workers computed the quantity Q, which is the ratio of the mass of solute between the meniscus and an arbitrary point x_p at time t, divided by the mass of solute present in the same volume at $t = 0$. Unlike the case previously discussed, the treatment is not restricted to short times of centrifugation but can be applied to any part of the transient state or to the equilibrium state. For given values of x_o and x_b,

$$Q = f(Z, \tau) \qquad (10)$$

$$\text{where} \quad Z = \frac{\omega^2 s x_p{}^2}{2D} = \frac{\sigma x_p{}^2}{2} \qquad (11)$$

$$\text{and} \quad \tau = 2\omega^2 s t \qquad (12)$$

Several methods of calculation are given as well as useful approximations for computing Q from σ and τ.

A single measurement of Q and t with known cell dimensions, centrifuge speed, x_o and x_p could be satisfied according to equation 10 by an infinite number of pairs of values of σ and τ—i.e., one curve in the s-D plane. If two values of Q are determined and the two curves constructed, then the intersection gives the unique values of s and D.

A convenient nomographic method perhaps would be to list values of Q for two different fixed times in two vertical columns and the corresponding unique values of σ and τ/t in two more vertical columns, placing the values in the latter two columns so that a straight line extended from two values of Q intersect the proper values of σ and τ/t.

Experimental values of Q could be found from refractive index gradient curves by a numerical or graphical integration or even more simply by fringe counting with the Rayleigh fringe method. Yphantis and Waugh (11), however, have devised a special cell for the measurement of Q. This cell, while otherwise similar to the usual sector-shaped centrifuge cell, has in addition a movable separation plate lying in a plane perpendicular to the direction of sedimentation. During centrifugation the plate is held by centrifugal force on the bottom of the cell. As the centrifuge is slowed down, two U-shaped pieces of rubber gently raise the plate to a rest position (x_p in equation 11) so that the cell contents are divided into two parts without mixing. At this position the plate fits snugly into the cell centerpiece because of the sector shape. The top layer can then be drawn off for analysis and the determination of Q. This technique has very great potentialities in studying mixtures containing one or more substances which can be detected by a chemical or bioassay. Since Q is a ratio of concentrations, the concentrations themselves may be expressed in arbitrary units—e.g., so and so many units of activity per ml. One can then determine both s and D for an enzyme, coenzyme, antibiotic, or other biologically active substance susceptible of quantative assay without ever having isolated the pure material. All this, of course, assumes that the molecules of the substance in question act independently and without being bound to any other component in the mixture. In view of the widespread occurrence of protein-protein and protein-ion interaction, this assumption must be examined in each case. It may be noted, however, that any change in s or D during the isolation of a biologically active material may furnish valuable clues concerning methods of purification and the state of molecule in the original tissue. The use of this separation cell was demonstrated on several materials including DPN and serum albumin (11).

It is frequently desired in interpreting sedimentation velocity diagrams to estimate the relative concentrations of two sedimenting components by integration of the refractive index gradient curves. Johnston and Ogston (12) first presented an accurate analysis of an anomaly which had been noted previously in this connection. Qualitatively, there is an apparent increase in the concentration of the slow component and a decrease for that of the fast component. This effect was correctly ascribed to the concentration dependence of s. Assuming one value of s for the slow component above the fast boundary (S_s) and a smaller value below the fast boundary, $S_{s,mix.}$—i.e., in the mixture of slow and fast— simple transport equations indicated that the slow component tends to "pile up" behind the fast boundary, thus raising the refractive increment due to the slow component and lowering that due to the fast.

This anomaly was further investigated by Harrington and Schachman (13) in connection with studies on mixtures of bushy stunt and tobacco mosaic viruses. Their result, also being based on simple transport equations assuming as before one value for S_s and another for $S_{s,mix.}$ can be written in the form:

$$\frac{c_s{}^{obs}}{c_s{}^o} = (e^{-2\omega^2 S_s t})\,(e^{2\omega^2 t[S_s - S_{s,mix.}]})\left(\frac{e^{S_F} - e^{S_{s,mix.}}}{e^{S_F} - e^{S_s}}\right)$$

(13)

The first factor on the righthand side of equation 13 is the classical time dependence of concentration which we have noted previously. The second factor is an additional time dependence increasing $C_s{}^{obs}$ with increasing time so long as $S_s > S_{s,mix.}$, which will ordinarily be the case. The third factor, independent of time, is a number greater than unity if $S_s > S_{s,mix.}$ and its effect will be evident from the very beginning of sedimentation.

This treatment has assumed that $S_{s,mix.}$ is independent of concentration. A better assumption would be that $S_{s,mix.}$ increases as C_F decreases due to dilution (chiefly radial). If, for example, $S_{s,mix.}$ is assumed to be a linear function of t, then the analog of equation 13 indicates that $C_s{}^{obs}$ decreases with time, still of course being too large. This behavior was actually observed with

mixtures of bushy stunt and tobacco mosaic viruses. The area of the first was increased over 200% in the presence of 1% of the second. This area decreased with time since $S_{s,mix}$. depends on C_F very strongly.

After the advent of the Spinco ultracentrifuge it soon became evident that the results from the Spinco were systematically 8 to 10% lower than those from the Svedberg oil turbine machines. The difference was traced to erroneous measurement of rotor temperature in the Svedberg centrifuge. Investigations precipitated by this discrepancy revealed an hitherto unsuspected cooling of the rotor of about 1 degree on acceleration and corresponding heating on deceleration (14). These workers employed radiation exchange to measure the rotor temperature. Confirmation of these results were obtained (15, 16) by direct melting point studies on diphenyl ether in the centrifuge cell.

The use of the sedimentation velocity method has been extended to very low molecular weights (∞200) by the device of boundary-forming cells (17, 18). These cells result in the formation of a boundary by layering of solvent over that of solution during the acceleration period of the centrifuge. In this way it is possible to measure s directly by the boundary velocity for rapidly diffusing substances whose boundaries would never separate from the meniscus in a conventional cell.

In addition, one can also measure in a boundary-forming cell the sedimentation constant of a slow-moving component in the presence of a fast component since the initial composition of the two layers is arbitrary. The resulting boundaries will be stable, of course, only if the lower liquid is the more dense.

LIGHT SCATTERING

A major advance in the theory of light scattering was made almost simultaneously by three independent groups of workers (19, 20, 21). In slightly different but equivalent forms, the three groups developed a general theory for the scattering from multicomponent systems of small particles. The general result for the excess scattering can be written in the form used by Edsall et al. (22)

$$2R_{90,u} = K'' \Sigma_i \Sigma_j \psi_i \psi_j A_{ij} / |a_{ij}| \qquad (14)$$

In equation 14 $R_{90,u}$ is the excess scattered intensity observed at 90° with unpolarized incident light and $K'' = 4000 \, \pi^2 n^2 / \lambda^4_0 N_0$ where n is the refractive index, λ_0 the wavelength in vacuo, and N_0, Avogadro's number; ψ stands for the molar refractive increment. The terms a_{ij} in the determinant $|a_{ij}|$ are the derivatives $\partial ln a_i / \partial m_j$ where a_i is the activity and m the molar concentration. The term A_{ij} in the double summation is the cofactor of the term a_{ij}; i.e., the determinant obtained from $|a_{ij}|$ by striking out the row and column in which a_{ij} occurs and multiplying the resulting determinant by $+1$ if $i + j$ is even and by -1 if $i + j$ is odd. The summation extends over all but one component (usually the solvent).

For a two-component system, equation 14 reduces to the well-known form

$$\frac{K''(dn/dc_2)^2 c_2}{2000 R_{90,u}} = \frac{1}{M_2} + \frac{1000 \beta_{22} c_2}{M_2} \qquad (15)$$

where dn/dc_2 is the refractive index increment, M_2 is the molecular weight of solute present at concentration c_2, and $\beta_{22} \equiv \partial ln \gamma_2 / \partial m_2$ where γ is the activity coefficient. For this system, an extrapolation of $c_2/R_{90,u}$ versus c_2 to zero concentration gives the molecular weight while the slope of the curve gives the necessary data for computing the activity coefficient of the solute.

Proteins are rarely studied in two-component systems since usually one or more low molecular weight electrolytes are present in addition. A two-component system, however, can be realized by deionizing a protein solution on ion-exchange resins. Bovine serum albumin has recently been studied in this way (23, 24). The latter workers interpreted the observed activity coefficients at low concentrations in terms of the charge fluctuation theory of Kirkwood and Shumaker (25). This theory predicts a long-range intermolecular attractive force for isoionic proteins in water resulting in a negative excess chemical potential and activity coefficients less than unity. At low concentrations, $c_2/R_{90,u}$ becomes linear in $c_2^{\frac{1}{2}}$ and from the slope in this region the root-mean-square charge fluctuation can be calculated. The value obtained in this way (3.5 protonic units) agreed closely with the value from titration data (3.4).

For three components, equation 14 becomes

$$2R_{90,u} = \frac{K''[\psi_2^2 a_{33} - 2\psi_2 \psi_3 a_{23} + \psi_3^2 a_{22}]}{a_{22}a_{33} - a_{23}^2} \qquad (16)$$

The derivative a_{23} represents the molecular interaction betweeen components 2 and 3 and could be evaluated provided a_{22} and a_{33} can be measured. All three derivatives, however, are functions of both m_2 and m_3 making the problem a very diffi-

cult one to handle. Edsall et al. (22) have carefully examined the system containing bovine serum albumin, water, and sodium chloride. These workers concluded that the contributions of the second and third terms in the numerator of equation 16 were negligibly small so that the behavior of the system reduced to that of two components. For cases where a_{23} is large or where ψ_2 and ψ_3 are comparable, the above approximations will not be satisfactory. For such cases an approximate treatment could perhaps be made by using the attractive but unjustified assumption that a_{22} and a_{33} are dependent upon m_2 and m_3 only, respectively. These derivatives could then be determined in separate experiments (with 2 components only), the behavior of the three component systems then fixing a_{23} which is a measure of "selective binding."

The problem of calibrating light-scattering instruments has until recently involved both theoretical uncertainties and many sources of error.

The advantages of using Ludox silica sols in calibration procedures have been evident for some time. Briefly, the procedure is as follows. An aqueous silica sol is prepared by diluting the commercial 30% sol used in floor waxes. The concentration is chosen so that the turbidity, τ, can be determined by transmission measurements in a spectrophotometer. These sols have been shown by electron microscopy to contain only particles small compared to the wavelength of visible light (26).

For Rayleigh scattering there is a simple relationship between τ and the scattered light:

$$\tau = \frac{16\pi R_{90,u}}{3} \qquad (17)$$

This equation is used to obtain a calibration factor by determining the instrument reading for a known value of $R_{90,u}$. The difficulties with this method arise from the rather high turbidities needed for direct measurement of τ by transmission. Two factors, negligibly small at low τ become very important at high τ. First the attenuation of the transmitted beam is appreciable and second, the value of $R_{90,u}$ is increased by secondary and higher orders of scattering. The first of these factors, the attenuation effect, was studied by Maron and Lou (27). The phenomenon of secondary scattering was investigated both theoretically and experimentally by Kraut and Dandliker (28). The latter workers showed that the

intensity of secondary scattering is proportional to τ^2 and depends also on cell geometry. A method is outlined to permit evaluation of the instrument calibration constant without inaccuracy due to either the attenuation effect or secondary scattering.

The problem of clarifying solutions prior to light-scattering measurements has in the past been rather unsatisfactory.

Most procedures have used either filtration or centrifugation of the solution followed by transfer to the light-scattering cell. These methods, in our laboratory, have never afforded perfectly clean solutions and are erratic and very laborious. An improvement has been made (29) by devising a thin-walled Erlenmeyer-shaped Pyrex cell which when filled with solution to the appropriate level floats upright in a glycerol water mixture. The cell is closed with a rubber cap and centrifuged at high speeds (up to 90,000 times gravity) in a Spinco swinging bucket rotor. After 0.5 to 3 hours, the cell is transferred to the light-scattering instrument and measurements made without opening the cell. In this way, extremely clean solutions are readily and reproducibly obtained. Even very weakly scattering liquids such as water, 0.15 M sodium chloride, or 2% raffinose when clarified in this way show only Rayleigh scattering at the lowest angles obtainable in our instrument (25.8°). When using this method with proteins, the resulting concentration gradients may be destroyed by diffusion, convection, or gentle mixing.

DISCUSSION

DR. HEINZ FRAENKEL-CONRAT: I would like to ask you, Dr. Dandliker, about the scattering curves for serum albumin. In some cases the curves for only partially clean solutions go below the curve for the clean solution. Why is that?

DR. DANDLIKER: Those curves were on a relative basis and normalized to unity at 90° so that whether the actual height is the same from curve to curve is not important. Only the relative scattering at different angles matters.

DR. EDMOND H. FISCHER: What information does one need to interpret the Johnston-Ogston effect in a two-component system?

DR. DANDLIKER: You need to be able to measure the sedimentation constant of the slow component alone and in the presence of the fast. This can be accomplished with the boundary-forming cell by layering a solution of the fast over a mixture of the slow and the fast. Until

the fast overtakes the slow, the sedimentation rate of the slow can be measured.

DR. HOWARD MASON: I would like to ask a rather elementary question. It seems to me that much of the treatment that you have discussed today is based on the assumption that proteins can be purified to simple, single molecules. Now in this pronounced effect of components upon one another it seems to me that a question is raised as to how simple a simple two-component system is. If instead of a single molecular type you actually had a distribution of molecular types within a narrow range, what would be the effect on this?

DR. DANDLIKER: It's difficult to say exactly except that some kind of an average answer would result from the scattering measurement. It must be admitted that until one is sure of the number of components the equations cannot be applied rigorously.

DR. HOWARD MASON: Would you care to express an opinion as to whether proteins might be true single components or optimum distribution systems?

DR. DANDLIKER: Well, I don't know the answer to that either. Again, all you can do is to apply an operational method here. If you think you have a pure component, or if you think you have isolated a pure component, all you can do is apply as many criteria as you can that you know a pure component must obey and see how many it obeys. Then if it obeys all of them, you are on pretty firm ground, but there is no guarantee.

DR. HUGO KRUEGER: In your early discussion, most of the values listed when you were considering serum albumin in the presence of urea were about 2.0, but the first one was 2.2. My question is whether the difference between 2.0 and 2.2 is significant?

DR. DANDLIKER: Scheraga and Mandelkern thought that this discrepancy was due to experimental error, since very exact data are necessary to get reliable values for β.

DR. HUGO KRUEGER: I'd like to ask another question, if I may, something that probably I'm getting myself tied up on, but in the development of your equation for the concentration, one of them was the integral of $X^2 \dfrac{dc}{dx} dx$ so that there is going to be X^2 times the concentration while the first integral was $\dfrac{dc}{dx} dx$.

DR. DANDLIKER: The first integral also had a factor of X^2 outside the bracket to take care of the dimensions so that they are the same in both cases.

DR. HOWARD MASON: With reference to the data on the urea swelling and the interpretation that it represents swelling of a central molecule: Is that consistent with Kauzmann's interpretation of his optical rotation data?

DR. DANDLIKER: Well, to date no theory has appeared to tell us what the optical rotation should be for a large number of asymmetric centers orientated, say on an α-helix.* Now this is something that is being worked on, and it seems that by knowing the complete geometry of a macromolecule, and knowing the characteristics of the optical centers, the contribution of each asymmetric center can be expressed by a single vector. Presumably by the proper summation of all these vectors for all the centers, the rotation of a given molecule can be worked out. I think at present it is pretty hard to say what might be the effect of an isotropic swelling or unfolding on the optic rotation. The current interpretation of the changes in rotation that one finds on denaturation is that the change in configuration of the helix is responsible. Whether or not the same sort of thing might apply to isotropic swelling, I am not sure.

DR. TSOO E. KING: Dr. Dandliker, would you care to comment about the determination of the diffusion coefficient and sedimentation constant of very asymmetric molecules like DNA where the results are very difficult to extrapolate to zero concentration by conventional methods?

DR. DANDLIKER: In the case of molecules like DNA where the concentration dependence is very large, difficulties are encountered in measuring either the sedimentation or diffusion constants. This results in skewed boundaries and a rapid change in the result at all attainable concentrations. Some improvement may be noted by carrying out differential diffusion experiments where the concentration is nearly the same on both sides of the boundary.

A more fruitful approach is to use more sensitive methods of detection, as for example, the ultraviolet absorption for nucleic acids† so that lower concentrations can be studied.

---

* See W. Moffit, J. Chem. Phys., 25, 467-478 (1956).
† Shooter and Butler (Trans. Faraday Soc., 52, 734, May 1956) have been able to use the ultraviolet absorption for determination of sedimentation constants of DNA at concentrations as low as 0.001 per cent. By using Philpot-Svenson optical system, 0.02 per cent is the lower limit of the concentration required. At 0.02 per cent of DNA, the self-sharpening peak is very marked.——Editor.

REFERENCES

1. H. A. Scheraga and L. Mandelkern. Consideration of the Hydrodynamic Properties of Proteins. J. Am. Chem. Soc., 75, 179-184 (1953).
2. J. W. Mehl, J. L. Oncley, and R. Simha. Viscosity and Shape of Protein Molecules. Science, 92, 132-133 (1940).
3. H. Neurath and A. M. Saum. The Denaturation of Serum Albumin. Diffusion and Viscosity Measurements of Serum Albumin in the Presence of Urea. J. Biol. Chem., 128, 347-362 (1939).
4. W. J. Archibald. The Integration of the Differential Equation of the Ultracentrifuge. Ann. N.Y. Acad. Sci., 43, 211-227 (1942).
5. H. Gutfreund and A. G. Ogston. A Method for Determining the Sedimentation Constant of Material of Low Molecular Weight; Studies on Oxidation Products of Insulin. Biochem. J., 44, 163-166 (1949).
6. R. L. Baldwin. Sedimentation Coefficients of Small Molecules: Methods of Measurement Based on the Refractive-Index Gradient Curve. The Sedimentation Coefficient of Polyglucose A. Biochem. J., 55, 644-648 (1953).
7. S. M. Klainer and G. Kegeles. Simultaneous Determination of Molecular Weights and Sedimentation Constants. J. Phys. Chem., 59, 952-955 (1955).
8. W. J. Archibald. A Demonstration of Some New Methods of Determining Molecular Weights from the Data of the Ultracentrifuge. J. Phys. and Coll. Chem., 51, 1204-1214 (1947).
9. D. F. Waugh and D. A. Yphantis. Transient Solute Distributions from the Basic Equation of the Ultracentrifuge. J. Phys. Chem., 57, 312-318 (1953).
10. D. A. Yphantis and D. F. Waugh. Ultracentrifugal Characterization by Direct Measurement of Activity. I. Theoretical. J. Phys. Chem., 60, 623-629 (1956).
11. D. A. Yphantis and D. F. Waugh. Ultracentrifugal Characterization by Direct Measurement of Activity. II. Experimental. J. Phys. Chem., 60, 630-635 (1956).
12. J. P. Johnston and A. G. Ogston. A Boundary Anomaly Found in the Ultracentrifugal Sedimentation of Mixtures. Trans. Faraday Soc., 42, 789-799 (1946).
13. W. F. Harrington and H. K. Schachman. Analysis of a Concentration Anomaly in the Ultracentrifugation of Mixtures. J. Am. Chem. Soc., 75, 3533-3539 (1953).
14. D. F. Waugh and D. A. Yphantis. Rotor Temperature Measurement and Control in the Ultracentrifuge. Rev. Sci. Inst., 23, 609-614 (1952).
15. S. Shulman. The Determination of Sedimentation Constants With the Oil-Turbine and Spinco Ultracentrifuges. Arch. Biochem. Biophys., 44, 230-240 (1953).
16. A. Biancheria and G. Kegeles. Thermodynamic Measurements of Ultracentrifuge Rotor Temperature. J. Am. Chem. Soc., 76, 3737-3741 (1954).
17. G. Kegeles. A Boundary Forming Technique for the Ultracentrifuge. J. Am. Chem. Soc., 74, 5532-5534 (1952).
18. E. G. Pickels, W. F. Harrington, and H. K. Schachman. An Ultracentrifuge Cell for Producing Boundaries Synthetically by a Layering Technique. Proc. Nat. Acad. Sci., 38, 943-948 (1952).
19. H. C. Brinkman and J. J. Hermans. The Effect of Non-Homogeneity of Molecular Weight on the Scattering of Light by High Polymer Solutions. J. Chem. Phys., 17, 574-576 (1949).
20. J. G. Kirkwood and R. J. Goldberg. Light Scattering Arising from Composition Fluctuations in Multi-Component Systems. J. Chem. Phys., 18, 54-57 (1950).
21. W. H. Stockmayer, Light Scattering in Multicomponent Systems. J. Chem. Phys., 18, 58-61 (1950).
22. J. T. Edsall, H. Edelhoch, R. Lontie, and P. R. Morrison. Light Scattering in Solutions of Serum Albumin: Effects of Charge and Ionic Strength. J. Am. Chem. Soc., 72, 4641-4656 (1950).
23. W. B. Dandliker, Light Scattering of Isoionic Bovine Albumin. J. Am. Chem. Soc., 76, 6036-6039 (1954).
24. S. N. Timasheff, H. M. Dintzis, J. G. Kirkwood, and B. D. Coleman. Studies of Molecular Interaction in Isoionic Protein Solutions by Light Scattering. Proc. Nat. Acad. Sci., 41, 710-714 (1955).
25. J. G. Kirkwood and J. B. Shumaker. Forces Between Protein Molecules in Solution Arising from Fluctuations in Proton Charge and Configuration. Proc. Nat. Acad. Sci., 38, 863-871 (1952).
26. M. Sveda. New Silica for Floor Wax, Soap and Sanitary Chemicals. July-August (1949).
27. S. H. Maron and R. L. H. Lou. Calibration of Light Scattering Photometers with Ludox. J. Polymer Sci., 14, 29-36 (1954).
28. J. Kraut and W. B. Dandliker. A Study of Secondary Scattering. J. Polymer Sci., 18, 563-572 (1955).
29. W. B. Dandliker and J. Kraut. A Centrifugal Light-Scattering Cell for the Ultraclarification of Liquids. J. Am. Chem. Soc. 78, 2380-2384 (1956).

Partition Column Chromatography of Proteins

Frederick H. Carpenter

When asked to present a discussion before this colloquium, I submitted several suggestions to your chairman, Professor King. Being a man of action, Professor King drew a collective title out of the suggestions and spared me at least that part of the birthpains associated with the selection of a title for the discussion. Unfortunately the title is a bit too broad in that a good deal of what I will talk about today will be limited in application to those proteins which are relatively stable in acidic solutions and organic solvents.

As a prologue to the experimental results which we have obtained in the study of the partition column chromatography of insulin, I would like to give a short discussion on a theory of partition column chromatography. The theory is one that we have adapted for applications to analysis of elution curves obtained in connection with our experimental work. At this point I would like to give recognition to Dr. George P. Hess, who initiated this work with me several years ago; to Dr. Jean Close, a visitor from the Belgian Congo, who added considerable stimulus to the problem recently; and finally to Miss Juanita L'Esperance whose able assistance has contributed greatly to the experimental work. I would also like to acknowledge the receipt of financial aid from both Eli Lilly and Co. and the United States Public Health Service, (grant a-608 (C2)).

THEORY

Although partition column chromatography may have been recognized and applied before the classical work of Martin and Synge reported in 1941 (1), most of the developments in the field stem from this paper which not only described a workable theory of chromatography but also laid down the experimental ground rules for its application. The big advance in the theory of chromatography made by Martin and Synge was the introduction of the "theoretical plate" concept which had previously met with so much success in fractional distillation. The "theoretical plate" treatment was adapted by Mayer and Tompkins (2) with considerable success to the explanation of the results of ion-exchange columns in 1947. Since this time a number of criticisms of the theoretical plate treatment have been promulgated and several more elaborate theories have been advanced. See Glueckauf (3). Despite its shortcomings, several of which will be enumerated presently, it is my opinion that the "theoretical plate" treatment of the data is still the simplest and most useful to the person interested in purification problems.

Martin and Synge developed their theory of partition column chromatography for substances which could be visualized on the column. Measurements were made upon the actual rate of movement of a visible band on the column. Until recently there were no adequate descriptions of equations to be used when solutes were completely eluted from the column. It is on this particular phase of the theory that I wish to dwell for a moment.

In a development of a theory of elution analysis for partition column chromatography, I prefer to approach the problem from the viewpoint of countercurrent distribution (CCD). In the CCD technique, introduced by Craig (4), one has a series of tubes containing two phases. A solute is introduced into the first tube and the mixture is shaken until the solute has distributed itself between the two phases according to its nature. When the equilibrium point has been reached the phases are allowed to separate and one phase, containing part of the solute, is transferred to the next tube. Fresh moving phase is added to the first tube, both tubes are mixed and the process repeated through a serial number of tubes. Now the point to make here is that each tube in a CCD train actually represents a "theoretical plate." The latter is defined as that section of a column wherein the solute is in equilibrium with the two phases and where the average concentration of the solute in the moving phase of that section is determined by the distribution constant of the solute. The definition and theoretical significance of the distribution constant will be taken up in a moment, but at the time I want to emphasize that for practical considerations a partition column may be considered to work in a fashion strictly analogous to a CCD train. If one approaches the theory from this point of view it is much easier to visualize the operation of a column and furthermore one can apply many of the excellent mathematical treat-

ments that have been worked out for CCD (4, 5) to the partition column.

If we consider the partition column to work like a CCD train, then when one elutes a solute completely from the column, this is similar to performing a CCD experiment by the "withdrawal" technique (4, 5). In this technique the moving phase is collected one layer at a time as it is transferred out of the last tube in the train. One continues to pass moving phase through the train until all of the solute has been carried out. Using this as an analogy one can consider a partition column to be made up of a number of tubes each with a fixed volume (v_s) of stationary phase and a volume (v_m) of moving phase (see figure 2A). The total number of tubes or "theoretical plates" (**N**) comprising the column would then be determined by the total volume of stationary phase on the column (V_S) divided by the volume of stationary phase in one theoretical plate (v_s) or by the total volume of moving phase in the column (the hold-up volume, V_H) divided by the volume of moving phase in a theoretical plate (v_m).

$$\mathbf{N} = V_S/v_s = V_H/v_m \qquad (1)$$

When a solute is introduced into the first "theoretical plate or tube" it will distribute itself between the two phases according to its nature and the nature of the solvents. Thermodynamics tells us that at equilibrium the fugacity (f) or escaping tendency of the solute in the two layers (s for stationary and m for moving) is equal.

$$f_s = f_m \qquad (2)$$

Since the activity of the solute (a) can be expressed as a ratio of its fugacity to a standard state fugacity (f^o) one can write the following expressions:

$$a_m = f_m/f_m^o \; ; a_s = f_s/f_s^o \qquad (3)$$
$$f_m = a_m f_m^o = f_s = a_s f_s^o \qquad (4)$$
$$a_m/a_s = f_s^o/f_m^o = K^o \qquad (5)$$

Furthermore, activities of the solutes can be expressed in terms of their concentrations (c) and their activity coefficients (γ).

$$\frac{\gamma_m c_m}{\gamma_s c_s} = K^o \qquad (6)$$

Now the expression c_m/c_s, or the concentration of the solute in the upper or moving phase divided by the concentration in the lower or stationary phase is called the *distribution constant* (K).

$$K \gamma_m/\gamma_s = K^o \qquad (7)$$

One can see from this equation that the distribution constant is not a true constant as its value depends upon the ratio of the activity coefficients. In dilute solutions the activity coefficients approach unity and the measured distribution constant approaches the true thermodynamic constant. The same may be true at even quite high concentrations, however, if the activity coefficients have approximately the same value in both layers so that their ratio approaches unity. I have gone into this rather long derivation of the meaning of the distribution constant because its true significance is generally ignored in most theoretical treatments and also because most treatments, as well as ours, assume that the distribution constant remains fairly constant under the conditions of the experiment. The form of this equation indicates that this is a reasonable assumption for dilute solutions.

Now if we return to the situation where a small amount of solute has been introduced into the first "theoretical plate" we will find, at equilibrium, the solute distributed between the two phases according to its distribution constant. When a moving layer is transferred to the next theoretical plate, however, the fraction of the solute present in the first plate that is transferred to the next plate will be determined not only by the distribution constant of the solute but also by the relative volumes of the two layers. In order to take the volume effect into account we have defined a new term which we call the *partition factor* (K').

$$K' = \frac{v_m c_m}{v_s c_s} = \frac{v_m}{v_s} K = \frac{V_H}{V_S} K \qquad (8)$$

K' represents the ratio of the total amount of solute in the moving phase to the total amount in the stationary phase. The rate at which a solute moves through a column and appears in the eluted fractions is actually determined by the *partition factor* which takes into account not only the distribution constant but also the relative volume of the two phases in a theoretical tube. These in turn are related to the total volume of moving phase on the column (V_H) and the total volume of stationary phase (V_S) on the column.

There is not time to give a complete development of the equations we have derived to aid in interpreting the action of a partition column, but I can indicate how they were obtained. If one makes the following assumptions (1) that the solute is introduced on the column in a volume equal to the volume of the moving phase in a

theoretical plate (v_m), (2) that the fractions are collected in volumes equal to the theoretical v_m, and (3) that the solute moves through the column according to its partition factor (K'), then one can show that a plot of the amount of material
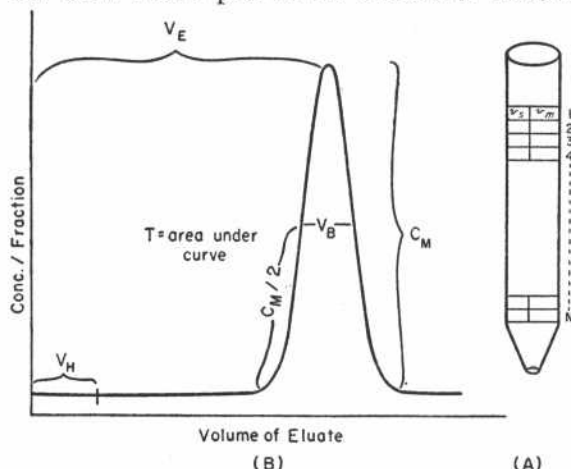

Figure 2. (A) Schematic representation of a partition column as a CCD train where v_s and v_m equal the volume of the stationary and moving phase of each theoretical plate and **N** equals the number of theoretical plates on the column.

(B) Schematic representation of the terms used in describing the behavior of the column.

in the eluted fractions will give a nearly symmetrical curve (figure 2B) which has a mean defined by $\mathbf{N}/K'$ and a standard deviation defined by $[\mathbf{N}(K'+1)/(K')^2]^{\frac{1}{2}}$. With this information one can relate the partition factor (K'), the distribution constant (K), and the number of theoretical plates on the column (**N**) to measurements that are ordinarily made in the course of performing an experiment. These measurements

Figure 3. Schematic representation of terms used in describing behavior of solute on the columns.

are illustrated in figure 3. These measurements are the volume of the stationary phase used in preparing the column (V_S), the hold-up volume of the column (V_H) which can be calculated as indicated, and the total volume of eluate collected from the time the solute is introduced on the column up to and including the eluted fraction of maximum concentration. This we will call the *elution volume* (V_E).

Making use of these measurements and of the value for the mean in the Pascal distribution curve, one arrives at the following relationships:

$$K' = V_H/(V_E - V_H) \qquad (9)$$
$$K = V_S/(V_E - V_H) \qquad (10)$$

One thing that is interesting to note here is that the elution volume for the maximum concentration is independent of the number of theoretical plates describing the column.

NUMBER OF THEORETICAL PLATES

The number of theoretical plates can be estimated by several means (see figure 2B), most of them based on the use of the standard deviation. One can derive relationships which use the concentration of the material in the maximum tube, or the total width of the elution curve. For reasons that will become apparent later, however, we prefer to estimate **N** from measurements made on the width of the band between the two points of one-half the maximum concentration (V_B). Using these measurements one derives the following expression for the number of "theoretical plates" on the column.

$$\mathbf{N} = 5.55 \ (V_E^2 - V_E V_H)/V_B^2 \qquad (11)$$

DEFICIENCIES

It was assumed in the derivation of the equations that the solute was put on the column and collected from the column in volumes equal to the theoretical v_m. This of course would rarely be realized in practice. The assumption, however, does not introduce particularly large errors in the results. Figure 4 illustrates this point. The solid line represents a theoretical elution curve for a substance with a K' value of 0.1 on a column with 400 theoretical tubes. The points represent the results that would have been obtained if the eluate was collected in volumes of 50 (●), 100 (△), and 200 (0)$v_m s$. The greatest deviation from the theoretical curve occurs at the peak and at the base of the band, and even here the deviations are not large. It is for this reason, however, that a formula using measurements on
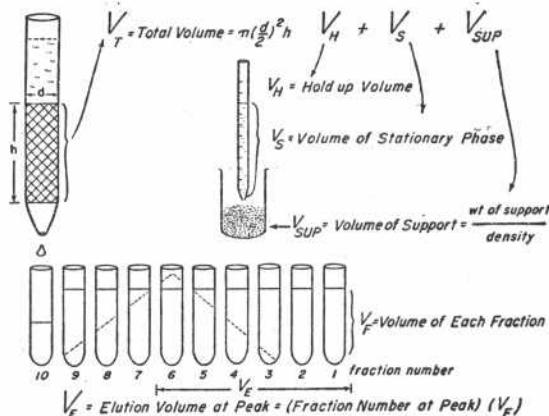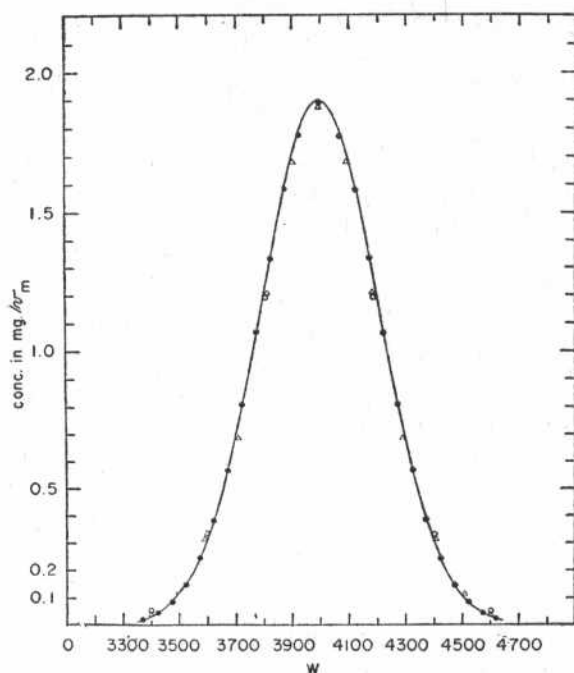
Figure 4. Plot of concentration of material in mg./v_m against number of v_m eluted after hold-up volume has come off (W) for a solute with $K = 0.1$ on a column with 400 theoretical plates. The solid curve is the result obtained when the eluted fractions are collected in volumes of size v_m. The points are the results obtained when the eluted fractions are collected in volumes which correspond to 50 (·), 100 (Δ), and 200 (0) $v_m s$.

the width of the band at one-half maximum concentration was selected for determining **N** rather than formulas making use of the height of the peak or the total width of the band at the base. It should also be noted that measurements made on the latter may also be complicated by impurities in the sample which would have the highest probability of being located at the edges of the band. Other deficiencies in the theory derive from the fact that the column is actually a continuous process rather than the stepwise process implied in the theory. This may lead to complications which have been pointed out by Glueckauf and other authors (3) but which we will choose to ignore at the present time.

SEPARATION EQUATION AND OPTIMUM CONDITIONS

In the use of partition columns in homogeneity studies, it is of some importance to know what relationship exists between the partition factors of closely related substances and the degree to which one can expect to separate them on any

particular column. Since the partition factor can be changed not only by changing the solvents (thus changing the K value) but also by changing the relative volumes of the two layers, it is highly desirable to have a theoretical approach designed at selecting the optimum partition factor.

One can arrive at such an expression quite simply as shown in figure 5. This shows the
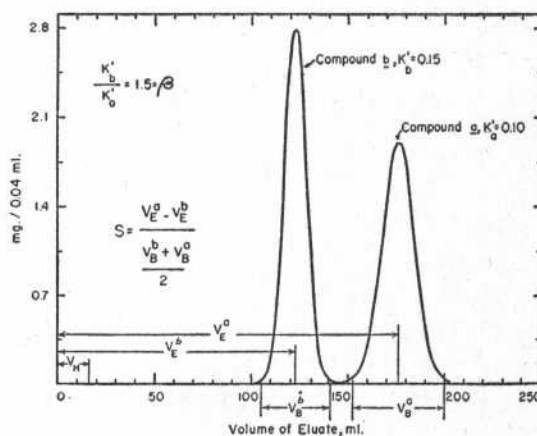
Figure 5. Theoretical diagram of the separation of equal amounts of compounds a and b with K' values of 0.1 and 0.15 respectively on a column with $V_H = 16$ and with 400 theoretical plates.

theoretical elution curve of two compounds with K' values of 0.1 and 0.15 or a ratio between the two (β value) of 1.5. These two components can be considered as being completely separated when the distance between the 2 peaks is equal to or greater than $\frac{1}{2}$ of the sum of the distance covered by their elution bands. Making use of the fact that better than 99.7% of the material falls within the limits of the mean ± 3 standard deviations one can arrive at the following expression for S:

$$S = \mathbf{N}^{\frac{1}{2}}(\beta-1)/3[\beta(K_a'+1)^{\frac{1}{2}} + (\beta K_a'+1)^{\frac{1}{2}}]$$
(12)

When S, the ratio of the distance between the two peaks to one-half of the sum of the distance covered by the elution bands, is equal to or greater than 1, the components will be completely separated. Figure 6 shows a plot of S as a function of K'_a for several β values in a column with 400 theoretical plates. It is obvious from this plot that as K_a' decreases in value, S increases. This rate of increase, however, falls off rapidly below 0.1 and approaches a limit value. This limit value becomes apparent when one realizes that the terms $(K_a'+1)^{\frac{1}{2}}$ and $(\beta K_a'+1)^{\frac{1}{2}}$ in the separation

Figure 6. Plot of the degree of separation (S) of two components against the K' value (on a logarithm scale) of the component with the lowest partition factor for several β values on a column with 400 theoretical plates.

equation both approach unity as K_a' approaches zero, giving the limit expression

$$\text{Limit}\atop K_a \to 0 \; S = \mathbf{N}^{\frac{1}{2}}(\beta - 1)/3(\beta + 1)$$

The pertinent lesson to be learned from these equations is that one should work with small K' values. The optimum K' value will change with **N** but for most columns values in the order of 0.1 would appear to approach optimum conditions.

GENERAL CONSIDERATIONS ON PARTITION CHROMATOGRAPHY

The remarks made in this rather prolonged, but even so, perhaps too sketchy, survey of the theory of partition column chromatography, apply to the chromatography of all sorts of compounds, not merely proteins. The factors that are unique in the chromatography of proteins are limitations imposed on the selection of solvent systems. In order to consider the use of partition column chromatography one must have two immiscible phases in which the protein will dissolve. This in itself is a fairly limiting requirement. For experience has taught us that there is very little possibility of dissolving proteins in any solvent that does not have a fairly high mole fraction of water. So one is limited to these fairly aqueous solvent systems. Unfortunately there are just not many solvent pairs that can be considered. In addition there is another serious limitation imposed by the proteins. Many are stable in a very limited pH

range, generally a few pH units on each side of neutrality, and/or are extremely sensitive to organic solvents. Use of extremes of pH or organic solvents may bring about a nonreversible change in the molecule which is often reflected by changes in its biological activity or its solubility.

It is obvious thus that use of partition column chromatography on proteins is going to be of much more limited applicability than a separation technique such as electrophoresis. If the protein is limited in stability to regions around neutrality, the only applicable solvent systems appear to be those introduced by Martin and Porter in 1951 (6). These consist of various ethers of ethylene-glycol which are equilibrated with water containing quite high concentrations (ca. 2 M) of neutral salts such as ammonium sulfate, potassium phosphate, and magnesium sulfate. The resulting phases have a high water content and will dissolve a number of proteins. Owing to the high salt content, however, there are probably an equal number of proteins that are insoluble in one or both layers. The glycol ethers do not appear to be particularly bad denaturing agents, and what effect they do exhibit can be controlled to a certain extent by working at low temperatures.

For those proteins which can stand extremes of pH, a variety of other solvent systems may be investigated. However, only a limited number of proteins are able to withstand such conditions. A number of the protein hormones and a few of the more stable enzymes may be expected to withstand the conditions. Also the large polypeptides being obtained in the course of structural studies on the various proteins may withstand extremes of pH and organic solvents.

The rest of the discussion I wish to restrict to proteins that will withstand quite acidic conditions. It had been noted empirically that proteins which would not go into an organic solvent from neutral solutions would do so if one incorporated a fairly strong acid into the system. Dichloroacetic acid, trichloroacetic acid, and toluenesulfonic acid have been found to be quite efficacious in this respect. Presumably these acids bring about this effect (1) by decreasing the dissociation of the carboxyl group, thus eliminating the zwitterion properties of the protein, and (2) by existing in quite high concentrations in the organic layer and thereby drawing the positively charged protein into the organic layer as follows:

At the risk of getting the cart before the horse, I am going to describe some recent experiments we have performed in an attempt to take some of the empiricism out of the selection of solvent systems. In this work we have studied the relationship that exists between the distribution constant of various substances between 2-butanol and water containing various mineral acids with the way in which the mineral acids themselves distribute between these two phases. In the butanol-water system each phase has a high mole fraction of water (the organic layer 0.67 and the aqueous layer 0.94 mole fraction of water) and this remains fairly constant on the distribution of dilute mineral acids. As a first approximation, we have assumed that the mineral acids are completely ionized in each layer. The distribution constant of approximately 0.1 N solutions of various mineral acids was determined in the 2-butanol-water solvent system. Then the distribution constant of *trace* amounts of various proteins and peptides was determined in the various 2-butanol-acid systems. The result of such a set of determinations with insulin is shown in figure 7. The distribution constant for the acids, designated K_{HX}, is shown on the abscissa while the distribution constant for insulin in each one of the acids (K_{Ins}) is shown on the ordinate. Note that the curve is asymptotic, the distribution constant of the insulin changing as a power of the distribution constant of the mineral acid.

The results of similar studies on several model compounds are shown in figure 8. It is of interest to note here that ammonia and alanine, both of which have one positive charge per molecule,
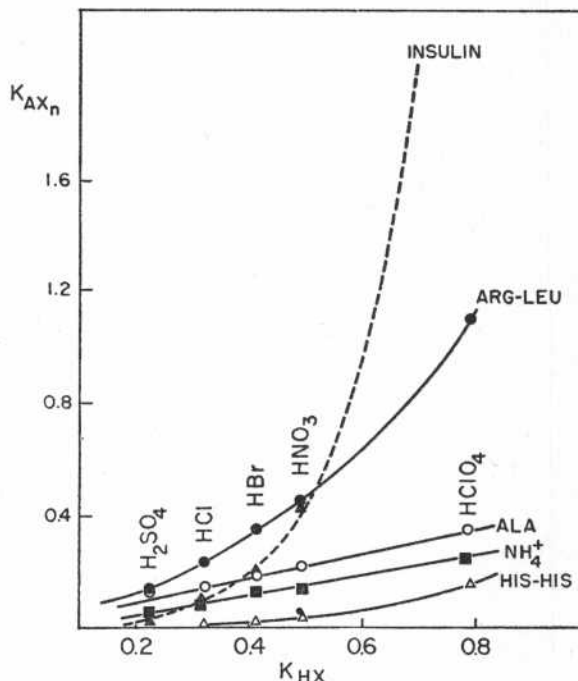
Figure 8. Distribution constant of various substances (AX_n) as a function of the distribution constant of various mineral acids.

gave a straight line relation between their distribution constant (K_{ax_n}) and that of the acids (K_{HX}). Arginyl-leucine however, with two positive charges and histidyl-histidine with three positive charges gave asymptotic curves. Before analyzing the mathematical significance of these curves, I would like to point out a few things about them that were of aid to us and also of practical significance to partition chromatography of positively charged compounds. In connection with some studies on insulin we desired to separate insulin from alanine and ammonia in a reaction mixture. At this time we were using a system containing hydrochloric acid. Inspection of figure 8 shows that the distribution constants of the three compounds lie very close together in the hydrochloric acid system and as a consequence they were difficult to separate. Had we

Figure 7. Distribution constant of insulin in various mineral acids as a function of the distribution constant of the acids.

used the nitric acid system, the resulting K values would have permitted a facile separation.

One point about these curves which is somewhat discouraging from the point of partition column chromatography is the fact that they come very close together at low K values for the acids. As pointed out earlier a partition column works most efficiently when the substance has a small K value. The curves in figure 7, however, show that the K values for all these substances begin to converge at small K values. As a consequence the difference between K values of different components is likely to become smaller and therefore the ease of separation more difficult.

Mathematical analysis of the curves shown in this figure indicated that they could be expressed by the following formulas:

$$K_{ala} = 0.45 K_{HX}$$

$$K_{arg-leu} = 1.66 (K_{HX})^{1.7}$$

$$K_{his-his} = 0.316 (K_{HX})^{3.3}$$

$$K_{ins} = 9.0 (K_{HX})^{4.2}$$

These results indicate that the power terms present in these formulas are related to the number of charges on the molecule: Alanine having 1 charge, arginyl-leucine 2 charges, histidyl-histidine 3 charges, and insulin 6 charges. As a loose approximation the distribution constant of the salt AX_n is related to the distribution constant of the acid HX by the following formula

$$K_{AX_n} \infty k (K_{HX})^n$$

where the distribution constant of the salt is approximated by a constant, k, determined by the chemical nature of the compound, times the distribution constant for the acid raised to the n power, where n represents the number of charged groups on the compound. The experimental values did not agree exactly with this equation. The lack of agreement is explainable on theoretical grounds as there are several variables which the equation does not take into consideration. Nevertheless the relationship appears approximately true and probably of considerable value in separation studies.

Now that I have put the cart before the horse by giving some results which we have obtained in recent months, I am going back to work which was started several years ago and in reality led up to the results just presented.

PARTITION CHROMATOGRAPHY OF INSULIN

Our interest in the use of partition column chromatography for proteins derived out of some earlier studies performed in collaboration with

Professor C. H. Li on the purification of the adrenocorticotropic hormone. In studies on the purification of this hormone we noted that information obtained from CCD experiments was extremely useful in the design of conditions for partition column chromatography of this material (7). As I have pointed out previously, the two techniques can be approached from the same theoretical considerations. In actual use, however, there are some practical differences between the two techniques that are worth mentioning. Partition column chromatography experiments often can be conducted with considerable savings in time and equipment costs over comparable experiments by CCD. Another difference is found in the nature of the solvent pairs which may be used in the two techniques. In CCD one frequently encounters solvent pairs which would be ideally suited for the technique on the basis of the distribution constant of the solute between the two pairs but which have to be abandoned owing to the formation of emulsions that are slow to separate. The presence or absence of stable emulsions is of little importance in the selection of solvents for the partition column technique.

We had become interested in separation methods for insulin in connection with a basic problem in the laboratory designed to elucidate some of the structural characteristics of the insulin molecule that are needed for biological activity. In order to obtain completely unambiguous results we needed a method by which we could separate insulin from closely allied derivatives and degradation products. Harfenist and Craig had made an exhaustive study of insulin by CCD (8) and had found this to be an admirable method of separating closely related insulin deriva-
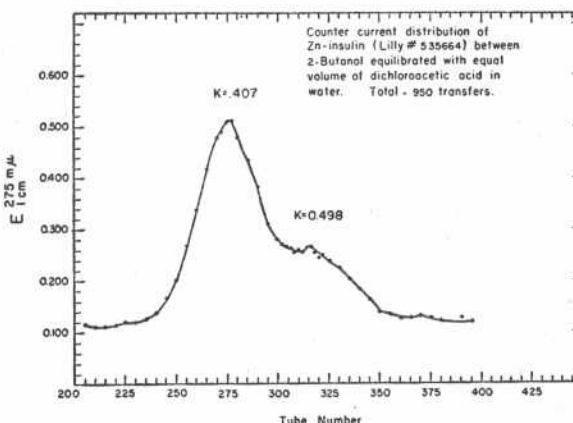
Figure 9. Countercurrent distribution experiment on insulin.

tives. Harfenist and Craig observed that most crystalline insulin samples studied by them were composed of two biologically active components, called insulin A and B. Figure 9 shows the results we have obtained on a crystalline sample after 950 transfers by CCD according to the method of Harfenist and Craig. The solvent system was 2-butanol-1% dichloroacetic acid. The result is very similar to that reported by Harfenist and Craig (8). This experiment was performed in a hand-operated instrument, several weeks being required to complete the experiment and work up the products. Because of the time involved we thought it worthwhile to determine whether similar results could be obtained by partition column chromatography (11).

Shortly after we had become well immersed in the problem, Porter reported on the usefulness of partition columns in the separation of a number of proteins including insulin (9). In his work Porter had used solvent systems made by mixing various cellosolves with various neutral, concentrated salt solutions. Using these systems Porter did not detect the A and B fractions of insulin which were observed by Harfenist and Craig in most of the samples they had studied. More recently Andersen (10) has reported on the partition column chromatography of insulin between 2-butanol and 0.01 N trichloroacetic acid, conditions more comparable with Harfenist and Craig who used 2-butanol-1% dichloroacetic acid (DCA), but also failed to detect the heterogeneity found in the CCD work.

EXPERIMENTAL CONDITIONS

Since the CCD work had indicated heterogeneity in the original crystalline insulin we decided to use solvent systems as nearly comparable as possible with those used in the CCD work. Figure 10 shows a diagram of the column evolved for this work. The "inert" support for the stationary phase was a form of diatomaceous earth, Hyflo Super-cel, which had been washed with acid, water, and 2-butanol and then air-dried before use. The solvent systems were 2-butanol equilibrated with equal volumes of various acids and are designated according to their composition before equilibration. The equilibrations as well as all subsequent operations were performed at constant temperature (25° C.). Known volumes of the lower, aqueous layer were ground into weighed amounts of Hyflo (generally 0.66 ml./g.) and the mixture poured into the column in the form of a slurry in the upper, organic layer. Im-
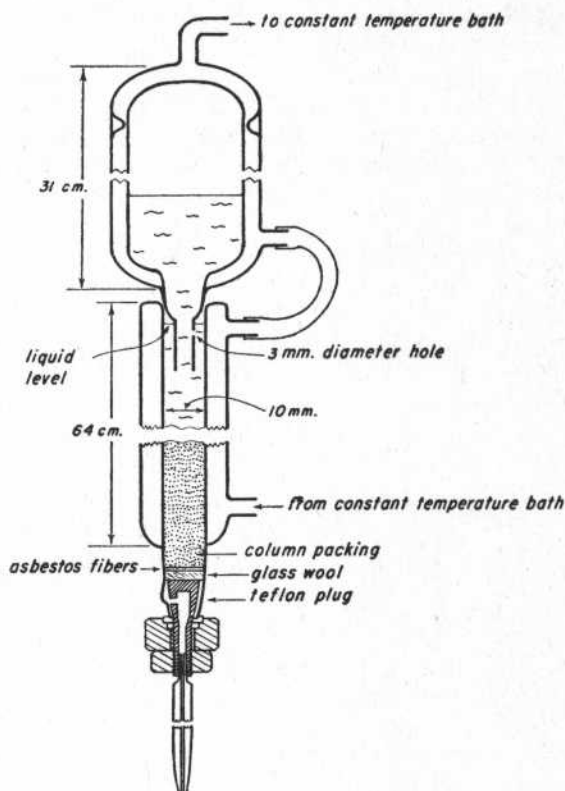
Figure 10. Diagram of the column apparatus used in the partition column chromatography of insulin.

mediately upon addition of the sample dissolved in the moving phase (generally 5 mg./ml.), collection of fractions was started at a flow rate of 0.1 ml./sq. cm./min. The arrangement of the column shown in figure 9 was derived from considerable experimentation. The arrangement was necessitated by our desire to determine the insulin in the eluted fractions by absorption of light at 275 mμ. In order to eliminate high and variable blanks, materials which would contribute impurities absorbing in the ultraviolet had to be avoided. This included the exclusion of commercially available stopcock greases and rubber and plastic tubing. The column is constructed from a commercially available buret containing a teflon plug. The arrangement of the latter makes it especially easy to dismantle and clean the column.

RESULTS

2-Butanol—0:5% DCA in 0.01 N Hydrochloric Acid

The solvent system of Harfenist and Craig (2-butanol-1% DCA) could not be used without some modification as the insulin came off the column shortly after the hold-up volume. This was

to be expected on theoretical grounds from the value of the distribution constant and partition factor in this system. Empirical experimentation showed that incorporation of 0.01 N hydrochloric acid in 0.5% DCA yielded solvent systems in which the K' value of the insulin was decreased appreciably. Figure 11 shows the results of a
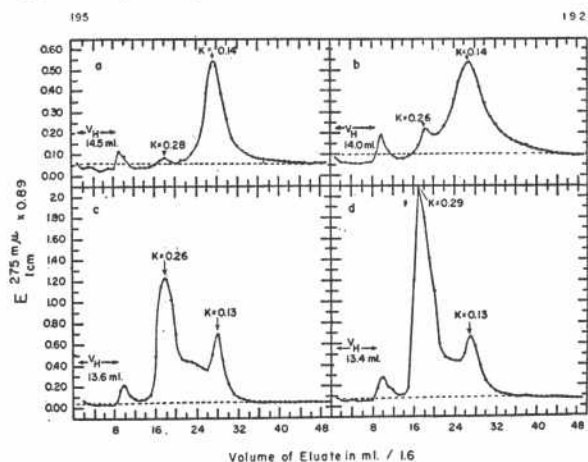
Figure 11. Effect of amount of insulin added to column on shape of elution curve. Solvent was 2-butanol-0.5% DCA in 0.01 N hydrochloric acid. Variable amounts of crystalline zinc insulin (Connaught 885) put on column in 1.5 ml. of moving phase; (a) 7.5 mg., (b) 11.25 mg., (c) 15.0 mg. (d) 20. mg.

series of experiments using this solvent system and illustrates some initial difficulties we encountered which should serve as a warning to others contemplating use of the technique.

The columns were all the same except for the amount of insulin added to the column which was varied from 7.5 to 20 mg. Note that in each case there was an indication of three components, one moving essentially with the solvent front, one with a distribution constant (K) of about 0.28 and another slower moving component with a K value of about 0.13. The relative amount of the total material that occurred in these peaks, however, varied with the load on the column. Obviously some sort of anomaly was present. This was further emphasized by the fact that if material was isolated from either one of the two slow moving peaks, and rechromatographed under similar conditions, the results were very similar to those found on the initial insulin. Further experiments showed that the anomalous results were due to the amount of insulin put on the column and not to the concentration in which it was added. Without going into the experiments which

delineated the cause for these anomalous results, I will simply state that they were traced to two factors: (1) the very small amount of hydrochloric acid present in the butanol-0.5% DCA in 0.01 N hydrochloric acid solvent system, and (2) the fact that isoelectric insulin was used as the solute. It appears that in solvent systems containing both hydrochloric acid and DCA, the more highly dissociated hydrochloric acid bears the primary influence in determining how the insulin will distribute between the two phases. When small amounts of isoelectric insulin are placed on the column, the insulin is largely converted to its hydrochloride which partitions between the two layers with a relatively small distribution constant. Increasing the amount of isoelectric insulin added to the column saturates the hydrochloric acid and the DCA salt begins to form. The latter has a larger distribution constant than the insulin and moves ahead down the column.

Chromatography of Insulin Hydrochloride Between 2-Butanol and 0.5% DCA in 0.1 N Hydrochloric Acid

In order to avoid these anomalies, insulin was converted to its hydrochloride and the amount of hydrochloric acid in the system was increased to 0.1 M. Figure 12 shows the results of a series of
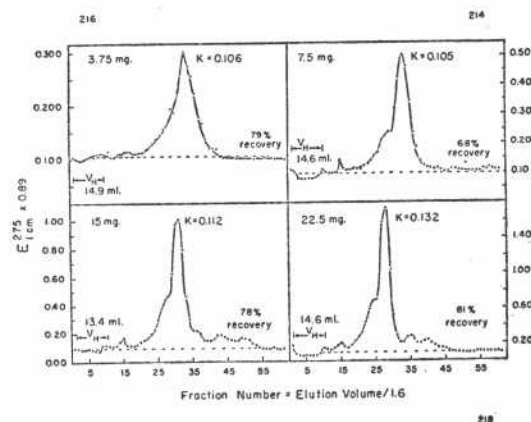
Figure 12. Partition chromatography of various amounts of insulin hydrochloride. Solvent was 2-butanol-0.5% DCA in 0.1 N hydrochloric acid. $V_s = 4$ ml., 6 gm. of Hyflo. Various amounts of insulin (Lilly 535664) as the hydrochloride were added to column in 1.5 ml. of moving phase.

experiments wherein the amount of insulin added as its hydrochloride was varied over a greater range than that shown in figure 11, but the results in each case were quite similar. Each had a main peak moving with a K value of about 0.11 which exhibited a small shoulder composed of material

moving with a slightly greater K value. Further evidence that the columns were performing satisfactorily was obtained when the amount of stationary phase used to prepare the columns was varied. Figure 13 shows the results obtained when
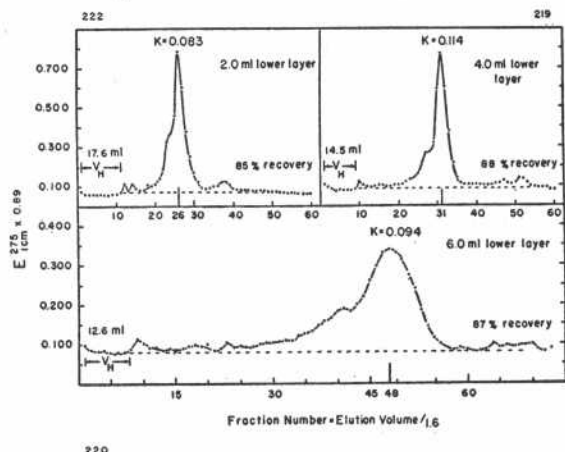
Figure 13. Partition column chromatography of insulin as a function of stationary phase used in preparing column. Solvent was 2-butanol-0.5% DCA in 0.1 N hydrochloric acid. 6 gm. of Hyflo. 7.5 mg. of insulin (Lilly 535664) as the hydrochloride added to column in volume of 1.5 ml. (a) $V_s = 2$ ml.; 85% recovery; *calcd.* $N = 276$. (b) $V_s = 4$ ml.; 88% recovery; *calcd.* $N = 337$. (c) $V_s = 6$ ml.; 87% recovery; *calcd.* $N = 160$.

2, 4, and 6 ml. of stationary phase were used. These changes in stationary phase were reflected by changes in the values for the partition factor for the main component as would be predicted from the theoretical considerations. K' values of 0.73, 0.41, and 0.20 were found for the columns with 2, 4, and 6 ml. of stationary phase. At the same time, the distribution constant remained practically unchanged a result predicted by the theory.

These last results would indicate that the columns are functioning as true partition columns, and therefore one should be able to attach some significance to the shape of the curves. Concentrating on the main band (figure 13), it is obvious from inspection that the band does not have the symmetrical shape expected for one component. It appears to be made up of a main component with a K value of about 0.1 and a faster moving component which appears as a shoulder on the main band. In this respect the results from the partition column are similar to those shown earlier by CCD. That the shoulder on the main peak was real and not an artifact was demon-

strated in a series of experiments wherein the material was isolated from various areas of the band and subjected to rechromatography (figure 14). Fraction I, from the right hand side of the band, gave a nearly symmetrical elution curve
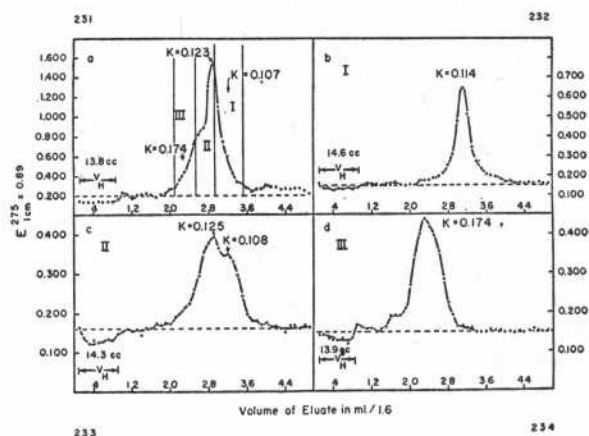
Figure 14. Partition column chromatography of an insulin hydrochloride sample and of fractions obtained from the original sample.

with a K value of 0.114. Fraction II gave rise to a broad elution curve with two distinct peaks corresponding to K values of 0.125 and 0.108, and fraction III gave a rather broad band with a K value of 0.174. Thus each fraction on rechromatography gave rise to an elution curve which was qualitatively in agreement with the result predicted for these fractions on the basis of their origins.

COLUMN PERFORMANCES

The results presented so far have demonstrated the potentialities of partition column chromatography in the separations of insulins. One problem which is as yet unsolved concerns the recovery of material from the columns. The recoveries rarely go as high as 90% and will average around 80%. As yet there is no explanation for these low recoveries. Recoveries of around 80% are good enough in making use of the columns for isolation of insulin from biological materials or reactions mixtures. However, the low recoveries restrict observations on the homogeneity of the material to that fraction of the sample which is recovered in the eluate. Remarks on the original sample would have to take into consideration that 15% to 20% of the sample was not recovered.

Another point which should be discussed is

the value for the partition factor (K') of insulin on these columns. The distribution constant for the main component was about 0.1. In most of the columns the ratio of moving to stationary phase was about 4 which yields a K' value of about 0.4. The earlier theoretical discussions indicated that much better performance would be expected with a K' value of about 0.1. There are two experimental approaches by which one can alter the K' value. One involves changing the amount of stationary phase used in preparation of the column. In this particular case, increasing the amount of stationary phase on the column decreased the value for K'. Unlimited increase of the amount of lower layer, however, is prohibited by the fact that there is a limit to the amount of stationary phase which can be held immobile by the inert support. The Hyflo that we used appeared to have a limit of about 1 ml./gm., but in order to obtain good reproduction of results it was found advisable to use a somewhat smaller ratio (about 0.66 ml./gm.). The other experimental approach to changing the K' value requires a search for a different solvent system which will yield a smaller distribution constant. It was this sort of a search that led to the previously mentioned results on the effect of various acids on the distribution constant of insulin.

As a final point I should like to call attention to the efficiency of the columns. Most of the work reported above was performed on columns 1 cm. in diameter with heights of about 25 cm. The number of theoretical plates describing the column varied somewhat from column to column but generally was in the range of 200-300. The degree of separation of the two components indicated in the elution curves was roughly comparable to that observed in about a 1,000 transfer CCD experiment while the time involved in running the column was only a fraction of that encountered in the CCD experiments. In theory, the resolution obtained on the column could be considerably improved by increasing the height of the column four to five times the height of those used here. In practice, we have not increased the height to this extent but have noted improved resolution on increasing the height from 25 to 37 cm. Of course increasing the height of the column increases the length of time it takes to perform an experiment.

In summation, I think it is obvious that partition columns, at least in their present state of development, are of limited applicability in the separation of proteins. Their greatest usefulness lies in the hormone and large polypeptide field.

In these areas, they compete with CCD as a separation technique. The columns have certain advantages over CCD from the point of equipment costs, time involved and choice of solvent systems. On the other hand, CCD has an inherently larger capacity and a much sounder theoretical basis than the columns and in addition has a practical advantage in the fact that substances cannot be lost—whatever goes in must come out which is certainly not always true with the columns. I am sure that as a result of further studies on the theory and application, the partition columns will become more generally useful to the investigator in biology, but even so the technique should not be entered upon lightly as there are numerous pitfalls to trap the unwary investigator.

DISCUSSION

DR. EDMUND H. FISCHER: From a practical point of view, what is the highest molecular weight protein that could be separated with this type of procedure?

DR. CARPENTER: In the acidic solvent systems described here one is restricted by the stability of the protein. In general, I think, stability rather than size is the important criterion. The solvent systems introduced by Porter, which consist of various ethers of ethylene glycol, have been used successfully on proteins of molecular weights up to about 100,000. The limiting factor would appear not to be so much size in itself but rather the effect of size on the stability of the protein. The large molecules as a rule would be expected to be less stable in the organic solvents than small molecules.

QUESTION: Do different proteins have different rates of equilibration?

DR. CARPENTER: Theoretically one would expect the time for equilibration to increase with size of the protein owing to change in the diffusion rate between the two phases. We have no comparative data. In our insulin work the optimum flow rate was determined experimentally by decreasing the flow rate by factors of 2 until further decreases had little effect on the shape of the elution curve. The flow rate determined for insulin was 0.05 to 0.07 ml./sq. cm. cross section/min.

QUESTION: Other things being equal, what would be an ideal ratio of length to cross section?

DR. CARPENTER: This is an experimental problem. According to theory, the number of

theoretical plates is a function of the length of the column. Doubling the length should double the number of theoretical plates. The cross sectional area determines the amount of material one can put on the column. In theory the cross sectional area has nothing to do with the number of theoretical plates. In practice, however, the cross sectional area has an indirect effect in that an overloaded column will perform as if it had fewer theoretical plates then an optimum loaded column. One point that is not very widely appreciated is the fact that doubling the length of the column is not nearly as good a way to improve a separation as figuring out some way to run the column so that one obtains double the number of theoretical plates on the same length column. Frequently this better way is found by decreasing the load or by the comparable operation of increasing the diameter of the column.

REFERENCES

1. A. J. P. Martin and R. L. M. Synge, Biochem. J., *35*, 1358 (1941).

2. S. W. Mayer and E. R. Tompkins, J. Am. Chem. Soc., *69*, 2866 (1947).

3. E. Glueckauf, Trans. Faraday Soc., *51*, 34 (1955).

4. L. C. Craig, J. Biol. Chem., *155*, 519 (1944); L. C. Craig and D. Craig, "Techniques of Organic Chemistry," Vol. 3, ed by A. Weissberger, Interscience Publishers, Inc., N. Y., N. Y., 1950, Ch. 4.

5. J. R. Weisiger in "Organic Analysis," Vol. 2, Interscience Publishers, Inc., N. Y., N. Y., 1954; E. Hecker, "Verteilungsverfahren in Laboratorium," Verlag Chemie, Gmbtt., Weinheim, Germany, 1955.

6. A. J. P. Martin and R. R. Porter, Biochem. J., *49*, 215 (1951).

7. G. P. Hess and F. H. Carpenter, J. Am. Chem. Soc., *74*, 4971 (1952).

8. E. J. Harfenist and L. C. Craig, J. Am. Chem. Soc., *73*, 877 (1951); *74*, 3083 (1952).

9. R. R. Porter, Biochem. J., *53*, 320 (1953).

10. W. Anderson, Compt. rend. trav. lab. Carlsberg, Ser. Chim., *29*, 49 (1954); Acta Chem. Scand., *8*, 359 (1954); *8*, 1723 (1954).

11. F. H. Carpenter and G. P. Hess, J. Am. Chem. Soc., *78*, 3351 (1956).

Reactions of Proteins

HEINZ FRAENKEL-CONRAT

Again we start out by discussing the title of this talk as chosen by Dr. King. I was very happy to be allotted this title because sometimes I wonder whether I am going to maintain a status as a protein chemist or slip altogether into virology which I very definitely do not intend to do. However, this very unlimited title "The Reactions of Proteins," covers everything, including fields such as immunology, enzymology, and so forth. I am sure Dr. King did not mean me to cover all that because he has other speakers on the program, and I do not know very much about 95 per cent of the subjects which the title covers. What he means is the chemical reactions of proteins, and I think since this is a biology symposium he wants me to stress the importance of chemical reactions in learning something about the biological mode of action of proteins. Chances are that almost all processes that proteins enter into biologically are interactions of chemical groups on the surface of molecules or near the surface of molecules, involving the same groups which react chemically with the various reagents.

This field which I will briefly review has been a stepchild all along, really, in the history of protein research. Physical chemists took over the proteins in a grand style about thirty years ago, and then the biologists entered into the field more actively and chemists were really just running along, sort of. In recent years amino acid and the structural analysis of proteins has come very much in the foreground, which is a very good thing because the first thing a chemist wants to do is to write down the complete structure and he can now thanks to Sanger do that for insulin and probably soon for ribonuclease and some other proteins. But the more we learn about amino acids sequence the more we become sure that that is not the whole story. Proteins are so big that they can obviously not be unorganized chains of amino acids; they must have definite 3-dimensional structures. And thus groups will come close together that are not neighbors along the peptide chain. Proteins are 3-dimensional on cineramic proportions and the chemistry of the native protein, the study of the reactions of the groups in the native protein, is probably going to be a key to an understanding of enzyme action. This field will come in the foreground after people

have finished the amino acid sequences and then ask themselves the question: so what? Therefore I think that I want to stay in this field so that I can witness its renaissance as an active participant.

In the past I was associated with an effort at the Western Regional Research Laboratory and several other laboratories to search specific reagents for the various groups that determine the nature of a protein—that is, the side chain groups of the various reactive amino acids. These groups may be briefly summarized as follows: There are the basic groups of lysine, arginine, and histidine. (The latter, however, is not ionized under physiological conditions and most of the reactions are performed under physiological conditions. If one wants to react the native protein chemically, one should try to stay away from techniques which denature it before they get started.) Then there are the acidic groups of the glutamic and aspartic acids. There is also one acidic and one basic group at the end of each chain. Then there are the nonionically charged yet polar groups such as the aliphatic hydroxyls and the tyrosine phenolic group. Less polar are the indole, amide, disulfide, and thioether groups. And finally the highly reactive —SH groups.

The ideal would be that for each of these groups we would have a selective reagent which would preferably react in stepwise and reversible manner. But this practically does not exist. Some of the reagents we have are relatively specific in that they attack only one type of group, others attack two or several groups. A lot of work has been done with such reagents, and I will summarize a little of it very briefly. Altogether the conclusions from that type of work have been limited. New methods and new ideas are desperately needed in this field. I will later give a few examples of recent progress, not really in finding new methods because new methods have not been found, but in a few interesting isolated leads where reactivities of protein groups, have turned up that were not suspected. At first let me review some of our older work.

Amino groups are very readily acetylated by acetic anhydride and this is a very gentle reaction that can be carried out under physiological conditions. Carboxyls can be pretty well esteri-

fied in cold methanol containing weak HCl, but already here we have nonphysiological conditions. Another more physiological method of esterification has been described in more recent years, but not yet in any detail by Dr. Wilcox who is now at Seattle. He used diazoacetamide, but the reaction does not go very far. In contrast, with methanolic HCl one can esterify all the carboxyls of a protein and the conditions are really surprisingly nonharmful to many proteins. For instance, serum albumin can be completely esterfied without denaturation. One thing that fits in with Dr. Carpenter's talk is that when you do esterify a protein, it dissolves in the methanol in the course of the reaction. When the carboxyls are blocked, then the proteins act like bases and their hydrochlorides are more soluble in organic media than the original protein.

Now for phenolic and imidazol groups, we have no reliable selective reagent. Iodine reacts very rapidly with the phenolic groups of many proteins, forming mono and diiodotryrosine. With some proteins histidine does not react at all, while all the tyrosine can be iodinated. With another protein lysozyme, it worked just the other way around. First iodine goes on the histidine residue which seems to be connected with the enzymatic activity because lysozyme is inactivated by this first atom of iodine, and if you are quick enough and add a reducing agent you can reduce iodine off again and regenerate the activity. The tyrosine of this protein reacts very sluggishly. So here you have an example of the fact that there are no rules in this game of playing with proteins. Whatever you establish as being a firm and certain fact about a certain protein does not necessarily hold for the next. Critical analytical control is required at each step.

Now for the aliphatic hydroxyl groups concentrated sulphuric acid has, surprisingly, turned out to be a selective reagent. If you tell that to biologists and they do not throw you out of the room it is really surprising. But actually you can perform this reaction with insulin and insulin will retain its activity. Thus you can transform all its hydroxyl into $-O-SO_3^-$ groups and get a completely different protein with an isoelectric point near pH 2.0 and it still retains the original insulin activity.

This may suffice as a review of some of the reagents we used five to ten years ago in studying the essential groups of a number of biologically active proteins. Now a few summarizing remarks on the type of result that we got. Insulin turned

out to be remarkably unaffected by chemical modifications. Only some of the phenolic groups appear essential, besides the disulfide bridges. Avidin from egg white is another protein of great stability. Its remarkable affiinity for the vitamin biotin seems to be practically independent of all groups that were tested. On the basis of this and some very indirect evidence we proposed a cyclic diimide group as the biotin-binding site. Now an example of quite different behavior is crotoxin, the rattlesnake venom neurotoxin which we worked with over the years quite a bit. Although this protein is chemically very similar to insulin, very high in cystine, and generally quite stable, it is one which was inactivated by almost every reaction. Another interesting pair of proteins that were studied are trypsin and ovomucoid. The latter is an egg protein which is a powerful trypsin inhibitor. It is thus a member of a whole class of proteins which will interact stoichiometrically with trypsin and inactivate the enzyme. Trypsin does not need its amino groups for its enzymatic activity. Thus one can acetylate trypsin and it is still active but it will then not be inhibited by ovomucoid. On the other hand, trypsin does seem to need indole and certain other groups; there was some evidence in that direction. We therefore postulated that the enzymatic site was surrounded by amino groups and that those amino groups would combine with carboxyl groups of the inhibitor. Actually it was found that in ovomucoid the carboxyls were essential for the activity, and it has definitely been established that the reaction of the two proteins involves ionic interactions. Another interesting protein is conalbumin, because of its affinity for iron. The mechanism of this binding of the iron is complex and involves a variety of groups. The finding that denaturation alone inactivates conalbumin indicates that the site that binds the iron is one that requires hydrogen bonding to hold certain groups together which form the active site. In all of this type of work it is important to realize that reactions which do *not* destroy the activity, particularly together with analytical proof for a high extent of reaction are of greatest value in establishing what groups are absent from the enzymatic site. In contrast, loss of activity with a given reagent is less conclusive, because it could be due to nonspecific causes, such as steric hindrance or denaturation.

One of our particular interests in reactions of proteins in recent years has beeen with sulfhydryl reactions and with one particular one,

which again has been a nice illustration of the way in which you cannot predict what is going to happen with a given protein. Tobacco mosaic virus (TMV) has one cysteine, one SH group per subunit or peptide chain, of which there are almost 3,000 in each virus particle. These subunits could be regarded as separate protein molecules except that they are all neatly arranged in a hydrogen-bonded structure. Thus their -SH groups, just like those of egg albumin, are masked and permanently stable; however, if one denatures the virus protein, the sufhydryl group becomes auto-oxidizable and disappears just as with egg albumin. So the nature of this -SH group has been of marked interest. When one treats most -SH proteins with iodine, the sulfhydryl groups get oxidized to disulfides, requiring one molecule of iodine for two -SH groups. $(2\text{-RSH} + I_2 \rightarrow \text{R-S-S-R} + 2\text{HI.})$ It was known, however, that tobacco mosaic virus required twice as much iodine for the removal of its sulfhydryl group. That was the thing that just stymied people and was noted as a fact and left there because it did not seem to make any sense. Similar behavior is known for the -SH group of serum albumin.

Now if you formulate how two iodine atoms can react with one -SH group $(\text{R-SH} + I_2 \rightarrow \text{R -SI} + \text{HI})$ you arrive at a sulfenyliodide, a derivative of sulfenic acid (R-SOH). And aliphatic sulfenyliodides are so labile they are as good as nonexistent. Rheinboldt, after considerable effort, succeeded in preparing tertiary butyl sulfenyliodide and this unusually stable sulfenyliodide turned out to have a half-life of half an hour in ether at minus 20°. In water it immediately decomposes. Yet when you treat the virus with iodine you find that for each -SH that disappears and each molecule of iodine that is used up, one iodine is found in the protein. The reaction product looks slightly yellow as do sulfenyliodides generally, as long as they exist. But this one exists a long time in the refrigerator—i.e., indefinitely. Thus it seems that the masked state of the -SH group of a protein can be transferred to a derivative of that group. Just as the -SH group of TMV does not show the reactivity of mercaptans until the virus is denatured, in the same way can a sulfenyliodide be protected by the native protein structure. If this is the mechanism, then upon denaturation of protein the -SI group should regain its customary reactivity. Thus cysteine or any other sulfhydryl compound added prior to denaturation would react with the -SI group and would form a disulfide bond as soon as the denaturant was added. It appears probable that this is the mode of oxidation of all sulfhydryls. Instead of imagining that two -SH's have to be brought together with oxygen to give a disulfide bond, it appears more probable that one -SH goes to an -SOH stage and the -SOH then reacts with a second SH group to give the -S-S- bond.

Thus we are quite convinced that really here we have a case where the protein through its native state, through its native foldings, is able to stabilize an organic intermediate which normally, to the organic chemist, is nonexistent. We naturally tried to find other similar cases, but we have so far found no other stable sulfenyliodides, in serum albumin or other proteins. This does not mean that such a reaction mechanism does not exist as an intermediate. But in no other protein was there a stable sulfenyliodide which could be isolated, and which then could be found to lose its iodine upon addition of a denaturing agent, and that was our criterion.

This finding points in an interesting direction. Protein chemistry can be a chemistry different from, and not obeying all the rules of, organic chemistry because of the 3-dimensional folding and ensuing stabilization of intermediates. Many enzyme reactions may involve similarly labile intermediates which may be stable within the protein. A particular system where this phenomenon would explain many observations is the activity and the inhibition of chymotrypsin and similar esterases by di-isopropylfluorophosphate, which will be discussed in detail by Dr. Fisher.

Now just a few words on some new reactions of the disulfide bonds of proteins which have interested me quite a bit, although I have not been actively engaged in this work. We used to assume that these bonds are quite stable. However, when an -S-S- protein is partly reduced it usually gets very insoluble and often sets to a gel. This occurs even if auto-oxidation is excluded. The explanation seems to lie in an interesting chain reaction. Thus if to a disulfide protein one adds a trace of a small molecular mercaptan, then this will add to a disulfide bond, generating an -SH group on the protein. This new -SH group can interact similarly with another disulfide group on the same or another molecule, each time forming new -S-S- bonds and thus one hydrogen atom can be handed on from one disulfide to the other. The end result is a 3-dimensional cross linking of the whole re-

action mixture and the whole test tube may set to a gel. This realization that very few sulfhydryl groups can start off the cross-linking of an entire protein solution has proven of very great usefulness in explaining various phenomena. One of these is the mechanism of blood clotting. It has been suggested that the change of fibrinogen to fibrin may involve such a chain reaction, initiated by a trace of -SH. Another instance is the elasticity of wool, which was found to be dependent on the presence of a few -SH groups. A number of other protein problems may be similarly explained when it is realized that it is only one SH group that is needed to cross link an unlimited number of disulfide bonds.

Another oddity in the behavior of some -S-S- proteins may have a similar explanation. It has been known that certain disulfide proteins are very sensitive to copper. For instance lysozyme and ribonuclease, both rich in cystine, both very stable proteins, are inactivated at pH 7-8 in the presence of copper ions. Recently it was found in Klotz' laboratory that serum albumin undergoes a complex series of reactions in the presence of cupric ions. The explanation is the interaction of the copper-substituted -SH group with a disulfide bond in the vicinity and a similar chain reaction may actually occur. It seems that traces of copper may be harmful to reactive disulfide proteins because every disulfide protein probably has a trace of -SH present through hydrolysis and if the copper can be handed on from one -SH to many disulfide groups it can cause havoc with the neatly folded disulfide cross link structure of such proteins. So these are just a few examples and isolated observations which point in interesting directions where chemical methods contribute to our understanding of the mode of enzyme action, and other protein functions. But as I said at the beginning, this field really needs more new blood and new ideas and less talking.

DISCUSSION

DR. HOWARD S. MASON: Your mentioning Klotz reminds me of a passage in "The Mechanism of Enzyme Action" symposium at Johns Hopkins in which Klotz discussed in length the focus of catalytic action at a single point. At the end of his talk, Rittenberg got up and asked him what was the function of a protein, that is in the rest of the protein. I think to a certain extent that question could be asked of you: If the activity is entirely localized at one point on the protein surface, what is the function of the rest of the protein in increasing the catalytic activity of that configuration?

DR. FRAENKEL-CONRAT: Well, I thought I made it clear that I do think that most enzymatic sites will be not two or three adjacent amino acids but a definite 3-dimensional structure, formed by the folding of several chains in the area, which the rest of the protein is necessary to stabilize. The size of this structure may vary very much. Some enzymes may have very small sites and others bigger active areas, judging from their inactivation by so many reagents. But whatever the area, it needs a lot of neat arranging to form a stable surface, and that is where much of the protein is apparently needed. But there are other examples I did not mention. Papain seems to be an enzyme which does not require over half of its protein. According to recent studies by Emil Smith, the enzymatic site is very close to the carboxyl end of the peptide chain and may be quite small. So I think we have to have a wide open mind. In some proteins you may find that only ten amino acids will do; thus in cytochrome C a small section of the molecule will simulate some of the enzymatic activity.

DR. VERNON H. CHELDELIN: Is it economical for an organism to synthesize a molecular weight of a 100 thousand in order to get the effect of 5 thousand?

DR. FRAENKEL-CONRAT: Well, nobody has done it with five thousand yet.

DR. HENRY BORSOOK: I think in answer to your question there is a clear indication in one case, from the work of Alberty on fumarase, that the locus of the fumarase action is really two histidines. The only way one can get those two histidines close enough together is by having the molecule big enough to properly fold. One histidine then can take one proton, the other gives a proton, both operations go on at the same time. So this is what the large size of a protein molecule makes possible. It looks as if in ribonuclease we have somewhat the same thing. I wish to ask you, Dr. Fraenkel-Conrat, what would you say to the wild guess that in every enzyme, where there is a big change in its activity around neutrality, then in the enzyme locus histidine is involved?

DR. FRAENKEL-CONRAT: Arguments based on the dissociation constant of histidine alone have always left me a little bit cold because one knows that in proteins the dissociation constants are often changed.

DR. BORSOOK: I take it that the dissociation constant of histidine in proteins is different from that of free histidine. But it is in proteins that histidine side chain has a pK of nearly 7.

DR. FRAENKEL-CONRAT: It may often be so. Therefore I would think one should look for other evidence, but I think that histidine does seem to play a role in quite a number of bioactive proteins.

QUESTION: The work on chymotrypsin?

DR. FRANKEL-CONRAT: Yes, and ribonuclease also. Histidine seems to be one of high bonding capacities and pliable in its activities, and I think whenever there is this particular pH relationship there is more justification in looking for histidine being involved. Papain has no histidine in its active parts.

DR. EDMOND H. FISCHER: It might be interesting to note in this respect that an imidazole group seems to be involved in the active site of nearly a dozen hydrolytic enzymes. Very recently there was a note by Larner indicating that histidine appeared to be the active group of two intestinal glucosidases: maltase and an oligo-1, 6-glucosidase.

DR. FRAENKEL-CONRAT: What is the pH on that?

DR. FISCHER: Close to neutrality, I imagine. In addition histidine has been found to be of primary importance, in lysozyme, histidase, ribonuclease, and chymotrypsin, fumarase, etc.

DR. OTTO BEHRENS: I am just questioning whether there is any evidence that in the protein some of the groups are activated. If I remember correctly, there are some of the amino acids that are more sensitive during the time of hydrolysis when they are present in separate forms, and I wonder whether this would not be an evidence that some groups may be reactivated when they are present sometimes?

DR. FRAENKEL-CONRAT: I think serine is one of them.

DR. EDMOND FISCHER: I would like to ask two questions. The first concerns the structure of the site on avidin which is responsible for the reaction with biotin. You assumed an imine group placed between two carboxyls. The hydrogens of such a group should be quite dissociable. Could this be shown by titration; could you block such group and prevent the combination of avidin to biotin?

DR. FRAENKEL-CONRAT: We tested only its rate of hydrolysis, and we made model compounds, diacetamide and so on. With only one or two active groups in a molecular weight of 80,000 the analytical methods used at the time were not good enough to prove anything convincingly. The destruction of the activity, under the same conditions under which diacetamide is hydrolyzed and several other experiments were suggestive, but nothing is definitely proven. No work was done on this for several years. It is a nice subject for somebody to work on again.

DR. FISCHER: The next question concerns the SH group TMV which reacts with iodine. Do you find any metal in the protein moiety of TMV?

DR. FRAENKEL-CONRAT: I do not know about trace metals. But the sulfhydryl group reacts like any free SH compound when you denature it. You can titrate it just as egg albumin and other SH groups. I mean it is not blocked by any metal, but maybe the catalytic effect of the metal, if that is what you are thinking about, cannot be excluded.

DR. FISCHER: Well, I do not know. I am thinking of the possibility of some metal ion being bound by the sulfhydryl group and which could accept iodine at the same time. An SI compound would be so unstable that I can hardly believe it could exist as such.

DR. FRAENKEL-CONRAT: The principle is no different from the stability of the sulfhydryl group as such which is also in a sense energy rich and will auto-oxidize when free at pH 7. I am pretty sure there are no metals in stoichiometric proportion. I believe that at various times people have analyzed tobacco mosaic virus quite thoroughly for practically everything and I am pretty sure that if there were one . . .

DR. FISCHER: They would have found it.

DR. FRAENKEL-CONRAT: Yes, yes. I am pretty sure it would have been detected and described.

DR. HOWARD MASON: Now that emphasis is being placed on biological continuum, I think we might start thinking in terms of part of the protein being involved in its binding or localization within the cell and part of it being concerned with the catalytical activity, so maybe the truth in this matter lies some place in the middle.

QUESTION: Are there any concentrations of

the sulfhydryl as they react in the chain reaction? There are some cases where a metal in sulfhydryl concentration does not give a gel, and other cases I know of where a gel is produced. So when you go down certain condensations, how do you explain that?

DR. FRAENKEL-CONRAT: In the case of wool, it has a very low sulfhydryl content, only about 0.2 per cent cysteine, and 12 per cent cystine. And yet that 0.2 per cent is enough to account for some of the typical physical properties of wool. I would expect that this is evidence for a high reactivity of some disulfide bonds. It would appear that the availability of the -S-S- bonds, whether there are more of them on the surface and their distance from the next sulfhydryl and from one another, would determine how far this chain reaction can go, what can initiate it, and how far it can go. So it would again be a problem to be solved individually for each protein.

Structure of Enzyme Loci*

Edmond H. Fischer

I would like to discuss today one or two points which are known about the structure of enzyme loci. If time permits, I shall tell you about the work which we are carrying out in Seattle on the activation of a muscle enzyme which exists in an inactive state but is transformed physiologically to an active form which can then control the metabolism of glycogen.

One usually understands the term enzyme loci as referring to those particular "essential groups" or "active centers" of an enzyme molecule which are responsible for its specificity in catalyzing a given chemical reaction.

The concept of the existence of such "active centers" arose gradually from a great variety of findings. I belive it began with Henri's, and Michaelis and Menten's theory of enzymic reaction, postulating the formation of a dissociable enzyme-substrate complex, according to the equation:

$$E + S \underset{k_2}{\overset{k_1}{\rightleftarrows}} ES \overset{k_3}{\rightarrow} E + products$$
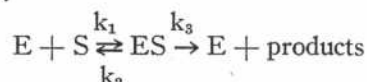
In this complex, stabilized by forces of interaction, the substrate molecule is supposed to be strained—i.e., become more chemically reactive (we say "activated" today), by some sort of intramolecular transformation. This, in turn, will enable a second reaction to take place, by which the enzyme-substrate complex will break down into regenerated enzyme and final products. As the steady state concentration of the complex can be strictly defined in terms of the law of mass action and rate of its transformation, the lifetime of the complex must have a finite value, so that it was logical to assume that certain chemical groups of the protein were specifically involved. This assumption was supported by evidence that in certain cases of enzyme inhibition by molecular species of a structure similar to that of the substrate, the inhibition could be reversed at high concentration of substrate. This phenomenon could be explained in terms of a competition between the two reactants for one or several

"essential groups" or "active sites" on the enzyme, and therefore, this type of inhibition was called "competitive."

A second argument pointing toward the existence of "active centers" resulted from the discovery of characteristic prosthetic groups in certain enzymes. These nonprotein groups act as mediators of the enzymic action, as they participate directly in the reaction catalyzed by the enzyme. In several cases, they can even act by themselves, although their activity is very much enhanced when they are combined with the protein or apoenzyme. A third argument, as mentioned by Dr. Fraenkel-Conrat, is that, in certain cases, one can chop off certain parts of enzyme molecules and still retain enzymic activity. This is true, for instance, with pepsin (1), papain (2), ribonuclease (3), etc. When chymotrypsin is formed from chymotrypsinogen by the action of trypsin, a single peptide bond is split (a) and we
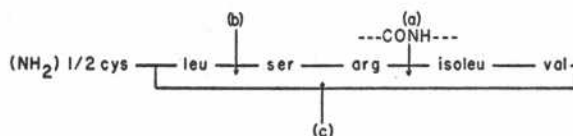
Figure 15. Conversion of chymotrypsinogen to chymotrypsins.

get what is called "π-chymotrypsin"; then by the splitting of a second peptide bond (b), seryl arginine, a dipeptide, is released and we get δ-chymotrypsin; finally, by the splitting of bond (c) we get α-chymotrypsin (4) as shown in figure 15. All these various protein molecules have approximately the same specific activity. Other enzymes could be greatly modified by chemical reagents with little or no loss of enzyme activity.

In all these cases, the suggestion of the concept of an "active center" on an enzyme molecule was certainly justified and helpful. The inevitable corollary to this concept, however, is an extremely dangerous one. This corollary is that if one assumes that the catalytic activity of an enzyme can reside in a limited portion of the molecule, one automatically infers that the large remainder of the molecule constitutes an inactive ballast which could be dispensed with, provided,

naturally, that no secondary ill effects would result for the "active center."

What then would be the role of the protein, as Dr Mason has asked? Well, since the studies of the early thirties, it is known that the substrate specificity of enzymes is due to the protein portion of the molecule. One could also assume, however, that the protein is necessary in stabilizing the "active center," that the principal function of the remainder of the molecule is to protect the active site in some way. Or one could imagine that it might be concerned with controlling the distribution of enzymes in an orderly fashion in or around cellular particles in order to enable metabolic processes to proceed according to well-established patterns. (This function, by the way, would not show up in experiments carried out on isolated systems.) Evidence has accumulated during the last five or ten years that many catalysts involved in given metabolic sequences are associated, within the cell, with well-differentiated structural units such as the nuclei, mitochondria, microsomes, chloroplasts, etc. (5).

Personally, I am not too satisfied with any of these hypotheses. The extraordinary fact that all enzymes are or contain a protein of specific structure suggests, from a purely teleological point of view, that the major function of the protein itself must be connected in some way with the catalytic property of the enzyme.* Indeed, there exists much experimental evidence that conflicts with the view that the activity of an enzyme is restricted to a discrete portion of its molecule. As you have seen, proteins contain many reactive groups which may participate in chemical catalysis (6): phenyl, p-hydroxyphenyl, carboxyl, hydroxyl, thiol, histidyl, amino, and guanido groups, etc. If one assumes that several of these groups can absorb some excitation energy, and that a transport (or a redistribution) of this energy can take place within the enzyme molecule, then it follows that activation of substrate groups may occur in the intermediate complex (7).

Various theories have been proposed by which this transport of energy could be achieved (8): tautomeric transposition of hydrogen bonds normal to the peptide chain and parallel to the helical axis; tautomeric displacement of protons in connection with resonance structures of the peptide

bond; a so-called "continuous theory" based on a hypothesis according to which proteins may be treated analogously to semiconductors. (In other words, a protein particle could be looked upon as "a formation in which electrons are shared by the whole system.") Finally, there is the Perrin-Fuerster theory according to which an excitation energy can be transmitted from a "donor" to an "acceptor." This last theory has been widely applied to photochemical reactions in which the electromagnetic energy of light is chemically stored before being used in a chemical reaction. A very interesting example of this can be seen in Bücher's (9) experiment in which he irradiates carbon monoxide myoglobin at wave lengths absorbed by the tyrosyl or tryptophan groups of the protein, and shows that the energy absorbed is transmitted to the heme prosthetic group, resulting in the dissociation of the carbonyl group from the iron.

I have thought it necessary to stress these last examples so that during the ensuing discussion, in which I will be dealing very much with what have been called "enzyme loci," you may bear in mind the limits of our discussion. I will speak of active centers and prosthetic groups, but remember, please, that in all cases, these particular groups are bound to proteins.

Let us now consider some specific examples in which attempts were made to establish the chemical structure of the reactive site of an enzyme. While extremely early advances could be made in the cases of conjugated proteins, due to the presence in them of the easily recognizable prosthetic groups or coenzymes, great difficulties are still encountered in the investigation of simple nonconjugated proteins.

Until very recently, when direct evidence was obtained on the formation of stable complexes between enzymes and substrates, or certain parts thereof, most of our information in this connection was of a completely indirect nature. At first, the type of work done was that described by Dr. Fraenkel-Conrat, in which an enzyme was treated with a number of reagents whose action on the various groups of proteins was supposed to be known (ketene, nitrous acid, phenylisocyanates, formaldehyde, iodine, iodoacetic acid, oxidants and reductants, heavy metals and metal chelating agents, etc.). If no inhibition occurred, the groups which were supposed to have reacted were declared "nonessential" for the enzyme activity, and vice versa. Now, in most instances, this approach was

*The term "enzyme" was introduced by Kühne in 1878 to designate a catalyst of biological origin, posessing certain properties of living matter in that it was usually destroyed under conditions which destroyed life. As this behavior is characteristic of proteins, and as, moreover, all enzymes thus far isolated in the pure form have turned out to be simple or conjugated proteins, the present accepted definition for these compounds is that they are "catalytically active *proteins.*"

perfectly valid, but unfortunately inconclusive regarding the nature of the active center of the molecule.

More pertinent information was gained through the study of the mechanism of action, substrate specificity, and competitive inhibition of various enzyme systems. It is mainly by this approach that Nachmansohn and Wilson performed their classical work on acetyl-choline esterase (10). It enabled them to postulate an *anionic* site upon the enzyme molecule (attracting the $(CH_3)_3N^+$ group of the substrate) and an *esteratic* one forming a covalent bond with the strongly polar carbonyl end of acetyl choline as shown in figure 16. These two charged groups contribute
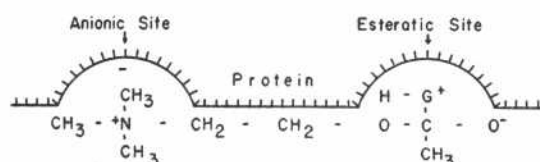
Figure 16. Hypothetical picture of interaction between the active groups of acetylcholinesterase and its substrate (H-G symbolizes the esteratic site). D. Nachmansohn and I. B. Wilson (10).

to the attraction, orientation, and fixation of the substrate upon the enzyme surface. The reaction would then proceed, after the formation of the acyl-enzyme intermediate and through resonance of this complex, by the hydrolysis of the ester bond (11).

The most recent and most direct approach to this problem was made by the use of certain phosphate esters such as tetraalkyl pyrophosphate, DFP, etc., which seem to be attacked by certain esterases or proteases in the same way as substrates, with the difference that the phosphorylated intermediate formed with the enzyme is stable and hydrolyzes only extremely slowly. In a sense, the enzyme commits suicide each time it attacks one of those molecules. In combining chymotrypsin with DFP^{32} followed by acid hydrolysis, Schaffer (12) et al. obtained di- and tri-peptides containing aspartic acid (or asparagine), O-serine phosphate and glycine*. In order to avoid possible degradation or transposition of the phosphoryl group during acid hydrolysis, Oosterbaan (13) et al. used an enzymatic degradation of the DIP-chymotrypsin and obtained a single peptide con-

taining pro (1), leu (1), ser (1), and gly (2 or 3) per DIP, the phosphate group being presumably on the seryl residue. In a similar approach, Dixon, Go, and Neurath (14) treated trypsin with C^{14} labeled DFP and again, after enzymatic degradation of the DIP-trypsin by α-chymotrypsin, obtained a "hot" peptide containing serine of the following composition:

cyst, asp, glu, gly, ser, ala, pro, and val.

A second peptide differing from the first only in the fact that it contained an additional lysine residue (presumably as a C-terminal) was also isolated.

Unfortunately, as usual, things are not as straightforward as they might seem, and there has been good evidence that serine is not the primary group reacting with DFP. Wagner-Jauregg and Hackley (15), investigating the spontaneous hydrolysis of DFP and diethyl fluorophosphonate in water, have shown that it is accelerated by imidazol, histidine, pyridine, and certain of their derivatives. They suggest a hypothesis according to which two essential centers might be involved in the DFP-enzyme reaction. A center containing imidazol would first react with DFP; then, functioning as a phosphorylation catalyst, would transfer the phosphoryl residue to a second acceptor, for instance serine (see fig. 17). This hypothesis was supported (16, 17) by studies of the reactivation of inhibited esterases which showed that, although inhibition could easily be reversed immediately following the addition of DFP, reactivation became more and more difficult the longer the inhibition was allowed to progress. This behavior was also interpreted in terms of a two-step reaction: a first step resulting in the formation of an unstable phosphorylated enzyme derivative followed by a second step in which the alkyl phosphate group was transferred to a second site on the enzyme molecule, forming a more stable derivative: Both enzyme derivatives are inactive. That serine might be the specific and ultimate phosphate acceptor in B seems to be proven by the fact that, out of a total of 27 seryl residues present in chymotrypsin, only one is esterified by DFP (18).
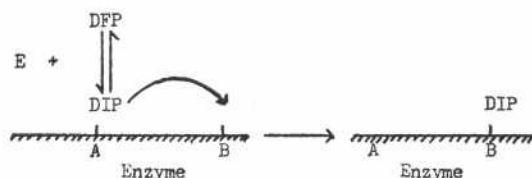
Figure 17.

---

* Schaffer et al., Fed. Proc. 15, 347, (1956): Several radioactive peptides were isolated from Sarin (Isopropyl methyl phosphonofluoridate) treated chymotrypsin and trypsin. Indications are that both trypsin and chymotrypsin contain identical residues (aspartyl and glycyl) adjacent to the serine residue which binds the alkyl-phosphate group.

As to the exact nature of group A, the answer is still obscure. By treating chymotrypsin with various chemical reagents, Sizer (19) has concluded that primary amino, sulfhydryl, and disulfide groups are not required for enzymatic activity, while tyrosine is essential. By studying the pH dependence curve of acetyl choline esterase, Wilson and Bergmann (20) concluded that, indeed, the basic group involved might well be imidazole. A similar conclusion was reached by Weil and Buchert (21), while investigating the photooxidation of chymotrypsin in the presence of traces of methylene blue. They found that complete inactivation occurred when 4 moles of oxygen were taken up per mole of enzyme, at which point only 1 mole of histidine (out of a total of 2) and 2.4 moles of tryptophane (out of a total of 6) were photooxidized.

Finally, Balls (22) has shown that p-nitrophenyl acetate can function as a substrate for chymotrypsin and that, during the reaction, an acylenzyme derivative is formed. The stability of this complex at an acid pH distinguishes it from N-acetyl-imidazole, as Stadtman (23) has shown that this compound is energy-rich and unstable below pH 5 in aqueous solution. But the possibility remains that an N-acetyl histidine derivative might be formed transiently during the enzymatic reaction, the acetyl group being then transferred by a N→O shift, as postulated by Wagner-Jauregg and Hackley.

Although it has not been my intention to give an exhaustive survey of the work which has been done on this problem, I might add nevertheless that, very recently, Larner (24) has indicated that histidine appears to be the active group in intestinal maltase and oligo-1,6-glucosidase. In addition to choline esterase and chymotrypsin, the same group was shown to participate actively in histidase (25), lysozyme (26), and ribonuclease (27). I think that, as brought out by Dr. Borsook, it is very interesting indeed that so many different types of hydrolases seem to owe their catalytic activity to the presence of the same "active" group.

Another noteworthy fact which follows from what we have said is that although we were dealing with purely hydrolytic enzymes, these enzymes can act nevertheless as group transfer agents in the formation of stable enzyme-substrate derivatives. This is important because the most striking advance made during these last few years in the field of enzyme action is the finding that in biosynthetic reactions in general, the enzyme parti-

cipates directly as a group transfer agent. In practically all cases that have been thoroughly investigated, an identical pattern in the enzymatic reaction was found: (1) There is a reaction of the enzyme with the substrate in which the latter is transformed, usually by the splitting of a covalent bond. (2) The bond energy, instead of being liberated as heat, is preserved in the formation of an "active" intermediate. This might involve, in certain conjugated proteins, either the whole enzyme, the coenzyme alone, or the protein alone. (3) The "active" intermediate will react, usually not with one, but with any number of acceptor groups, so that the same enzyme will be catalyzing a great number of different reactions.

As can be seen, we are very far from our classical concepts of the specificity and mechanism of enzyme action. In the second part of this discussion, I would like to deal with the structure of some of these active intermediates, as we visualize them today.

Probably the first clear example of the formation of a stable active intermediate complex was demonstrated by Hassid and Doudoroff in the case of the sucrose phosphorylase system (28). This is an enzyme which splits the glucosidic bond of sucrose, to produce fructose and an active glucosyl-enzyme complex as shown in figure 18. This was proven very elegantly. Hassid

Glucosido-fructoside + E ⇌ [Glucosido-E] + fructose
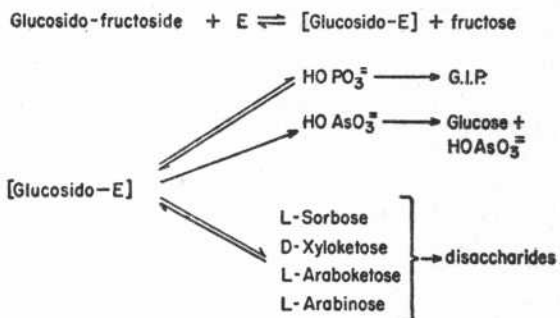
Figure 18. Mechanism of action of sucrose phosphorylase. W. Z. Hassid and M. Doudoroff.

and Doudoroff incubated sucrose with radioactive fructose and showed that in the absence of any acceptor the labeled fructose appeared in the disaccharide molecule. This is represented in the first reaction: splitting of glucosidic bond, but conservation of the energy residing in this bond through the formation of an "active intermediate" between the enzyme and part of the substrate. In a second reaction, this activated glucose resi-

due can react either reversibly with phosphate to give glucose-1-phosphate, or irreversibly with arsenate to give a hypothetical glucose-1-arsenate which is exceedingly unstable in water and hydrolyzes spontaneously to give glucose and arsenate. Or the active glucose can further react reversibly with any number of sugars to give the corresponding disaccharides. We see that this enzyme can act as a phosphorylase, a hydrolase (if catalytic amounts of arsenate are present), or as a transglucosidase, the enzyme itself acting as the group transfer agent.

Table 2. COENZYMES AS GROUP TRANSFER AGENTS

Coenzyme	Abbr.	Group carried
Diphosphopyridine nucleotide	DPN	H^+, e^-
Triphosphopyridine nucleotide	TPN	H^+, e^-
Flavin mononucleotide	FMN	H^+, e^-
Flavin adenine dinucleotide	FAD	H^+, e^-
Lipoic acid	——	H^+, e^- acyl
Thiamine pyrophosphate	TPP	aldehyde glycolaldehyde
Coenzyme A	CoA	acyl
Tetrahydrofolic acid	THFA	hydroxymethyl formyl
Pyridoxal phosphate	——	amino amino acid

In table 2 we see how practically all the coenzymes known today can and do function as group transfer agents. First listed are the well-known coenzymes involved in oxidative reactions: DPN, TPN, FMN, and FAD, which carry hydrogen and electrons. Lipoic acid also carries hydrogen and electrons, but in addition can transfer an acyl group. Thiamine pyrophosphate carries an aldehyde group which we will call an "active aldehyde." Coenzyme A carries an active acetyl group; tetrahydrofolic acid carries hydroxymethyl or formyl groups, either on the five or ten position, and therefore, functions in one-carbon metabolism; pyridoxal phosphate carries an amino or an amino-acyl group, etc.

I will only examine a few typical examples here in which we can see first, the formation of the active intermediate, and then, in a second reaction, how this active intermediate can react with a great variety of acceptors.

To begin, let us examine the case of thiamine pyrophosphate, the coenzyme containing vitamin B_1 as a pyrophosphate ester and involved in the reactions catalyzed by pyruvic oxidase and transketolase. Figure 19 represents a hypothetical structure through which thiamine pyrophosphate
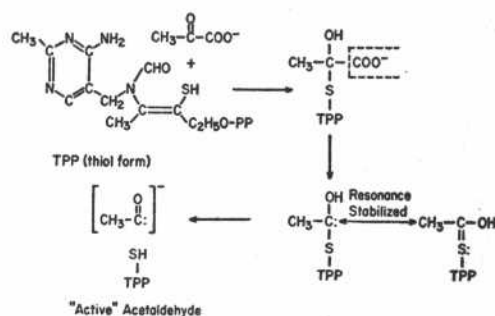
Figure 19. Proposed mechanism for the formation of "active" acetaldehyde.

can react. We do not know how this prosthetic group is bound to the enzyme, although it presumably is linked through its pyrophosphate group and a nitrogen with a catonic and an anionic site on the protein.

There are two schools of thought as to the way the coenzyme might react with pyruvic acid: one, that it reacts in its thiol form (29) in a "thioclastic" type of attack; the other, that it reacts as a thiazolic tertiary nitrogen (as a TPP-pseudo base) (30) to form the intermediate compound. I lean toward the first type of reaction in which there is an attack by the thiol group forming a thio hemi-acetal in the same way Lynen (31) proposed the formation of acetyl-CoA from pyruvate. When this compound is formed, the carboxyl group of pyruvate is automatically released, then, through various rearrangements or resonance, results in the "active acetaldehyde." In a second reaction this active aldehyde will react again with many different potential acceptors (32) as shown in figure 20. Pyruvate itself could be the first acceptor (as obviously it is always present in the pyruvic oxidase reaction) to

PYRUVATE DEHYDROGENASE MECHANISM
I.C. Gunsalus
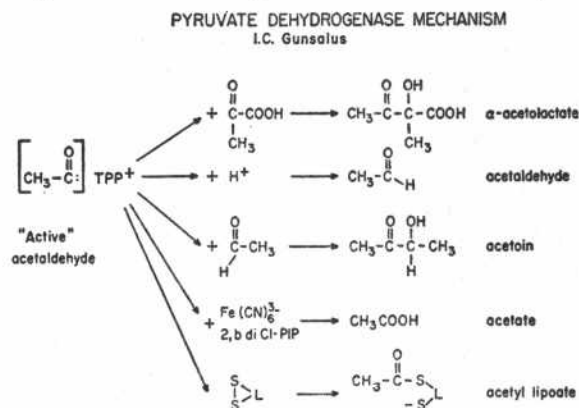
Figure 20. Pyruvate dehydrogenase mechanism. I. C. Gunsalus.

produce acetolactate; however, this reaction does not occur in most systems, as shall be seen in a moment. The active acetaldehyde can also be attacked by a proton to release acetaldehyde, and the acetaldehyde produced will be able to react in turn with the active intermediate to give acetoin. This reaction was known for a very long time, but it was originally assumed that it resulted from the action of a second enzyme called "carboligase." The active intermediate can also react with various electron acceptors, such as ferricyanide or 2-6-dichlorophenolindophenol, in which case the acetaldehyde residue will be simply oxidized to acetic acid. Finally, the last acceptor considered, the one which is physiologically active in mammalian metabolism and in most bacterial systems is lipoic acid. Lipoic acid is a fatty acid containing 8 carbon atoms with a disulfide bridge linking carbon 6 and 8. In the presence of lipoic acid, presumably, the reaction goes almost quantitatively to acetyllipoate, which is the second "active intermediate" in the pyruvic oxidase reaction. Oxidation of the aldehyde group takes place at the expense of lipoic acid which is reduced concurrently to the thiol form.

Reed (33) et al. showed that a mutant of *E. coli* contained lipoic acid in a conjugated form, which was identified as lipothiamide pyrophosphate (LTPP). The presence of LTPP could also be demonstrated in a few other cases. Interestingly, in the presence of lipoic acid, a steric hindrance prevents pyruvic acid or acetaldehyde from attacking the "active aldehyde," and therefore, the reaction is entirely shifted in the direction of the formation of acetyllipoic acid (33).

To be complete, here is how the whole reaction proceeds in the presence of lipoic acid and CoA (32, 33) as shown in figure 21. In the first reaction, pyruvate reacts with TPP to give
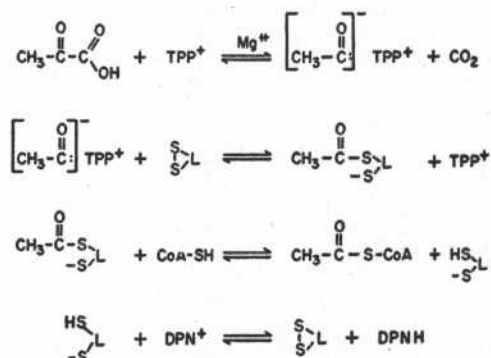
Figure 21. Group transfer functions of lipoic acid. I. C. Gunsalus—1954.

the active intermediate; this intermediate reacts with lipoic acid to give acetyl lipoate. In a third reaction, the acetyl group is transferred to CoA, the reduced lipoic acid being finally reoxidized to the disulfide form by DPN. Although two distinct reactions occur in steps 1 and 2, there are indications that they both are catalyzed by a single enzyme.

Thiamine pyrophosphate also participates in the formation of an "active glycolaldehyde," an intermediate in the transketolase reaction, as shown by Horecker (34) and Racker (35, 36) (cf. figure 22). Here again it was shown that
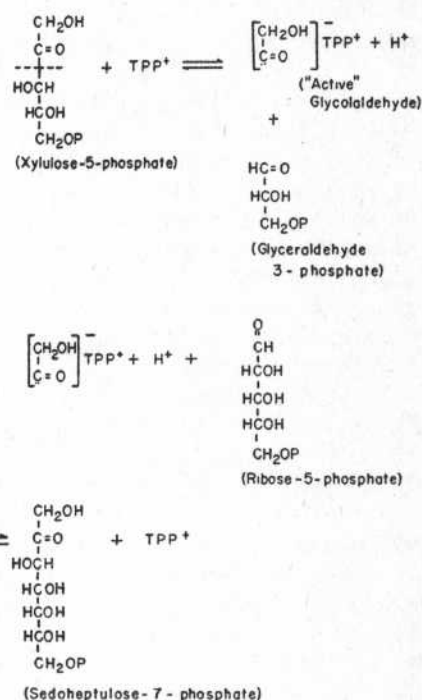
Figure 22. Formation of "active" glycolaldehyde in transketolase reaction. E. Racker—1955.

a number of compounds could react as donors or acceptors of the glycolaldehyde group in the reaction catalyzed by Racker's crystalline transketolase, as summarized in table 3.

Table 3. DONORS AND ACCEPTORS OF "ACTIVE" GLYCOLALDEHYDE

Donors	Acceptors
Xylulose-5-P	D, L, DL Glyceraldehyde-3-P
L-Erythrulose	Ribose-5-P
Sedoheptulose-7-P	Glycolaldehyde
Hydroxy pyruvate	D-Erythrose-4-P
Fructose-6-P	

Let us now consider the reaction catalyzed by triose phosphate dehydrogenase (TDH), one of the most extensively studied conjugated proteins. According to Racker (37), the reaction proceeds as in figure 23. In the first reaction, 3-phospho-
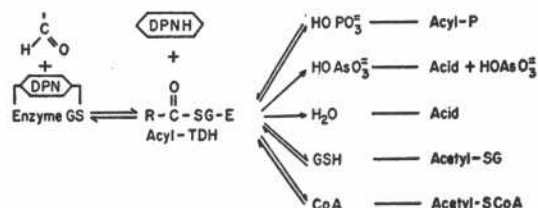
Figure 23. Mechanism of action of triose phosphate dehydrogenase (TDH). E. Racker—1955.

glyceraldehyde is oxidized at the expense of DPN which is reduced to DPNH. The reduced prosthetic group is displaced and an acyl-TDH intermediate is formed, stable in the absence of DPN. And again, this "active intermediate" will be able to react in a number of ways with a number of acceptors (38). It will react reversibly with inorganic phosphate as a phosphorylase or a phosphotransacetylase to give a high energy acyl-phosphate derivative (incidentally, it was the exchange of inorganic P^{32} with the acyl phosphate of 1, 3-diphosphoglyceric acid in the presence of TDH that pointed toward the existence of an acyl-TDH intermediate). Or it will react irreversibly as a hydrolase (acylase, phosphatase) in the presence of arsenate or in acid medium. Finally, it will react as a transacetylase in catalyzing an acyl transfer to glutathione or CoA. Indeed, as stated by Racker (38), the enzyme "looks more like a medusa head than a simple double-headed enzyme."

Recently, Vallee (39) has shown that muscle TDH, and yeast and horse liver alcohol dehydrogenases (ADH), (also pyridinoproteins) were in reality zinc metalloproteins. Indeed they found that yeast ADH, which contains 4 moles of DPN per mole of enzyme, also contains 4 gram-atoms of zinc per mole (40). TDH, which has a molecular weight of 137,000 and contains 3 DPN per mole (41), also possesses zinc (42). The possibility was advanced that all DPN containing enzymes might be zinc metalloproteins and that this metal might be directly involved in the binding of the nucleotide prosthetic group by the protein (42).

Let us now consider a third type of conjugated enzyme, namely the flavoproteins, which catalyze a reaction somewhat more complicated.

As is known, the flavoproteins are involved in the transfer of hydrogen or electrons from a reduced substrate, possibly DPNH, finally to oxygen, usually through the cytochrome system. Green and coworkers have shown that the linking of the electron scheme of the flavin nucleotides (which are two-electron acceptors) to the cytochromes or ferricyanide (which are one-electron acceptors) was mediated by a metal, according to figure 24. Under certain conditions, the metal can be re-
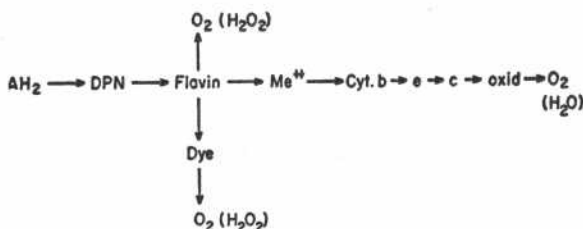
Figure 24. Role of flavoproteins in electron transport.

moved or lost, in which case we are left with a half knocked-out enzyme. This beaten up system will no longer be able to transfer the electrons to oxygen through the cytochromes, but will be able nevertheless to catalyze a two-electron transfer to oxygen, sometimes directly, more generally through the intermediate of certain dyes. In this case, H_2O_2 will be produced instead of water. At the present time, most of the flavoproteins have beeen shown to contain a specific metal ion, although the particular way in which the metal is bound to the protein is still unknown. Table 4 lists a few of these metallo-flavoproteins.

Table 4. PARTICIPATION OF METALS IN FLAVOPROTEINS

Enzyme	Flavin	Metal
"Green" Acyl CoA dehydrogenase (butyryl CoA dehydr.)	FAD	Cu^{+2}
"Yellow" Acyl CoA dehydrogenase	FAD	Fe^{+3}
DPN-cytochrome C reductase	FAD	Fe^{+3}
Xanthine oxidase	FAD	Mo^{+6}
Aldehyde oxidase	FAD	Mo^{+6}
TPNH-nitrate reductase	FAD	Mo^{+6}
Hydrogenase (Cl. pasteurianum)	FAD	Mo^{+6}

Butyryl-CoA dehydrogenase contains copper, which accounts for its green color; a "yellow" acyl-CoA dehydrogenase (concerned with the oxidation of higher fatty acids) and a DPN-cytochrome-C reductase, both contain iron; and finally we have four molybdenum containing enzymes for which the charge of the metal ion is still open to discussion.

There is fairly good evidence that the whole system starting with the flavin might be regarded as a single enzyme, or better as a single enzymatic unit. This view seems to be supported by several arguments. First of all, it is well known that if one mixes together the separate members of the system—the flavoprotein, the various cytochromes, cytochrome oxidase—the rate at which oxidation of a substrate proceeds will be extremely slow in comparison to that at which it would proceeed in the mitochondrion, or in what Green has called the "Electron Transport Particle" (43). Another argument rests in the fact that when such an electron transporting particle is isolated under mild conditions, it can be shown to possess little if any of the activity of one or the other of the individual components of the enzyme sequence (e. g. cytochrome C activity). If however, the particle is disrupted with a substance such as deoxycholate, then the various fractions obtained will display specific activities characteristic of a flavoprotein, cytochrome C, cytochrome oxidase, etc. Finally, crystalline yeast lactate dehydrogenase contains both a flavin (F) and a heme (H) group as depicted in figure 25.
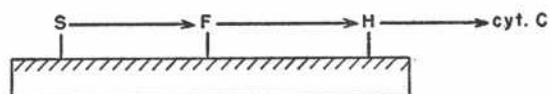
Figure 25. A diagram for yeast lactate dehydrogenase.

The enzyme therefore contains at least three active sites, carrying the substrate (S) through three successive reactions (44). These are all indications that the various catalytic activities of such systems must be linked together in a sort of organizational or physiological unit (43).

In passing, I would like to mention the very interesting work performed on the structure of the "active site" of cytochrome C. Tsou (45) showed that through peptic digestion, a peptide could be isolated from purified cytochrome C whose absorption spectrum was closely similar to that of the original enzyme. Tuppy and Paleus (46) determined that this peptide had the following structure:
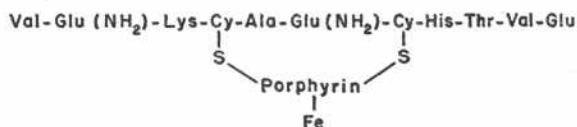
When a scale model of this hemopeptide was constructed (47) as a left-handed α-helix, the imidazol group from the histidine residue could form an unstrained bond with the iron, perpendicular to the plane of the heme, in confirmation of Theorell's (48) earlier hypothesis.

In a similar way, Singer et al. (49) after tryptic and chymotryptic digestion of purified beef heart succinic dehydrogenase, obtained several flavin peptides of related amino acid composition. As all the fractions contained 3 or 4 amino end groups, these authors concluded that the prosthetic group might be bound to peptides, and hence to the protein, through esteratic linkages involving the ribityl hydroxyl groups of riboflavin.

Before terminating this rapid survey of the active sites of enzymes, I would like to mention the fact that lately a number of nucleotides have been found to exist in a conjugated form, rendering likely the assumption that these compounds could play an active part as group transfer agents as shown in table 5. Naturally, that they do par-

Table 5. NUCLEOTIDES AS GROUP TRANSFER AGENTS

Nucleotide	Abbr.	Group carried
Adenosine monophosphate...	AMP	P or PP acyl amino acyl methyl sulfate carbonate
Guanosine monophosphate..	GMP	P or PP mannose-1-P
Uridine monophosphate.......	UMP	P or PP glucose-1-P galactose-1-P glucosamine-1-P
Cytidine monophosphate......	CMP	P or PP choline-P

ticipate in the transfer of phosphate or pyrophosphate groups was known for quite some time, but since the discovery of uridyl diphosphoglucose by Leloir and of S-adenosyl methionine by Cantoni, a great number of compounds of this type have been reported. There is little doubt that they represent some "active form" of the group involved, but it is not yet quite clear whether these active forms should be looked upon as obligatory intermediates in a normal sequence of reactions, or as products of side reactions. I heard a moment ago from Dr. Borsook that he will have very interesting things to tell us tomorrow on this particular matter.

There is a last group of coenzymes of which I have hardly spoken. These are simply metal

ions. It is a fact however, that a compound as ubiquitous as ATP for instance (each time a biochemist has a system which refuses to work, he begins by dumping in some ATP) is not only completely inactive in the absence of divalent metal ions but is, in most cases, a potent inhibitor for many enzyme systems. ATP becomes a very dangerous reactant in a metal free system. I believe that in many processes involving this nucleotide, phosphorylations, energy transfers, muscle contraction, etc., where people are so much concerned with what happens to the phosphate groups and with the fate of ATP, they should be a little more concerned with the role of the metals. Dr. Krebs and myself, at the University of Washington, have obtained good evidence in a study concerning the enzymatic activation of a muscle enzyme by ATP, that in the crude extract it is not the nucleotide which is the limiting factor, nor the enzymes, but only the concentration or the availability of the metal ions. It is unfortunately too late for me to speak to you of the mechanism by which phosphorylase, which is found in resting muscle in a physiologically inactive form, is converted to an active enzyme. Discussion of this mechanism might have provided an interesting example of how a metal ion alone can trigger a very complex chain of reactions.

I hope that in speaking of what is known today of the active sites of enzymes, I have succeeded in creating in your minds the utmost confusion. If so, I will be satisfied to believe that I have given you a rather truthful, though incomplete, account of our present knowledge of the structure of enzyme loci.

DISCUSSION

DR. VERNON H. CHELDELIN: I wonder if it would be appropriate at this point to request a comment regarding the peptidases reported by Binkley. Is it true that certain enzymes may not necessarily be proteins but that their activity may reside in low molecular weight compounds, resistant to heat and to proteolytic degradation?

DR. FISCHER: Well, this is a matter of semantics. I mean, what is an enzyme, after all? If you use hydrochloric acid as a catalyst to hydrolyze starch, it works perfectly well, provided you wait long enough or heat the solution. Now hydrochloric acid is not an enzyme, by definition. If we want to extend the term "enzymes" to various other catalysts of high or low molecular weight, it is quite all right with me, but we should first

agree on the definition of the term. With respect to Binkley's report, yes, I do believe in principle that polynucleotides may be vested with catalytic activity, and I think this may even be the basis of your discussion tomorrow, Dr. Borsook, won't it?

DR. HENRY BORSOOK: Well, this is a good example of how a biochemist drops dead, bound by a chain.

DR. FISCHER: I don't know what else I could add. I do believe, for instance, that if DNA or RNA can act as templates in protein synthesis, in controlling the polymerization of amino acids in specific sequences, that this undoubtedly constitutes a catalytic type of activity. So, if a molecule such as DNA can function in this way, I don't see why it could not function in various other ways, especially in light of what I have said a moment ago. I wonder if I have answered your question?

DR. CHELDELIN: Yes, I think so. I just asked for your comment. Thank you.

DR. HEINZ FRAENKEL-CONRAT: This is the same question to be a little more blunt. Has anybody heard of anyone's having ever been able to repeat any part of this type of work completely? Now, if I may make some remarks on your talk. First, when we have been dealing here with why so much protein was wasted on enzymes, I wonder whether we were not being a little bit too anthropomorphic about it. We assume the cell to be full of proteins, and that certain sites on these proteins, protoplasmic proteins, will develop enzymatic activties. But it does not mean that their sole function is to catalyze reactions. Your idea of structure, that they are really threaded along a chain, is a good one, but they may be at the same time bulk proteins; they may be proteins which in certain areas have become specialized to do certain things, just as adenosine triphosphatase in myosin.

DR. FISCHER: Like the PZ protein of Monod or something like that?

DR. FRAENKEL-CONRAT: Yes. So I think that just because we have put labels on those proteins, it does not mean that they are made for, and exist for, no other purpose than splitting a given substrate. And then, just for the sake of generating more confusion, Dr. Fischer—you tried to bring out the nonspecificity of enzymes. The same is also true for the hydrolytic enzymes where proteases turned out to be esterases. The enzyme needs the right substrate, first for its secondary linkage probably, and then for the

stripping and grinding of a half of it. It is the second decomposition reaction which is non-specific, isn't it?

DR. FISCHER: No, not quite. Look, for instance, at the reaction catalyzed by triose phosphate dehydrogenase: many reactions which were written on the right side of the slide were reversible, so that you can form your active enzyme-substrate intermediate complex (I call it thus, but in reality, only a portion of the substrate is involved) from acetyl phosphate, from acetyl glutathione or acetyl CoA. In addition, the formation of acyl-phosphate was demonstrated with many different aldehydes. So that indeed, you may have a great number of donors, provided they can donate the right group. The same is true for the transketolase reaction, where approximately four sugars have been shown to date to be able to act as donors in the formation of "active glycolaldehyde." And there is little doubt that many other compounds will be found capable of reacting in an analogous way.

DR. HOWARD S. MASON: I was reminded of another class of enzymes, really, that do the same things, and that would be the peroxidases. On the one hand, they react with hydrogen peroxide in a specific manner (at least the hydroperoxidases) and on the other hand, they react with a very wide variety of donors.

DR. EDMOND H. FISCHER: That's right.

REFERENCES

1. G. E. Perlmann, Nature, *173*, 406 (1954).
2. R. L. Hill and E. L. Smith, Biochim. et Biophys. Acta *19*, 376 (1956).
3. S. M. Kalmar, K. Linderstrom-Lang, M. Ottesen, F. M. Richards, Biochim. et Biophys. Acta *16*, 297 (1955); G. Kalnitsky and E. E. Anderson, ibid, *16*, 302 (1955).
4. H. Neurath and W. J. Dreyer, Disc. Faraday Soc. *20*, 32 (1955).
5. D. E. Green in "Pathways of Metabolism" *1*, 27 (D. M. Greenberg, ed.) Acad. Press, New York, 1954.
6. W. Langenbeck, Adv. Enzymol, *14*, 163 (1953).
7. T. A. Geissman, Quarterly Rev. Biol. *24*, 309 (1949).
8. Th. Bücher, Adv. Enzymol. *14*, 1 (1953).
9. Th. Bücher, Angew. Chem. *62*, 256 (1950).
10. D. Nachmansohn and I. B. Wilson, Adv. Enzymol. *12*, 259 (1951).
11. I. B. Wilson in "Mechanism of Enzyme Action" p. 642 (W. D. McElroy and B. Glass, eds.) The Johns Hopkins Press, Baltimore, 1954.
12. N. K. Schaffer, S. C. May, and W. H. Summerson, J. Biol. Chem. *206*, 201 (1954).
13. R. A. Oosterbaan, P. Kunst, and J. A. Cohen, Biochim. et Biophys. Acta *16*, 299 (1955).
14. G. Dixon, S. Go, and H. Neurath, Biochim. et Biophys. Acta *19*, 193 (1956).
15. T. Wagner-Jauregg and B. E. Hackley, J. Am. Chem. Soc. *75*, 2125 (1953).
16. I. B. Wilson, Disc. Faraday Soc. *20*, 119 (1955).
17. G. J. Jandorf, H. O. Michel, N. K. Schaffer, R. Egan, and W. H. Summerson, Disc. Faraday Soc. *20*, 134 (1955).
18. N. K. Schaffer, S. C. May, Jr., and W. H. Summerson, J. Biol. Chem. *202*, 67 (1953).
19. I. W. Sizer, J. Biol. Chem. *160*, 547 (1945).
20. I. B. Wilson and F. Bergmann, J. Biol. Chem. *186*, 683 (1950).
21. L. Weil and A. R. Buchert, Fed. Proc. *11*, 307 (1952).
22. A. K. Balls and E. L. Aldrich, Proc. Nat. Acad. Sci. U.S.A. *41*, 190 (1955).
23. E. R. Stadtman in "Mechanism of Enzyme Action" p. 581 (W. D. McElroy and B. Glass, eds.) The Johns Hopkins Press, Baltimore, 1954.
24. J. Larner and R. E. Gillespie, Arch. Biochim. Biophys. *58*, 252 (1955).
25. A. C. Walker and C. L. A. Schmidt, Arch. Biochem. *5*, 445 (1944).
26. L. Weil, A. R. Buchert, and J. Maher, Arch. Biochem. Biophys. *40*, 245 (1952).
27. L. Weil and T. S. Seibler, Arch. Biochem. Biophys. *54*, 368 (1955).
28. W. Z. Hassid and M. Doudoroff, Adv. Enzymol. *10*, 123 (1950).
29. P. Karrer and M. Viscontini, Helv. Chim. Acta *29*, 711 (1946); P. Karrer, Bull. Soc. Chim. *149*, 100 (1947).
30. S. Mizuhara and P. Handler, J. Am. Chem. Soc. *76*, 571 (1954).
31. F. Lynen and E. Reichert, Agnew. Chemie *63*, 47 (1951).
32. I. C. Gunsalus in "Mechanism of Enzyme Action" p. 545 (W. D. McElroy and B. Glass, eds.) The Johns Hopkins Press, Baltimore, 1954.
33. L. J. Reed, Physiol. Revs. *33*, 544 (1953).
34. B. L. Horecker, P. Z. Smyrniotis, and H. Klenow, J. Biol. Chem. *205*, 661 (1953).
35. E. Racker, G. de la Haba, and I. G. Leder, J. Am. Chem. Soc. *75*, 1010 (1953); Arch. Biochem. Biophys. *48*, 238 (1954).
36. E. Racker, et al., J. Biol. Chem. *214*, 409 (1955); Arch. Biochem. Biophys. *59*, 535 (1955).
37. I. Krimsky and E. Racker, Science *122*, 319 (1955).
38. E. Racker, Physiol. Revs. *35*, 1 (1955).
39. B. L. Vallee, Adv. Prot. Chem. *10*, 317 (1955).
40. B. L. Vallee and F. L. Hoch, Nat. Acad. Sci. *41*, 327 (1955); J. Am. Chem. Soc. *77*, 821 (1955).
41. J. B. Fox and W. B. Dandliker, J. Biol. Chem. *218*, 53 (1956), and *in press.*
42. B. L. Valee, reported at the 1956 Federation Meetings in Atlantic City, New Jersey.
43. D. E. Green in "Chemical Pathways of Metabolism" *1*, 27 (D. M. Greenberg, ed.) Acad. Press, New York, 1954.
44. Mentioned in M. Dixon, Disc. Faraday Soc. *20*, 9 (1955).
45. C. L. Tsou, Biochem. J. *49*, 367 (1951).
46. H. Tuppy and S. Paleus, Acta Chem. Scand. *9*, 353, 365 (1955).
47. A. Ehrenberg and H. Theorell, Acta Chem. Scand. *9*, 1193 (1955).
48. H. Theorell and A. Akesson, J. Am. Chem. Soc. *63*, 1804 (1941).
49. E. B. Kearney, V. Massey, and T. P. Singer, Fed. Proc. *15*, 286 (1956).

Chemistry and Biology of Glucagon

Otto K. Behrens

In this symposium Dr. Borsook has presented a stimulating picture of the signficance of proteins. Such a presentation draws on information of many kinds obtained through a variety of experimental techniques. Among these are the numerous published descriptions of purification of individual proteins, the recent development of methods for determination of structure of simpler proteins, and many studies on the biological activities of various proteins. I shall deal with work of this sort for a specific protein, glucagon, the hyperglycemic factor of the pancreas. Banting and Best were the first to prepare reproducible extracts of pancreas that contained insulin activity (1). Within a few years Murlin reported that some of these pancreatic extracts also contained a hyperglycemic substance as well as the hypoglycemic material, insulin. It was Murlin who proposed the name glucagon for this principal. Over the course of years, Burger in Germany maintained a continued interest in glucagon. He accomplished some purification and reported that the material appeared to have some physical properties similar to insulin. Comparison with insulin was made repeatedly by various investigators. Now that glucagon has been isolated in crystalline form a revaluation of such comparisons is of considerable interest.

Burger reported that glucagon had a glycogenolytic effect in the liver. Sutherland, in this country, performed extensive work; he found that the glycogenolytic effect on liver slices was observed in preparations in which the insulin had been inactivated with alkali. Sutherland and de-Duve studied the distribution of glucagon in tissues and found that it was present in the pancreas of all species studied. They also reported that glycogenolytic activity was present in the upper two-thirds of the intestinal mucosa of some species but not of others. The identity of this substance with glucagon was postulated but has not been proved by isolation. Glucagon was present in the islets of the alloxan diabetic animal. On the basis of this observation it was assumed that this substance could not be produced in the beta cell, the cell of origin of insulin. Most of the information now available indicates that glucagon is formed in the alpha cell of pancreas, although agreement on this point is not universal.

Our own interest in glucagon was stimulated by the possible relationship to insulin, both biologically and chemically, for glucagon tended to accompany insulin in purification procedures and indeed small amounts of the substance were present in insulin that was used for treatment of the diabetic. In the work on purification and structure, I wish particularly to acknowledge my collaborators: Dr. A. Staub, who exercised initiative on the isolation and undertook the early work on the structure; Mr. L. Sinn, who has been associated with the project throughout; and Dr. W. W. Bromer, who has brought the work on structure to successful completion.

The early work on purification encountered difficulties that tended to reinforce the idea that glucagon and insulin may be similar. Fractionation procedures generally caused precipitation of both substances. Eventually, sufficient information was accumulated concerning the properties of glucagon, and we were successful in devising methods that led to its purification and crystallization (2). The crystals belong to the isometric system and were recognized as rhombic dodecahedra. The isoelectric point of glucagon, approximately 8, indicated that its properties should be distinctly different from insulin, as the isoelectric point of insulin is approximately 5.3. An interesting similarity between the two proteins is found in the property of fibril formation in acid solution. Glucagon fibrils can be converted back to the dissociated protein in alkaline solution. As a matter of fact, recovery is easier than with insulin because of the greater alkali stability of glucagon.

Several methods of characterization have been applied to supply information on the purity of the crystallized product. Insulin was absent as measured biologically and on the basis of amino acid content, for certain amino acids that are present in insulin are absent from glucagon. Zone electrophoresis also provided evidence of the purity of the crystalline material. Separation into two fractions is apparent (figure 26), the larger one being approximately 90 per cent of the total substance and the smaller one about 10 per cent. At first we thought this was an indication of an impurity; however, when material from the main peak was separated and used as the starting

Table 10. THE ANALYSIS AND RECOVERY OF PEPTIDES FROM THE 50-HOUR TRYPTIC DIGESTION OF GLUCAGON

Peptide number	Peptide composition and analysis	Approx. percentage recovered*
LT-1	Leu (met, asp, thr) Moles 1.0 0.7 1.0 1.0	80
LT-2	Ala (glu₂, asp, phe, val, try) Moles 1.0 2.1 1.1 0.9 1.1 1.0	50
LT-3	Thr (ser₂, asp, tyr, lys) Moles 1.0 1.7 1.1 0.5 1.0	100
LT-4	Try (leu, asp, ser, arg) Moles 0.8 1.0 1.2 1.0 0.6	60
LT-5	His (ser, glu, gly, thr, phe) Moles 0.5 1.0 1.0 0.9 1.0 1.2	60
LT-6	Arginine	40

*Amount of peptide recovered/amount of peptide liberated by enzyme.

Again the peptides isolated from each of these digestion procedures account for all 29 amino acids in the molecule. The information obtained may be integrated with that previously presented from chymotrypsin to locate additional parts of the molecule.

A third proteolytic enzyme, subtilisin, provided important additional evidence. A total of

Table 11. ANALYSIS AND RECOVERY OF PEPTIDES FROM THE SUBTILISIN DIGESTION OF GLUCAGON

Peptide number	Peptide composition and analysis	Approximate percentage recovered*
S-1	Thr. ser Moles 0.8 1.2	50
S-2	Asp. thr Moles 1.1 1.0	90
S-3	Asp. (try, ser) Moles 1.0 1.0 1.1	60
S-4	Asp. phe Moles 1.1 1.0	50
S-5	Leu. meth Moles 1.1 0.9	60
S-6	Gly (thr, phe) Moles 1.1 0.9 1.0	70
S-7	Leu (asp, ser, arg) Moles 1.0 0.9 1.0 1.0	80
S-8A	His (ser, glu) Moles 0.5 1.0 1.3	30
S-8B	Val (glu, try) Moles 1.0 1.0 1.0	30
S-9	Arg (ala, glu) Moles 1.2 0.9 1.0	50
S-10	Lys. tyr Moles 1.0 1.0	70

*Amount of peptide recovered/amount of peptide liberated by enzyme.

10 peptide bonds were split; fortunately all 11 peptides were isolated and were characterized (table 11). Addition of this information to that from the other two enzymes helped to clarify many aspects of the structures and showed no inconsistencies.

A few parts of the molecule were still incompletely characterized. Partial acid degradation of peptide LT-4 followed by fractionation and characterization of the products produced the information presented in table 12. Similar work led to characterization of C-4 (table 13) and,

Table 12. PARTIAL ACID DEGRADATION OF PEPTIDE LT-4, TYR (LEU, ASP, SER, ARG)

Major Fragments Isolated

Tyr
Tyr (leu, asp)
Try. leu
    leu. asp
      asp. ser
        ser. arg

Sequence: Tyr. leu. asp. ser. arg

Table 13. PARTIAL ACID DEGRADATION OF PEPTIDE C-4, HIS (SER, GLU, GLY, THR, PHE)

Major Fragments Isolated

His
  ser (glu, gly)
    glu. gly
        thr. phe

His [ser. glu. gly] [thr. phe]

His (ser, glu) and gly (thr, phe) isolated from subtilism digest, therefore *sequence* is: } His. ser. glu. gly. thr. phe

Table 14. PARTIAL ACID DEGRADATION OF PEPTIDE LT-3, THR (SER, ASP, TYR, SER, LYS)

Major Fragments Isolated

Thr
  ser (asp, tyr)
  ser. asp
    asp. tyr
        ser. lys

Sequence: thr. ser. asp. tyr. ser. lys*

*Carboxypeptidase.

coupled with carboxypeptidase data, led to determination of the amino acid sequence of LT-3 (table 14). The elucidation of the sequence of peptide S-8B, val. glu. try. was accomplished by means of a time study with carboxypeptidase.

sis, and identification of the DNP-amino acids. Two DNP amino acids were found, ε-DNP-lysine and di-DNP-histidine; one mole of the latter per 3,500 mol. wt. was present. This finding led to the conclusion that histidine is the N-terminal amino acid, and the ε-amino group of lysine is free in the protein.

The Akabori hydrazinolysis procedure was applied to glucagon to identify the C-terminal amino acid (6). The method is based on the formation of hydrazides of all amino acids except the C-terminal one. Threonine was identified as the only C-terminal amino acid. Carboxypeptidase was also used to determine the amino acids that are located at the carboxyl end of the chain. This proteolytic enzyme rapidly released a number of amino acids from glucagon. Rate studies showed that the yield of threonine was 5 to 10 per cent greater than that of any other amino acid during the early stages of digestion. We concluded that these data are consistent with threonine being the carboxyl terminal amino acid. Upon prolonged digestion approximately one mole of each of the following amino acids was released: threonine, asparagine, methionine, leucine, tryptophane, glutamine, and valine. Smaller amounts of phenylalanine, aspartic acid, alanine, and a second glutamine were also released.

The next information was obtained by subjecting the molecule to splitting by certain proteolytic enzymes. Since these enzymes selectively split certain peptide bonds, their use led to the liberation of a few peptides. These protein fragments were separated on Dowex 50 chromatography columns and were characterized. Digestion with chymotrypsin caused splitting of 5 peptide bonds; the resulting 6 peptides were separated. By applying the DNP method, the N-terminal amino acids were identified. Additionally, the quantitative amino acid composition of each

peptide was determined. A summary of the chymotryptic information is presented in table 8.

Table 8. THE ANALYSIS AND RECOVERY OF PEPTIDES FROM THE CHYMOTRYPTIC DIGESTION OF GLUCAGON

Peptide number		Peptide composition and analysis	Approx. percentage recovered*
C-1	Leu (met, asp, thr) Moles 1.0 0.8 1.2 1.0		45
C-2	Thr (ser, asp, tyr) Moles 1.0 0.9 1.1 1.0		85
C-3	Val (glu, try) Moles 1.3 0.6 1.0		60
C-4	His (ser, glu, gly, thr, phe) Moles 0.8 1.0 1.1 1.1 1.1 1.0		55
C-5	Leu (asp₂, ser, arg₂, ala, glu, phe) Moles 0.9 1.9 1.0 1.8 1.0 1.0 0.8		50
C-6	Ser (lys, tyr) Moles 0.8 1.3 1.0		40–50

*Amount of peptide recovered/amount of peptide liberated by enzyme.

The order of the amino acids within the parenthesis is undetermined at this point. It is of interest to note that the 6 peptides contain 29 amino acid residues, the correct number and distribution according to the quantitative amino acid determinations. Since there is only one histidine in the molecule, and we know it is located at the N-terminal end of the molecule, we can locate the C-4 peptide. On the basis of the hydrazinolysis and carboxypeptidase data and the knowledge that only one methionine occurs in the molecule, the C-1 peptide may be located at the C-terminal end. Similarly, the tryptophane-containing peptide, C-3, must be adjacent to C-1, and C-5, which contains alanine, must be adjacent to C-3.

A similar study was performed with trypsin. Some bonds were split quite rapidly, others more slowly; consequently this enzyme was used with two different periods of digestion, 2¼ hours and 50 hours. Separation of the peptides and characterization as described for chymotrypsin led to the information presented in tables 9 and 10.

Table 9. THE ANALYSIS AND RECOVERY OF PEPTIDES FROM THE 2¼ HOUR TRYPTIC DIGESTION OF GLUCAGON

Peptide number		Peptide composition and analysis	Approximate percentage recovered*
ST-1	Tyr (leu, asp, ser, arg) Moles 1.0 1.0 1.2 1.0 0.8		70
ST-2	Arg		45
ST-3	His (ser₃, glu, gly, thr₂, phe, asp, tyr, lys) Moles 0.8 3.2 1.2 1.0 2.1 1.0 1.0 1.0 1.1		32
ST-P	Ala (glu₂, asp₂, phe, val, try, leu, met, thr) Moles 1.0 2.2 1.8 1.0 0.9 1.0 1.0 0.6 0.9		80

*Amount of peptide recovered/amount of peptide liberated by enzyme.
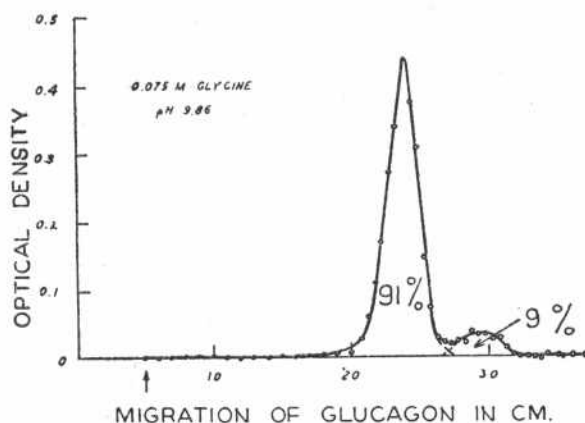
Figure 26. Zone electrophoresis of crystalline glucagon on starch. The arrow indicates the point of application of the sample. Optical density is obtained from quantitative tyrosine analysis.

Table 7. AMINO ACID COMPOSITION OF GLUCAGON AND INSULIN

Amino acid	Number of residues	
	Glucagon	Beef insulin
Lysine	1	1
Alanine	1	3
Histidine	1	2
Valine	1	5
Phenylalanine	2	3
Leucine	2	6
Glycine	1	4
Tyrosine	2	4
Threonine	3	1
Serine	4	3
Glutamic acid	3	7
Arginine	2	1
Aspartic acid	4	3
Methionine	1	
Tryptophane	1	
Proline		1
Cystine		3
Isoleucine		1
Totals	29	48

material for a new run, separation into two fractions of similar proportions are again observed. Furthermore, the small fraction contains the same amino acids as the large fraction; additionally, it is biologically active. So we have concluded that the two fractions probably are interconvertible, and the zone electrophoresis experiments provide evidence of the high purity of the protein. Other indications of the purity of the substance were obtained through determination of the N-terminal and C-terminal amino acids. Only one amino acid could be detected as the terminal group at each end of the molecule. Repeated crystallization failed to give an increase in specific activity, and the fraction that remained in the mother liquors was of the same high specific activity as that in the crystals.

The chemical analysis of the crystalline material gave values for carbon, hydrogen, nitrogen, and sulfur, typical of those to be expected for proteins (table 6). The value for sulfur

Table 6. CHEMICAL ANALYSIS OF GLUCAGON

C 50.11%	H 6.43%	N 17.45%	S 0.77%
amino N		amide N	
0.83% (2.1 groups)*		1.6% (4.0 groups)*	

*Calculated on the bases of a mol. wt. of 3500.

of 0.77 sharply differentiates the material from insulin as the sulfur content of insulin is about 3.2 per cent. Table 7 provides a comparison of

the quantitative amino acid composition of glucagon and insulin. Glucagon contains two amino acids, methionine and trytophan, that are absent from insulin. There are a considerable number that are common to the two proteins, and three amino acids are present in insulin and are not present in glucagon. This information again shows that there are major differences between the two substances. The quantitative amino acid composition of glucagon was determined by application of the method of Moore and Stein (3). Information obtained by the Levy method (4) of quantitative chromatography of dinitrophenyl (DNP) derivatives also was consistent with these results. Hydrolysates of the protein were prepared at various time intervals so that we could determine the amount of destruction of amino acids that was occurring and could correct for such destruction. There are a total of 29 amino acid residues in the molecule. This compares with a total of 51 for insulin. This information leads to a minimum molecular weight of approximately 3,500. Ultracentrifugation data and determinations of partial specific volume are consistent with a molecular weight in this range (5).

Having determined the amino acid composition, we were ready to investigate the arrangement of the amino acids in the protein. The identity of the N-terminal residue was determined by treatment with dinitrofluorobenzene, hydroly-

At this point the information available was sufficient to determine the correct order of the amino acids, but the location of amide groups on the dicarboxylic acids was not adequately characterized. Earlier work had shown the presence of 7 dicarboxylic amino acids, but only 4 amide groups. Thus, we may expect that the second carboxyl of 3 of the dicarboxylic acids must be free. From the series of peptides released by the proteolytic enzymes we chose a suitable representative to contain each of the dicarboxylic acids. Each of these peptides was acid digested and the released ammonia was determined. The data shown in table 15 provided information on the location of each of the amide groups.

Table 15. ANALYSIS OF SELECTED PEPTIDES FOR AMIDE LINKAGES

Peptide number	Composition	Moles of amide found per mole of peptide
C-3	NH₂ $\mid$ Val (glu, try)	1.1
C-4	NH₂ $\mid$ His (ser, glu, gly, thr, phe)	0.6
S-2	NH₂ $\mid$ Asp. thr	0.9
S-3	Asp (tyr, ser)	0.03
S-4	Asp. phe	0.03
S-7	Leu (asp, ser, arg)	0.0
S-9	NH₂ $\mid$ Arg (ala, glu)	.4–.5

All of the data presented may be used to solve the amino acid sequence of glucagon (table 16). Several points of interest may be emphasized. Although several methods of degradation and structure determination were utilized, no conflicting information was encountered at any point. The information obtained concerning glucagon as well as that obtained by Sanger and others for insulin is consistent with the interpretation that these small pancreatic proteins as obtained from one animal source represent specific molecular entities rather than related families of compounds of differing lengths or sequences.

Glucagon, in contrast to insulin or the pituitary hormones, oxytocin and vasopressin, does not contain a ring structure. Thus, a small peptide ring cannot be the sole structural characteristic associated with the biological activity of proteins.

Glucagon may also be contrasted with insulin in that the sequence of the amino acids in the structure does not repeat that of insulin in any significant portion of the molecule. Indeed no peptide sequence larger than a dipeptide is common to the two proteins.

Following the discovery of suitable purification methods for a substance, studies on the biological charcterization can be prosecuted. We have supplied material to many laboratories for such a biological characterization. I wish to emphasize the information that has been obtained in our laboratories largely through the initiative of Dr. Mary Root in the Department of Pharmacology. The presence of glucagon in some commercial preparations of insulin had been recognized for many years and had led to speculation regarding its effect on both the biological assay and the hypoglycemic action of insulin. This problem was studied in mice using solutions of glucagon free insulin and adding crystalline glucagon in amounts varying from 1.5 to 100 per cent by weight. Table 17 shows that there was no effect

Table 17. EFFECT OF CRYSTALLINE GLUCAGON ON MOUSE CONVULSION ASSAY OF INSULIN

Added glucagon	Glucagon-free insulin	Insulin and glucagon
1.5%	97/200*	102/190
3.5%	43/80	36/60
100.0%	403/728	401/728

*Number of convulsed mice/total number of mice injected.

on the mouse convulsion tests even when the ratio of insulin to glucagon was one to one (7). When the blood sugar response of mice treated in this way was determined, it was observed that the curves were different during the first 30 minutes. In the presence of 50 per cent by weight of glucagon, the blood sugar did not fall as rapidly as when insulin alone was injected, but the level reached at 30 minutes was the same. Apparently the glucagon hyperglycemia is so brief that it is without effect during the period between 30 and 90 minutes in which the convulsions occurred.

Similar results were obtained in rabbits. Upon addition of 1 or 10 per cent glucagon, the hyperglycemic effect was noted during the first 30 to 60 minutes but the maximum hypoglycemic response of the insulin was still observed.

The glucagon content of a number of samples of crystalline insulin has been determined on a quantitative basis (8). The method for measuring

Table 16. Amino Acid Sequence of Glucagon.

Peptide number	Hydrolytic agent	Structure of glucagon
ST--3	Trypsin (2.25 hr.)	His(ser,glu,gly,thr,phe,thr,ser,asp,tyr,ser,lys)
S--8A	Subtilisin	his(ser,glu)
C--4 / LT--5A	Chymotrypsin / Trypsin (50 hr.)	his.ser.glu(NH$_2$).gly.thr.phe
S--6	Subtilisin	gly(thr,phe)
LT--3	Trypsin (50 hr.)	thr.ser.asp.tyr.ser.lys
S--1	Subtilisin	thr.ser
C--2	Chymotrypsin	thr(ser,asp,tyr)
S--3	Subtilisin	asp(tyr,ser)
C--6	Chymotrypsin	ser(lys,tyr)
S--10	Subtilisin	lys.tyr
LT--1 / LT--4	Trypsin (2.25 and 50 hr.)	tyr.leu.asp.ser.arg
S--7	Subtilisin	leu(asp,ser,arg)
C--5	Chymotrypsin	leu(asp,ser,arg,arg,ala,glu,asp,phe)
ST--2 / LT--5B	Trypsin (2.25 and 50 hr.)	arg
S--9	Subtilisin	arg(ala,glu(NH$_2$))
ST--P	Trypsin	ala(glu,asp,phe,val,glu,try,leu,met,asp,thr)
LT--2	Trypsin (50 hr.)	ala(glu,asp,phe,val,glu,try)
S--4	Subtilisin	asp.phe
C--3 / S--8B	Chymotrypsin / Subtilisin	val.glu(NH$_2$).try
S--5	Subtilisin	leu.met
C--1 / LT--1	Chymotrypsin / Trypsin (50 hr.)	leu(met,asp,thr)
S--2	Subtilisin	asp(NH$_2$).thr
Amino acids from carboxypeptidase		asp(NH$_2$).thr
Summary		His.ser.glu(NH$_2$).gly.thr.phe.thr.ser.asp.tyr.ser.lys.tyr.leu.asp.ser.arg.arg.ala.glu(NH$_2$).try.leu.met.asp.thr

51

the biological activity is an adaptation of the procedure for observing the hyperglycemic response in cats described by Olsen and Klein in 1947 (9). Interference in the bioassay by the hypoglycemic action of insulin was avoided by incubating the samples with cysteine. The crystalline Zn-insulin preparations were found to contain 0.3 to 0.5 per cent of the factor. In view of the lack of effect of much larger quantities of glucagon on the insulin assay in mice and rats, it appears unlikely that these small quantities could significantly influence the clinical hypoglycemic response.

Additional studies were undertaken to determine whether chronic administration of glucagon to normal animals may have a diabetogenic or any other effect (10). Rabbits and rats were treated with glucagon daily for six months. The rats were injected intraperitoneally and the rabbits intravenously once a day with a dose of amorphous glucagon equivalent to 140 mcg. of the crystalline preparation. This is a relatively large dose for these animals. There was no indication of a permanent alteration of blood sugar, and no histological alteration of the pancreatic islets or other tissues was observed. The only effect that was observed was an increase in liver glycogen concentration as shown in figure 27.
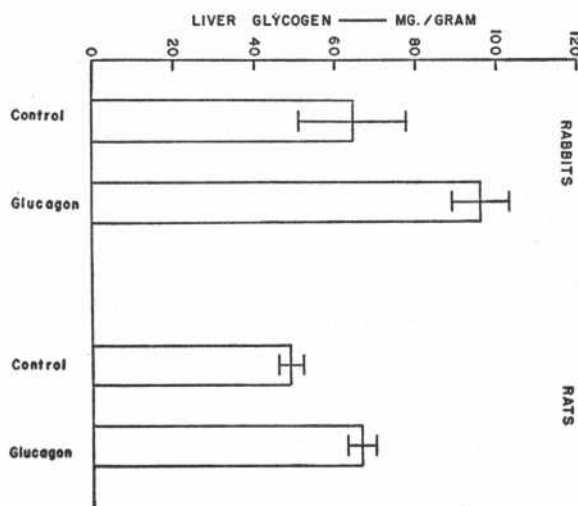
Figure 27. Effect of daily administration of glucagon (140 mcg.) for 6 months on the liver glycogen of rabbits and rats.

The increase was greater in rabbit liver than in rat liver but was statistically significant in both species.

Since glucagon stimulates increased glycogenolysis, it seemed likely that the increased liver

glycogen stores in animals that had received repeated doses of glucagon might also lead to an increased hyperglycemic response to a single dose of glucagon. Rabbits were injected subcutaneously twice daily with amorphous glucagon equivalent to 100 mcg. of crystalline product. Control animals received saline injections. Blood sugar responses to the morning injection were measured at various intervals. A typical response for one of the glucagon-treated animals is illustrated in figure 28. A
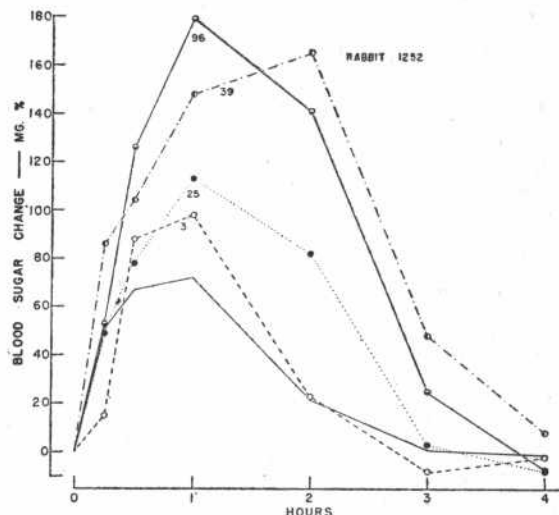
Figure 28. Blood sugar response to the administration of 100 mcg. of glucagon. The rabbit received 100 mcg. of glucagon twice daily. The number under each curve indicates the day of the experiment on which the blood sugar response to the morning dose was determined.

striking increase in the hyperglycemia is readily apparent.

Additional studies in rats were undertaken to investigate further this increase in liver glycogen concentration. To eliminate variations in food consumption as a factor in liver glycogen concentration, rats were force-fed twice daily with the medium carbohydrate diet prescribed by Ingle (11). The animals received twice daily a subcutaneous dose of amorphous glucagon equivalent to 100 mcg. of the crystalline substance. The results of the experiment are shown in figure 29. Within 3 days the liver glycogen concentration was increased several-fold. The effect was still greater at 7 days but did not continue to increase at 14 days. The number of rats in each group was small, usually 5 saline-controlled and 10 glucagon-treated animals.

In a similar experiment alloxan diabetic rats
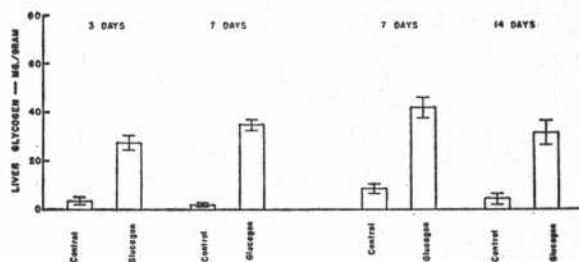
Figure 29. Liver glycogen of force-fed rats after twice-daily subcutaneous administration of 100 mcg. of glucagon.

were treated with glucagon or with insulin, and the liver glycogen was determined after three weeks. The response of the diabetic animals to glucagon was definite (figure 30), but not as
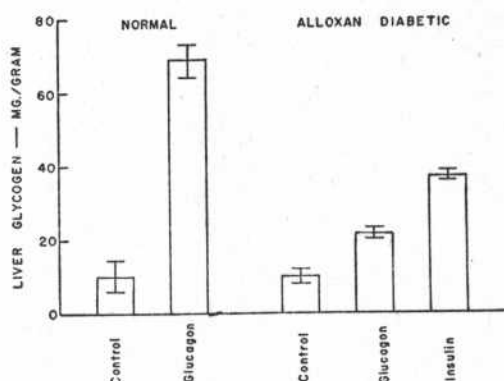
Figure 30. Liver glycogen of force-fed rats after twice-daily subcutaneous administration of 100 mcg. of glucagon or daily injection of insulin (PZI) for 3 weeks.

great as in normal animals. The diabetes in the glucagon-treated animals was not changed; urinary excretion of sugar was similar to the control group. Liver glycogen of the insulin-treated animals was increased; however, in these animals the urinary excretion of sugar was decreased markedly and the animals gained weight.

These data support the viewpoint that glucagon is not strictly an insulin antagonist. Its point of action in the carbohydrate metabolism scheme is quite different from that of insulin. The fact that it has no influence on the insulin assay, as well as the fact that it does not have a diabetogenic-type action in these animals is consistent with this interpretation. This view is not fully shared by all workers at the present time, and we may expect to see additional data brought to bear. A number of workers have presented evi-

dence which they interpret as supporting the viewpoint that glucagon may be an insulin synergist. Superficially considered, it seems rather peculiar to consider a substance that has a hyperglycemic action as synergistic to one that has a hypoglycemic action. The release of glucose from liver glycogen by the action of glucagon, however, makes additional glucose available for the peripheral action of insulin and in this manner may increase the effectiveness of insulin.

In conclusion we may say that glucagon is an interesting hormone-like substance. The physiological significance is still being explored. At the same time the possible importance of glucagon in protein structure work is evident, since it is one of the few proteins for which the amino acid sequence has been determined. The structural work has given no clue concerning the structural requirement for physiological activity. Possibly an understanding of the three dimensional structure, referred to by other participants in the symposium will be required before there is a basis for correlating structure and activity.

DISCUSSION

DR. CHOH HAO LI: First, I would like to congratulate Dr. Behrens and his coworkers for their beautiful work on glucagon. I wonder if he would care to comment on a few points: (1) From the slides just shown, it looks as if the specificity of trypsin does not hold true in this case. (2) Has any study been made to see whether or not the whole sequence of amino acids is required for the biological activity of glucagon? (3) From your data, glucagon apparently contains four amide bonds. Are they biologically essential?

DR. BEHRENS: We are aware of the apparent discrepancy between our observations of the bonds split by trypsin and the previous work on trypsin specificity. We are undertaking some work with trypsin inhibitors, and in addition are attempting a differential destruction of enzymes in hope that we can produce some evidence, but it is not available yet. In answer to your second question, some evidence is available but there is room for additional work. Glucagon is rapidly destroyed by trypsin and there seems to be little hope that a smaller active fragment can be prepared by use of this enzyme. Carboxypeptidase also appears to destroy activity; however, careful rate studies have not been performed, and it is possible that some amino acids can be removed with retention of activity. We have no informa-

tion on the importance of the amide groups for biological activity. Since there is no cystine in the molecule, glucagon is much more stable to alkali than most proteins. It can survive pH 10.5 or 11 for rather long periods of time. Possibly some of the more labile β-or γ-amids may be selectively hydrolyzed by careful application of alkaline conditions.

DR. HEINZ FRAENKEL-CONRAT: (1) Trypsin, even the crystalline preparations, is usually contaminated with chymotrypsin. Is it possible that the trypsin used in the glucagon studies was contaminated with a little chymotrypsin which could account for its attacking bonds not usually susceptible to trypsin? (2) No appreciable breakage of amide bonds has been noted by us at pH 11.5 to 12 with a variety of proteins, so that I doubt that such conditions will be useful for selective release of amide groups.

DR. BEHRENS: Chymotrypsin is certainly a common contaminant of trypsin. However, several peptide bonds that are rapidly split by chymotrypsin were not split by our trypsin sample. For this reason we haven't been convinced that a chymotrypsin contamination is an adequate explanation for our observations.

REFERENCES

1. The interested reader may be referred to one of the reviews on glucagon for more detailed information concerning earlier work and for the specific references, i.e., C. deDuve, Lancet 2, 99 (1953).
2. A. Staub, L. Sinn, and Otto K. Behrens, J. Biol. Chem. 214, 619 (1955).
3. S. Moore and W. H. Stein, (ibid.) 192, 663 (1951).
4. A. L. Levy, Nature 174, 126 (1954).
5. Determined by E. O. Davisson and H. W. Fisher, Physicochemical Research Division, Lilly Research Laboratories, Indianapolis.
6. S. Akabori, K. Ohno, and K. Narita, Bull. Chem. Soc. Japan 25, 214 (1952).
7. Mary A. Root, J. Ellis, and A. Staub, Proc. Soc. Exptl. Biol. Med. 85, 507 (1954).
8. A. Staub and Otto K. Behrens, J. Clin. Invest. 33, 1629 (1954).
9. Norman S. Olsen and J. Raymond Klein, Proc. Assoc. Research Nervous Mental Disease 26, 118 (1947).
10. Mary A. Root, Proc. Soc. Exptl. Biol. Med. 87, 108 (1954).
11. D. Ingle, J. Endocrin. 46, 505 (1950).

Recent Studies on ACTH and Hypophyseal Growth Hormone: A Brief Discourse

Choh Hao Li

The pituitary is a very complex structure, and it would be very difficult in 60 minutes or less to touch on all of the established hormones. The pituitary gland consists of two major portions: the anterior and posterior lobes, and a third portion, the very small intermediate lobe. I shall not go into any detail about the hormones of the posterior pituitary, except to note the beautiful work done recently by Professor du Vigneaud (1) and his coworkers on the isolation and synthesis of oxytocin and the vasopressins.

I should like to mention briefly the principle, known as intermedin, or the melanocyte-stimulating hormone, secreted by the intermediate lobe. This substance has very recently been isolated in pure form in our laboratory from porcine pituitaries. It is apparently a polypeptide containing 18 amino acids, with an isoelectric point at pH 5.7. I do not have the time here to go into any detail about our structural studies on this peptide hormone (2).

At least six different hormones are known to be secreted by the anterior lobe. These include the following: the thyrotropic hormone; the two gonadotropic hormones, the follicle-stimulating hormone and the interstitial cell-stimulating hormone; lactogenic hormone, or prolactin; ACTH, or adrenocorticotropic hormone; and growth hormone, or as it is alternatively called, somatotropin. All these substances are actually growth promoting in one way or another, generally with respect to a specific target organ; for example, thyrotropic hormone controls the growth and functioning of the thyroid gland. Growth hormone, or somatotropin, however, is a general growth-promoting substance, not only responsible for bone and tissue growth, but also affecting all the metabolic processes of the body, including carbohydrate, fat, and protein metabolism.

When a young animal, let us say 6 days of age, is hypophysectomized; that is, when the pituitary is removed, the growth of the animal will continue for a short time, but will stop when the animal reaches 30 days of age. After that, there will be some slight growth, but the animal will die at about 70 days of age. If it is hypo-physectomized at 13 days of age, the growth will again continue until the animal is about 30 days old, when it will cease. However, when the animal is hypophysectomized on the 28th day, the growth curve immediately levels off. The death rate of animals hypophysectomized at 6 days of age continues to mount steadily until the 74th day, when no survivors are left. The reason for these deaths is very interesting: it was found that even without a pituitary and with no treatment with growth hormone, the brain of the hypophysectomized animal continues to grow at the same rate as that of the normal animal. On the other hand, when the average cranial length of animals hypophysectomized at 28 days of age is compared with that of normal animals, it is evident that almost immediately upon operation the cranial growth ceases. Thus it is clear that the deaths are the result of brain damage, caused by the continued growth of the brain within a cranium the growth of which has been arrested; and indeed, the behavior of these animals shows signs of neuro-damage, such as convulsions and the like, with, as we have noted, a total mortality by the end of the 74th day. When growth hormone is administered to the hypophysectomized animals, however, their cranial growth continues at a normal rate, so that even when hypophysectomy is performed as early as 6 days of age, there are no deaths (3, 4). This influence of growth hormone on the skeleton represents only one of its functions.

ACTH is another substance which has an effect on general body growth, but in the opposite direction (5). The inhibiting effect that ACTH exercises on the rate of body growth, moreover, is counteracted by the simultaneous administration of growth hormone. In a typical experiment with young hypophysectomized rats (operated on at 28 days of age), the uninjected controls gained only 1 gm., whereas the animals treated with growth hormone, in a dosage of about 300 gamma per day given for 10 days beginning 2 weeks after operation, gained 29 gm. When α-corticotropin (ACTH from sheep glands) was administered, in a dosage of only 18 gamma, the animals lost 11 gm. When the two hormones were injected in combination, the two tendencies neutralized

each other, so that the weight increase approximated that of the controls with a net gain of 3 gm. These same effects are evident on the weight of the spleen.

From all the data available on these two hormones (5), it seemed reasonable to postulate that protein synthesis is inhibited by ACTH and stimulated by growth hormone. To clarify this concept, an experiment was carried out in which S35-labeled albumin was injected into animals which 3 weeks previously had received either growth hormone or ACTH, or a combination of the two. In the normal controls, the half-life of the S35-labeled albumin is 3.1 days, and the replacement rate, calculated on the basis of turnover time, is about 114 mg. per day. In the uninjected hypophysectomized controls, the replacement rate decreases to 60. When the hypophysectomized animals were treated with growth hormone, the replacement rate was restored from 60 to 164. The prior injection of ACTH alone did not produce any further inhibition of the replacement rate, but when the ACTH was administered in combination with growth hormone, the strongly inhibiting effects of the former nullified to a great extent the restoration that would be effected by the latter alone, so that with the combination the replacement rate reached only 84. These results (6) demonstrate that the process of protein synthesis is influenced in opposite directions by the injection of either growth or adrenocorticotropic hormones.

Another aspect of this problem of the inhibition of protein synthesis by ACTH and its enhancement by growth hormone is tumor growth. We know that the growth of a transplanted tumor in mice is inhibited by ACTH or cortisone; on the other hand, the tumor-bearing animal will die, even though the tumor growth is practically completely inhibited, because of the "toxic" effect of the ACTH or cortisone. It was also of interest to note that the injection of the ACTH or cortisone inhibited not only tumor growth but body growth as well. We thought that if the mutually antagonistic effects of growth hormone and ACTH in these respects could be utilized to balance each other, the "toxic" effect of the ACTH or cortisone might be counteracted without eliminating the beneficial effects in connection with tumor inhibition. Consequently, an experiment (7) was carried out, in which mammary adenocarcinoma were transplanted to C3H mice. In the controls, at the end of a 34-day period the tumor had grown to 6 gm. When ACTH

was injected, the tumor weight was only 0.1 gm., but the ultimate survival rate was very low. Growth hormone alone produced only insignificant alteration of the rate of tumor growth, but the survival rate with that hormone was much higher than with ACTH. When the two hormones were combined, the survival was almost complete, and the tumor grew only to half a gram.

The opposition of effects between these two hormones has been investigated further with another biological common denominator, their effect on mortality in rats that had been challenged with highly toxic doses ($4 \times LD_{50}$) of the organism *Pasteurella pestis* (8). At the end of a twenty-four hour period, the mortality of the controls was 27%; ACTH or cortisone, as you well know, enhances the susceptibility of the animal to infection, so that with ACTH the mortality at the end of the same period was as high as 73%. Growth hormone alone reduced the mortality to 8% and, injected along with ACTH, counteracted the effect of the latter, so that the mortality was only 23%. This then furnishes another example of the counteraction operative between growth hormone and ACTH.

An extension of this observation is furnished by the effect of the two hormones on antibody formation (8). It is known that ACTH depresses the formation of serum antibodies against a specific antigen. An experiment was carried out to see whether the anabolic effect of growth hormone would counteract this depression of antibody levels if the two hormones were administered together. ACTH alone, growth hormone alone, and the two hormones in combination, were administered to normal female rats for a 2-week period, and then a single dose of an antigenic substance known as Fraction 1A, a protein antigen extracted from the *Pasteurella pestis* organism, was given. It was found that a marked depression of the antibody level was effected by ACTH, with a mean titer of 1:21. Growth hormone alone gave a slightly elevated mean titer over the controls (1:425 compared with 1:150), although this does not represent a highly significant elevation. The interesting result, however, is that the marked depression of antibody levels caused by ACTH was effectively counteracted by the simultaneous administration of growth hormone, so that the average titer of the group given the combination was 1:84, a value not significantly different from that observed for the controls.

All these biological studies have given an idea

of the very complex interrelationship between these two pituitary hormones in body metabolism. Now I would like to talk a bit about structural studies. The growth hormone preparation that we have isolated has been shown by various purity tests to behave as a homogeneous protein; it has a molecular weight of 45,000 and a total amino acid content of 396. We have reported the following empirical formula (9) for the hormone:

$$Glu_{50}Asp_{35}(Cys-)_8Ser_{22}Thr_{26}Gly_{20}Ala_{31}Pro_{14}$$

$$Val_{14}Met_7(Leu + ILeu)_{76}Phe_{27}Tyr_{11}$$

$$Lys_{23}His_7Arg_{26}Try_3(-NH_2)_{30}$$
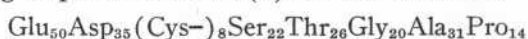
Several years ago, on the basis of the analytical data that had been obtained up to that time, we postulated that there were three possible structures for the growth hormone molecule. Phenylalanine and alanine had been established by studies with the fluorodinitrobenzene and phenyl*iso*thiocyanate procedures as the N-terminal residues of the molecule. It seemed likely that the molecule would be composed of 2 peptide chains, with 2 C-terminal residues as well. When the hormone was allowed to react with carboxypeptidase, phenylalanine (along with smaller amounts of leucine and alanine) was liberated completely, in an amount corresponding to two residues per mole of growth hormone (10). Thus, according to one possible structure, phenylalanine could occur as a C-terminal residue in each of two constituent peptide chains. A second possibility would be that the two phenylalanine residues might occur in the sequence Phe.Phe at the C-terminus of one of the chains, with the other C-terminal residue an unknown amino acid highly resistant to the action of carboxypeptidase. It was not clear from these studies, however, whether there actually were two C-terminal residues or only one. If either of these two structures were the true one, moreover, then oxidation of the -S-S- bonds should give rise to two fragments, whereas oxidation with performic acid still yields a single component.

Two chemical methods for C-terminal analysis—reduction of the esterfied protein with lithium aluminum hydride, and hydrazinolysis—which have subsequently been applied to growth hormone, have both given evidence that the molecule has only one C-terminal residue, identified as phenylalanine. Hence, at this point we favor the third possibility for the structure of the growth hormone molecule: a branched or Y-shaped polypeptide chain with two N-terminal residues (Phe. and Ala. and only *one* C-terminal residue (Phe.):

H Phe.——CONHCHCONH——Phe. Phe OH
|
(CH₂)₄
|
(S-S)₄   NH
|
CO
|
H Ala.———CONHCHR
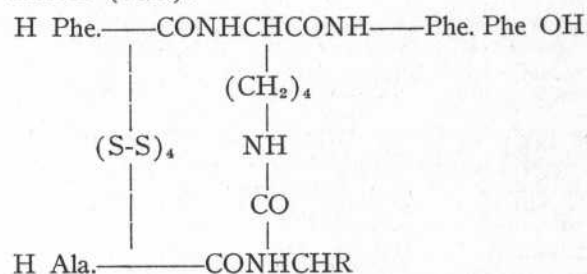
This structure requires the formation of a peptide bond by which the ε-NH₂ group of lysine would be joined with the α-COOH group of an amino acid residue. From the results of C-terminal residue analysis by the two procedures mentioned above, it appears unlikely that the γ-COOH group of glutamic acid or the β-COOH group of aspartic acid is involved in the peptide linkage through the ε-NH₂ group of lysine. Although there is no evidence in the literature for the existence of branched chains from the ε-NH₂ group of lysine, the possibility of such a bond cannot be discounted. (I was very pleased this morning when Dr. Borsook mentioned having found in his laboratory that the epsilon-amino group of lysine is capable of forming peptide bonds in naturally-occurring substances; this is evidence that lends support to our theory.) Particularly since the molecule has as many as 23 lysine residues, there seems a high probability that one of these would form a branched chain that would give a Y-shaped molecular structure, as postulated. Another factor in favor of this structure is the difficulty we have been having during the past few years in reducing the cystine residues in the hormone protein. This would be in accord with the necessity for rigidity on the part of the disulfide bonds in this theoretical arrangement.

Another extremely interesting observation was made in connection with the investigations on the "active core" in this protein. It had been found in earlier studies on the effect of enzymes that, although when growth hormone was submitted to extensive digestion with trypsin or pepsin, its biological activity was destroyed; after limited hydrolysis with chymotrypsin there is no loss of potency. We carried out some further experiments in which growth hormone was digested with chymotrypsin, in a ratio (by weight) of 1 : 300, at pH 9.5 and a temperature of 25°, for

various intervals. It is evident from the bioassay data, based on the increment of tibia width, that there is no loss of activity until the extent of digestion exceeds 30%. This means that the protein molecule can be degraded to a smaller unit, a "core" which is still active although it possesses different physiochemical properties from the original molecule. This active core can be isolated by either TCA precipitation or dialysis; the components which are nondialyzable and insoluble in 5% TCA contain the growth-promoting activity (11). We believe that the core contains no undigested protein hormone; also, according to all evidence, it is more acidic than the native growth hormone. The results of these enzymatic studies indicate to us that the activity must reside in a portion of the molecule, and that there must be some parts of the molecule that are not essential to the activity. It is most likely that the inessential parts are either the C-terminal or N-terminal portions, or both. Although we have been unable to remove the N-terminal residues in a stepwise manner without causing radical changes in the other amino acid groups, there is evidence that the N-terminal amino acid is not essential for the activity. Although this is as yet entirely hypothetical, I visualize the core as the central area, compact and resistant to enzymatic action, which contains all the -S-S- bridges; this would be the area entirely responsible for the growth-promoting activity.

I will talk only very briefly about ACTH. Active peptides have been isolated in pure form from glands of two different species by three laboratories. White and his colleagues (12) of the Armour Laboratories, and Bell (13) and coworkers of the American Cyanamid Company, have described the preparation of corticotropins from unhydrolyzed extracts of pig glands; the former preparation was designated Corticotropin-A, and the latter, β-corticotropin. An active polypeptide has been isolated in this laboratory (14) from sheep glands, also unhydrolyzed; it has been designated α-corticotropin, to differentiate it from the porcine preparations. This polypeptide has been found to have an isoelectric point at pH 6.6 and a molecular weight of 4,541, and it is made up of 39 amino acids. Although the corticotropins derived from glands of the two species are similar in many respects, they are not chemically identical and they behave very differently in countercurrent distribution.

When α-corticotropin was isolated several years ago, we embarked on investigations of the N- and C-terminal sequences by means of various stepwise degradation procedures. Reaction of the hormone with carboxypeptidase released three amino acids, phenylalanine, glutamic acid, and leucine, successively from the C-terminus. No further digestion took place after the third step, suggesting that the fourth amino acid might be proline, which is resistant to the action of the enzyme. This was confirmed when fragments, obtained by peptic digestion, with this same sequence were identified chromatographically. Thus, the C-terminal sequence was established as . . . Pro.Leu.Glu.Phe. The N-terminal sequence was elucidated by means of the paper-strip modification of the phenyl*iso*thiocyanate procedure. We were able to identify 7 successive amino acids by this method: Ser.Tyr.Ser.Met.Glu.His.Phe. Then, by fitting together overlapping sequences of various peptide fragments obtained by tryptic, chymotryptic, and partial acid hydrolysis of the protein molecule, we were able to propose a complete sequence of 39 amino acids for α-corticotropin, as follows:

Ser.Tyr.Ser.Met.Glu.His.Phe.Arg.Try.Gly.Lys.
1   2   3   4   5   6   7   8   9   10 11

Pro.Val.Gly.Lys.Lys.Arg.Arg.Pro.Val.Lys.
12 13 14 15 16 17 18 19 20 21

Val.Tyr.Pro.Ala.Gly.Glu.Asp.Asp.Glu.
22 23 24 25 26 27 28 29 30

Ala.Ser.Glu.Ala.Phe.Pro.Leu.Glu.Phe.
31 32 33 34 35 36 37 38 39

When the α-corticotropin molecule is submitted to enzymatic digestion with pepsin for a short time, within a few minutes there occurs a change in its physicochemical nature, but no amino acids are split off. When digestion proceeds further, then C-terminal amino acids begin to be cleaved. It is very interesting that the C-terminal amino acids in the corticotropins are not essential for the adrenal-stimulating activity. Limited hydrolysis with pepsin of both α-corticotropin (15) and corticotropin-A releases 11 amino acid residues from the C-terminus with no loss of activity. If, however, peptic digestion is continued for a longer period, around 24 hours, then amino acids are split off from the N-terminus, and after the fifth bond has been broken the activity is lost. I would like to mention something about our terminology in this connection. We call the unhydrolyzed polypeptide α-corticotropin; after it has been submitted to hydrolysis, we use

the designation β-corticotropin. β₁-corticotropin represents α-corticotropin minus the 11 amino acids split off by the action of pepsin. We also use the designations β₂-, β₃-, etc., according to the number of amino acids that have been split off. It has been indicated that β₁-corticotropin can be further hydrolyzed to split off a few more amino acids without loss of activity.

The chief biological properties of α-corticotropin can be outlined as follows:

1 Depletion of ascorbic acid in the adrenal of hypophysectomized rats.
2 Stimulation of adrenal weight in both normal and hypophysectomized animals.
3 Repair of histological structure in the adrenal of hypophysectomized rats.
4 Lowering of the level of circulating eosinophils in hypophysectomized rats one day after operation.
5 Reduction of the weight of thymus and spleen.
6 Inhibition of body growth and antagonism to the action of growth hormone.
7 Mobilization of fat into the liver of fasted animals.
8 Correction of the anemia that occurs after hypophysectomy; erythropoietic activity.
9 Melanocyte-stimulating activity in frogs.

These hormonal studies in all their diverse aspects, both chemical and biological, raise questions about the actual nature of the hormone as it occurs in the body, in relation to the substance that is extracted in the laboratory. It may be that the hormone as it is obtained in the laboratory is not the same thing as the hormone in action in the body, and that the chemists, when they extract a hormone from the body, may be getting an artificial form of the substance. The body itself is probably a better chemist than we are, so that it may start with a larger molecule than the one we derive, and it may be able to degrade it to a smaller unit as the circulating or "actual" hormone. For example, the molecular weight of growth hormone is around 45,000, but inside the cell it might be larger, and in circulation it might be much smaller, than we can demonstrate (11). The body as chemist has, after all, the enzymes that are specific for the degradation of a particular hormone, and we as yet have no knowledge of precisely what enzymes it uses for its degradation "procedures."

DISCUSSION

DR. HUGO KRUEGER: I would like to ask Dr. Li about some things he mentioned when he added pepsin to his hormone α-corticotropin, and observed a change in physicochemical properties. Were any of the amino acids split off? What was the nature of the change in physicochemical properties?

DR. LI: Digestion with pepsin changes the chromatographic behavior of α-corticotropin in Amberlite IRC-50 resin columns, although there is no indication that the enzyme has split off any amino acids. Apparently the physicochemical behavior is altered with no concomitant changes in chemical structure. The same thing is observed with growth hormone; its electrophoretic behavior is altered by limited enzymatic action which does not cause any hydrolysis of peptide linkages in the hormone protein.

QUESTION: I wonder if Dr. Li would comment on whether he feels there is a separate factor with erythropoietic activity emanating from the pituitary different from ACTH.

DR. LI: I would like to refer this question to Dr. Borsook.

DR. HENRY BORSOOK: I was going to ask the same question. We have been in fairly close touch by correspondence and telephone with the group at Berkeley. The Berkeley group have isolated from the pituitary a substance which, when injected into hypothosectomized rats, corrects the anemia which is characteristic of hypothosectomy. We have got out a crude fraction from the blood of rabbits, made anemic with phenylhydrazene, which when injected to normal rats at sea level makes them behave as if they were up 18,000 feet, as far as their red blood cells and hemoglobin are concerned. Naturally both the Berkeley group and we are interested as to whether we are getting out the same substance or not. The Berkeley group had been concerned whether the substance they had got out of the pituitary was or was not ACTH. It did not look as if our stuff could be ACTH. They tested it for ACTH and found too little to account for the erythropoietic activity. The last time I saw them, and this is what I wish to ask you, Dr. Li: You no doubt have later information than I have, whether they felt satisfied that they had found a method which destroyed the ACTH in their extracts and left the erythropoietic activity. The question I wish to ask you is what is the status of this work now? So that I will not have to get up again I will add some small comments. It is possible, though it seems unlikely, that the substance we were working with and the substance the Berkeley group got out of the pituitary are quite unrelated. It may be that they are related as was indicated by Dr. Li's remarks. He told us how one can degrade a hormone and still retain much of the activity in one fragment. What we may have is an effective

fragment of the pituitary erythopoietic factor. Another possibility is that different hormonal activities reside in different parts of the same molecule. Activity in different organs may reflect cleavage in different ways.

DR. LI: Dr. Borsook, this is a somewhat embarassing question, since I can not really give you any information about the nature of this factor that my colleagues in the Institute for Experimental Biology have been studying. I do know, however, that α-corticotropin does possess erythropoietic activity.

DR. HUGO KRUEGER: I like to sometimes oversimplify some of these complexities and try to give most statements a biological significance, and in this respect I very often get into difficulties in terms of dealing with nonreality. But then at least I am dealing with nonrealities which I can understand, and that is sometimes a little satisfying. Thus, the fact that the isoelectric constant is 6.6, suggests some rather interesting things to me, and I would like to have Dr. Li comment on them. On general principles I would estimate on the basis of the bicarbonate content and CO_2 of tissues and of the blood that the average pH throughout the body may be about 7.0. Now in certain areas you know definitely that the pH has some other value; and particularly wherever epinephrine or acetylcholine is liberated, you could get rather rapid changes in pH, and with the proper enzymes you could in time knock off certain amino acids. I wonder if when you knock off certain amino acids from this protein, there is much change in the isoelectric point? And then I would wonder if you would not automatically have a partial mechanism by which the distribution of these different fractions could be so directed that molecules of different size would get to particular tissues, and in particular, since the first compound that would be liberated would be the longer one, that particular compound might be sent to the cells which would need the large molecules to direct them. There is implied that binding of the hormone to cells depends upon cellular pH and upon the isoelectric point of the remaining molecule. All this is very highly theoretical and also highly simplified, but I would like to have Dr. Li comment if he would care.

DR. LI: Yes, when α-corticotropin is converted to β_1-corticotropin by peptic digestion, the isoelectric point also changes from 6.6 to a value of about pH 9. Apparently the enzyme removes some acidic residues from the parent molecule, causing the product to be more basic.

DR. KRUEGER: I feel that a change of that magnitude might aid in selecting certain areas in the blood vascular system from whence these proteins could move (I won't say diffuse).

REFERENCES

1. V. du Vigneaud, Science, 123, 3205 (1956).
2. I. I. Geschwind, C. H. Li, and L. Barnafi. J. Am. Chem., Soc., 78, 4494 (1956).
3. D. G. Walker, C. W. Asling, M. E. Simpson, C. H. Li, and H. M. Evans. Anat., Rec., 114, 19 (1952).
4. C. W. Asling, D. G. Walker, M. E. Simpson, C. H. Li, and H. M. Evans. Anat. Rec., 114, 49 (1952).
5. C. H. Li. Harvey Lectures, Series XLVI, 1950-51. Springfield, C. C. Thomas, 1952, P. 181.
6. F. Ulrich, H. Tarver, and C. H. Li. J. Biol. Chem., 209, 117 (1954).
7. L. L. Sparks, T. A. Daane, T. Hayashida, R. D. Cole, W. R. Lyons, and C. H. Li. Cancer, 8, 271 (1955).
8. T. Hayashida and C. H. Li. J. Exper. Med., in press.
9. C. H. Li and D. Chung. J. Biol. Chem., 218, 33 (1956).
10. J. I. Harris, C. H. Li, P. G. Condliffe, and N. G. Pon. J. Biol. Chem. 209, 133 (1954).
11. C. H. Li, H. Papkoff, P. Fønss-Bech, and P. G. Condliffe. J. Biol. Chem. 218, 41 (1956).
12. W. F. White and W. A. Landmann. J. Am. Chem. Soc., 77, 1711 (1955).
13. P. H. Bell. J. Am. Chem. Soc., 76, 5564 (1954).
14. C. H. Li, I. I. Geschwind, A. L. Levy, J. I. Harris, J. S. Dixon, N. G. Pon, and J. O. Porath. Nature, 173, 251 (1954); C. H. Li, I. I. Geschwind, J. S. Dixon, A. L. Levy, and J. I. Harris. J. Biol. Chem., 213, 171 (1955); C. H. Li, I. I. Geschwind, R. D. Cole, I. D. Raacke, J. I. Harris, and J. S. Dixon. Nature, 176, 687 (1955).
15. R. D. Cole, C. H. Li, J. I. Harris, and N. G. Pon. J. Biol. Chem., 219, 903 (1956).

Transforming Factors and Bacterial Genetics

Rollin D. Hotchkiss

I am going to talk about genetic experiments with a kind of organism that has not entered into this colloquium so far, the bacteria. Only a few years ago it did not seem very reasonable to do genetics with bacteria. Around the corner here, in the library, I saw a great many books giving the herd registers for Ayrshire cows, and Guernseys, Holsteins, and so on, and was really impressed as I never had been before with the amount of data you have to accumulate to give a pedigree for a sexually breeding organism. In higher forms you are interested in the individuals and how they come to have the properties they do, and you have to consider their genetic background individual by individual. It is altogether different with bacteria. A bacteriologist just puts a label on his culture, he lets it grow overnight, and then probably will leave it in the icebox for many weeks. He has a certain amount of faith that they will always come out about the same after many weeks in the icebox or indeed after many weeks of subculture over and over again, growing steadily in the laboratory, dividing every half hour or so. It is really because of this that there can be the science of bacteriology. Any bacteria that do not behave a few weeks later the same as when first observed are not considered to be very orthodox bacteria, are not given a proper name, and are generally left outside the science. The idea that genetics can be done with these creatures began to appear when certain experiments—the recombinations, transformations, and transductions—began to be done in the last ten years, and their recent popularity may be one reason why farm boys are leaving the herds behind and going on to work in bacterial genetics. I think it my task this afternoon, however, to tell you how difficult this can get if you get a little bit inquisitive, and perhaps I can help to keep some of the boys back on the farm.

The transformation phenomenon is the modification of the inheritable properties of one bacterial strain through the action of an extract from related but different organisms. I shall anticipate somewhat and say that it is genetic material that is transferred, that this genetic material chemically is deoxyribonucleic acid, and that transformation amounts to a partial genetic cross.

CLASSICAL BACTERIAL TRANSFORMATION

Transformation was first accomplished by Griffith in some animal experiments with pneumococcus about 25 years ago, and it was worked through by Dawson and by Alloway into a test tube phenomenon. Finally in the classic paper of Avery, MacLeod, and McCarty published in 1944, it was shown that the material extracted from encapsulated pneumococci which made unencapsulated strains become partly encapsulated was deoxyribonucleic acid. It was a dramatic fact that isolated material in the extract creating what seemed to be genetic changes could be described as a deoxyribonucleic acid.

Bear in mind that the usual behavior of pneumococci is just to make more of the same, and this goes on until there are millions of organisms, and you very seldom find an odd one, a mutant. When you start with an encapsulated organism, the kind you find in infections in man or animals, grow it to give more of the same, then extract the "DNA," the deoxyribonucleic acid from them, you have a "transforming agent." If you treat unencapsulated organisms with this material under proper conditions, a few per cent of them will become encapsulated. To recognize these you have to do something to get them away from the untransformed part of the population— for example, spread them out on an agar surface until you have them separated, or you can grow them in an antibody medium which will help you collect them at the bottom of a liquid culture. The latter procedure was used in the classic demonstration with type III encapsulated organisms by Avery, MacLeod, and McCarty (1). They also showed that it was highly specific in that, if DNA was extracted from type II organisms, the transformants were always of type II, while DNA from type I cells induced only type I organisms. They brought out for the first time that DNA, deoxyribonucleic acid, must be present in nature in a wide variety of forms, capable of causing specific inheritable effects. Now, the inheritable part is that any one of these organisms can be separated out and will grow to produce a culture true to type—encapsulated type III or type I, etc., each making a special polysaccharide all having the same capsule, and furthermore, all making a

DNA with which you can repeat this process on more unencapsulated organisms.

This classic demonstration of type transformation in pneumococcus has been followed by a number of other examples. There were first of all many demonstrations of various capsule type changes in pneumococcus, a few appeared in *Hemophilus influenzae* in the work of Alexander and Leidy at Columbia University, and certain other antigenic surface factors were shown to be changed. We introduced about 1951 a series of drug resistances which show that characters of an entirely different kind can be transferred, and recently there have been transformations of properties which may in a more orthodox sense be called changes in enzymatic properties. All of the transformations of course are attributable to biochemical metabolic changes in the organism and are probably basically enzymatic.

There is another interesting manifestation of heredity here. With the exception of the work on *Hemophilus* from Alexander and Leidy, which was inherited through the literature, almost all of this work was inherited through a very special chain stemming from association with the late Dr. O. T. Avery. The classical capsule work (1) was done in his laboratory, and was carried further in MacLeod's laboratory, partly by Austrian who has continued elsewhere work of this kind. Dr. Ephrussi-Taylor and your speaker who have looked at other aspects of the problem, both also began their work with Dr. Avery. Much of his motivation and a great deal of know-how was all brought through even to a third scientific generation by some kind of inheritance from the stimulus of that great man.

CHEMICAL NATURE OF ACTIVE AGENT

What about the nature of the agent? We have called it a deoxyribonucleic acid. The striking demonstration that Avery, MacLeod, and McCarty made was that the material was specifically inactivated by the enzyme, deoxyribonuclease. The enzyme was only then being worked on, and was not available in pure form, but they showed rather conclusively that inactivation was a specific breakdown that could only be obtained when the deoxyribonuclease was active in the material. All kinds of other enzymes such as proteases, ribonucleases, which damage proteins, and ribonucleic acids, did not affect the biological activity of the crude extracts or of the purified transforming agent. Furthermore, the purified material behaved like a deoxyribonucleic acid. In

some work begun with Dr. Avery, we wanted in particular to have a little more data showing that it was an orthodox deoxyribonucleic acid, not only in its elementary carbon, nitrogen, and phosphorous composition, but also in its content of those more complex units which are characteristic of DNA, such as deoxyribose and the purine and pyrimidine bases. Nitrogen and phosphorous analyses are not adequate to show that we have a pure deoxyribonucleic acid, simply because the nitrogen-to-phosphorous ratio is almost unchanged when there is as much as 5 per cent protein or 10 to 15 per cent of ribonucleic acid. The absorption at 260 mμ is characteristic of the purine-pyrimidine bases as well as other ring compounds and in relation to phosphorous gave figures helping to support the idea that we were dealing with DNA. The deoxyribose color value in relation to phosphorous was similarly like that of calf thymus deoxyribonucleic acid. As purification proceeded there was one step in which we removed ribonucleic acid and that gave a big increase in activity per unit weight, with no loss of deoxyribose nor of thymine, a characteristic base component of DNA (2). Even at that early time we did some of the orthodox things like making a preliminary measure of the base composition but we soon realized that we were probably dealing with a large variety of kinds of DNA and that, as they were presumably present in mixtures, we would not be able to find out much about any particular one.

We also made attempts to look a little more deeply for traces of protein because, at the time (about 1947) when this work was done (2), it seemed rather possible that there could still be enough protein present to give us the specific genetic effects which we were able to find. A study on the hydrolysis of highly purified material showed that there is amino acid liberated, not in as large a quantity or coming out at as great a rate as that from typical proteins. Under carefully standardized conditions of hydrolysis, proteins delivered their characteristic ultimate total high content of amino acid, while under the same conditions the transforming agent in purified form gave only a very small amount, but nevertheless some.

We knew that adenine, one of the orthodox constituent purine bases, could give amino acid on hydrolysis, from a degradative decomposition, for which an organic chemist could write a beautifully balanced equation. Now although the intense acid degradation of adenine gives an

amino acid, this is only a simple one, glycine. Even an organic chemist can not easily rearrange adenine any way to make phenylalanine, tryptophane, or arginine, and we only had to show that essentially all of the amino acid which is being produced from the purified transforming agent was glycine. Our microanalyses came out actually with figures like this: 99, 101, 99 per cent of glycine in that small amount of amino acid which is present in the hydrolysate. Therefore within the limits of error of our methods we could say that there could not be more than two hundredths of a per cent of protein containing a variety of amino acids in the transforming agent. So that together with enzyme information from purer enzyme available by this time, we were, even more positively than before, able to ascribe the action to DNA, deoxyribonucleic acid.

This name, deoxyribonucleic acid, sounds rather precise and impressive, but I must remind you that it does not really begin to tell us the variety of forms, configurations, and compositions that we suspect can reside in molecules that have a molecular weight of over a million and probably several million. I was thinking again today how much I admire the name, protein, which does not seem to specify or give false security of knowing something chemical, knowing something "scientific" about the structure and composition. I rather like the symbol DNA which is not quite so presumptuous as the full name.

QUANTITATION OF TRANSFORMATIONS

As mentioned, we have begun to depart from the traditional transformations to capsule type and one of the most useful things we were able to do was to develop transformations to drug resistance (3). Now, there is no medical interest in going from drug sensitivity to drug resistance; if we could go the other way, we might in some cases think it practical to do it, but it would still presuppose getting the organism in hand first, like salting a bird in order to catch it. We are mainly interested in the utility of picking up an organism that has been changed because it has been changed in such a way that we can recognize it in the middle of the large population that may not have been changed. Drug resistant transformants can be counted to give quantitative data on the rate and extent of specific pneumococcal transformation.

In a typical growing culture there is an occasional resistant mutant that appears spontaneously when a large number of drug sensitive organisms is assembled. Now pneumococcus is naturally sensitive to a number of the common drugs, penicillin, streptomycin, and so on. But in every few million of them we will find sooner or later one of the odd resistant ones. These can be selected by the simple expedient of putting the whole culture into drug media. When we do that we have only the drug resistant ones growing and soon there is a culture of this pure type. Here again we make a DNA extract and let it act upon the sensitive kind. There will be a rather large number of transformants formed which are resistant to the very same drug to which the donor strain serving as source of DNA was resistant.

In this type of transformation there is an especial advantage. We can put the whole group of cells into drug medium and only work with the transformed resistant cells. We can find out just when they begin to be resistant, how many of them are there at any time, and we do not have to carry along the troublesome group of untransformed cells. This way one can find out just what influences give an increase in yield, and in this way we have been able to increase the yields so that we not only do not carry along our sensitive cells but are able to find a number of influences and expedients that will give us quantitatively more transformants. Again, any of these transformants can be selected out as a single cell and can become the ancestor of a large population, all resistant to the very same specific drug (and no other drug), all giving a DNA which is able to make further transformations to the same drug resistance.

GENETIC MEANING OF TRANSFORMATION

The same system allows us to make a test of whether transformation is a genetic transfer. In the first place the gene is a concept; it has been implied increasingly through the last two decades or so that a gene is going to turn out to be a substance, but it was essentially originally a concept. As we have come in these last several years to talk about genes as though they are something like hormones—something that exists as a substance in a material form—we have come to think of them as chemical units of influence and one of the things that we can test now is whether these transforming agents operate and act as units of influence. The character of penicillin resistance in pneumococcus is acquired in stepwise form as a culture undergoes gene mutations away from its native sensitive state. One organism appears

which is resistant to a small additional concentration of penicillin. If this single organism is selected out and grown up to large populations, several million, and put into a stronger concentration of penicillin, most of them die. But again one odd organism or so may turn out to be resistant to a higher level, and we can reach high levels of resistance in this way. The important point is that these steps appear one at a time and it is only on the background of the first kind of mutation that the second superimposed mutation is recognized. We can right away inquire whether the DNA shows these facts in the life history of a strain, and what a resistance level amounts to in the transformation experiment.

By checking against several levels of penicillin concentrations we inquired how many and what kind of transformants are produced after exposure to nucleic acid from a highly penicillin resistant donor source (3). There appeared a group of about 75,000 organisms which were resistant not only to a penicillin level to which the original strains entirely succumbed but to a level about 5 times higher. So this group is one class, one resistance level and in further experiments we learned about higher successive resistance levels. The data are basically quantitative in nature and what happens in transformation indicates that each mutation led to the production of a different transforming agent for each of the indicated resistance levels.

At one mutation step we may have a fivefold jump in resistance as the first step. Such a first step population can give rise to a new mutant and this will be resistant to more penicillin, approximately two or three times as much penicillin as before. After opportunity to collect a number of organisms this culture will serve as the basis for a new mutation to a still higher level of resistance. Such a donor strain was made and its DNA was used as a source of transforming agent, and we obtained resistance of the first level. Our first exposure gave us a culture which "remembered" that same first fivefold resistance jump which the now highly resistant donor strain had once undergone. Subsequent transformations also happened stepwise. The DNA also reflected the second step, since when we took organisms which were produced by the first step transformation or spontaneous mutants having that level of resistance we could treat them with the very same DNA and then they would jump to the higher second level. The DNA "contained" in fact all three of the steps of resistance, but it could only

transfer them one at a time to suitable recipient organisms.

In other words, the DNA traced out in transformation the history that the cell strain had gone through in arriving by mutation at its high level of resistance. This kind of result indicates that we are really dealing with those basic units which had been modified in a mutation step and which can now be transferred in a kind of genetic crossing experiment. In other words, for an asexual organism like bacteria, these units satisfied the definition of genes. A gene should be that entity which is heritably modified in the single mutation step, and it is desirable if possible to demonstrate that it can be outcrossed to another strain and still behave as a unit. There were no ways of outcrossing bacteria until transformation and recombinations came along.

So the analysis in 1951 came to the point where it could be said that there were units of mutation in bacteria and that these could be transferred. It further became known that it is principally DNA from the donor cell which is transferred. This is now even what is believed to occur in bacterial recombination and the new kind of bacterial cross called transduction. In this last the heritable factors are carried by a virus which will infect bacteria with high efficiency and transfer genetic properties from the previous host at a lower efficiency.

All of these things are bringing to bear on bacterial genetics concepts that had developed from higher organisms, showing that we can think of bacteria much as we do of the higher organisms even though the genetic mechanism seems to be somewhat simpler. Some of the things that can be done with factors capable of transforming enzymatic properties have led us to experiments that cannot easily be done with higher organisms. The character of mannitol utilization is a hereditary property of certain pneumococcal strains which was studied by Dr. Marmur in our laboratory (4). Its characteristic is that the mutant type, the rare type of pneumococcus which can use mannitol like glucose for energy will convert the sugar alcohol, mannitol, to a sugar, fructose. Mannitol has two more hydrogens than an ordinary sugar, and the extra oxidation can be observed in a respiration experiment or by the reduction of the coenzyme diphosphopyridine nucleotide (DPN) in extracts of the bacteria. The fructose formed is then used in the regular energy economy of the cell.

Only this specific kind of mutant pneumo-

coccus can oxidize the mannitol to fructose, but even they have to be adapted to do so. If one has the mutant strain unadapted to mannitol, it takes quite a while before it will be able to grow in the absence of glucose and use mannitol as a source of energy. This is the criterion for an adaptive system and it is a sign that the mannitol dehydrogenase is not actually present all the time but is called forth when the mannitol requires it to be formed. The same thing can be shown in respiration experiments.

There is something very useful for our purposes about the difference between adapted and unadapted cells. Unadapted cells are not able to use mannitol and do not have this active enzyme system at all. But they do have the genetic capacity. You can not get this adaptation with the ordinary nonmutant pneumococcus strains because in the first place they do not have the genetic capacity to adapt. So we tested the DNA from adapted and unadapted mutant cells. If it is a genetic property that the DNA is carrying, it should be equally present whether they are adapted or not. If it is an enzymatic property it should reflect the enzymatic state of the cell. It turns out that it does not make any difference in the amount of transformation obtained no matter whether the DNA is extracted from adapted or unadapted cells.

So, then, this particular kind of DNA can make it possible for the cells which acquire it, by transformation or by mutation, to make a certain mannitol phosphate dehydrogenase. But the DNA only provides the genetic potentiality; the enzyme is only there in active form if the substrate is provided and if the cells are starved for other substrates like glucose. The presence of the DNA brings the know-how which allows it to be made, indicating very nicely its essentially gene-like function.

It is a little too early these days to infer that any particular DNA brings about the production of some particular protein from start to finish. It actually appears possible that this DNA is only bringing about the production of a system and maybe only part of a system which forms mannitol dehydrogenase. Maybe it rearranges pieces of enzymes which are usually used for something else, altering the specificity of the enzyme to permit the near approach to the active sites of phosphorylated mannitol where ordinarily phosphorylated glycerol, for example, fits. We really do not know for any case; this is just to indicate that

there are many simple possibilities that could be envisioned.

It is convenient for the chemist to think of protein as made up of such and such amino acids. We have seen how interesting and important it can be to know these structures and ponder how such a molecule can in some way bring about an enzymatic reaction, but we still do not know much about how it does it. I would like to emphasize that we must continue to consider other properties such as the 3-dimensional properties and learn to think not only in terms of the convenient amino acids but also of the inconvenient larger complex groupings, for these last seem in some cases to be concerned in biological activity or specificity.

GENETIC INTERRELATIONSHIPS OF DNA-BORNE GENES

When other transferable markers than capsule type became available, we were able to study the effects of DNA from doubly marked strains (3). At the time the first experiments were done it was common to think that the transformation was going to make a transformant just like the donor. It did not; two characteristics transferred independently in the transformation process. For example, DNA from drug resistant encapsulated cells produced encapsulated but still drug sensitive cells and also drug resistant but still unencapsulated cells. It appeared as though the two marker substances behaved independently, and we infer the existence of two independent kinds of particles transferring these two specific kinds of effect, capsulation or drug resistance.

Other pairs behaved similarly and in fact some nine markers have been compared in many different pairs and they have always with one exception given only preparations behaving as though they contained independent, different kinds of DNA's responsible for the different kinds of marker. This of course is the way genes behave, too. Genes behave as units, only one does not ever catch them by themselves except in transformation. In higher organisms one has to study a lot of different individuals to see that they act as units.

But our DNA preparations have molecules, apparently big ones of molecular weight 1 to 5 or 6 million and it seems quite possible that they are still bigger than single gene units. It does seem that the particle we are dealing with in solution is in fact essentially a kind of molecule, and if so we have signs of independent molecules, one

having for example encapsulation factor and one having streptomycin resistance factor. One case, however, turned up which is different and does not show completely independent factors. When the donor strain is mannitol using and streptomycin resistant, the DNA having both of these properties makes mannitol utilizing cells and streptomycin resistant cells, but it makes also quite a large number of cells which are like the donor, both mannitol using and streptomycin resistant (5).

This was the first case of the various pairs we studied showing association between factors. It might mean that the DNA has two properties in a molecule, and this was a little supported by the fact that if we mixed the two kinds of DNA having these two separate factors they did not make doubly marked transformants. If both of these two factors travel together we get this kind of cell, but they only travel together when they are originally coming out of the same cell. We have not seen any tendency for M and S to stick together, or react together artificially as chemical substances. This little cluster of MS only forms when they are carried by the same cell; then it can sometimes stay together during transfer and give a high proportion of doubly marked MS transformants.

Another kind of experiment greatly strengthens this conclusion. The donor strain which is mannitol using but streptomycin sensitive gives a DNA having a mannitol factor. If you put that on an ordinary pneumococcus you will just detect a mannitol transformation. You will not know that there is anything having to do with streptomycin, but if you put it on a streptomycin resistant recipient cell you will also pick up cells which are like the donor again, streptomycin sensitive and mannitol using. You push streptomycin resistance out of some cells because you are putting mannitol in. This is a sign that there is actually something there in a place that "belongs" to streptomycin, and whether it is in the sensitive phase or the resistant phase it will stick together with mannitol. And the same happens when you do the opposite in that streptomycin DNA factor on a mannitol cell can "push" mannitol out. So it does not matter whether the DNA represents the mannitol utilizing or mannitol not-utilizing form, whether streptomycin resistant or sensitive form, the DNA responsible for these two properties behaves partly as a unit showing what seems to be genetic linkage.

There is, therefore, reason for thinking that the molecule of DNA is something a little bigger than a gene and capable of making several genetic changes. Yet the bacterial transforming DNA's give us genetic material subject to chemical and physical study in vitro, in a form far simpler than the nucleus, as yet the smallest isolatable functional genetic unit of higher organisms. Here we seem to be dealing at a molecular level with one of the great biological problems. I know of nothing more dramatic than to make a stable DNA preparation from a strain of mannitol using bacteria, then perhaps many months later to continue the interrupted inheritance by putting this highly specific material into organisms that have never until that moment been able to solve the problem of converting this sugar alcohol into energy and cell substance.

DISCUSSION

DR. HENRY BORSOOK: I would like to take this opportunity to make a few comments on the very remarkable story we have just heard. The magazines on the newsstands tell fantastic stories going on in fantastic worlds, all made up. What we have heard is much more fantastic and it is all true. I would like to go back a little. I remember hearing physicists talk about the structure of matter as follows: When we were young there were just molecules, these were depicted as solid balls, and then now there was a big black ball which was the nucleus and little black balls around which were the electrons. We have lived through a similar time in biology. Within the memory of everybody in this audience bacteria were the solid black balls, all black, we knew nothing of their internal structure. No one dreamt they had an hereditary apparatus remotely resembling that of the higher animals or plants. We have seen and heard that not only do bacteria have an hereditary apparatus but have the actual kind of apparatus as in the higher animals. This is one of the most remarkable advances in science of the last 20 or 30 years. We see again and again that if nature has a certain job to do, as a rule it does it always in the same way, whether it is a biochemical operation or a genetic operation. There is one other very remarkable and fantastic feature of this story. The geneticist in very early days when mutants were first discovered was bedeviled by the fact that they had to wait for them to happen. There were few mutants and they could do so few experiments. Then it was discovered that they could make mutants by bombarding cells with x-ray, ultraviolet, and so on.

Still later it was found that substances like mustard gas do the same thing. But it was always a matter of hit and miss. One smashed the hereditary apparatus, got a lot of pieces, and if one was lucky and patient and skillful one picked out a few bits here and there that were useful. In the very beautiful work that Dr. Hotchkiss told us about, a genetic factor is put into the cell. This is a new world, a fantastic new world. Hereditary material is introduced, under control. Nothing is smashed up. Experiments can be done on such classical genetic phenomena as crossing over and the interaction of genes, on adaptive enzyme formation. This work is of the highest importance.

I have two small questions that I would like to ask Dr. Hotchkiss. One is, when you introduce the mannitol transforming factor, is there any information on whether more than one enzyme eventually appears which was not there before? I should like to preface my second question with some general explanatory remarks. The biochemist hopes eventually to get out the hereditary apparatus. Work such as that of Dr. Hotchkiss has encouraged us to believe that each hereditary unit is a DNA molecule. I had better put "molecule" in quotes because once the size gets up to 6 millions, what does a molecule mean? Nevertheless it is a packet. The biochemist's hope is that eventually he will be able to work out the structure, with all of its foldings and cross linkages. Then we may be able to say that "this combination of purine and pyrimidine nucleotides, folded in this way, this distance apart, is the genetic locus." My question is this: Is the purified DNA in your experiments all the DNA of the background, or is it the DNA pertaining to one or only a few genes? If it contains the factors for a whole lot of genes the biochemist is stuck. To work out the nature of the genetic locus he has to have a DNA pure that is a DNA for one factor and one factor only. My question is, what are the chances of getting out the DNA for one gene?

DR. HOTCHKISS: As to the chances of separating single pure DNA's, a good many people are interested in this problem now. We have cooperated with three groups, and there are two others working also, and the result is uniform so far. Inactive material has been separated from typical preparations so a fraction of the DNA is apparently not effective. It is easy to make such inactive material by slight degradations, inactivations with acid or base and so on, and we think

that this may be what it is. The active material has not yet been definitely separated into two kinds of active material, although of course some of the groups have worked with only 1-marker material and so their active and inactive fractions may have seemed at first to have shown separation. We have worked with 4-marker material and we find all our 4 markers going together. In fact, we hope to satisfy the question whether there really is only one long continuum of DNA, and we just get relatively uniform pieces of this, but constituting different fragments, or whether they are preformed and natural.

As to whether there is more than one enzyme formed in the mannitol induction, we know of course our normal pneumococcus can use fructose-6-phosphate, so we need assume no new enzyme from that point on. We know the transformant learns to use phosphorylated mannitol and it seems likely that they must phosphorylate it, but whether this last is due to an enzyme which is only present in the mutant or is present in all pneumococci we have not been able to determine. These two enzymes, the phosphokinase and the dehydrogenase, would seeem to be all that there is any reason to suspect, at least.

DR. TSOO E. KING: As we have agreed in our conversation last night, the appearance of deoxynucleic acid isolated from regardless of what sources is not as that in situ because chromosomes just cannot house the molecules of such size. Apparently many bonds are broken and perhaps some new bonds are formed during isolation in chemical laboratory. All right. Now in your last slide you have showed us the isolated DNA which possesses the transforming activity is in the shape of nice long threads. Will you care to comment on whether the swollen DNA as shown on the slide as such is biologically active or whether the bacteria have first to convert it back to, shall we say, the natural form, then the transforming action can be exhibited? If the latter is the case, do you imply the bacteria possess the ability to restore the original configuration of DNA as existed in situ?

DR. HOTCHKISS: Well, here we are facing the eternal biochemical problem that we are trying to work with chemicals and are attempting also to work back toward the cell but are not getting very far back. Since we are insisting on the workability of extracting the DNA we are obliged to take it as present in the form in which we get it. Now as we have agreed in conversation this form is a rather swollen, ex-

68 PROTEINS

panded, and elongated molecule—so long that it is actually almost as long as the cell from which it comes. Anybody who has watched the process can see at the time it is happening how much swelling there is as it comes out from a form which is compact, water insoluble, is not sticky, it is perhaps not hydrophobic, but it is not very highly solvated. When it begins to be liberated from whatever it is that holds it in a chromosome or whatever form it exists in the cell, it swells up. But the same would be true of other substances released from cells. Gelatin becomes very unlike the natural precursor collagen, and so on. It is good to remember that this is the time when DNA begins to get the nice double helix that can be seen in oriented fibers and x-rays, the time when it begins to get this more or less elongated rod form that the physical chemist can measure. Back in that little cell it may have a more condensed form. I think it should also be pointed out that when it reacts with the cell it has to find a suitable receptive cell. When you try taking DNA from a rather distant strain, it does not make genetic transfers so successfully in an unrelated organism.

REFERENCES

1. Avery, O. T., C. M. MacLeod, and M. McCarty, J. Expt. Med. *79*, 137 (1944).
2. Hotchkiss, R. D., in Phosphorus Metabolism, vol. II, p. 426 (McElroy and Glass, eds.) Johns Hopkins Press, Baltimore, 1952.
3. Hotchkiss, R. D. Symposia on Quantitative Biology, Cold Spring Harbor, *16*, 457 (1951).
4. Marmur, J, and R. D. Hotchkiss, J. Biol. Chem. *214*, 383 (1955).
5. Hotchkiss, R. D., and J. Marmur, Proc. Natl. Acad. Sciences, U.S., *40*, 55 (1954).

Reconstitution of Tobacco Mosaic Virus

Heinz Fraenkel-Conrat

Practically all the terms which I usually have to define when talking about this work on tobacco mosaic virus have been discussed by previous speakers, and I will have opportunity to refer to practically every preceding talk and thereby to shorten my talk a little bit. I think you will all appreciate this, since it is now late in the meeting and late in the day. I hope I do not get carried away and forget about the promise I have made you.

There was a little poem cut from a magazine and pinned in the elevator of our building recently which read:

> Spring is here, again life focuses
> On grasses and on crocuses,
> On rows of deep blue irises,
> And on one or two new viruses.

This is as close as we have come in our laboratory to new viruses, and anybody who expects startling disclosures about the "synthesis" of new viruses can leave now.

This talk is entitled "Reconstitution of Tobacco Mosaic Virus" because this is what has caused some interest and attention, but I would like to review a little of the general chemistry of what we know about tobacco mosaic virus because it does fit in with many of the preceding talks in a way and shows that also in somewhat bigger and complex proteins similar laws and rules can be applied and similar studies can be made. When one considers that the molecular weight of tobacco mosaic virus, 50×10^6, is almost a thousandfold that of hemoglobin, a chemist might easily become discouraged from looking at it. Fortunately, however, this super-molecule is built up of subunits which are probably identical, and which are no more complex than many simple proteins, hormones, etc. I believe I detected a slightly disappointed note in Dr. Behren's conclusions concerning the structure of glucagon. It sounded like: This is the amino acid sequence of this protein—so what? Chances are it will be the same when we have the complete amino acid sequence of TMV, but that day is still far off. Nevertheless, the realization that the virus consists of small subunits has greatly stimulated chemical attack on virus structure. This realization was the result of the accumulation of different types of evidence. It was known that alkali, urea, guanidine, or detergents rapidly degrade the virus into smallish stuff. All of these agents lead to a separation of the nucleic acid from the protein. All viruses contain nucleic acid and protein, although in greatly varying proportions. The nucleic acid in TMV is ribonucleic acid exclusively. In contrast, some other viruses contain only deoxyribonucleic acid, so there does not seem to be any general rule. All we know is that all infectious virus particles contain nucleic acid, either RNA or DNA. In amount the protein is by far the major component, 6 per cent only being nucleic acid and 94 per cent protein. Therefore TMV a priori appeals to a protein chemist and to a protein colloquium. All methods of degradation of TMV (mentioned above) which separate nucleic acid from proteins also degrade the proteins into the small units. Consequently the building blocks are held together, not by primary chemical bonds, but by a different kind of cement, secondary valences, hydrogen bonds, ionic forces, etc. The molecular weight of the smallest unit is very poorly defined in physicochemical terms because of its horrible tendency to reaggregate to bigger complexes, particularly under conditions of ionic strength, suitable for physicochemical study. According to H. K. Schachmann, a preliminary value of 10 to 20 thousand seems to delimit the molecular weight range of the protein subunit. Then there came the remarkable finding of Harris and Knight who treated the virus with carboxypeptidase, the enzyme that has been previously mentioned as a means of degrading proteins from the carboxyl end. They found about 2,800 threonine residues to be split off from the virus and nothing else. You heard how carboxypeptidase will eat up glucagon half the way down the chain before the chemist can stop the reaction. In many other cases there is similar degradation so that many amino acids are split and it is difficult to establish end group or sequence. In contrast, in TMV, the enzymatic attack is remarkably clear cut. One can perform the reaction either with the isolated protein or with the whole virus. If one uses the whole virus, the particle appears unchanged though of slightly different electrophoretic mobility, and its infectivity is not decreased. If one

assumes that all the threonine is located at the end of peptide chains, then the molecular weight of these chains is about 18,000. And that was a very nice figure which appealed to us because it was known that there was one sulfur atom, due to cysteine, as discussed yesterday, per 18,000, and there are two lysine groups per 18,000. Thus, prey to the fascination of numerology, we and others became convinced that 18,000 was a good molecular weight approximation for the minimal subunits which build up the virus. Some support for this value comes from x-ray diffraction studies, the most recent of which favor a similar weight for the x-ray crystallographic unit, on the basis, however, of a lower particle weight (40×10^6).

To establish definitively that all the threonine was C-terminal on so many peptide chains, it appeared necessary to confirm this by a chemical method. The method we used, hydrazinolysis, has already been mentioned. It degrades the whole protein to amino acid hydrazides, leaving only the C-terminal amino acid intact. When this was carried out with the virus or its proteins, threonine alone was found, in quantitative accord with the enzymatic results. Hydrazinolysis is a method only for end group analysis, but if the threonine were first removed enzymatically then one should get the second amino acid. By this approach, we found alanine to be the second amino acid from the carboxyl end. All this work was done largely by Dr. C. I. Niu in our laboratory. Through a fortunate coincidence he was able to establish also the amino acid next to the alanine. Thus after short periods of hydrazinolysis, he obtained not only alanine, but also some prolylalanine in the free form. Thus the C-terminal tripeptide sequence was now known to be -proline-alanine-threonine. This sequence was further extended by three more residues from this carboxyl end, which were established by a different technique, again one that has already been mentioned here—namely, selective enzymatic degradation of the entire protein. Thus, trypsin and chymotrypsin were again used to degrade the protein and many peptides have been isolated, about 40 I think by now, and their amino acid composition determined. The sequences are being slowly worked out, a more difficult job than amino acid composition. When enough peptide sequences will be known, and enough overlaps established, then it should be possible to write the complete structure. This goal is being approached for ribonuclease by Moore, Stein, and coworkers at the Rocke-feller Institute, a protein chain three times longer than any yet elucidated, and two-thirds of that of the TMV subunit.

In contrast to the ease with which the C-terminal amino acids could be identified, no corresponding N-terminal amino acids were found by either of the laboratories studying this problem. Only after the protein was treated under controlled hydrolytic conditions (boiling trichloroacetic acid solution), did N-terminal proline appear together with small amounts of other N-terminal and C-terminal groups. It seemed probable that this finding might fit in with the previously postulated ring peptide.

Thus the proline may be one that closes the ring with a glutamic or aspartic acid residue. The resulting linkage is not a real peptide linkage, but a γ- or δ-carboxyl imino linkage. If such a bond is more labile than the typical peptide bond, then trichloroacetic acid might split it in preference to random splitting and one would get N-terminal proline upon so maltreating the protein. This seems a reasonable assumption, but is not yet definitely proven.

Most of this structural work was done with a protein preparation isolated by such methods that it was completely denatured. Denatured proteins have a great tendency to aggregation and precipitation owing to random hydrogen bonding, and this is true also for the virus protein. It therefore becomes of interest to compare the behavior of this protein with a more native preparation. This is obtained by gentle alkaline treatment (around pH 10.5) in the cold. Nucleic acid and protein are split, and the protein precipitated with ammonium sulfate. It is then still native in the sense that it retains its masked -SH group, and its solubility in water. When the solutions are brought to pH 5-6, a remarkable phenomenon occurs. And here now we go one order of magnitude higher from the scale of protein chemistry than we have been talking about. The native protein has a molecular weight of about 100,000, and is apparently built up of 5 or 6 of the peptide chains we have previously discussed. At the lower pH, however, it has a tendency to aggregate in a very specific manner: There are formed circular objects with a central hole which have been compared to doughnuts, and then rods of the same diameter, which may at the proper pH be almost undistinguishable, in electron microscopical appearance, from TMV virus. Thus a simple protein invisible in the electron microscope aggregates helically to a long rod presumably owing to affinities of

autocomplementary sites on its sides. I think that most biological activities of most proteins are due to their ability to specific binding of something or of themselves. In this particular case, I believe that the biological activity of the virus protein resides in its ability or its tendency to reform rods of the shape of the virus, a property that is inherent in its native protein structure.

Having this protein on hand we wondered what would happen if we let it aggregate in the presence of the nucleic acid. This curiosity is not new; lots of people have tried it before. In previous attempts, however, either the protein or the nucleic acid were not native enough, and no interaction was detected. I think the reason is that the nucleic acid is damaged by the alkali treatment, which yields good protein. On the other hand the detergent treatment that I first mentioned gives denatured protein, but good nucleic acid. We know very little what "good" and "bad" nucleic acid means, but use those terms only to differentiate between preparations that "work" and those that do not. Well, it happened that we had at the same time both a native protein and a fresh preparation of good nucleic acid. When we mixed the two at pH 6-7 we got particles that again looked like virus, both in appearance in the test tube and in the electron microscope, while the protein alone did not give long rods at that pH. When plants were inoculated with such solutions virus-lesions resulted, and that was something new and somewhat surprising and exciting. We did not believe it for a long time, and we have been busy trying to convince ourselves beyond any element of doubt that it is real ever since.

Electron microscopy has supplied good evidence that the nucleic acid is incorporated into the same position in the core of the protein helix which it occupies in the original virus. The infectivity obtained upon combining protein and nucleic acid is never 100 per cent in the sense that all the protein will give active virus. From the electron micrographs, one can calculate that about $\frac{1}{3}$ of the mass forms rods of the right length, so one would expect a maximum activity of 30 per cent. Actually our best activities have beeen 3 per cent and often they are appreciably lower, so some of these rods that look right still do not act right. It was a quite definite possibility that only a small fraction of the nucleic acid is really good in the sense of making active rods, and we toyed with the idea that when we degrade the virus and strip the protein off, part of the nucleic acid might not fall apart into subunits as does the bulk of the material. The molecular weight of the bulk of the material of the nucleic acid is about 200,000 to 250,000. And the molecular weight of the total nucleic acid in one particle must be about 3 million (6 per cent of 50 million). It seems possible that 10-12 subunits of nucleic acid form a central possibly pleated strand. Thus we thought of the possibility that actually our active material was the small fraction that stayed big and that was not detected in ultracentrifuge and similar tests because it was only of the order 1 to 3 per cent. But recent indications are that this is not at all the case. We now think that the small nucleic acid and small protein units recombine and that the protein influences the nucleic acid aggregation and vice versa in getting the right shaped particles. Yet the length of the nucleic acid chains may be a critical factor, for the 250,000 molecular weight units have the right length to fit into and possibly determine the length of 300 millimicrons which is definitely the upper limit for reconstituted virus particles.

Now in recent months we have extended our work to include strains or mutants of TMV which have led us a good bit further. Dr. Hotchkiss already has defined some of the genetics of bacteria to you and justified the use of terms such as mutants. These terms have long been used with virus, and a lot of cases are known where viruses have been found to change suddenly so that a one-step mutation was assumed, or in other cases a virus has been found which seemed very closely related to another and yet not quite identical. In the case of TMV many such variations on the theme are known and these strains have been studied off and on. Much work on their composition has been done by Knight, both at the Rockefeller Institute and in our laboratory. Surprisingly, all mutants of TMV that he studied had the same nucleic acid composition. Almost all strains, however, differed in their amino acid composition. And obviously they differed in their symptoms. For a mutant is generally detected by the nature of plant lesions or disease it gives to test plants. If one prepared nucleic acid and protein from different strains of TMV and combined the two, one should be able to learn something new about the importance or the role of protein and nucleic acid in determining the genetic characteristics of the virus. The virus is in a way the model of a gene. It consists of nucleic acid and protein, and all it can do is to direct a cell to make more of itself. The question

is: does it achieve this through its nucleic acid, or its protein, or both? And the answer we have gotten was pretty clear cut. We have obtained good nucleic acid out of four different strains, all four that we have tried. The proteins show sufficient differences to render each a different preparative problem, so that we had to work out conditions for each, and have so far only gotten 3 out of the 4 in the native form out of the virus. But yet we had enough with 4 nucleic acids and 3 proteins to get 7 or 8 combinations, and in each of these combinations the disease symptoms were always those characteristic of the strain supplying nucleic acid. So we convinced ourselves somewhat to my regret as a protein chemist that beyond any doubt the genetic material in the virus is nucleic acid. This is quite in line with recent work on bacteriophages. It is only surprising because this particular virus contains ribonucleic acid and ribonucleic acid is believed to be so different in structure from DNA that one would not expect the two to be able to fulfill the same complicated function.

A byproduct of the preparation of such mixed virus was the convincing proof that the activity in reconstituted preparations was due to particles formed from the two components. The strain most dissimilar from common TMV in amino acid composition, termed HR (Holmes ribgrass) is also serologically differentiable. One can make antiserum to TMV and antiserum to this HR virus strain. The two will cross react, but if each serum is pretreated with the heterologus virus reasonably selective sera are obtained which neutralize the infectivity of the homologous virus with little or no effect on the other virus. When the two reconstituted mixtures of the components of these strains were treated with the antisera, the activity of each was inhibited by the antiserum to its protein component, while the symptoms of the activity were clearly characteristic of the nucleic acid component of each. Thus it seems definitely proven that the active particles have a dual nature, deriving their protein coats from one and their nucleic acid innards from the other strain.

There remained one important question to settle. How free from virus activity were our protein and nucleic acid preparations? Almost all protein preparations were remarkably free from virus activity, even when tested at 10^6 fold concentration. In contrast, almost all fresh preparations of nucleic acid produced virus lesions at about a thousandfold concentration of standard

TMV. Indications were soon obtained that this activity was not due to contaminating virus. It was destroyed by a trace of ribonuclease (10^{-12} M) and decreased rapidly in 0.1 M pH 6.8 phosphate whereas TMV is completely stable under these conditions. The activity in nucleic acid preparations is not sedimented upon ultracentrifugation, whereas added TMV is sedimented to over 99 per cent under the same conditions. Many other tests and criteria contributed, and we got more and more convinced that our nucleic acid was not contaminated with virus but that the nucleic acid had an intrinsic infectivity, similar to the DNA of Dr. Hotchkiss. Thus it appears that viral activity is really inherent in the nucleic acid but is extremly labile and sensitive to a variety of agents, and that what the protein does is to stabilize and protect it, by forming a tightly packed coat around it.

The final subject to be discussed is the chemical nature of the progeny of nucleic acid or mixed reconstituted virus preparations. Do the symptoms mean that the virus which is multiplying will be identical with the strain which supplied the nucleic acid? Will the nucleic acid which determines the symptoms also determine the nature of the protein in the next generation even if it is wrapped in another protein? Does the nucleic acid carry enough genetic information to tell the plant what it used to be wrapped in before it was forced into the wrong coat and inoculated? We have again concentrated our efforts on the two dissimilar strains mentioned above. The most notable difference is the presence of histidine and methionine in HR, in contrast to all other strains of TMV. Well, progeny of HR nucleic acid, even if applied in TMV coats, contained both amino acids. Progeny of the reverse cross (TMV-nucleic acid + HR – protein) contained neither of these amino acids. But whenever you think something is settled, then you do some more detailed and thorough experiments and you begin to wonder whether you really have the whole story. So we have done complete amino acid analysis now on the first progeny and have compared it with both HR and TMV. The complete amino acid composition is dissimilar from TMV and is very much like HR, but we get in about two or three amino acids minor differences which seem too consistent to be shrugged off, and we continue to do amino acid analysis over—we have done it about four or five times and it is a lot of work—and when we have done it ten times more and have got these same differences,

then we might begin to wonder whether maybe the protein after all did influence the progeny to the extent of modifying the progeny protein a little bit. That would then be a new type of virus produced through the influence of adding the wrong protein to a nucleic acid. At the moment, however, all we can say for sure is that the nature of the progeny is very similar to that of the original virus supplying the nucleic acid.

There is one other difference between the progeny and the original virus. HR virus is only about one twentieth as active as TMV. Well, the progeny virus has generally shown about five times higher specific activities than the HR. In regard to activity, therefore, the progeny of the hybrid has been halfway between the parent strains. But there too there is a possibility of error in that the progeny was isolated always in smaller batches than the very big scale preparation of pure strains that we get for starting material, and it is possible that such factors influence the extent to which virus rods are broken. We are therefore at the moment isolating progeny and both parent strains under the same conditions of batch sizes, etc. In conclusion, we are in no way certain that the differences between HR and the progeny of the mixed virus are real.

A different approach to the production of virus mutants was then attempted. On the assumption that the nucleic acid reacted in the form of subunits, a mixture of nucleic acid from two different strains was treated with protein, and assayed. It was hoped that new virus particles which contained nucleic acid from both parents might be formed and that these might replicate in that way. These experiments are only in their beginning stages, but it does not at present seem possible to produce stable mutants in that manner.

To summarize then: we are quite convinced that active particles form with great ease in solutions containing small-molecular virus protein, nucleic acid, and salt at pH about 6-7. Active particles form also from nucleic acid and protein derived from different strains. The formation of new types of self-replicating particles has not yet been demonstrated. The nucleic acid has been shown to have infectivity of its own. The progeny of such nucleic acid appears the same as the strain supplying the nucleic acid. No practical significance can be attributed to this work at the moment. One lesson one can learn which one did not need this work for is that proteins are antigenic. If one wants a nonharmful

antigen to immunize against pathogens, such as polio, why use polio virus, why not use the polio virus protein? A new approach is suggested by this work, that people should separate viruses into nucleic acid and protein and use the protein only when they want protein properties. Perhaps one thing that this work has pointed out is that the ever disappearing frontier between living and nonliving matter is further pushed back. We are dealing not with anything alive in the first place, but with one property of life, namely the self-replicating, and it seems that the property of self-replication does not require a structure of several million molecular weight, but only of about two hundred thousand. So, in short, life has neither been taken nor given in the course of this research.

DISCUSSION

QUESTION: Have serological studies been made with the progeny of this mixed virus?

DR. FRAENKEL-CONRAT: Again we are trying to refine the techniques, but so far it has not been possible to distinguish progeny derived from HR nucleic acid from the HR protein which it chemically resembles so closely. It is quite different from TMV and quite similar if not identical with HR.

DR. ROLLIN D. HOTCHKISS: I like the way Dr. Fraenkel-Conrat has carefully said that he has not really put life together. I think it is one of the unfortunate things that tend to be read into this work and also into the transformation work, that something has beeen created. We have taken nucleic acid away from one bacterium (or virus) and we have put it back into a new cell, and I think we can not call it synthesis as long as we are not able to define any more properly the nature of the materials that go into the reaction. But obviously a very great operational step has been made, and this step means that there is not quite as much mystery in the interaction as there used to be, and we know more about it.

I am thinking, too, about the other part, the interaction between the protein units in forming these rods. I was struck, too, by the fact that glucagon which Dr. Behrens was telling us about, and insulin, as we know, and some others, tend to form rather characteristic organized fibers or particles, and these interactions may really be one of our interesting clues to protein-protein interactions. I am asking now whether Dr. Fraenkel-Conrat has been tempted to polymerize different virus proteins together and see whether they can tell each other apart and build up rods that com-

bine the elements of the two kinds of virus. I assume that something could be done with antiserum to show whether such a mixed material had been made.

DR. FRAENKEL-CONRAT: To your question, no. For some reason we kept thinking it seemed more timely to mix the nucleic acid, and we have not gotten around to the other. I think you are quite right and we should do that. We have talked much about protoplasm at this meeting. I do think that it is a very important part of getting from protein chemistry to the biological structure of the cell to think of such protein aggregation reactions as being of great importance. And here would be an interesting case to study the specificity of such an interaction. One would think the two proteins would not form mixed rods, but it should be done.

QUESTION: I would like to ask a question. Dr. George Beadle has just finished his series of lectures here on the campus where he talked about the shell of the virus being left on the outside of the bacterial cell and the nucleus being fired into the bacterial cell. Do you think that is similar in action, or do you have any indication?

DR. FRAENKEL-CONRAT: No, I have no ideas on that. Plant viruses are introduced into the cell by damage of the cell walls. A very inefficient way of infection, apparently much less neat and clean than bacteriophage infection for which this mechanism holds as you describe it.

DR. EDMUND H. FISCHER: In connection with the experiments in which you treated your purified nucleic acids with the specific antisera, do the antisera react also with the protein subunits, and if so, with how small subunits?

DR. FRAENKEL-CONRAT: Yes, the antisera react specifically with the proteins, I think at all levels of aggregation. But definitely with the native protein (molecular weight of 100,000), they do react. Precipitation may be more efficient for the big particles and intact virus than for the subunit. I don't think a good quantitative study has been made.

QUESTION: I wonder if the protein can bind nucleic acid after it has become aggregated to rods?

DR. FRAENKEL-CONRAT: The protein rods are much more labile, and apparently in an equilibrium with small material. They form at a low pH, and the solution is never free of small particles. We know that we can aggregate a protein and then add nucleic acid and get activity out of it, but we are not sure whether it is the same protein rod and the nucleic acid slips in or whether it is a rearrangement, with the protein rods serving as a storehouse for small particles.

Biological Synthesis of Proteins*

HENRY BORSOOK

The problem of the biological synthesis of protein has two aspects, at least in thinking about it. One is the provision of the energy by which a peptide bond is made. I had better be a little more precise and say that to make a dipeptide from two amino acids requires in the neighborhood of 2,500 to 4,000 calories per mole of peptide bond made. Such a bond cannot be made by itself. This is the reason that enzymatic hydrolyses of peptide bonds go to completion. This may require some qualification; to hook together two large peptides may require less energy. This probably is easier, but I shall not go into this aspect at all. I will merely say that in the synthesis of protein there is an energetic problem, which is how the energy is conveyed. And then there is the problem of the amino acid pattern. How is the inherited pattern of the protein molecule made? We shall leave out of consideration altogether such problems as the specific folding. These may very well solve themselves once the problems of the provision of energy and the amino acid pattern are solved.

We do not yet know whether these two sides of the problem are experimentally separable into two steps. As we shall see, the tendency nowadays is to try to separate them experimentally. This may be a mistake. They really may have to go together. Most of the knowledge that we actually have of the possible ways in which proteins are made has come not from protein synthesis at all, but from the study of the synthesis of quasi-peptides. Energetically and to some extent chemically their synthesis resembles that of small peptides. They are so small—there are two or three amino acids in them at the most—that there is no problem of pattern. There are only the problems of enzyme specificity and of the energetic considerations. And so we may say these are rudimentary proteins. But on the basis of the way in which these rudimentary proteins are synthesized, we may then make some more or less informed guesses and then see how far we can get toward the explanation of a number of other phenomena which attend the synthesis of quasi-peptides and small peptides.

Synthesis of Hippuric Acid

$$E + ATP \rightleftharpoons E\text{-}AMP + P\infty P$$
$$\text{(enzyme)}$$

$$E\text{-}AMP + HSCoA \rightleftharpoons E\text{-}SCoA + AMP$$

$$E\text{-}SCoA + \underset{\text{(benzoic acid)}}{HOOCC_6H_5} \rightleftharpoons CoAS\text{-}\overset{\overset{O}{\|}}{C}C_6H_5 + E$$

$$CoAS\text{-}\overset{\overset{O}{\|}}{C}\text{-}C_6H_5 + \underset{\text{(glycine)}}{H_2NCH_2COOH} \rightleftharpoons \underset{\text{(hippuric acid)}}{HOOCCH_2NH\,\overset{\overset{O}{\|}}{C}\text{-}C_6H_5} + CoASH$$

$$C_6H_5\text{-}COOH + H_2NCH_2COOH + ATP \rightleftharpoons C_6H_5\text{-}\overset{\overset{O}{\|}}{C}NHCH_2COOH + AMP + P\infty P$$
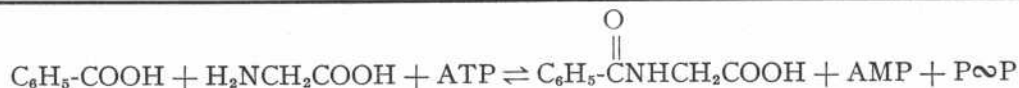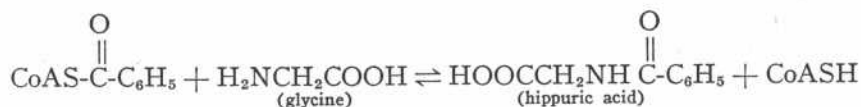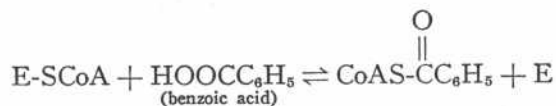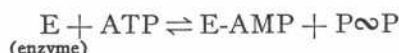
In the synthesis of hippuric acid, which was historically the first one studied in vitro, the carboxyl group of benzoic acid is condensed with

the amino group of glycine. This condensation requires about 2,500 calories of free energy. The formation of alanylglycine and leucylglycine from their constituent amino acids requires about 3,500. So the problems are in this respect the same. The first step appears to be reaction between enzyme

and ATP to make an enzyme-AMP compound. It is at this stage that the ATP is split with the formation of pyrophosphate. The enzyme-AMP then reacts with CoA to form enzyme-CoA and AMP splits out. Enzyme-CoA reacts with benzoic acid to form benzoyl-CoA. Then the benzoyl CoA reacts with the glycine to make the hippuric acid, and the CoA is regenerated. The energy required for the condensation has been introduced by activation of the carboxyl group; there has been no such activation of the amino group. Everything that I will say from now on will either be based on explicit proof or will infer that the amino group that enters into the formation of a peptide bond does not need activation. But this is not to say that the enzyme that makes the peptide bond does not have a specificity for the amino group; it does. But there does not appear to be any necessity for its activation. Nothing needs to be attached to the amino group to enable it to condense with the carboxyl group.

Synthesis of Pantothenic Acid

$$HOCH_2C(CH_3)_2CH(OH)COOH + H_2NCH_2COOH + ATP \rightleftharpoons$$
$$\text{(pantoic acid)} \qquad \text{(}\beta\text{-alanine)}$$

$$HOCH_2C(CH_3)_2CH(OH)CONHCH_2CH_2COOH + AMP + P\infty P$$
$$\text{(pantothenic acid)}$$
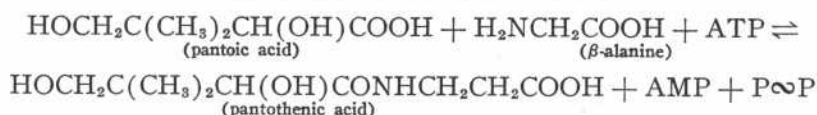
In the synthesis of pantothenic acid from pantoic acid and β-alanine ATP is required and is split into AMP and pyrophosphate, as in the synthesis of hippuric acid. But in this case CoA does not enter into the story.

Synthesis of Glutamine

$$E + ATP + HOOC(CH_2)_2CH(NH_2)COOH \rightleftharpoons ATP\text{-}E\text{-}OC(CH_2)_2CH(NH_2)COOH$$
$$\text{(glutamic acid)}$$

$$ATP\text{-}E\text{-}OC(CH_2)_2CH(NH_2)COOH + NH_3 \xleftarrow[Mg^{++}]{Mn^{++}} H_2NOC(CH_2)_3CH(NH_2)COOH + ADP + P + E$$
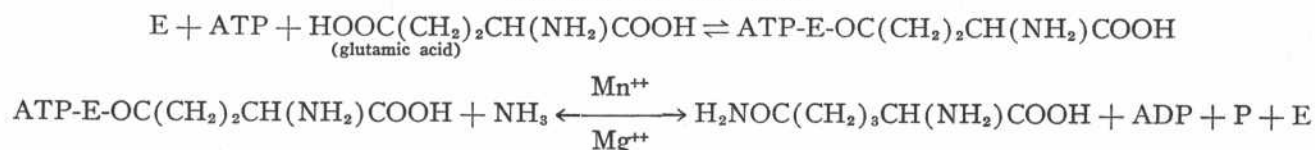
In the synthesis of glutamine in which glutamic acid and ammonia are condensed to form glutamine the cleavage of ATP is different. The products are not AMP and pyrophosphate but ADP and inorganic phosphate.

Glutamine Exchange Reactions

In exchange reactions between the amide nitrogen of glutamine and free ammonia (or hydroxylamine) catalytic amounts of ATP are required.

$$H_2NOC(CH_2)_2CH(NH_2)COOH + \begin{array}{c} \overset{*}{N}H_3 \\ \text{or} \\ NH_2OH \end{array} \rightleftharpoons \begin{array}{c} H_2\overset{*}{N}OC(CH_2)_2CH(NH_2)COOH \\ \text{or} \\ HOHNOC(CH_2)_2CH(NH_2)COOH \end{array} + NH_3$$
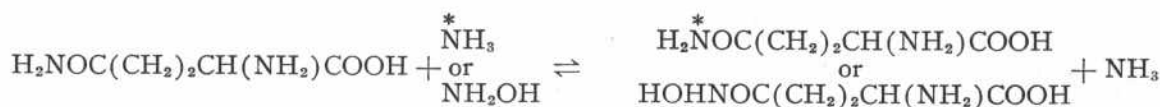
Synthesis of Glutathione

So far we have dealt only with quasi-peptide bonds; in glutathione a genuine peptide bond is synthesized. Its synthesis occurs in two steps. Glutathione is glutamylcysteinylglycine. The first step is again the condensation of the γ-carboxyl of glutamic acid with the amino group of cysteine to form cysteinylglycine. Again the reaction is driven by energy derived from the breakdown of ATP into ADP, and inorganic phosphate. There is no CoA requirement.

First step

$$HOOC(CH_2)_2CH(NH_2)COOH + H_2NCH(CH_2SH)COOH + ATP + E_1$$
$$\text{(glutamic acid)} \qquad\qquad \text{(cysteine)}$$

$$\rightleftharpoons HOOCCH(CH_2SH)NHOC(CH_2)_2CH(NH_2)COOH + ADP + P + E_1$$
$$\text{(glutamylcysteine)}$$

Second step

$$HOOCCH(CH_2SH)NHOC(CH_2)_2CH(NH_2)COOH + H_2NCH_2COOH + ATP + E_2$$
$$\rightleftharpoons HOOCCH_2NHOCCH(CH_2SH)NHOC(CH_2)_2CH(NH_2)COOH + ADP + P + E_2$$
(glutathione)
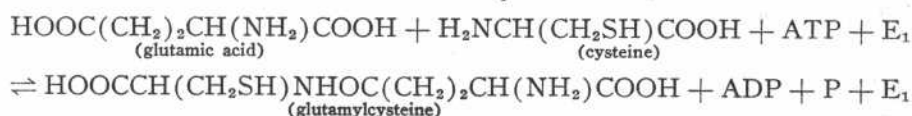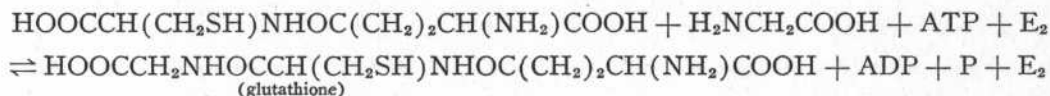
In the second step we come for the first time to the synthesis of a peptide bond comparable to those in proteins. The reaction is glutamylcysteine plus glycine to give glutamylcysteinylglycine. The enzyme for this step is different from that which catalyzes the first step. ATP is again the source of the energy; it is split in this activation reaction to give ADP and inorganic phosphate. No free phosphorylated intermediates have been found, and one wonders how ATP contributes its activation energy. In the case of the synthesis of hippuric acid the intermediate has been identified, it is benzoyl CoA; in none of the cases have phosphorylated intermediates so far been found. One is inclined to think that the negative evidence is significant. It is dangerous, as every experimenter knows, to draw conclusions from a negative finding, but the methods are very sensitive, and also there is indirect evidence that there are no free phosphorylated intermediates in these cases, that the intermediate is on the enzyme.

A general way of testing for the formation of activated carboxyl groups is by means of hydroxylamine. Activated carboxyl groups form hydroxamates which give a color reaction, or one can isolate them. If it is an amino acid hydroxamate which is formed, and its formation is dependent upon ATP in the same way as the synthesis of the peptide under consideration, one is on fairly safe ground in considering that activation of the carboxyl group of the amino acid is involved in the synthesis. In all of these reactions, synthesis of hippuric acid, of pantothenic acid, of glutamine, of both steps in the synthesis of glutathione, one can demonstrate by the formation of the corresponding hydroxamate that there is an activation of the carboxyl group.

Carboxyl Group Activation*

$$(C_6H_5CO)_2OPO(OH) + H_2NCH_2COOH \rightleftharpoons$$
(dibenzoyl phosphate)

$$C_6H_5CONHCH_2COOH + C_6H_5COPO(OH)_2$$
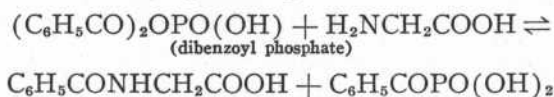
* Reaction goes rapidly at 37° and pH 7.4. Monobenzoylphosphate is inactive. N-phosphoglycine is inactive.

The direct proof that it was the carboxyl group which needs to be activated and not the amino group was provided by Chantrenne long before this story began. He synthesized dibenzoyl phosphate and found that, simply by warming it with glycine at 37 degrees, hippuric acid was formed without an enzyme. On the other hand N-phosphorylated glycine heated with benzoic acid does not go. This was the first clear-cut evidence that it is the carboxyl group and not the amino group which needs to be activated. All of the later enzymatic work has confirmed this general conclusion.

Table 18. FIVE PEPTIDE SYNTHESES

	CoA required	ATP required	ATP → AMP + P P	ATP → ADP + P	COOH activated
Hippuric acid	+	+	+	—	+
Pantothenic acid	—	+	+	—	+
Glutamine	—	+	—	+	+
Glutamyl-cysteine	—	+	—	+	+
Glutamyl-cysteinyl-glycine	—	+	—	+	+

Table 18 is a summary of what has been said so far. Of the 5 different quasi-peptide bonds whose enzymatic synthesis has been studied in vitro, CoA is required in only 1, ATP is required for all 5, the splitting of ATP into AMP and pyrophosphate occurs in only 2, into phosphate

and ADP in the other 3, and the correlative evidence indicates that carboxyl group activation is common to all 5.

About a year ago, in order to get on further with the real problem, which is the synthesis of protein, I felt that some hypothesis had better be made regarding the mechanism of protein synthesis based upon what we knew, and in the hope, then, that the picture would lead us to some intelligent experiments.

Amino Acid Activation

$$E + RCH(NH_2)COOH + ATP \rightleftharpoons E\text{-}NH\text{-}CH(R)COOAMP + P\infty P$$
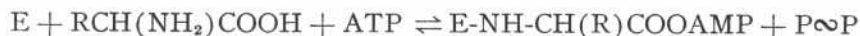
This hypothesis was made as clear, as definite and unequivocal as possible, in order that it could be proved wrong as quickly as possible. It was built on the basis of what was known about the synthesis of small peptides and quasi-peptides and of the little general information we had of the synthesis of proteins.

In this hypothesis the first step is the activation of the carboxyl group of free amino acids. The next step is the transport of the amino acids so activated to something we call a template (just a word) where the pattern is made. It is fashionable, it is more than fashionable, it is reasonable to take it, that the place where the pattern is made is on nucleic acid. The logic of the situation calls for a place where the pattern of the protein molecule is made, and there is evidence that nucleic acids are involved in protein synthesis. We might as well, then, go the whole hog and say that the place where the pattern is made, the template, is nucleic acid.

In this simple hypothesis no place or function is assigned to peptide intermediates. We used to think it more reasonable that amino acids were first linked into small peptides and then these were arranged into a specific pattern and eventually joined to form a protein molecule. All the evidence is against this. With the methods available it would have been easy to find peptide intermediates; nobody had ever found any in spite of a lot of hard looking by everybody concerned. So we say there are not any peptide intermediates. This may be wrong.

Once the amino acids are on the template, they do not need to have but may have an enzyme that zippers them off. The carboxyl groups are already activated, all that is needed is some mechanism to bring the amino groups close enough to the carboxyl group which, energetically, is ready to react. Peptide bonds are thus formed. Because the nucleic acids are rather shallow spirals, and peptides are much tighter spirals with a much steeper pitch, it could very well be that this alone is enough to take the protein off the nucleic acid

template. Specific proteins, such as antibodies, on the basis of this picture, would be formed at the time the protein is coming off the template. The antigen directs the folding as the protein comes off the template. Then one depends upon local pH and salt effects to account for the dissociation of the antigen from the antibody so that the antigen can function in this manner again. In the case of adaptive enzyme formation the assumption is that there is a template for every enzyme that the cell can make. In those cases where the substrate is not present the template is filled and the process stops at this stage. When the substrate appears it combines with the protein on the template and by so doing takes it off. This process continues until there is sufficient enzyme to be detected.

E-Amino Acid-AMP complex is stable. Separate enzymes are required for each amino acid. There is no evidence yet of free Amino Acid-AMP. Problem: How is activated amino acid transferred to the template?

Every cell contains enzymes that activate amino acids according to the first step in our hypothesis. This finding at first seemed to be a brilliant confirmation of the basic hypothesis, but as we shall see in a moment, I think we had better go slow. There is evidence that a complex is formed involving the enzyme AMP and the amino acid. Pyrophosphate is split out. If one adds labeled pyrophosphate, one can demonstrate its interchange with pyrophosphate in the free ATP; for such an exchange an amino acid is required for which the enzyme is specific. No such exchange occurs if the amino acid is absent. AMP on the other hand does not promote pyrophosphate exchange, nor does it exert any inhibition. Yet there is no measurable net cleavage of ATP. This is strange. To get a net breakdown of the ATP it is necessary to break the enzyme-AMP-amino acid complex with hydroxylamine in high concentration. The amino acid hydroxamate is formed, which may be viewed as a quasi-peptide. There appears to be a differ-

ent enzyme for each amino acid. If carboxyl group activation of amino acids by these enzymes (and ATP) is the first step in protein synthesis, then there must be a physiological analogue of hydroxylamine which carries the activated amino acid onto the template retaining the activation. So far no such transporting substance has yet been found. This raises a doubt whether the amino acid activating enzymes so far found are responsible for the first step in protein synthesis. On the other hand the hypothesis that carboxyl group activation of free amino acids is the first step in protein synthesis is not weakened yet.

An alternative is that the whole enzyme-AMP-amino acid moves over onto or is always part of the nucleic acid template. This seems too difficult. It would require a large enzyme molecule for each kind of amino acid to be fitted on to the RNA molecule so that each amino acid is fitted into its special place to make the specific pattern. This alternative is most unwelcome, because the difficulty of demonstrating it appears to be very great.

The tables below show two processes in which carboxyl groups are activated without involving ATP. In my view these have nothing to do with protein synthesis.

Reactions not Requiring ATP in Which COOH is Activated: Hydrolyses

A-B-C + E

$$A \overset{\downarrow}{+} C$$
$$\overset{|}{E}$$

$+ H_2O$		$\to A + B\text{-}C + E$
$+ NH_2HOH$		$\to A\text{-}NHOH + B\text{-}C + E$
$+ D$		$\to A\text{-}D + B\text{-}C + E$
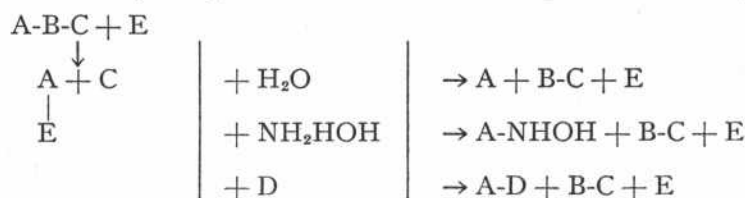
A, B, C, and D are amino acids.

One such process is hydrolysis. Let us consider a polypeptide A-B-C. In the course of its hydrolysis there is formed a compound consisting of A-enzyme. Most, if not all, of the energy in the A-B bond is conserved here. What will happen next will depend upon what is in the solution. There will be a competition for the A-enzyme bond. If water intervenes there will be hydrolysis. If hydroxylamine, the hydroxamate of amino acid A will be formed. If amino acid D intervenes the dipeptide A-D is formed. These three possibilities have been demonstrated to occur during hydrolysis before hydrolysis is complete.

Reactions not Requiring ATP in Which COOH is Activated: Transfer Reactions

Glu.-CONH$_2$ + E
(glutamine)
$\uparrow\downarrow$

Glu.COE + NH$_3$

$+ \overset{*}{N}H_3$		$\to Glu.CO\overset{*}{N}H_2 + NH_3$
$+ NH_2OH$		$\to Glu.CONHOH + NH_3$
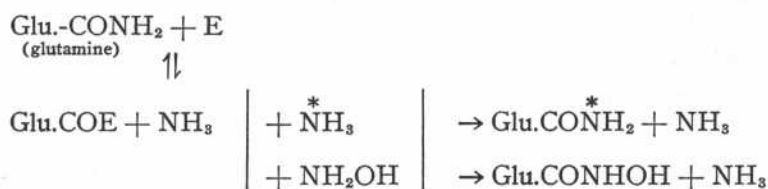
An analogous exchange occurs in transfer reactions, as for example between the amide nitrogen of glutamine and either free ammonia or hydroxylamine. In these exchange reactions there is no net increase in peptide bonds.

We have postulated that the net increase in peptide bonds which occurs in protein synthesis is obtained through activation of the carboxyl groups of amino acids at the expense of ATP. What is the evidence which leads one to believe that this view of the first stage is even approxi-

mately right; that ATP is required, that it is the energy source, and that there is carboxyl group activation? Greenberg and Peterson found that liver mitochondria show an ATP dependence, at least an acceleration by ATP, in the incorporation of amino acids in protein. Siekevitz and later Zamecnik and his coworkers confirmed and extended this observation.

That nucleic acids were directly implicated in protein synthesis was first suggested by Caspersson and Brachet, who collected a great deal of cir-

cumstantial evidence to support this suggestion. In more recent times more indirect evidence has been obtained. The microsomes in liver which have the highest content of RNA of any liver component also incorporate amino acids the fastest. Similarly with immature red cells, amino acid incorporation and the RNA content go parallel. The same was found in bacteria. The ultraviolet inhibition spectrum of protein synthesis in yeast is that of nucleic acid and not of protein. More direct evidence was obtained from the use of RNAase. When RNA was depolymerized by this enzyme, protein synthesis was inhibited. This was observed in particles from *Staphylococcus aureus,* from particles derived from *Micrococcus lysodeikticus,* in onion root tips, and in microsomes. The most direct evidence is in the work of Dr. Fraenkel-Conrat, which we heard this afternoon. It is the RNA and not the protein of the infecting particle which determines the kind of protein which is synthesized.

It must have been a little surprising to find in this work that DNA is not directly involved in protein synthesis. The most active system in a cell in the synthesis of protein is not the nucleus but the microsome fraction that contains no DNA, but is richest in RNA. Various means which destroy DNA leave the protein synthesizing mechanism intact. These findings need to be reconciled with the fact that the amino acid pattern in proteins is an inherited characteristic and that DNA and not RNA is in the genetic apparatus. One might say that DNA is involved in making the specific RNA template and it is on such a template that protein is synthesized.

Template: Incompletely Fitted
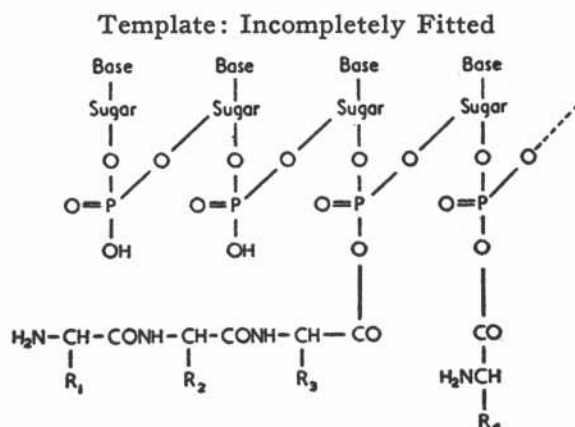
This figure is a naive picture of the template. It is a picture of a polynucleotide with a small peptide attached by a covalent bond to the phosphate. It is in this bond that the activation energy of the carboxyl group is conserved. There is another amino acid attached to the template. The peptide is shown dangling from the template; it takes time for the template to be filled. The amino group of the adjacent amino acid, if brought close enough, will react with the carboxyl group attached to the template and so make the next peptide bond, and then this peptide will dangle until the protein is completely made. There may be an enzyme that brings the amino group close enough to combine with the carboxyl group, but it may not be necessary. To envisage a template only partly filled helps to explain some facts otherwise difficult to explain—for example, the findings of Anfinsen and his collaborators that there may be unequal labeling of an amino acid in different parts of the protein molecule. If there are no significant concentrations of peptide or other intermediates, and if the amino acids go into protein directly from the free amino acid pool, there should be no unequal labeling unless there is already some of the amino acid (unlabeled) on the template. This amino acid might exchange with that in the pool so long as it is not tied up in a peptide bond.

Adaptive enzyme formation has been used as a means of studying protein synthesis. What is known about this phenomenon is in accord with the picture given: with the inducer there must be present all the necessary amino acids, an energy source, and the template. Dr. Hotchkiss' experiments show that DNA leads to the formation of the template, which presumably is RNA. Adaptive enzyme formation is inhibited if ATP formation is inhibited or if the template is masked by amino acid analogs.

Table 19. FACTORS WHICH STIMULATE HEMOGLOBIN SYNTHESES IN VITRO

Amino acid mixture	Iron 5 μgm./ml.	Glucose 1 mg./ml.	Siderophilin 50 μgm./ml.	Rate of hemoglobin synthesis
—	—	—	—	100
—	+	—	—	100
—	+	—	—	110
—	+	+	—	111
+	+	—	—	170
+	+	—	—	351
+	+	+	—	170
+	+	+	—	497
+	+	+	+	597

We have been studying hemoglobin synthesis in vitro. Table 19 summarizes the different substances which we have found to stimulate the synthesis of hemoglobin. Iron occupies a central posi-

tion in this story. This effect of iron was discovered by Dr. Fischer while he was with us. The addition of iron alone has no stimulating effect. Neither has glucose, nor iron plus glucose. Amino acids plus iron have a synergistic effect. Amino acids plus glucose act as if only amino acids had been added. Amino acids plus iron plus glucose have a synergistic effect. Proper chelating agents of iron are stimulating, when added with amino acids and iron. Such chelating agents as fructose-amino acids are stimulating with low concentrations of iron.

It has been found in experiments in animals that an indispensible amino acid must be fed with all the other amino acids and sugar for the indispensible amino acid to make itself felt. Table 19 shows an analogous effect in hemoglobin synthesis in vitro. All the foregoing observations are in accord with the general hypothesis regarding protein synthesis.

Hemoglobin is a conjugated protein consisting of globin and heme. What is the relation between the rates of synthesis of these two very dissimilar substances? On chemical grounds a priori one would expect the rates to be very different, and accordingly that different stimulating agents would stimulate these dissimilar processes differently.

Table 20. STIMULATION OF HEME AND OF GLOBIN SYNTHESES BY IRON

	Rate of syntheses of	
	heme	*globin*
Blank	100	100
+ Fe : 5 μgm./ml................	165	183

Table 20 shows that iron stimulated the synthesis of heme. This was to be expected because iron is a constituent of heme. It was not to be expected that iron would also stimulate the synthesis of globin and to about the same extent.

Table 21. COMPARISON OF CONCURRENT RATES OF SYNTHESES OF HEME AND OF GLOBIN LABELED AMINO ACID: GLYCINE

Duration of experiment	Heme synthesis	Ratio of rate of glycine incorporation into globin to rate of heme synthesis
	mM/M Heme	
1 hour	1.67 — 2.87	1.14 ± 0.24
2 hours	2.48 — 7.51	0.98 ± 0.15
4 hours	2.53 — 9.36	1.06 ± 0.14

Table 22. RELATIVE RATES OF INCORPORATION INTO GLOBIN OF LABELED GLYCINE, HISTIDINE, LEUCINE, AND LYSINE

Amino acid pair	Ratio of incorporation rates
His/Gly	1.09 ± 0.35
Leu/Gly	1.01 ± 0.27
Lys/Gly	0.84 ± 0.18
His/Leu	0.98 ± 0.05

Tables 21 and 22 show that under a wide difference in rates of synthesis of hemoglobin, produced in a number of different ways, the rates of synthesis of heme and of globin were the same. This would not have been observed under our experimental conditions, where the measure of the rate was the C^{14} labeling, if the globin and the heme had been synthesized at different rates, or if there had been in the red cell any significant accumulation of unlabeled globin or heme or of unlabeled precursors other than free amino acids. Evidently there is a mechanism within the reticulocytes which equalizes the rates of synthesis of heme and of globin.

In view of the case of hemoglobin synthesis, the question arises regarding the synthesis of other conjugated proteins; is there a similar self-regulation in the synthesis of virus nucleoprotein so that there is no accumulation of the parts?

The hypothesis proposed regarding the synthesis of protein suggests the following as a possible mechanism whereby the globin and heme of hemoglobin are synthesized at equal rates. Globin is synthesized on but cannot separate from its template by itself. The separation occurs when heme combines with the globin. Then the template is freed for the synthesis of another molecule of globin. Further the presence of the free heme inhibits the formation of any more. Its removal by combination with globin permits the synthesis of more heme. The same general mechanism is applicable to adaptive enzyme formation. In this picture it is inferred that all the templates for making all the enzymes that a cell can make are present in the cell. They are determined by heredity of the cell. The amino acids and energy source are there. The template is filled and remains occupied until the inducer, the substrate for the adaptive enzyme is produced. Then the protein—i.e., the adaptive enzyme—is removed from the template, and then more of the enzyme can be formed. Dr. Hotchkiss' experiments provide direct evidence in support of this picture of adaptive enzyme formation.

The simple picture I have given of the mechanism of protein synthesis accounts for major phenomena of nitrogen metabolism in animals such as the effects of feeding biologically incomplete proteins—i.e., of feeding a protein dificient in a single amino acid. Other phenomena accounted for are the necessity of feeding carbohydrate (energy source) with protein, and the specific dynamic action of proteins.

DISCUSSION

DR. EDMOND FISCHER: You spoke of a possibility of an exchange of labeled amino acid with unlabeled amino acids in a peptide combined to an "incomplete template." I think Gale has suggested that this could also take place in a complete template, in other words that the "unzipping" reaction could be a reversible process. This would mean that a completely synthesized protein could recombine with its template to exchange an amino acid. What do you thing of this interpretation?

DR. BORSOOK: This is possible. According to the hypothesis proposed, however, the exchange would take place only with activated amino acids—i.e., energy is required for the exchange with free amino acids. Simpson at Yale and more recently Steinberg at the N.I.H. found that the same factors which influence incorporation are precisely those which affect excorporation, and

they do so in the same way. Energy is required for excorporation as for incorporation. If a protein is put back on the template, its amino acids might exchange with activated amino acids, but not with free amino acids. Such an exchange could occur at different loci on the template at different rates. This will account for the unequal labeling but will not account for the fact that this unequal labeling is greatest at the very beginning of the experiment and gets less and less as time goes on. It is because of this fact that I suggest that most of the amino acid exchange on the template occurs before the template is completely filled. When the template is filled, the protein comes off and from then on there will be equal labeling. Putting a protein back on the template must be a pretty slow process, probably slower than amino acids going on, but there must be a balance between these two, and this is the explanation of the remarkable phenomenon of the nitrogen balance.

QUESTION: What do you think of the suggestion that has been made that a nucleic acid polyphosphate might be an activating agent for the amino acid?

DR. BORSOOK: I see nothing against it. I have indicated my reasons for preferring the view of a separate amino acid activating enzyme. We have direct experimental evidence for such enzymes in all cells.

INDEX OF NAMES

Number = page number; a = left column of page; b = right column of page; * = bibliography.

SUBJECT INDEX

Number = page number; a = left column of page; b = right column of page.

Nitrous acid: 37b
Nuclei: 37a
Nucleotides, as group transfer agents: 43b
Nucleus: 8a

Oligo-1, 6-glucosidase: 34a, 39a
Osmotic pressure: 10a
Ovomucoid: 31b
Oxalacetic acid: 8b
Oxalosuccinic: 8b
Oxidative phosphorylation: 8b
Oxytocin: 55a

Pancreas: 46a
Papain: 33b, 36b
Partial specific volume: 10a
Partition factor: 19b, 20a
Pasteurella pestis: 56b
Penicillin: 63b, 64a
Pepsin: 36b
Peptidases of Binkley: 44a
Peptide: 5a
Peroxidases: 45a
Perrin-Fuerster theory: 37b
Perrin's equations: 11a
Phenylisocyanates: 37b
Phosphatase: 42a
3-Phosphoglyceraldehyde: 42a
Phosphorus: 5a, 5b
N-Phosphoglycine: 77b
Phosphorylase: 42a, 44a
Phosphotransacetylase: 42a
Photochemical reactions: 37b
Pituitary: 55a-60b
Pneumococci: 61b, 63a, 63b
Prolactin: 55a
Proteases: 44b, 62a
Proteins
    biological synthesis: 75a-82b
    classification: 4a-4b
    3-dimensional topography: 5a
    osmotic effect: 6a
    partition column chromatography: 18a-29b
    role in biology: 3a-9b
    semiconductors: 37b
    solvent systems: 22a-23a
    synthesis: 44b
Protoplasm: 7b, 8a, 8b
Pyridoxal phosphate: 40a
Pyruvate: 8b, 41b
Pyruvate dehydrogenase: 40b
Pyruvic oxidase: 40a, 41a
PZ protein of Monod: 44b

Raffinose: 15b
Refractive index gradient: 13a
Refractive index increment: 12a, 14b
Reproduction: 8b
Resonance of peptide: 37a
Ribonuclease: 33a, 34a, 36b, 39a, 62a
Ribonucleic acid: 44b, 69b
Ribose-5-phosphate: 41b
Rotary diffusion: 11b
    constant: 11b

Sarin (isopropyl methyl phosphonofluoridate): 38a
Secondary scattering, light: 15a
Sedimentation constant: 12a

Sedimentation equilibrium: 10a
Sedimentation velocity: 10a, 11b, 13b
Sedoheptulose-7-phosphate: 41b
Serine: 34a
O-Serine phosphate: 38a
Serum albumin: 15a, 31a, 33a
Shape factor: 10b
Simha's equation: 11a
Somatotropin: 55a
L-Sorbose: 39b
Spinco ultracentrifuge: 11b, 14a
Staphyloccocus aureus: 80a
Steam engine: 3a, 3b
Stokes law: 10b
Streptomycin: 63b, 66a
Substrate specificity: 38a
Sucrose: 39b
Sucrose phosphorylase: 39b
Sulfenyliodide: 32a
Sulfhydryl group: 32a, 32b, 33a, 34b, 35a, 39a
Svedberg equations: 11a
Synthesis
    glutamine: 76
    glutamylcysteine: 77
    glutathione: 76
    hemoglobin *in vitro*: 80b
    pantothenic acid: 76

Template: 44b, 78a, 80b
Tetraalkyl pyrophosphate: 38a
Tetrahydrofolic acid: 40a
"Theoretical plate": 18a, 20b
Thermodynamics: 3a, 10a
Thiamine pyrophosphate: 40a, 40b, 41b
Thioclastic action: 40b
Thiol: 37a
Thyrotropic hormone: 55a
Tobacco mosaic virus: 32a, 34b
    alkaline treatment: 70b
    chymotryptic digestion: 70a
    detergent treatment: 70b-71b
    electron microscopy: 71a
    end-groups: 70a-70b
    genetic implications: 72a
    hydrazinolysis: 70a
    mixed reconstituted: 72b
    molecular weight: 69a
    mutants: 71b, 72a
    nucleic acid content: 69a
    progeny of reconstituted virus: 72b-73b
    protein content: 69b
    reconstitution: 70b-73a
    reconstituted
        electron microscopy: 71a-71b
        molecular weight: 71b
        size: 71b
    resolution to nucleic acid and protein: 70b-71a
    serology: 73b, 74b
    sub-units: 70a
    threonine content: 69b-70a
    tryptic digestion: 70a
Toxin: 7a
Transfer reactions not requiring ATP in which car-
    boxyl group is activated: 79
Transforming factors, bacterial: 61a-67b
    chemical nature: 62a-63a
    "double marked strains": 65b