

1936

Max Bergmann, 1935

The Rockefeller University

Follow this and additional works at: <http://digitalcommons.rockefeller.edu/harvey-lectures>

Recommended Citation

The Rockefeller University, "Max Bergmann, 1935" (1936). *Harvey Society Lectures*. 20.
<http://digitalcommons.rockefeller.edu/harvey-lectures/20>

This Book is brought to you for free and open access by Digital Commons @ RU. It has been accepted for inclusion in Harvey Society Lectures by an authorized administrator of Digital Commons @ RU. For more information, please contact mcsweej@mail.rockefeller.edu.

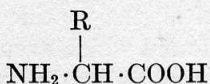
PROTEINS AND PROTEOLYTIC ENZYMES¹

DR. MAX BERGMANN

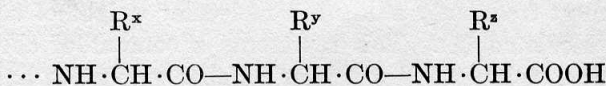
Associate Member of The Rockefeller Institute for Medical Research, New York

OUR knowledge regarding proteins has developed much more slowly than in the case of many other substances of physiological importance. The discovery of vitamins and hormones which occur in plants and animals in fantastic dilutions has repeatedly been followed in a short time by the elucidation of their structure and by their synthesis. There are many proteins which occur abundantly, and they have constantly been investigated since the beginnings of organic chemistry. However, up to the present time, the structure of no single protein is known. The reason for our incomplete knowledge becomes apparent if one examines the general molecular structure of the proteins.

The proteins are substances the molecules of which are built up of hundreds, or even thousands, of amino acids. The amino acids are linked in proteins in the following manner: The carboxyl group of *one* amino acid has reacted with the amino group of the *next* amino acid with the elimination of water. The resulting linkages between the amino acids, the CO-NH groups, we call peptide bonds.



General formula of amino acids



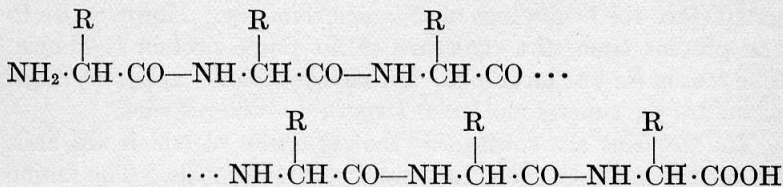
General formula of proteins

(R, R', R'', etc. = Side chains)

¹ Lecture delivered October 17, 1935.

Every protein is an extended, chain-like molecule which has at one end an amino group, at the other end a carboxyl, and, in addition, hundreds of peptide bonds. Between every two peptide linkages there is always a CH group, which carries the so-called side chain. Thus, the outstanding characteristic of the protein molecule is the hundred-fold repetition of analogous groups, peptide bonds, and side chains. The chemist must find his way in this bewildering abundance of similar atomic groups in order to clarify the molecular structure of a protein.

Protein chemistry would be simple if there were proteins which are composed of only one type of amino acids: for example, entirely of leucine residues.



Formula of an ideal protein, not found in nature

Then, all the side chains and the majority of the peptide bonds would be equivalent. It would be necessary only to determine the molecular weight and to discover which amino acid is formed upon the hydrolytic scission of all the peptide linkages. However, so far no protein has been found which even approaches this simple scheme. The metabolism of living matter, which usually prefers the simplest possibilities, evidently does not have any use for such ideal proteins.

How the actual proteins appear may be shown by the analyses of several examples (table 1). For gelatin, one of the proteins most frequently studied, a molecular weight of 35,500 has been calculated (1). This represents a content of about 360 amino acid residues and a similar number of peptide linkages. About fifteen different amino acids have been obtained as split-products from the hydrolysis of gelatin. A similar multiplicity of constituents has been found in the case of other proteins. The amounts of the different amino acids vary widely in the individual

proteins. The fact that no arithmetical regularity could be discerned has always been taken to indicate the extremely complicated structure of the proteins.

It was easy to recognize from a comparison of various proteins that they show wide differences in the content of the individual amino acids. Thus, for example, egg albumin gives, according to

TABLE 1

*The amino acid content of gelatin, casein, egg albumin, and silk fibroin**

	GELATIN	CASEIN	EGG ALBUMIN	SILK FIBROIN
Glycine.....	25.5	0.4	0.0	40.5
Alanine.....	8.7	1.8	2.2	25.0
Valine.....	0.0	7.9	2.5	
Leucine-isoleucine.....	7.1	9.7	10.7	2.5
Aspartic acid.....	3.4	4.1	6.2	
Glutamic acid.....	5.8	21.8	13.3	
Hydroxyglutamic acid.....	0.0	10.5		
Serine.....	0.4	0.5		1.8
Proline.....	9.5	8.0	3.6	1.0
Hydroxyproline.....	14.1	0.2		
Phenylalanine.....	1.4	3.9	5.1	1.5
Tyrosine.....	0.3	6.5	4.0	11.0
Cystine.....	0.2	0.3	0.9	
Arginine.....	9.1	5.2	6.0	
Histidine.....	0.9	2.6	2.3	
Lysine.....	5.9	7.6	3.8	
Tryptophane.....	0.0	2.2	1.3	
Totals.....	92.3	93.2	61.9	83.3

* Mitchell and Hamilton, *The Biochemistry of the Amino Acids*, New York, 1929, American Chemical Society Monograph Series, The Chemical Catalog Co., Inc.

the best analyses available, no glycine on hydrolysis (2); gelatin gives 25.5 per cent (3); and silk fibroin, as much as 40.5 per cent (4). Until recently, one of the principal tasks of protein chemistry was the determination of the percentages in which the individual amino acids could be obtained from the various proteins. There were developed by Fischer, Kossel, Dakin, Van

Slyke, and other investigators special, unusually artistic methods to differentiate the complicated mixture of similar amino acids found in protein hydrolysates. With this analysis of the constituents and with the proof that they were linked in the proteins by means of peptide bonds, the chemical investigation of the proteins had reached the limits of its possibilities. To step from the analysis of the constituents to the analysis of a protein molecule was impossible. We have no method to establish the structure of molecules which are composed of tens of thousands of atoms.

There can be no doubt that it is most desirable to overcome this difficulty and to seek methods which will permit an insight into the molecular constitution of the individual proteins. Many fundamental problems in biochemistry, physiology, and pathology require a more exact knowledge of the structure of the individual proteins.

The problem of the structure of a protein is identical with the problem of the nature and sequence of the side chains. The side chains of the protein molecules may be aliphatic, aromatic, or heterocyclic; they may contain a hydroxyl or sulfur; or they may be neutral, basic, or acidic. According to the distribution of these side chains, variations will be observed in the polar character of the molecule, in the solubility, in the colloidal behavior, and in many biological properties.

There is the complicated process of protein digestion. Here the individual proteins behave differently. Often proteins are only partially degraded in metabolism, and the primary digestion products have a very different value for the activities of the tissues and cells. Apparently, the different arrangement of the constituents plays a decisive rôle. It is essential to know this in order to understand the digestion, metabolism, and nutritional value of the individual proteins.

Or let us examine the crystalline pepsin of Northrop (5). Up to the present, the only constituents which have been found in it were amino acids; it appears, therefore, to be a protein. What is the basis for the fact that the protein pepsin possesses enzymic properties while other proteins, composed of the same amino acids, lack these enzymic characteristics? The same question may be

posed in regard to the hormone character of insulin. All evidence points to the fact that the enzymic properties of pepsin, as well as the hormonal properties of insulin, are due to particular arrangements of amino acids in these proteins. Or let us take as another example the sharply defined specificity of the serological reactions of proteins and protein constituents. The well-known experiments of Landsteiner (6) leave no doubt that in these reactions the arrangement of the amino acids in the peptide chain is of great significance.

The enzymes and serological phenomena have been mentioned not merely to illustrate the importance of the protein problem; I hoped also to indicate that the metabolism of living matter is able to differentiate between various proteins and peptide linkages by means of chemical reagents of high specificity. We must learn to imitate the specificity of metabolic processes and to develop experimental methods of high selectivity, which will enable us to differentiate between the analogous linkages and split-products of the large protein molecules. It is necessary, further, to study the specific protein reagents occurring in nature, the proteolytic enzymes, in order to learn their specificity and to apply them to the investigation of proteins.

For this reason, the first part of my discussion will be concerned with the specificity and reaction mechanism of proteolytic enzymes.

Our first knowledge regarding the specificity of the proteolytic enzymes we owe to Willstätter, Waldschmidt-Leitz, and Grassmann. In accordance with the findings of these investigators, the proteolytic enzymes are classified into three principal groups: The enzymes such as pepsin, the trypsin, and papain, which hydrolyze genuine proteins to more or less large split-products, are called proteinases; the smallest peptides, the dipeptides, are split by the enzyme dipeptidase into two amino acids; the intermediate products between proteins and dipeptides are attacked by the polypeptidases (table 2).

This schematic classification expresses the fact that protein digestion passes through three stages, for each of which there are appropriate enzymes. Thus, the proteolytic enzymes, like other

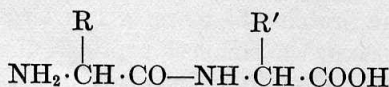
types of enzymes, have a high degree of specificity. Fortunately, the investigation of proteolytic enzymes has given us a great deal more information about their specificity than has been the case with the other types of enzymes. We already know some of the chemical and physical properties of the substrate, which are specific requisites for the action of the individual enzymes. Moreover, it has been possible in several cases to develop, on the basis of these specificity requirements, a general picture of the mechanism of the reaction between enzyme and substrate.

Let us begin with the simplest enzyme, dipeptidase, which is present, for example, in the erepsin of the small intestine, in the kidneys, in the liver, and in the cells of yeast. A number of ex-

TABLE 2

ENZYME		SUBSTRATE
Proteinases	Pepsin Tryptic proteinases Papain Cathepsin	Genuine proteins
Polypeptidases	Carboxypolypeptidase Aminopolypeptidase	Polypeptides
Dipeptidase		Dipeptides

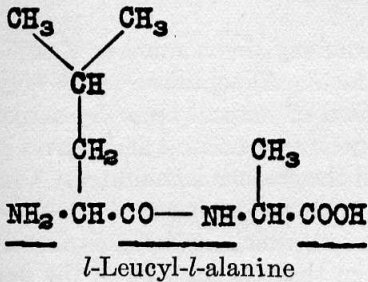
periments with synthetic substrates has shown that dipeptidase splits only those substances which have a hydrogen in the peptide bond, a free carboxyl in α -position to the peptide nitrogen, and a free amino group in α -position to the peptide carbon. Dipeptidase requires, therefore, the following atomic arrangement.



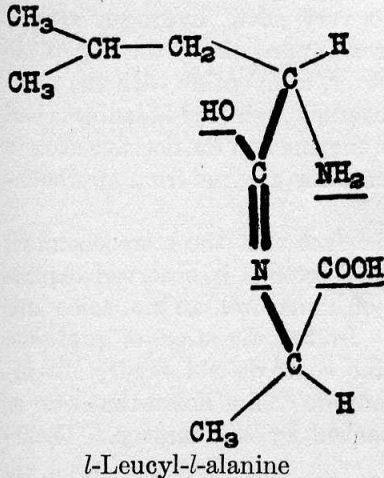
In order to connect the nature of this atomic complex with the specificity requirements of the enzyme, it should be recalled that the several necessary groups have a definite electropolar character, and that the enzyme itself as a protein must also have electropolar groups. When the enzyme acts on the dipeptide, it approaches the substrate and there results a sort of polyaffinity between the polar groups of the enzyme and the substrate. The

polar groups of the enzyme act on the polar groups of the substrate and force the latter to assume a definite spatial position. By the combination of dipeptidase and dipeptide at several points of the molecules, the dipeptide is inhibited in its intra-molecular mobility and held fixed in a definite spatial arrangement. This spatial position is such that the amino and carboxyl groups lie close together and that the peptide hydrogen has wandered from the nitrogen to the oxygen. There results a kind of hexagon the corners of which are formed by the carboxyl, the α -carbon, the peptide-nitrogen, the peptide-carbon, the α -carbon, and the amino group. The six corners of the hexagon lie nearly in a plane. However, the two α -hydrogen atoms and the two side chains are not situated in the plane of the hexagon (7, 8).

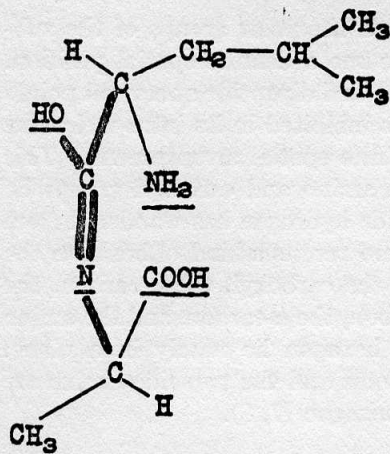
Action of Dipeptidase



When the enzyme approaches, this dipeptide is transformed into the spatial arrangement shown below



The enzyme finds no obstacle in its approach to the underlined groups from this side

*d*-Leucyl-*l*-alanine

The enzyme is inhibited in its approach to the underlined groups

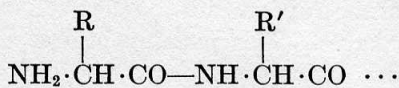
This simple stereochemical scheme explains a number of well-known but hitherto inexplicable facts. Dipeptidase splits only those dipeptides which are composed of natural optically active amino acids. As soon as the antipode of a natural amino acid is present in a dipeptide, the action of the enzyme is inhibited. Our scheme shows the dipeptide of two natural amino acids, such as *l*-leucyl-*l*-alanine, in the arrangement which it assumes in enzymatic hydrolysis. It will be seen that on one side of the flat molecule there are only the two very small hydrogen atoms. Nothing prevents the enzyme from approaching the dipeptide from this side. Let us take, however, a dipeptide with the antipode of a natural amino acid, for example, *d*-leucyl-*l*-alanine; then the isobutyl group lies between the enzyme and the decisive atomic groups of the dipeptide and prevents the enzyme from approaching the substrate.

It is necessary to emphasize the fact that the stereochemical influence of the isobutyl and similar groups is observed experimentally only when dipeptides of unnatural amino acids are subjected to dipeptidase action. In the digestion of proteins, such dipeptides of unnatural amino acids do not occur. In dipeptides of natural amino acids the side chains lie on that side of the molecule which is not approached by the enzyme. There-

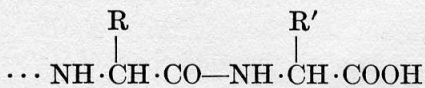
fore, dipeptidase splits all dipeptides of natural amino acids without exception. But it splits only dipeptides, not polypeptides. The enzyme requires a carboxyl and an amino group at a definite distance from each other. The specificity of dipeptidase is determined for the chain length of the substrate. One may also say that dipeptidase is specific for end groups, but not for side chains.

Our stereochemical interpretation of dipeptidase action leads to several unexpected consequences which have been confirmed by experiment (9). This indicates that the theory approximates the actual mechanism of dipeptidase action. It is thus possible for the first time to develop a picture of the complicated specificity phenomena of an enzyme only on the basis of the polyaffinity relationship between enzyme and substrate. Later on we shall see that the theory of polyaffinity is of value for other proteolytic enzymes.

Of the known polypeptidases—aminopeptidase and carboxypeptidase—we need only mention the fact that aminopeptidase requires an amino group and attacks that end of the peptide chain which bears the amino group. Carboxypeptidase requires instead a carboxyl and attacks that end of the peptide molecule which bears the carboxyl. Both enzymes are therefore end-group specific, just as is dipeptidase.

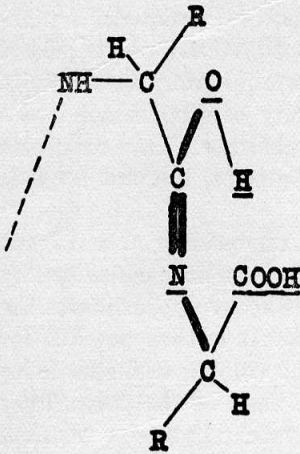


Essential groups for aminopeptidase action



Essential groups for carboxypeptidase action

In the case of carboxypeptidase it was possible to construct a spatial model for the interaction between enzyme and substrate (10).



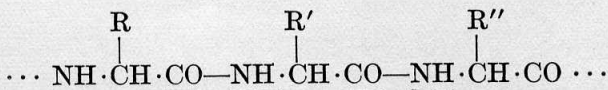
The enzyme finds no obstacle in its approach to the underlined groups from this side

Spatial arrangement of substrate for carboxypeptidase

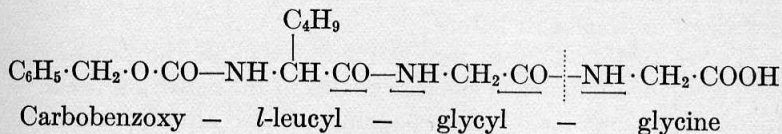
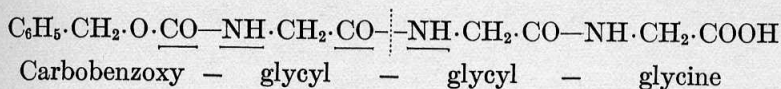
We shall not discuss it in detail. It might be pointed out, however, that this enzyme already shows distinct indications of side chain specificity. This type of specificity is much more distinct in another enzyme which represents a hitherto unknown class of proteolytic enzymes.

In coöperation with Doctors Zervas, Fruton, and Ross (11, 12), we recently demonstrated the presence of a new type of enzyme in the sap of the Papaya fruit. When activated, the enzyme splits peptide linkages. This papainpeptidase does not require for its action a free amino group nor a carboxyl in the substrate, but does require two adjacent peptide bonds. The chemical properties of the active enzyme indicate the presence of an aldehyde group which is indispensable for the enzyme action. It is especially significant that the specificity of this enzyme is independent of the ends of the peptide chain of its substrate.

Papainpeptidase



Essential groups for papainpeptidase action



The specificity is, however, strongly influenced by the nature of the peptide side chains. This may be illustrated by means of a simple example. We compare the action of the enzyme on two similar compounds such as carbobenzoxy-diglycylglycine and carbobenzoxy-*l*-leucylglycylglycine. They differ from each other only in the fact that the peptide chain of the second compound bears a non-electropolar aliphatic side chain. And yet, this slight difference suffices to shift the point of enzyme action. The two compounds are split at different bonds in the molecule as is shown in the illustration by means of the dotted lines.

We encounter here a significant difference between papainpeptidase and the enzymes which we mentioned before. Dipeptidase, aminopeptidase, and carboxypeptidase always hydrolyze their substrates at a very definite point in the molecule, namely at the outermost peptide bond. Papainpeptidase is, in contrast, not confined to a definite point of the chain, but chooses its place of attack under the influence of the nature of the side chains.

If we now recall the classification of proteolytic enzymes according to their specificity, it will be seen that it is necessary to add a new class of peptidases, of which papainpeptidase is a representative. In contrast to the other peptidases, the papainpeptidase resembles the true proteinases in the character of its specificity (table 3).

For all the peptidases it has been possible to synthesize a large number of artificial substrates. By means of these synthetic compounds the specificity of each enzyme could be studied in detail. From these results it was seen that, although each of these enzymes has its individual specificity requirements, they all follow the common rule of polyaffinity. The polyaffinity rule, when applied to the individual enzyme, gives a stereochemical

picture of the reaction between the individual enzyme and its substrate, and accounts for the stereochemical specificity requirements of the individual enzyme. The dipeptidase action has been described on this basis before.

We see further that as we ascend the scale of the proteolytic enzymes, the influence of the side chains in the substrate becomes increasingly important. The papainpeptidase already resembles the true proteinases in that it does not split the peptide chains at

TABLE 3

ENZYME	SUBSTRATE	SPECIFICITY INFLUENCE	
		End groups	Side chains
Proteinases	?	—	+
Papainpeptidase	$\begin{array}{c} \text{R} \\ \\ \text{—CO—NH}\cdot\text{CH}\cdot\text{CO—NH—} \end{array}$	—	+
Carboxypeptidase	$\begin{array}{c} \text{R} \\ \\ \text{—CO—NH}\cdot\text{CH}\cdot\text{COOH} \end{array}$	+	(+)
Aminopeptidase	$\begin{array}{c} \text{R} \\ \\ \text{NH}_2\cdot\text{CH}\cdot\text{CO—NH—} \end{array}$	+	—
Dipeptidase	$\begin{array}{c} \text{R} \qquad \qquad \text{R}' \\ \qquad \qquad \\ \text{NH}_2\cdot\text{CH}\cdot\text{CO—NH}\cdot\text{CH}\cdot\text{COOH} \end{array}$	+	—

the ends but attacks the more central peptide bonds. We may conclude, therefore, from the knowledge of papainpeptidase that the proteinases are strongly affected by the side chains of their substrates.

In the physiological digestion of proteins the first attack by the proteinases is not determined by the length of the protein molecule. A protein of say 360 amino acids is not split into two chains of 180 each nor into four chains of 90 amino acids. On the contrary, the digestion of a protein by a given proteinase is determined by the

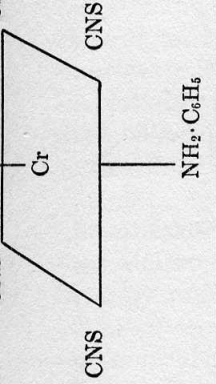
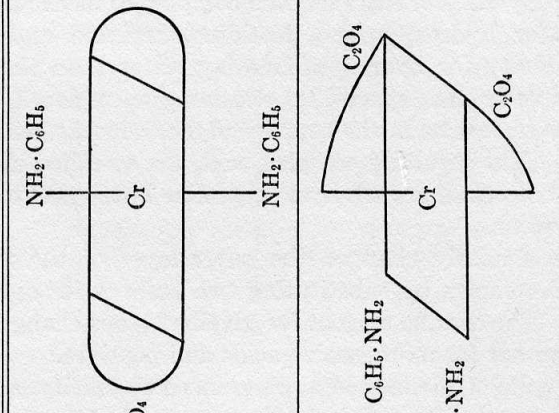
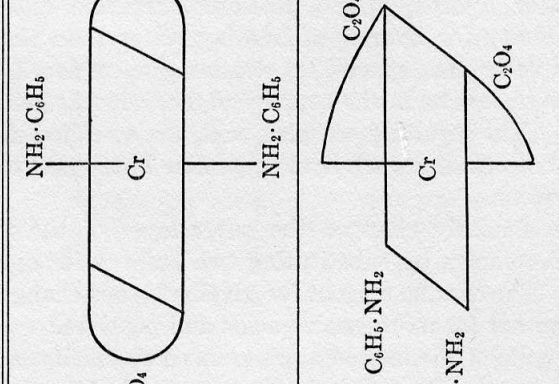
nature and sequence of the component amino acids. Each protein therefore at first gives split-products, which vary for different proteins. Only after the action of the simpler proteolytic enzymes does the individuality of the digested proteins gradually disappear and there result the non-typical amino acids.

We are faced now with the question: How can we apply our knowledge of the specificity of proteolytic enzymes in the study of protein structure? Some of these enzymes give us typical split-products of the proteins. It might be hoped that the isolation and study of these split-products would give us some information regarding protein structure. Such studies have often been attempted. The goal could not be reached, however, because of the difficulty in the isolation of homogeneous substances from the digestion mixture. Here again we encounter the problem which is a characteristic of protein chemistry—that of separating a complicated mixture of analogous substances with unpleasant physical and chemical properties. Such separations are constantly being performed in the metabolism of plants and animals, and the ability to differentiate has reached its highest degree in the specificity of serological reactions. The example of nature shows that it must be possible to find substances which are capable of selectively precipitating single constituents of a mixture of protein split-products. It is obvious that we cannot use nor copy antibodies for this purpose. What we need are compounds of known structure, which possess the necessary selectivity. They must be readily available and permit a wide range of variation in structure.

The search for such selective reagents has shown that several classes of substances fulfill these requirements. The class which we have studied first was the metal complexes (13,14). We have found many of them to give wonderfully crystallized compounds with amino acids and peptides. Another advantage of this type of reagent is the fact that by a slight alteration of the complex it is possible to alter its selectivity. Several such metal complexes are shown in table 4. These few complexes already give us an idea of the mechanism of the selective precipitation of amino acids. First, we have the well-known trioxalato complexes of chromium, iron, and cobalt. Each of them forms a series of analogous com-

TABLE 4

REAGENT		SUBSTANCE PRECIPITATED
Formula	Configuration	
$[\text{Cr}(\text{C}_2\text{O}_4)_3] \cdot \text{K}_3$ $[\text{Fe}(\text{C}_2\text{O}_4)_3] \cdot \text{K}_3$ $[\text{Co}(\text{C}_2\text{O}_4)_3] \cdot \text{K}_3$		Glycine Glycine Glycine
$[\text{Cr}(\text{CNS})_4 \cdot (\text{NH}_3)_2] \cdot \text{NH}_4$ Reinecke salt		Hydroxyproline, proline, histidine, arginine, etc.
	$\text{NH}_2 \cdot \text{C}_6\text{H}_5$	

$[\text{Cr}(\text{CNS})_4 \cdot (\text{NH}_2 \cdot \text{C}_6\text{H}_5)_2] \cdot \text{NH}_4$		<p>Proline, several peptides</p>
$[\text{Cr}(\text{C}_2\text{O}_4)_2 \cdot (\text{NH}_2 \cdot \text{C}_6\text{H}_5)_2] \text{Na}$	<p><i>trans:</i></p>  <p><i>cis:</i></p> 	<p>Amino acids, peptides</p>

pounds with glycine but with no other amino acid. It has exceeded our expectations to find in these simple substances so pronounced a specificity for a single amino acid.

For the composition and the properties of the resulting glycine compounds, it makes no difference whether the complex contains chromium, iron, or cobalt as central atoms, thus demonstrating that the outer layer, and not the central atom, of the metal complex is responsible for the specificity. This fact is of importance for the further development of selective precipitation. It makes it possible to change the selectivity of a complex by changing its outer layer. For instance, the complex Reinecke salt is a precipitant for hydroxyproline, histidine, arginine, and proline (15). Its affinity for hydroxyproline is greater than that for proline. To reverse this specificity relation, we replaced the two ammonia molecules in the complex of the Reinecke acid with two anilines. The resulting complex acid, the so-called rhodanilic acid, does not combine with hydroxyproline but is an excellent reagent for proline.

We similarly changed the outer layer of the trioxalatochromium complex by substituting two anilines for one oxalato residue. The specific reagent for glycine is thus changed to a general precipitant for many amino acids and peptides.

In spite of the lack of a pronounced specificity, such complexes have a well differentiated affinity toward different amino acids, making possible a fractionated isolation.

For the last mentioned complex, the theory provides for two spatial isomers, a *cis* and a *trans* form. Both forms have been obtained, and it is of interest that only the *trans* form precipitates protein split products, the *cis* form being rather unstable. This shows the influence of spatial configuration on the selectivity.

These few examples show how easy it is to find many new reagents for protein split products. For example, the substitution of aniline in our complexes by other amines leads to new specific reagents.

Finally, I should like to mention that in all these cases of selective precipitation we are dealing with salt formation between the complex ion and the amino acid or peptide. It is of some in-



Fig. 1. Proline rhodanilate

terest with respect to biological specificity phenomena to find on very simple models that salt formation can be accompanied by selectivity and specificity.

We first applied the method of selective precipitation to the amino acids of gelatin. We hydrolyzed gelatin completely, and, employing three different metal complexes, we estimated the content of glycine, proline, and hydroxyproline. Proline, for instance, was precipitated by means of rhodanic acid out of the mixture of about 15 amino acids in the gelatin hydrolysate. The purity of the precipitate is striking and demonstrates best the high selectivity which is attainable.

Gelatin has often been used to test new methods of protein

TABLE 5

PROTEIN	AMINO ACID	PER CENT	PER CENT/ MOLECULAR WEIGHT	RATIO
Gelatin	Glycine	25.5	0.34	6
	Proline	19.7	0.17	3
	Hydroxyproline	14.4	0.11	2
	Alanine	8.7	0.10	2
	Leucine-isoleucine	7.1	0.055	1
	Arginine	9.1	0.052	1

analysis. It belongs, therefore, to the few proteins the constituents of which are rather well-known. Table 5 is a combination of the figures obtained by Dakin (3) and by ourselves (14). The complete hydrolysis of water-free gelatin should yield 119 per cent of split-products. Until recently, 92 to 93 per cent had been accounted for. Our method has made it possible to account for 103 per cent.

We have begun to apply this method to enzymatic hydrolysates of proteins, and in certain cases obtained selective precipitates of peptides. Without going into detail regarding these peptides, it might be pointed out that the possibility of isolating homogeneous peptides in large amounts is a good argument for a regularity in the arrangement of the amino acids in the protein molecule.

Let us examine now our knowledge of protein structure in the

light of the information obtained on gelatin with the newer methods. To this end we consult again our table for the content of glycine, proline, and hydroxyproline in gelatin. The figures are given in percentages of weight. In order to find out how many molecules of these three amino acids there are in 100 molecules of all the split products, we must divide the weight percentages by the respective molecular weights. This calculation shows the remarkable result that glycine represents one-third of all the amino acids of gelatin, while proline represents one-sixth, and hydroxyproline one-ninth. The numbers of molecules for the three amino acids, which together represent 60 per cent of the gelatin constituents, are thus found to exist in a simple arithmetical ratio. This relationship can not be fortuitous and must represent a regularity in the structure of the protein itself. Other constituents of gelatin also appear to conform to a similar regularity. But their percentage is relatively small and the consequent experimental error greater. It may be of interest to look for other proteins which show similar regularities. In the case of another scleroprotein—silk fibroin—, glycine, alanine, and tyrosine form more than 60 per cent of the total split-products. These three amino acids occur, according to the best available analysis (4), in the molecular ratio 8 to 4 to 1 which means that in the protein every second amino acid residue is glycine, every fourth is alanine, and every sixteenth is tyrosine. Finally, a long time ago, Kossel and Dakin (16) found that several protamines give 87 per cent of arginine. The calculation has shown that two out of three amino acids in these protamines must be arginine residues.

The relatively simple composition of protamines and the prevalence of arginine as a constituent led to the opinion that the protamines differ from the other proteins in having a relatively simple structure. The protamines were therefore thought to be composed of sets of three amino acid residues, two of which were always arginine. It was later found that the fiber proteins gave x-ray diagrams, and it was attempted to explain these by the periodic repetition of simple structural elements in the proteins. Such hypotheses were disproved by the complicated analytical figures for the proteins. Today, however, it is possible to extend

the Kossel-Dakin hypothesis in a modified form to the more complicated proteins. We do not expect to find a simple elementary constituent which by manifold repetition confers upon the protein molecule a simple periodicity. The analysis of gelatin indicates that each of its predominant amino acids has its individual periodicity, but all of them have the figure three as the common factor. Similarly, in the case of silk fibroin the periodicity of the main amino acids has as the common factor the figure two. Thus, we find in both these proteins the single amino acids occurring in periodicities which are expressed by arithmetical series.

The presence of so complicated a periodicity explains why the problem of x-ray photography is so different for proteins than in the case of substances with simple periodicity.

The immediate task of the chemist is to test the validity of this concept of protein structure with regard to other proteins. The difficulty of this task cannot be underestimated; we hope, however, that with the aid of proteolytic enzymes and the method of selective precipitation we may succeed in penetrating further into the structure of these most complicated substances of nature—the proteins.

REFERENCES

- (1) ATKINS, W. R.: *J. of Leather Trades Chemists*, **17**, 575, 1933.
- (2) OSBORNE, T. B., JONES, D. B., AND LEAVENWORTH, C. S.: *Am. J. Physiol.*, **24**, 252, 1909.
- (3) DAKIN, H. D.: *J. Biol. Chem.*, **44**, 499, 1920.
- (4) ABDERHALDEN, E.: *Z. physiol. Chem.*, **120**, 207, 1922.
- (5) NORTHROP, J. H.: *J. Gen. Physiol.*, **13**, 739, 1930.
- (6) LANDSTEINER, K.: *Die Spezifität der serologischen Reaktionen*, Julius Springer, Berlin, 1933.
- (7) BERGMANN, M.: *Science*, **79**, 439, 1934.
- (8) BERGMANN, M., AND ZERVAS, L.: *Z. physiol. Chem.*, **224**, 11, 1934.
- (9) BERGMANN, M., ZERVAS, L., FRUTON, J. S., SCHNEIDER, F., AND SCHLEICH, H.: *J. Biol. Chem.*, **109**, 325, 1935.
- (10) BERGMANN, M., AND ZERVAS, L.: *Z. physiol. Chem.*, **224**, 45, 1934.
- (11) BERGMANN, M., ZERVAS, L., AND FRUTON, J. S.: *J. Biol. Chem.*, **111**, 225, 1935.
- (12) BERGMANN, M., ZERVAS, L., AND ROSS, W. F.: *J. Biol. Chem.*, **111**, 245, 1935.
- (13) BERGMANN, M., AND FOX, S. W.: *J. Biol. Chem.*, **109**, 317, 1935.
- (14) BERGMANN, M.: *J. Biol. Chem.*, **110**, 471, 1935.
- (15) KAPFFHAMMER, J., AND ECK, R.: *Z. physiol. Chem.*, **170**, 294, 1927.
- (16) KOSSEL, A., AND DAKIN, H. D.: *Z. physiol. Chem.*, **41**, 414, 1904.