1954

Studies on the Structure of Pepsin : An N-Terminal Sequence and the C-Terminal Groups

John Martin Passmann
Loyola University Chicago

STUDIES ON THE STRUCTURE OF PEPSIN: AN N-TERMINAL SEQUENCE AND THE C-TERMINAL GROUPS

by

John Martin Passmann

A Thesis Submitted to the Faculty of the Graduate School of Loyola University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

June

1954

APPROVAL SHEET

The dissertation submitted by John M. Passmann has been read and approved by five members of the faculty of the Stritch School of Medicine, Loyola University.

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated, and that the dissertation is now given final approval with reference to content, form, and mechanical accuracy.

The dissertation is therefore accepted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

May 21, 1954
DATE

Martin B. Williamson
Signature of Adviser

LIFE

John M. Passmann was born in Chicago, Illinois, January 29, 1927.

He was graduated from St. George High School, Evanston, Illinois in June, 1944.

He entered upon undergraduate studies in September, 1944 at Loyola University, Chicago. These studies were interrupted for a period of eighteen months during which time he served with the United States Navy. Undergraduate studies were resumed in September, 1946 at Loyola University. The degree of Bachelor of Science was conferred in June, 1949.

During the period, June, 1949 to September, 1950, he was employed as an analytical chemist in the control laboratories of a national food manufacturer and retailer.

In September, 1950 he began his graduate studies under a grant from Research Corporation under the sponsorship of Doctor Martin B. Williamson. The degree, Master of Science, was awarded in June, 1952. From June, 1953 to June, 1954 he was the recipient of a Standard Oil Company (Indiana) Fellowship.

He is co-author of the following publications:

"Structure of Pepsin", Fed. Proc., 11, 311 (1952); "The

Terminal Amino Group of Pepsin", J. Biol. Chem., 199, 121 (1952).

Portions of the work in this thesis have been presented by the author at the national meetings of the American Chemical Society in 1952 and 1953.

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to Doctor Martin B. Williamson for the original suggestion of the problem and for his advice and helpful suggestions during the course of the work. To the other members of the faculty of the Biochemistry Department who have from time to time offered helpful advice and encouragement the author also expresses his thanks.

The crystalline pepsin and carboxypeptidase preparations were obtained through the kindness of Doctors L. Lachat and W. White respectively, of Armour Laboratories, Chicago, whose generosity the author acknowledges here.

Thanks are also due to Doctor Marian Jutisz for details of his method of inverse phase chromatography of N-2,4-dinitrophenylamino alcohols prior to their publication. The author also wishes to thank Doctor Jutisz and Doctor Michel Privat de Garilhe for gifts of the Rhodorsil 240 and silicone treated Hyflo Super-cel used in these studies and for helpful suggestions on the use of their method.

Finally, the author wishes to acknowledge here the inspiration and encouragement given him by relatives and friends(three in particular) during the course of his academic career. Theirs is no small contribution to whatever success the author has achieved or may achieve in the future.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

CHEMICAL CHARACTERISTICS OF PROTEINS

The classical experiments of Fischer and his coworkers at the beginning of the century have given much impetus to the investigation of the chemical nature of proteins. Knowledge in this field has accumulated steadily. Particularly rapid advances have been made in the last decade. The past five years have seen the development of many ingenious methods of attacking the problem of protein structure and the advent of these techniques has brought intensified interest in the field. This problem is one of outstanding interest, for ultimately the explanation of the physiological and biochemical characteristics of proteins must lie in the complete elucidation of their chemical structures. The widespread occurrence and varied functions of proteins in the animal organism make the study of this class of compounds one of the most important in the field of biochemistry. Enzymatic activity, hormonal action and the conversion of chemical energy to mechanical work in muscle are only a few of the more important phenomena whose mechanisms may be more clearly understood when

the structures of individual proteins are known.

In this chapter, a few of the more important characteristics of proteins will be discussed. A consideration of some of the chemical characteristics is essential for an appreciation of the rationale of structure studies to be described later.

The major products of hydrolysis of proteins are amino acids, although a few proteins yield products such as polysaccharides, lipids, porphyrins, or the phosphate radical in addition to the amino acids. The number of amino acids which have been isolated from proteins thus far is quite small, about twenty-one. Certain amino acids, e.g., hydroxylysine, hydroxyglutamic acid, norleucine etc., although they have been found in a few proteins or free in biological fluids, are no longer considered to be of general occurrence in protein molecules. The amino acid compositions of most proteins are roughly comparable, the largest variations coming in the content of the various polyfunctional amino acids such as aspartic acid, serine, lysine, etc. There are a few proteins which have strikingly different amino acid compositions; gelatin, for example, which has an extremely high content of glycine. Only a few properties of proteins can be explained on the basis of amino acid analysis alone. For example, the low "biological value" of gliadin, a protein of wheat, is known to be due to lack of the essential amino acid lysine in this protein.

One would expect that if the total number of ionizable groups and their pK's in a protein were known, the isoelectric point of the protein could be calculated. With many proteins this calculated value is not far from the experimentally determined value. But there are some exceptions. One notable exception is pepsin, the isoelectric point of which has never been experimentally determined. It is apparently well below the calculated value of 3.5. Proteins with very similar amino acid compositions possess varied functions and characteristics. This is very striking in the case of the serum albumins from different species of animals. The amino acid compositions of these proteins are almost identical and yet the proteins are species specific. This specificity may lie in the arrangment of the amino acids in the molecules.

The fact that the biological activity of proteins is dependent on the free functional groups of the amino acids has been recognized for some time. For example, the free phenolic hydroxy groups of tyrosine are essential for the activity of the leutinizing hormone of the pituitary. This has been demonstrated by a loss of activity upon acetylation of the molecule. Several other examples could be cited which also support the role of functional groups in the biological activity of proteins. The present theories concerning the mechanism of enzyme action postulate certain "active sites" on the surface of the protein

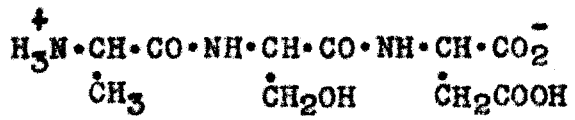
molecule. These sites are probably associated with the functional groups essential to the activity of the protein. The arrangement of the functional groups on the surface of the molecule, in turn, depends on the arrangement of the amino acids in the peptide chains. Therefore, it is evident that before the characteristics and functions of proteins can be explained, it will be necessary to develop definite procedures for determining the chemical and physical configurations of the protein molecules.

In 1902, Fischer (39) and Hofmeister (60), independently and simultaneously, proposed a theory concerning the nature of the bond joining amino acids in peptides and proteins. This is the only theory of the nature of this bond, thus far proposed, which has survived critical investigation. It is now universally accepted and all techniques developed for determining the sequence of amino acids in proteins and peptides are based on this theory. Fischer postulated that the amino acids are joined by amide linkages, the alpha amino group of one amino acid joined with the carboxyl group of the next amino acid in the sequence. The amide bond in proteins is referred to as a peptide bond or link and the polymer of amino acids is known as a peptide chain. The repeating unit in the peptide chain is the grouping

-CO·NH·CHR-.

Attached to each unit is a side chain, R, characteristic of the

amino acid involved. Thus, for the tripeptide, alanylserinyl-aspartic acid, the peptide chain could be represented by the following structure:

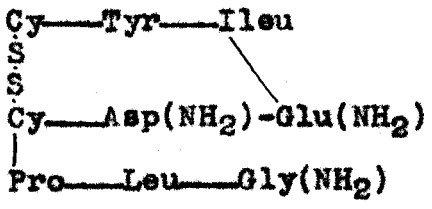


X-ray diffraction patterns of stretched keratin have been interpreted by Astbury to indicate a repeating unit of 3.3Å. Studies with silk fibroin have shown a unit of 3.5Å in length. These values are in close agreement with the figure expected from the -NH·CH·CO- distance (3.6 Å) calculated from the known atomic dimensions and bond angles that are involved. It is conceivable that the terminal amino group at one end of a peptide chain could condense with the free carboxyl group at the other end of the chain to form a cyclic peptide, and indeed the presence of such peptides in nature has been demonstrated. Sanger (102) in 1946 presented evidence of the cyclic nature of gramicidin S, and more recently Mueller *et al* (78) and Turner *et al* (117) have shown the pituitary hormones, oxytocin and vasopressin, to be cyclic peptides. Both gramicidin S and the pituitary hormones mentioned are small peptides, a pentapeptide and two octapeptides. Thus far no conclusive evidence of larger peptides or proteins with cyclic chains has been found. A bond between the epsilon amino group of lysine and the beta or gamma carboxyl group of aspartic or glutamic acids would lead to branching of the peptide

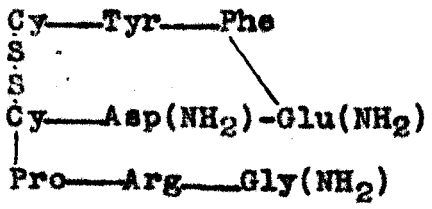
chain. Thus far there is no evidence of that type of bond existing in proteins. By chemical means it has been possible to demonstrate that one of the amino groups of lysine is free in most proteins and in those where an attempt has been made to identify this group, it has always been the epsilon amino group.

The presence in proteins of polyfunctional amino acids could lead to the existence of bonds other than peptide bonds. Three of these bonds involve the carboxyl group. This group could interact with the sulfhydryl group of cysteine, the phenolic hydroxyl group of tyrosine or the hydroxyl group of either serine or threonine. These ester and thioester bonds could form bridges between peptide chains or might form bonds within a single chain. Insufficient work has been done to determine if these bonds are of any significant occurrence in proteins.

The presence of cystine in proteins introduces another type of bond, the disulfide bond. If half of a cystine residue occurred in each of two peptide chains, then these peptides would be joined by a disulfide bridge. Sanger (103) has recently shown, by methods to be discussed later, that insulin consists of four open peptide chains joined by this type of bond. The disulfide link can also participate in cyclic structures in proteins and peptides. This has recently been demonstrated by duVigneaud *et al.* (33). These authors have proposed structures for oxytocin and vasopressin which may be represented as follows:



oxytocin



vasopressin.

It has been shown that in order to conform to the dimensions demonstrated by X-ray analysis, globular proteins must consist of folded long chain peptides or short parallel chains. This folding is probably in part responsible for the biological functions of proteins by bringing about a certain required spatial arrangement of the functional groups. Since proteins are quite labile with respect to their biological activity, it appears that the forces holding the peptide chain in the required configurations must be rather weak. These weak forces could result from ionic salt linkages, from mutual interaction of dipoles and from hydrogen bonding. Thus, the destruction of charged centers by the action of acid or alkali can be explain in part the denaturation brought about by pH changes.

The presence of the peptide bond in proteins makes possible the formation of a large number of hydrogen bonds. These bonds are possible between the imino nitrogen of one peptide bond and the carboxyl oxygen of another. Since the peptide bond is the predominant link in proteins, one might expect that the hydrogen bond is the one of the major forces responsible for the specific configuration of proteins. Hydrogen bonds are considera-

bly weaker than the ordinary chemical bonds, e.g., 8 kcal per mole for the hydrogen bond between the imino nitrogen and carboxyl oxygen of peptide bonds. The relatively small energy of these bonds together with the small changes required for denaturation, are two factors which have led to the hypothesis that these bonds are involved in holding the protein in its specific configuration. Pauling (85) has recently published an hypothesis of the helical arrangement of the folded peptide chain. The energy necessary to hold the protein in this configuration is supplied by hydrogen bonds. He has discovered several other types of configurations and in addition, other authors have proposed hydrogen bonded folded peptide chains of helical and nonhelical configurations. Most of these hypothetical configurations fall short of meeting theoretical criteria or X-ray defraction data. One configuration suggested by Pauling, however, meets most of the theoretical criteria set up by himself and other authors and also is in agreement with X-ray diffraction patterns obtained on several proteins. This configuration is known as the 3.7 residue or alpha helix. Each peptide bond is hydrogen bonded to the third peptide link beyond it in the chain and forms a spiral around a straight central axis, as shown in Figure 1. The bond angles and lengths (Figure 2) are those determined by Corey and Donahue from studies on alanine, threonine, N-acetyl glycine and beta glycylglycine.

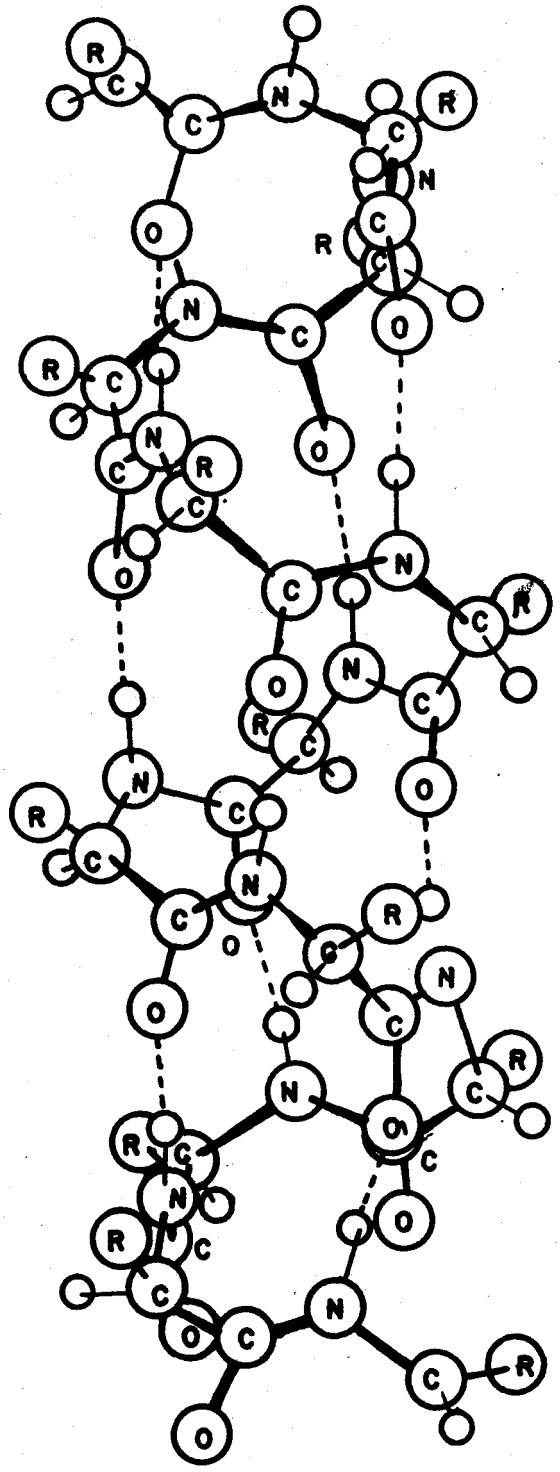


FIGURE 1

The α Helix (After Pauling, Corey, and Branson)

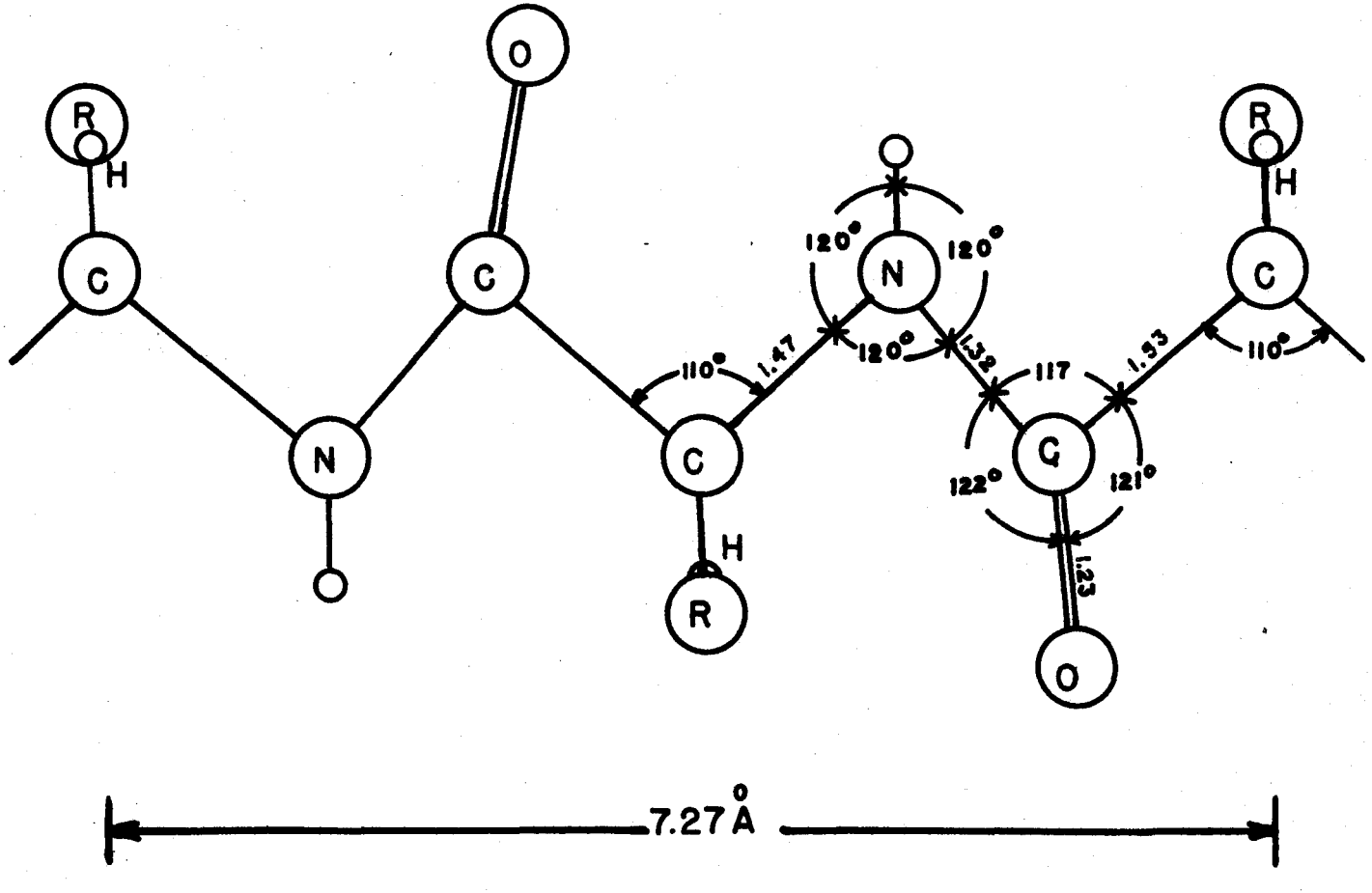


FIGURE 2

Bond Angles And Bond Lengths For The α Helix

Sanger (101) in 1945 began a series of studies which culminated in the complete determination of the amino acid sequences in the peptide chains of insulin (101,103,104,105,106, 107). Thus far, this is the only protein which has been completely characterized in this respect. A few naturally occurring peptides have been fully characterized and work is now under way on several proteins. These studies will be discussed later under the specific methods being employed. Prompted by the success of Sanger and other workers, the study of the amino acid sequence in the very interesting protein, pepsin, has been undertaken with a view toward explaining some of its unique characteristics.

CHAPTER II

PROTEIN STRUCTURE STUDIES

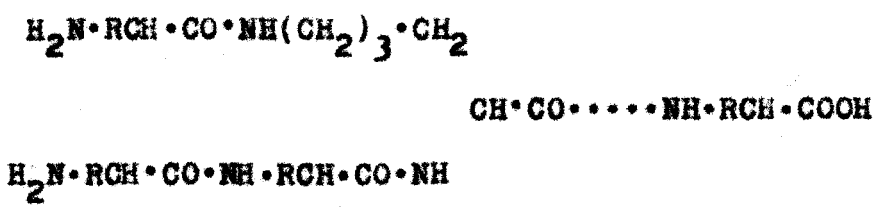
The development of methods for the study of the amino acid sequences in proteins has occurred very recently although the rationale of the methods was established as early as 1907 by Fischer and Abderhalden (40). Active application of these methods has taken place in the past 6-8 years. Most of the methods proposed have proven to have definite limitations although a few have yielded very excellent results, notably, the method of Sanger (101).

The objective of most of the attempts has been to develop methods which would enable the investigator to identify the amino acids situated at the ends of peptide chains or fragments of chains. These amino acids possess free alpha amino groups or free carboxyl groups. These groups are now termed N-terminal and C-terminal residues respectively. The determination of the number of these groups in a protein gives information as to the structure of peptide chains in the molecule. Some possible configurations are illustrated in Figure 3.

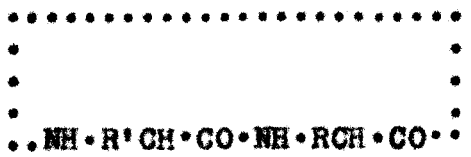
If analysis showed one N-terminal residue and one C-terminal residue, the protein would be considered as a single open peptide chain, disregarding the possibility of the presence of cyclic peptides. This is illustrated in Figure 3a. If more than one of each type of terminal residue were found, in equal amounts, i.e., 2,3, or n of each, the protein can be considered to be composed of 2,3, or n open peptide chains joined together by some secondary linkages. Figure 3b illustrates the case where two N-terminal groups and one C-terminal group are found. These results indicate branching of a peptide chain. In the case illustrated the branch comes off the epsilon amino group of lysine. The opposite type of branching could be conceived where the branch would come off the second carboxyl group of one of the monoamino dicarboxylic amino acids. In this case the analysis would reveal two C-terminal residues and one N-terminal residue. Cyclic peptides of the type shown in Figure 3c would show no terminal free residues. There is a type of cyclic peptide, involving a cystine disulfide link, which would show N- and C-terminal residues of cysteic acid after oxidation with performic acid. This type is shown in Figure 3d. Various combinations of the above configurations are also possible.

A second objective of the methods to be described is the determination of amino acid sequences by marking the N- or C-terminal group, selectively removing the marked residue and

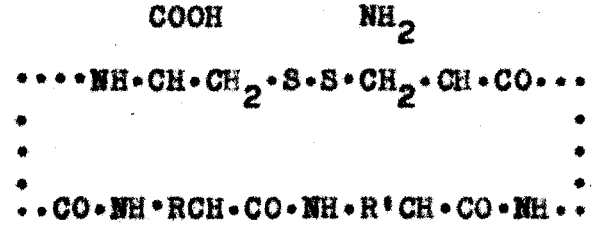
a.



b.



c.



d.

FIGURE 3

POSSIBLE PEPTIDE CHAIN CONFIGURATIONS IN PROTEINS

then marking the next residue in succession. Thus a stepwise degradation of the peptide or protein would be accomplished. Abderhalden and Brockmann (1) in 1930 described one of the first degradative methods. Previous to this time the emphasis had been on the determination of the terminal residues.

The first method for the determination of N-terminal residues was introduced by Fischer and Abderhalden (40). The peptide was treated with naphthalene-2-sulfonyl chloride converting the free amino group to the N-naphthalene-2-sulfonyl derivative which was isolated after hydrolysis of the peptide. The derivatives were separated from the hydrolysate by differential solubility which entailed large losses. Abderhalden and Stix (2) investigated the N-terminal groups of partial hydrolysates of silk fibroin by converting the terminal residue to this N:2, 4-dinitrophenyl derivative using 2,4-dinitrochlorobenzene as the labeling agent. The N:2,4-dinitrophenyl amino acids can be isolated from the subsequent hydrolysate by extraction. Both of the above reactions required a total hydrolysis of the protein to obtain the derivative of the terminal residue. Abderhalden and Brockmann (1) utilizing the reaction described by Bergmann in 1927 reacted the terminal free amino group with phenylisocyanate. By controlling the hydrolysis conditions they found it possible to remove the terminal group as the phenylhydantoin leaving the remainder of the peptides essentially intact.

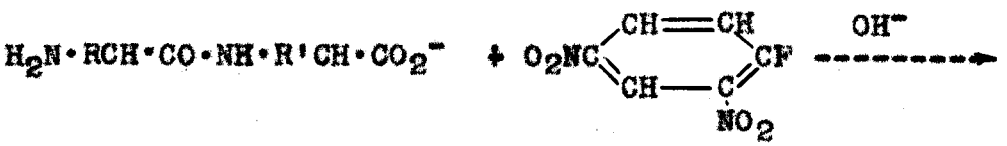
However, there was some spliting of peptide bonds during the hydrolysis. These authors then proposed that the process could be reapplied to the peptide obtained from the first step and then to that obtained at the second step and so on. This was the first description of a stepwise degradation. Jensen and Evans (65) used phenylisocyanate and α-naphthaylisocyanate to identify the free amino groups of insulin as the phenylhydantoins and naphthylhydantoins. Benzenesulfonyl chloride was used by Gurin and Clark (51) to identify the free amino groups in gelatin.

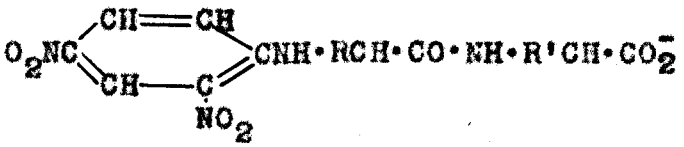
Within the past ten years there have been several very interesting methods proposed. These methods were devised with the end view of utilizing micro quantities of peptides and proteins, as is usually necessary in the case of naturally occurring products.

The derivatives of the N-terminal residues are not isolated in crystalline form for melting points, elementary analysis etc., but are identified usually by chromarography.

The first, and one of the two most widely used, of the newer methods is that of Sanger (101). The N:2,4-dinitrophenyl amino acids described by Abderhalden and Stix (2) are yellow compounds and lend themselves well to chromatography. Sanger wished to take advantage of this but the method of preparing them from 2,4-dinitrochlorobenzene required heating at rather high temperatures which would lead to partial destruction of the

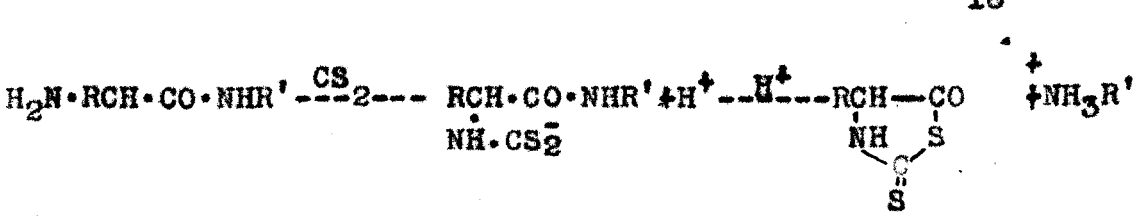
proteins. Therefore, he investigated the corresponding dinitrofluorobenzene (DNFB) and found that it reacted in sodium bicarbonate solution at room temperature.





Acid hydrolysis of the resulting DNP-derivative (protein or peptide) yields the free DNP-amino acids which are separated from the hydrolysate by ether extraction. The yellow DNP-amino acids are then separated and identified by chromatography and measured spectrophotometrically. This method served as the basis for Sanger's structural studies on insulin.

In 1950 three methods were introduced each of which utilized a different reagent capable of reacting with the free amino groups of proteins. Levy (73) utilized carbon disulfide as the reagent. Leonis in 1948 had used this reagent to titrate free -amino groups. The reaction leads to the formation of a thiocarbamyl peptide or protein and the terminal residue is split off as a 2-thiothiazolid-5-one.



The reaction with carbon disulfide is carried out at room temperature in 0.3 N Ba $(OH)_2$. The terminal group is split off by mild acid hydrolysis (pH 3-4) at room temperature. The derivative of the terminal group can be measured spectrophotometrically or the amino acid can be liberated from the derivative by acid hydrolysis. Further modifications have been reported by Leonis and Levy (72).

Ingram (62) altered the terminal free amino group in such a way that it would not appear among the ninhydrin positive spots of the hydrolysate of the treated protein. The peptide is methylated with formaldehyde and hydrogen in the presence of a palladized charcoal catalyst producing the corresponding N:dimethylamino acid of the N-terminal residue. These substituted amino acids do not give a positive ninhydrin test. The difficulty with the method lies in the detection of diminished intensity of ninhydrin spots. If the peptide contains two residues of alanine, for example, and one of them is a terminal residue, a diminished intensity would be expected. If, however, a protein contains twenty residues of alanine, one of which is a terminal residue, the detection of a diminished intensity is very unlikely. For this reason the method is limited to peptides with

only a small number of residues of like kind to the terminal residue. The method has been applied to three peptides which were obtained as autolysis products of pepsin (63). The results showed the peptides each contained one terminal free alanine residue. The difficulties described above were not encountered since the peptides contained only one residue of alanine each.

A method receiving wide attention at the present was proposed by Edman (36) in this same year. Two major modifications of Abderhalden and Brockmann's method were made by Edman, First, he reasoned that the ease of the reaction should parallel the ease of ring closure to hydantoin and in this respect the use of phenylisothiocyanate (PITC) should be an improvement over phenylisocyanate. He found, using various peptides, that the reaction was essentially complete within fifteen minutes at 40°C. Secondly, in order to avoid the peptide bond hydrolysis encountered under the hydrolysis conditions of Abderhalden and Brockmann, Edman developed an anhydrous medium. Thus, while water is necessary to break the peptide bond, the terminal bond was labile and could be broken by HCl in anhydrous nitromethane. Later investigators (37,44,84) found that the hydrolysis could be carried out in the presence of up to 30 per cent water without causing hydrolysis of peptide bonds other than the terminal bond. The reactions of the method are as follows:

$$H_3^+N \cdot RCH \cdot CO \cdot NH \cdot R'CH \cdot CO_2^- + C_6H_5NCS \longrightarrow$$

$$C_6H_5N \cdot CS \cdot NH \cdot RCH \cdot CO \cdot NH \cdot R'CH \cdot CO_2^- \xrightarrow{H^+} \text{RCH-NH (OC, CS, N, } C_6H_5\text{)} + {}^+H_3N \cdot R'CH \cdot CO_2^-$$

The phenylthiohydantoin can be identified chromatographically or the parent amino acid can be obtained by hydrolysis of the thiohydantoin. Sjoquist (110) and Landmann et al (70) have developed methods for paper chromatography of the phenylthiohydantoins.

In 1951, Udenfriend and Velick (119) reported a method using I^{131} labeled p-iodophenylsulfonyl chloride (pipsyl) to label the terminal amino group.

$$H_2N \cdot RCH \cdot CO \cdot NH \cdot R'CH \cdot CO_2H + I^*C_6H_5 \cdot SO \cdot Cl \longrightarrow$$

$$I^*C_6H_5 \cdot SO \cdot NH \cdot RCH \cdot CO \cdot NH \cdot R'CH \cdot CO_2H$$

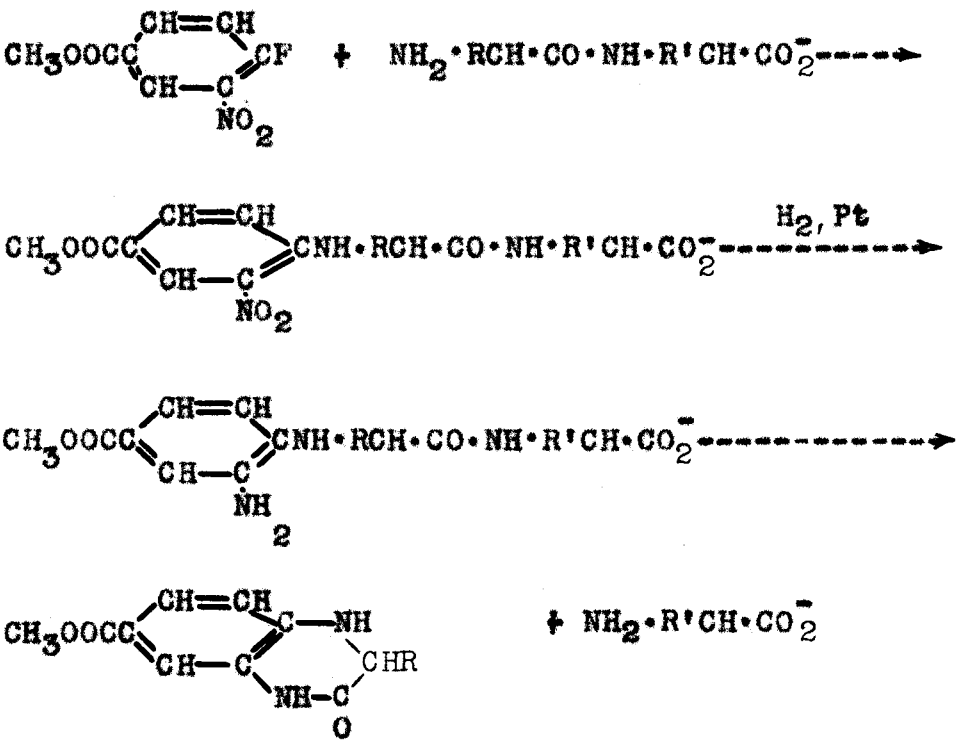
After hydrolysis of the protein derivative, the pipsyl derivative of the terminal amino acid is isolated and identified by countercurrent distribution and paper chromatography. For quantitative determination and also as a check of the purity of the isolated derivative, a known amount of the derivative labeled with S^{35} is added to a known amount of the protein

derivative labeled with I^{131}. The derivative isolated from the hydrolysate is chromatographed on paper. The area containing the derivative is located, cut in sections perpendicular to the direction of migration, and the derivative eluted from the sections. Differential determinations of I^{131} and S^{35} using filters are made on each eluate. If the ratio S^{35}:I^{131} remains constant along the direction of migration, purity is assumed. The quantitative calculation of end groups is made from known amounts of added S^{35} indicator and protein derivative employed and the specific activity of the labeling reagent. This method has a few advantages over the DNP technique when applied to amino acids and peptides. The method is more sensitive and therefore, smaller quantities can be used. In the former method, some breakdown of the DNP-derivatives does occur on hydrolysis, however, the S-N bond of the N:pipsyl derivative is very resistant to the hydrolytic conditions necessary to break the peptide bond.

However, when the technique was applied to insulin, very gross quantitative discrepancies were found between the results obtained by Udenfriend and Velick and those reported by Sanger (101), Whereas, Sanger found two terminal glycine and two terminal phenylalanine residues in insulin, the pipsyl method reveals only one of each of these residues. Porter and Sanger (92) found six terminal valine residues in horse hemoglobin but the pipsyl method shows two valine residues. The authors

suggested that perhaps the pipsyl reagent does not reach free amino groups in the interior of the folded peptide chains. Until further work is done on this method, its applicability to proteins is doubtful. If the author's explanation that residues in the interior of folded peptide chains are not reached by the reagent, is correct, it is conceivable that qualitative as well as quantitative discrepancies might arise. For example, if in the case of insulin, the glycine residues were "interior" and the phenylalanine residues were "exterior" the results would show only two phenylalanine terminal residues in the molecule.

In 1952, Holley and Holley (61) proposed a degradative method to be applied to the amino terminus which makes use of the following series of reactions:



The method has been applied only to small simple peptides. Difficulties are envisioned by the authors when amino acids such as methionine, cystine, systeine, lysine, tyrosine, proline and tryptophane are encountered. Therefore, the method is probably only of limited usefulness.

Ingram (64) describes an improvement of his method of 1950 in which the dimethylamino acid can be identified by paper chromatography. The dimethyl amino acid need not be separated from the hydrolysate of the protein since the derivative can be detected by reagents other than ninhydrin. Also, dimethylglycine and monomethylproline are more stable to acid hydrolysis than the corresponding 2,4-dinitrophenyl derivatives. This is an advantage over Sanger's technique.

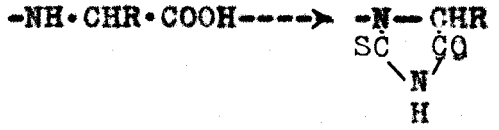
A preliminary report by Reith (95) in 1953 described the use of 3,5-dinitro-4-dimethylaminophenylisothiocyanate as a reagent for the stepwise degradation of peptides from the amino end. The reagent has the advantage that the resulting 3,5-dinitro-4-dimethylaminophenylthiohydantoins are colored derivatives, facilitating their chromatography and colorimetric measurement.

A series of four papers (38,42,43,96) with the general title "Studies on the Determination of the Sequence of Amino Acids in Peptides and Proteins" has very recently come from this same laboratory. In the first paper, the applicability of

azobenzene-p-sulfonyl chloride to the problem was studied. It was concluded that the resulting N-p-azobenzene-p-sulfonyl derivatives obtained were not sufficiently stable to the conditions of hydrolysis of the protein and therefore the method was not suited to end group analysis. The second method investigated was the isolation of methyl esters of DNP-amino acids from DNP-proteins. It was found however, that the stability of the methyl esters of the amino acids varied considerably, thus negating their use. The next paper describes the use of 3,5-dinitro-4-dimethylaminophenylisocyanate as the reagent. The conclusion drawn was that the method was good for N-terminal studies but was not applicable as a degradative method. Finally, a detailed account of the preliminary report mentioned above was published. 80-90 per cent yields at each step were obtained with small peptides.

Methods for the determination of the terminal free carboxyl group have not been so numerous, nor have the quantitative aspects been as favorable as at the amino end. Most of the methods developed before 1949 are of only historical importance for none of them are used at the present time.

The first method developed for the identification of the terminal free carboxyl group was that of Schlack and Kumpf (108). In this technique the peptide was treated with ammonium-thiocyanate, and acetic anhydride to form the thiohydantoin:



The thiohydantoin was split from the peptide, isolated and identified. In 1951, Tibbs (115) reported a micro method based on the above technique in which the isolated thiohydantoin is hydrolyzed in aqueous barium hydroxide to liberate the parent amino acid. The amino acid is then identified by paper chromatography. The method worked well on small peptides but in preliminary experiments on insulin and wool keratin it became evident that modifications would be necessary. In 1953, Baptist and Bull (9) modified the method in an attempt to make it useful in quantitative studies. The isolated amino acid was determined by a spot dilution technique developed by Bull et al (17). Certain of the amino acids however, did not give the expected results and the method has not been used extensively.

The method does not detect C-terminal residues of aspartic or glutamic acids, lysine or arginine and it gives ambiguous results with serine and threonine. In addition, the yields from peptide studies usually are below 50 per cent for most of the amino acids which can be detected. Waley and Watson (123) had described a similar modification in 1951 and recoveries of added thiohydantoins from protein solutions, such as are encountered in the method, were less than 20 per cent.

Bettzieche (14) investigated the free carboxyl groups of peptides by means of the Grignard reagent. Lichtenstein (75) used β-naphthol to identify the C-terminal residues. Bergmann (10,11), developed a complicated scheme for the stepwise degradation of peptides from the carboxyl end. The carboxyl group was first converted to the methyl ester; the ester was treated with hydrazine to give the hydrazide which in turn was converted via the Curtius reaction to the azide; the azide was reacted with benzyl alcohol to give the benzyl urethane which was converted to the aldehyde characteristic of the terminal carboxyl group by catalytic reduction. The reasons that this process has never become popular are apparent. Akabori (5) also employed hydrazine for the identification of the terminal free carboxyl group. The method however, was more straightforward than Bergmann's. The protein was heated with anhydrous hydrazine at 125°C. The resulting mixture consists of the hydrazides of all amino acids except the C-terminal residue which is present as the free amino acid. The amino acid hydrazides are converted to insoluble dibenzal compounds by treating the residue with an excess of benzaldehyde in water. The precipitate is filtered off and the supernatant containing the terminal free carboxyl amino acid is examined by paper chromatography. Experiments on insulin, glutahione, and tyrocidin are in agreement with other reports found in the literature.

Lens (71) in 1949 applied carboxypeptidase to the problem of identifying the terminal free carboxyl groups of insulin. Since that time several other authors (24,46,50,53,54,113) have used this method with success and in some instances even for determination of sequences adjacent to the C-terminal residue (124).

The reduction of the terminal free carboxyl group to a primary alcohol by means of lithium aluminum hydride was reported by Fromageot et al (47) in 1950. This is an excellent method for identification purposes but only one other author has published on the technique (19,20). The quantitative aspects of this method are however, very unfavorable. This involves oxidation of the amino alcohol to ammonia, formaldehyde and an aldehyde characteristic of the amino alcohol involved. Measurement of the ammonia or formaldehyde gives the total number of C-terminal residues. If however, there is more than one kind of terminal residue, the determination of the number of each type becomes difficult. Jutisz (67) has developed the method into a photocolorimetric one. The amino alcohol is converted to the N:2,4-dinitrophenylamino alcohol and measured in a photoelectric-colorimeter. The development of the technique was simultaneous and independent of the same method described later as a part of the original work of this dissertation.

Ohno (82) recently has described a similar modification

of Akabori's hydrazinolysis method. The hydrazinolysate is treated with DNFB and the DNP-amino acid corresponding to the C-terminal residue is extracted, identified and measured photocolorimetrically.

Boissonnas (15) has carried out anodic oxidations of small neutral DNP-peptides in methanol. The C-terminal is converted to a methoxyalkylamide. When the peptide is then hydrolyzed, the terminal group will not appear on a chromatogram of the hydrolysate. If, instead of complete hydrolysis, the peptide derivative is heated for a short time with dilute HCl, the methoxyalkylamide group is split off leaving a peptide with a new C-terminal group. This new peptide can then be subjected to a second anodic oxidation. These reactions can be summarized as follows:

$$\text{DNP-NH}\cdot\text{CHR}'.\text{CO}\cdot\text{NH}\cdot\text{CHR}\cdot\text{COOH} \longrightarrow \text{DNP-NH}\cdot\text{CHR}'.\text{CO}\cdot\text{NH}\cdot\text{CHR}\cdot\text{OCH}_3 + \text{CO}_2$$

$$\text{DNP-NH}\cdot\text{CHR}'.\text{OCH}_3 \xrightarrow{\text{HCl}} \text{DNP-NH}\cdot\text{CHR}'.\text{COOH} + \text{NH}_3 + \text{RCHO} + \text{CH}_3\text{OH}$$

Several reviews (66,98,127) on the newer techniques are now available.

Tables I and II give, at a short glance, an indication of the amount of work that has been done on proteins using the methods that have just been described. Most of this work has been done within the past five years. A few of these studies may be

TABLE I

N-TERMINAL GROUPS OF PROTEINS AND NATURALLY OCCURRING PEPTIDES

Protein or peptide	N-terminal Residues	Method	Ref.
Aldolase (rabbit)	2 Pro	Pipsyl	119
Avidin	3 Ala	DNP	45
α-Chymotrypsin	1 Ileu, 1 Ala	DNP	29
Chymotrypsinogen	none	DNP-	29
Conalbumin	1 Ala	DNP	45
Corticotropin A	1 Ser	ØNCS	69
Fibrinogen	none	DNP	8
G-3-P dehydrogenase	2 Val	Pipsyl	122
Glutathione	1 Glu	ØNCS	69
Glucagon	His	DNP	112
Gramicidin S	none	DNP	102
Growth Hormone	1 Phe, 1 Ala	DNP	74
Growth Hormone	1 Phe, 1 Ala	DNP	94
Hemoglobins			
Human, adult	5 Val	DNP	92
Human, adult	2 Val	DNP	55
Human, fetal	2.6 Val	DNP	92
Human, sickle	2 Val	DNP	55
Horse, Donkey	6 Val	DNP	92
Horse, Donkey	2 Val	Pipsyl	119
Cow, Sheep, Goat	2 Val, 2 Met	DNP	92

TABLE I (Cont.)

N-TERMINAL GROUPS OF PROTEINS AND NATURALLY OCCURRING PEPTIDES

Protein or peptide	N-terminal residues	Method	Ref.
Insulin	2 Phe, 2 Gly	DNP	101
Insulin	Phe, Gly	ØNCS	22
Insulin	1 Phe, 1 Gly	Pipsyl	119
Insulin	Phe Gly	ØNCS	69
Insulin	2 Phe, 2 Gly	DNP (Me Ester)	42
Keratin	4 Val, 2 Ala, 8 Gly,		
	8 Thr, 2 Ser, 2 Glu,		
	1 Asp	DNP	8
Lysozyme	1 Lys	DNP	109
Lysozyme	1 Lys	ØNCS	69
Myoglobin, horse	1 Gly	DNP	92
Myosin	none	DNP	8
Ovomucoid	1 Ala	DNP	45
Papain	Ileu	DNP	114
Pepsin	1 Leu	DNP	125
Ribonuclease	1 Lys	DNP	7
Salmine	1 Pro	Pipsyl	121
Salmine	Pro	ØNCS	69
Serum albumins			28

TABLE I (Cont.)

N-TERMINAL GROUPS OF PROTEINS AND NATURALLY OCCURRING PEPTIDES

Protein or peptide	N-terminal residues	Method	Ref.
Human	1 Asp	DNP	28
Bovine	1 Asp	DNP	
Porcine	1 Asp	DNP	
Horse	1 Asp	DNP	
Tropomyosin	none	DNP	8
Trypsin	1 Ileu	DNP, ØNCS	98,99
Trypsinogen	1 Val	DNP, ØNCS	98,99
Tyrocidin	none	DNP	21
Vasopressin (oxidized)	1 $CySO_3H$	DNP, ØNCS	3

TABLE II

C-TERMINAL GROUPS OF PROTEINS AND NATURALLY OCCURRING PEPTIDES

Protein or peptide	C-terminal residues	Method	Ref.
α-Chymotrypsin	1 Ileu, 1 Tyr	Carboxy.	49,50
α-Chymotrypsin	Gly	Hydraz.	6
Chymotrypsinogen	none	Carboxy.	49,50
Chymotrypsinogen	Gly	Hydraz.	6
Corticotropin A	Phe	Carboxy.	124
Glutathione	Gly	Hydraz.	5
Insulin	1 Gly, 2 Ala 1 unidentified	Reduct'n	19
Insulin	2 Ala, 2 Asp (NH_2)	Reduct'n	48
Insulin	2 Ala, 2 Gly	Reduct'n	47
Insulin	x Ala	S Hydan.	123
Insulin	Ala	Carboxy.	71
Insulin	2 Ala, 2 Asp (NH_2)	Carboxy.	53
Insulin	Ala, Gly	Hydraz.	5
Insulin	x Ala	S Hydan.	9
Insulin	2 Ala, 2 Asp (NH_2)	DNP,Carb.	107
Glucagon	Thr	Carb.,Hydraz.	112
Lysozyme	1 Leu	Carboxy.	113
Lysozyme	1 Leu	Carboxy.	53
Lysozyme	1 Leu	DNP-Hydraz.	82

TABLE II (Cont.)

C-TERMINAL GROUPS OF PROTEINS AND NATURALLY OCCURRING PEPTIDES

Protein or peptide	C-Terminal Residues	Method	Ref.
Ribonuclease	1 Val	Carboxy.	
Tobacco Mosaic Vir.	3400 Thr	Carboxy.	54
Tobacco Mosaic Vir.	2500 Thr	Carboxy.	46
Trypsin (denat.)	Lys	Carboxy.	24
Trypsinogen	none	Carboxy.	24
Trypsin (native)	none	Carboxy.	24
Tyrocidin	none	Hydraz.	5
Cvomucoid	Phe	S Hydan.	118
Ovalbumin	Ala	S Hydan.	118

mentioned here because the results are more than mere identification of terminal groups or short sequences.

The first extensive studies were carried out by Sanger and co-workers (101,103,104,105,106,107) on insulin. Sanger first established with his DNP technique that there were two free amino groups per insulin submolecule of molecular weight 6000. These were phenylalanine and glycine. He then showed that by oxidation with performic acid, the insulin molecule could be split into two chains, one, chain A, having an N-terminal residue of glycine and the second, chain B, having a free phenylalanine group. By partial hydrolysis of the chains, which had been converted to the DNP-derivatives, it was possible to determine the sequence of amino acids near the N-terminal groups. This is accomplished by isolating, by chromatography, the DNP-peptides from partial hydrolysate. The DNP-peptides were then hydrolyzed and the amino acid compositions determined by paper chromatography. Thus in the A chain, the peptides DNP-glycylisoleucine, DNP-glycylisoleucylvaline, DNP-glycylisoleucylvalylglutamic acid and DNP-glycylisoleucylvalylglutamylglutamic acid were isolated. The terminal sequence can be visualized from these peptides in the following manner,

DNP-gly.ileu

DNP-gly.ileu.val

DNP-gly.ileu.val.glu

DNP-gly.ileu.val.glu.glu

Removing the DNP residue, the sequence as it occurs in the insulin molecule is gly.ileu.val.glu.glu-. The abbreviations using the first three letters of the amino acids are those suggested by Brand et al. (16). The next step in the studies was to partially hydrolyze the individual chains, isolate the small peptides by chromatography and other methods and determine the sequences of amino acids in these peptides. This was accomplished in the manner described above for the terminal sequence. When all the sequences were determined the peptides were "pieced together" to obtain the sequences of higher peptides. It was found that certain peptide bonds, those involving the amino groups of serine and glycine, were absent in the partial hydrolysate, leaving gaps in what otherwise would be the complete sequences in the chains. Therefore, enzymic hydrolyses were employed and these bonds were obtained intact, thus allowing the reconstruction of the entire sequences in the chains, In the partial hydrolysis studies on the B chain in the sequences of twenty-three dipeptides, fifteen tripeptides, nine tetrapeptides, two pentapeptides and one hexapeptide were determined. Similar efforts were put forth in determing the sequence in the A chain. The sequences may be represented as follows:

Chain A - Gly.Ileu.Val.Glu.Glu.Cys.Cys.Ala.Ser.Val.CyS.Ser.Leu.Tyr.Glu.Leu.Glu.Asp.Tyr.CyS.Asp.

Chain B - Phe.Val.Asp.Glu.His.Leu.CyS.Gly.Ser.His.

Leu.Val.Glu.Ala.Leu.Tyr.Leu.Val.CyS.Gly.Glu.Arg.Gly.Phe.Phe.
Tyr.Thr.Pro.Lys.Ala.

It is of interest to note that there is apparently no simple periodic arrangement of amino acids as has been postulated by Bergmann (12) and others.

In addition to the sequences determined by Sanger on insulin, Table III lists several other partial sequences which have been determined.

TABLE III

PARTIAL TERMINAL SEQUENCES IN SOME PROTEINS

Protein	Sequence	Terminus	Method	Ref.
Carboxypeptidase	Asp(NH_2).Ser-	N	DNP	
Corticotropin A	Ser.tyr-	N	DNP,ØNCS	69
Corticotropin A	-leu.glu.phe.	C	Carboxy.	124
α-Chymotrypsin	Ileu.val-	N	DNP	100
γ-globulin(rabbit)	Ala.Leu.Val.Arg.	N	DNP	
Insulin (A chain)	Gly.Ileu.Val.Glu. Glu-	N	ØNCS	69
Insulin (B chain)	Phe.Val.Asp.Glu. His-	N	ØNCS	69
Lysozyme	Lys.Val.Phe.Gly-	N	DNP	109
Lysozyme	Lys.Val.Phe.Gly. Ser-	N	ØNCS	69
Papain	Ileu.Pro.Glu-	N	DNP	114
Trypsinogen	Val.$(Asp)_4$.Lys.	N	DNP	31
Trypsin	Ileu.Val.Gly-	N	DNP	31

Another interesting series of studies are those which have been undertaken on the pancreatic enzymes, -chymotrypsin and trypsin. Desnuelle et al. (30) have established, using DNFB and PITC, that -chymotrypsin contains one Nterminal residue of isoleucine and one N-terminal alanine. The zymogen,

chymotrypsinogen, on the other hand, contains no terminal free amino groups. The authors have interpreted this to mean that the activation process involves a proteolytic cleavage of peptide bonds in the zymogen, engendering the N-terminal residues in the enzyme. Gladner and Neurath (49,50) have investigated the C-terminal residues in this zymogen-enzyme system and have found no C-terminal residues in the zymogen and one leucine and one tyrosine C-terminal residues in the enzyme. Thus both investigators agree that in the activation of chymotrypsin there is a cleavage of peptide bonds in the cyclic zymogen thus producing two open peptide chains in the active enzyme. Trypsinogen and trypsin appear to be single open chains. The N-terminal residues have been identified by Rovery et al. (99) as isoleucine for the zymogen and valine for the active enzyme. Native trypsin and trypsinogen do not possess C-terminal residues which are reactive toward carboxypeptidase. However, acid denatured trypsin possess a free carboxyl group of lysine (24). Whether this same amino acid is at the carboxyl terminus or trypsinogen is not known. Thus, it appears that the activation of trypsin involves the splitting of a peptide from the amino end of trypsinogen. Desnuelle and Fabre (31) have established a sequence at the amino terminus of trypsinogen which contains isoleucine (1 mole), aspartic acid (4 moles) and lysine (1 mole). This composition is identical with the peptide isolated by Davie and Neurath (25) after the activation of trypsin.

In a most interesting series of studies (33,34,76,78,87, 188,89,90,91,117) employing Sanger's technique, Edman's degradation, and performic acid oxidation, duVigneaud's group from Cornell University Medical College have elucidated the structures of vasopressin and oxytocin illustrated in Chapter I. Acher (4) simultaneously published the same structure for vasopressin. Having deduced the structure of oxytocin, duVigneaud's group proceeded to synthesize this interesting hormone (34). The synthetic product produced the expected oxytocic activity in the isolated rat uterus. Further, the compound was fully effective in stimulating labor in the human. It likewise possessed milk-ejecting activity in the human. The synthetic material also possessed physical and chemical properties identical with the best preparations of the natural hormone.

CHAPTER III

THE CHARACTERISTICS OF PEPSIN

Pepsin, the proteolytic enzyme of the stomach, was first isolated in crystalline formed by Northrop (1930). Preparations of crystalline pepsin, which are 90-99 per cent pure by electrophoretic analysis, have been made by salt and ethanol precipitation.

The reaction catalyzed by pepsin is far from a complete hydrolysis of the protein. The end products of peptic hydrolysis are high molecular weight peptides. Desneulle et al. (1950) have shown that the average size of the peptides obtained from horse globin and egg albumin is that of dedecapeptides. Since only a relatively few peptides are hydrolyzed, it has been concluded that pepsin attacks only specific types of peptide bonds. In an attempt to determine the exact nature of the specificity of pepsin several investigators have prepared synthetic substrates which are hydrolyzed by pepsin. The first group of synthetic substrates were prepared and reported by Bergmann and Fruton (1941). This group includes carbobenzoxy-L-glutamyl-L-

tyrosine, glycyl-L-glytamyl-L-tyrosine, carbobenzoxyglycyl-L-glutamyl-L-tyrosine, carbobenzoxy-L-glutamyl-L-tyrosylglycine, and carbobenzoxy-L-glutamyl-L-phenylalanine. The peptide bond split in all cases was that between the aromatic residue and dicarboxylic amino acid residue. The presence of tyrosine or phenylalanine in the substrates was not in itself sufficient for peptic activity. Both free carboxyl groups and the aromatic residue were found to be essential.

Results which did not agree with the generalizations of Bergmann and Fruton were later obtained by Harrington and Pitt Rivers (1944). These investigators have shown that pepsin catalyzes the hydrolysis of cysteinyl-tyrosine and tyrosyl-cysteine and the corresponding cystine compounds.

Studies on proteins, however, do not reveal such a narrow specificity. With the development of techniques for identification of the amino acids which possess free amino groups in proteins, it has been possible to identify the bonds which are attacked by pepsin in the intact protein molecule. Desnuelle et al. (1950) in their studies on the action of pepsin on horse globin and egg albumin, found that pepsin attacks peptide links in which the amino groups of alanine, phenylalanine, leucine, and serine are involved. Sanger and Tuppy (1951) have used pepsin in studying the amino acid sequence in insulin. These investigations found the following bonds to be the major sites of action

TABLE IV

PEPTIDE BONDS SPLIT BY PEPSIN

Peptide	Reference
CBO-L-glutamyl-L-tyrosine	Bergmann and Fruton (1941)
glycyl-L-glutamyl-L-tyrosine	"
CBO-glycyl-L-glutamyl-L-tyrosine	"
CBO-L-glutamyl-L-tyrosylglycine	"
CBO-L-glutamyl-L-phenylalanine	"
cysteinyl-tyrosine	Harrington and Pitt Rivers (1944)
tyrosyl-cysteine	"
cystinyl-tyrosine	"
tyrosyl-cystine	"
R-alanine	Desnuelle et al. (1950)
R-phenylalanine	"
R-leucine	"
R-serine	"
leucyl-valine	Sanger and Tuppy (1951)
tyrosyl-leucine	"
phenylalanyl-phenylalanine	"
phenylalanyl-tyrosine	"
phenylalanyl-valine	"

of the enzyme: leucylvaline, tyrosylleucine, phenylalanylphenylalanine, alanylleucine, leucyltyrosine, and glycylphenylalanine. A summary of the bonds known to be split by pepsin is presented in Table IV.

The molecular weight of pepsin has been determined by a number of different methods. The results of these determinations are recorded in Table V.

TABLE V

THE MOLECULAR WEIGHT OF PEPSIN

Molecular Weight	Method of Determination	Reference
35,000 ± 1000	Sedimentation	Philipot and Ericksson-Quensel, 1933
34,000	Monomolecular Layers	Dieu and Bull, 1949
34,000	Phosphorus Analysis	*
35,100	Methionine Analysis	*
32,500	Lysine Analysis	*
35,000	Diffusion	Neurath et al, 1941

* See Table VI, Page 44.

Crystalline pepsin has a very low isoelectric point. No actual determination of the isoelectric point has been made

TABLE VI

AMINO ACID COMPOSITION OF PEPSIN
(Brand)*

Analysis	Percent of the Protein	Residues Mols per Mol of Protein
Glycine	6.4	29
Alanine		
Valine	7.1	21
Leucine	10.4	27
Isoleucine	10.8	28
Proline	5.0	15
Phenylalanine	6.4	13
Cysteine	0.5	2
Half-cystine	1.64	4
Methionine	1.7	4
Tryptophan	2.36	4
Arginine	1.0	2
Histidine	0.9	2
Lysine	0.9	2
Glutamic Acid	11.9	28
Aspartic Acid	16.0	41
Serine	12.2	40
Threonine	9.6	28
Tyrosine	8.5	16
Total	115.15	306
Total, Calculated as Amino Acid Residues	97.47	
Total, Amino Acid N as % of Total N	99.2	
Total N	14.6	
Total S	0.94	
Total P	0.09	
Free alpha amino N	0.081	
Total amino N	0.162	

*Unpublished data cited by J. H. Northrop, Crystalline Enzymes. Columbia Univ. Press, New York, 1948, pp. 26, 74.

for this reason. Tiselius, Henschen, and Svensson (1938) have shown that pepsin migrates as if negatively charged even in 0.1N HCl. These results confirm those of Northrop (1925) who studied the distribution of pepsin between egg albumin particles and the surrounding solution. It was found that pepsin behaves like the chloride ion in this respect. This result was taken to indicate that pepsin is a negatively charged, monovalent acid, even at pH 1.0. Ionographic studies indicate the pI to be well below pH 2 (Williamson 1951). An examination of the amino acid content of pepsin does not suggest an explanation for the unique electrochemical behavior of the enzyme since the calculated isoionic point is Ca. 3.0, using the pK's of the free amino acids. As yet no explanation for this phenomenon has appeared in the literature. Other investigators have reported isoelectric points near pH 3.0 but these have been found to be due to the presence of non-protein nitrogen.

The amino acid composition of pepsin is given in Table VI. A few points in this analysis are worthy of note. The content of dicarboxylic amino acids is high while that of basic and sulfur containing amino acids is exceptionally low. Free alpha amino nitrogen is also low. Pepsin contains one atom of phosphorous per molecule. This phosphorous is evidently an integral part of the molecule (Northrop et al, 1948). Flavin (41) has recently reported that the phosphorus occurs as phosphate and is in the form of O-phosphoserine. This phosphorylated

serine residue occurs in the sequence, threonylphosphoserylglutamic acid.

Pepsin is known to have such free groups as carboxyl, amino and phenolic hydroxyl. The introduction of twenty acetyl residues per molecule of pepsin causes a great loss of peptic activity. Iodination of the tyrosine residues causes a similar inactivation. Controlled acetylation of the free amino groups produces only a slight decrease in activity. From this data it has been concluded by investigators (Herriot, 1935,1937, Herriot and Northrop, 1934) that the phenolic hydroxyl groups are closely associated with the activity of pepsin.

Several of the characteristics cited above have led to the particular choice of pepsin as a subject for structural studies. The results of these studies may clarify the question of the specificity of pepsin in particular and proteolytic enzyme specificity in general. The unique isoelectric point of pepsin, which is unexplainable in terms of mere chemical analysis, may be explainable when the structure of the molecule is known. The low molecular weight of the protein will make the study relatively less difficult than if a larger molecule were studied. Finally, the enzyme is well characterized and readily available in a highly purified crystalline form.

CHAPTER IV

OUTLINE OF EXPERIMENTAL PROCEDURE

During the course of the present studies on pepsin, several techniques have been applied to the problem. The use of most of them has come about as a result of efforts to find a satisfactory quantitative method for the determination of the terminal free carboxyl group.

At the time these investigations were initiated, the only adequately studied chemical method for the C-terminal identification was the reduction of these amino acids with lithium aluminum hydride. Carboxypeptidase had been used earlier by Lens but the specificity requirements of this enzyme are such that certain amino acids cannot be detected. Therefore, the first method which was employed in the investigation of the C-terminal residue of pepsin was a reduction with lithium aluminum hydride. The technique used was a modification of that proposed by Fromageot. No quantitative studies were made but the amino alcohols were isolated and identified by chromatography. There was a degree of uncertaintity in the chromatographic identifica-

tions and so another method was employed to improve the identification.

The technique resorted to next was an enzymatic digestion with carboxypeptidase. This enzyme splits off terminal amino acids which possess free carboxyl groups. The tentative identifications made with the reduction method indicated that carboxypeptidase could be employed since the specificity requirements would be favorable. This study corroborated the conclusions drawn from the reduction experiments and in addition allowed a semiquantitative estimation of the number of C-terminal groups.

In an attempt to improve the quantitative aspects of the study, the thiohydantoin methods reported by Waley and Watson (123) and Baptist and Bull (9) were employed. Instead of recovering the parent amino acids of the C-terminal hydantoin as these authors did, the thiohydantoin itself was isolated and measured at its absorption maximum, in a spectrophotometer. Recoveries were poor as had also been reported by Waley and by Baptist. In addition, interfering substances made the measurements unreliable and without further attempts to remove these materials a fourth method was investigated.

This method made possible spectrophotometric measurements in conjunction with the lithium aluminum hydride reduction. The reduction was carried out as indicated previously, but the hydrolysate was treated with 2,4-dinitrofluorobenzene, converting

the amino alcohols and amino acids to the DNP-derivatives. The amino alcohol derivatives were then isolated, identified and measured in a Beckman spectrophotometer. At the time this technique was being developed there were only a few reports of the use of spectrophotometric methods in the study of terminal residues. At the present time the literature shows an increasing use of spectral measurements, where possible, in preference to chemical methods.

In addition to the identification of the C-terminal residue, the sequence of amino acids adjacent to the terminal free amino residue was investigated. This study is a continuation of previous work (125) on this end of the pepsin molecule. In the N-terminal sequence studies, a modification of the method of Edman was developed. This method is less laborious than Sanger's technique of partial hydrolysis of the DNP-protein and in addition served to corroborate the previous studies on the N-terminus in which the terminal group was identified as the DNP-amino acid.

Pepsin Preparation The pepsin used in these experiments was Armour's three times crystallized pepsin (Lot. No. 128-191) of porcine origin. Electrophoretic analysis by Armour Research Laboratories produced the pattern shown in Figure 4. The analysis was made in acetate buffer ($\Gamma/2$=0.1) at pH 4.3. The largest component represents 99 per cent of the total protein.

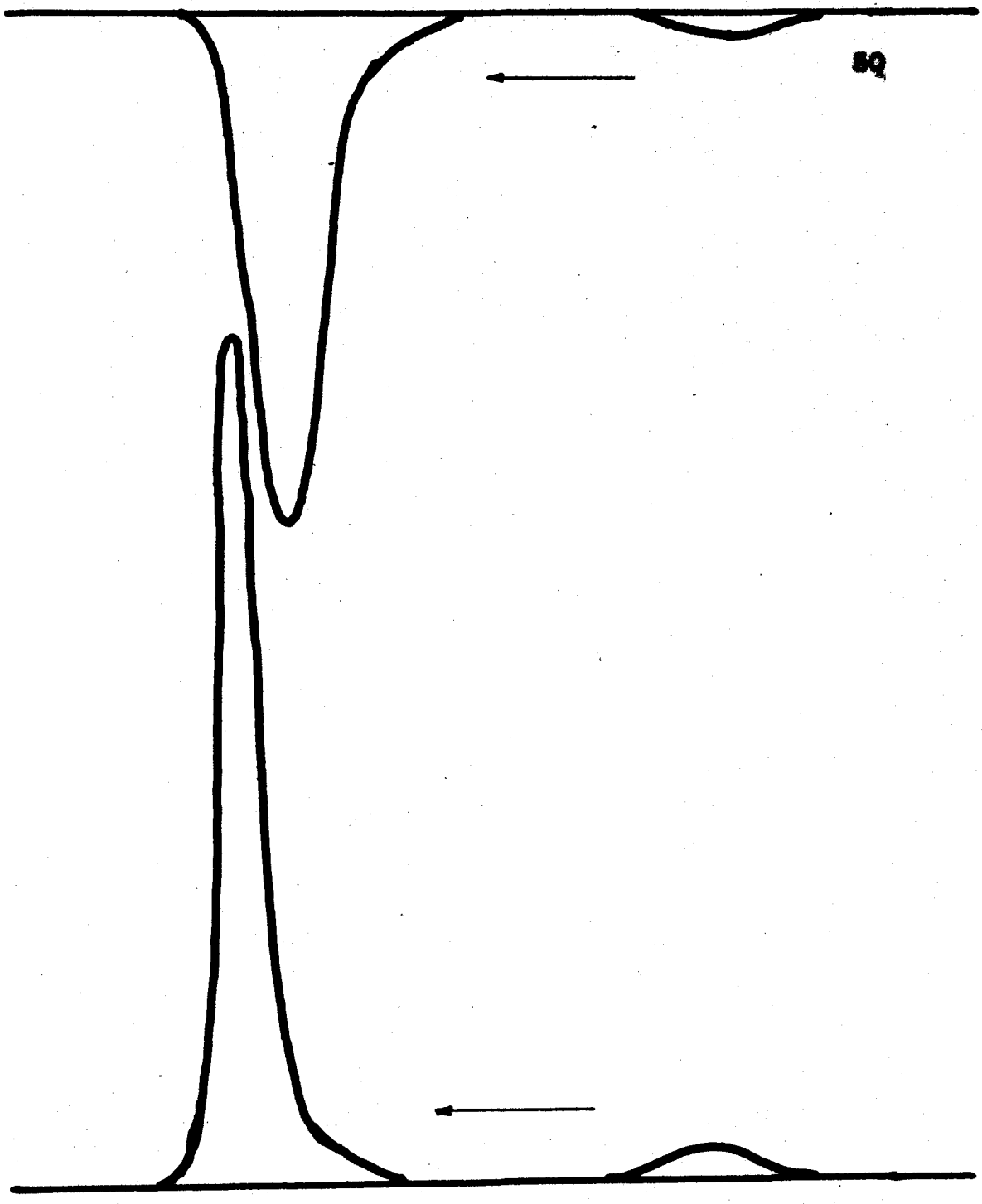


FIGURE 4

THE ELECTROPHORETIC PATTERN OF PEPSIN

The preparation contained 1.62 per cent moisture, 5.8 per cent sulfated ash and 15.2 per cent nitrogen on a moisture free basis. The pepsin was prepared by the method of Northrop (81).

Carboxypeptidase Preparation The carboxypeptidase used in these experiments was a six times recrystallized preparation from Armour Research Laboratories (Lot. No. R 381-169), containing 24 mg of enzyme per milliter. The product was claimed to be essentially free of tryptic and chromotryptic activity.

CHAPTER V

REDUCTION EXPERIMENTS ON PEPSIN

When a protein is treated with lithium aluminum hydride there is a reduction of all the free carboxyl groups. The reduction of the terminal free carboxyl group gives rise to a primary amino alcohol in the hydrolysate of the reduced protein. The reduction of β-and γ-carboxyl groups of aspartic and glutamic acids leads to the corresponding β-and γ-hydroxy-amino acids in the hydrolysate. These compounds are obviously analagous to the amino acids serine and threonine and hence do not interfere in the determination. The amino alcohols are extracted from the hydrolysate of the reduced protein at alkaline pH with ether while the amino acids remain behind.

The work of Fromageot on the reduction technique was the first to appear in the literature. He used unesterified protein in his procedure, which was based on the work of Nystrom and Brown (83). In addition he used the methods of Karrer, Suter and Portmann for the preparation of the amino alcohols from the amino acid esters. Chibnall and Rees (19) have used

esterified protein in their investigations. There appears to be no special advantage of one method over the other. Nystrom has indicated that 80-90 per cent yields can be obtained by the reduction of the unesterified amino acids. Karrer et al. have prepared the amino alcohols corresponding to most of the common amino acids. The present investigations were started prior to the publication of the work of Chibnall and Rees. Since there is no reported advantage of the use of esterified over non-esterified protein, no changes were made in the procedure described below:

Reduction Technique The apparatus used is the same as that described below for the preparation of the amino alcohols. As finally developed the procedure is as follows: 100 mgm of pepsin are suspended in 50 ml of anhydrous ethyl ether. To this suspension are added 50 ml of an ether solution of 75 mgm of lithium aluminum hydride. After addition is complete, the mixture is stirred for eight hours at about 40°C. At the end of this time the excess lithium aluminum hydride is destroyed by the cautious addition of water. The mixture is stirred for an additional half-hour. The mixture is then acidified with 9 N HCl. The ether is evaporated off at about 50°C. After evaporation of the ether, 35 ml of 6 N HCl are added and the reduced protein is hydrolyzed under reflux for twenty-four hours. At the end of this time the solution is taken to dryness three times to remove HCl. Sufficient water is added to make a syrup

which is cooled in an ice bath and brought up pH 9.0 with 15 M sodium hydroxide. The syrup is absorbed onto filter-cel and extracted with ethyl ether for thirty hours in a Soxhlet. The ether extract is dried over anhydrous sodium sulfate and evaporated under reduced pressure. The residue is taken up in 0.5 ml of acid-ethanol and used for chromatographic identification.

Chromatography of Amino Alcohols Identifications of the reduction product isolated from pepsin were made by unidimensional paper chromatography. Whatman No. 1 paper was used and the chromatograms were developed by descending solvent flow. The distance of migration was always at least 30 cm. The time required for development was about twelve hours. Chromabographic cabinet temperature was maintained at 20 ± 2°C. The color developing reagent for the amino alcohols was a 0.4 per cent solution of ninhydrin in 90 per cent butanol-10 per cent phenol. The dried chromatograms were sprayed with the reagent and heated in an oven at 100°C for five minutes to develop the color. The color obtained with amino alcohols is about ten to twenty times less intense than that obtained with the amino acids.

The solvent used in these chromatographic studies was n-butanol-acetic acid-water (77:6:17). This is a single phase system and is used as both developing solvent and to saturate the atmosphere of the chromatographic chamber.

Preparation of Amino Alcohols Authentic amino alcohols were prepared for use in the chromatographic comparisons. At first a few with widely different R_f's were prepared to get a general idea of the identity of the unknown alcohols. Then samples of those alcohols with R_f's close to that of the unknown were prepared.

The amino alcohols corresponding to alanine, leucine valine, tyrosine, aspartic acid and glutamic acid were prepared from the ethyl esters of the amino acids. The amino alcohols were characterized by chromatographic behavior only. The following description of the preparation of valinol (2-amino-3-methylbutanol-1) is illustrative of the method used. This is essentially the method of Karrer, Portmann and Suter (68). The general reaction may be written as follows:

$$4H_2N\cdot RCH\cdot COOEt + LiAlH_4 \rightarrow 1/2(H_2N\cdot RCH\cdot CH_2O)_4\,LiAl + 1/2\,(ETO)_4LiAl$$

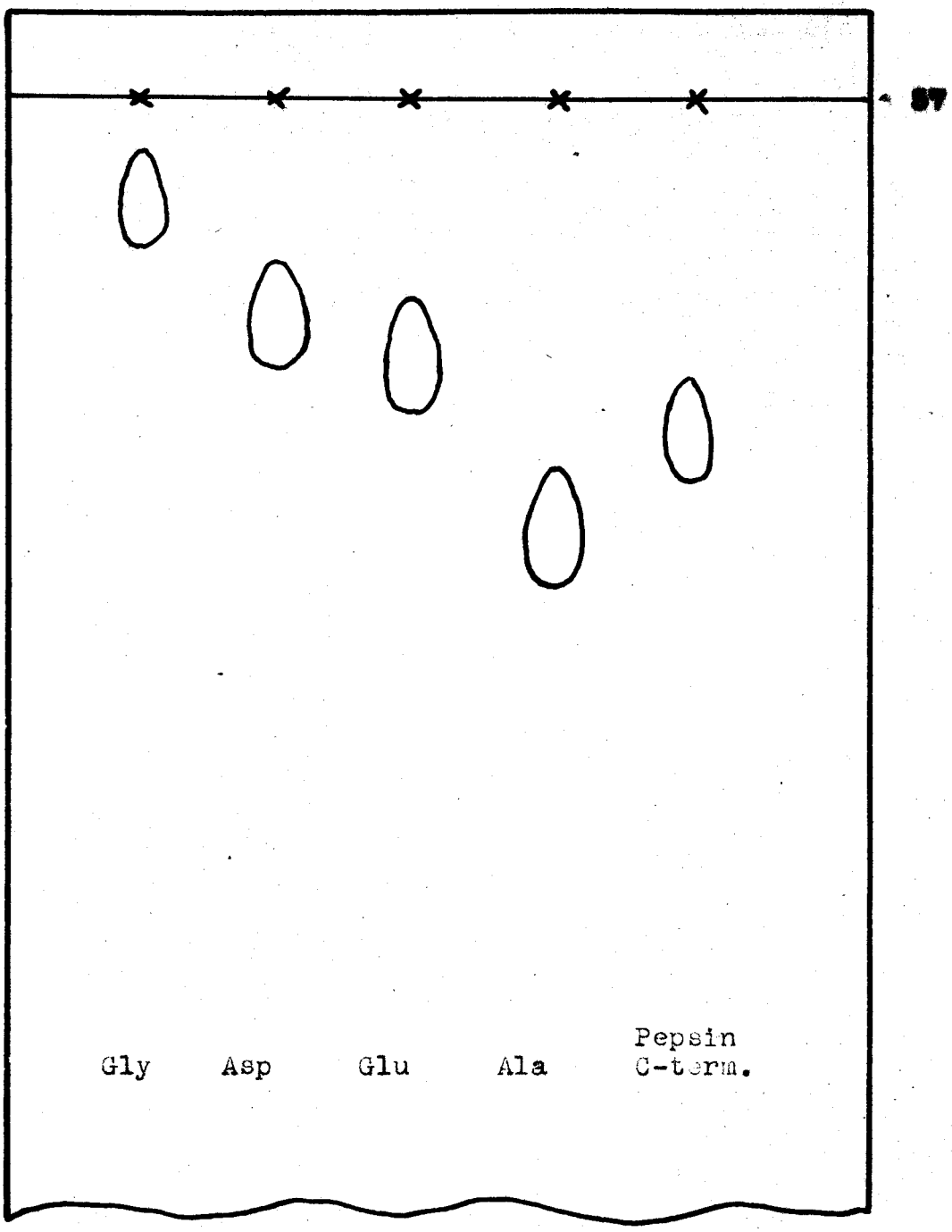
$$\downarrow$$

$$2H_2N\cdot RCH\cdot CH_2OH + 2\,EtOH$$

1 gm of valine was suspended in 50 ml of absolute ethanol. Dry HCl was bubbled into the suspension until all the valine was dissolved. The ethanol was then evaporated off under reduced pressure. The residue was taken up in a little water, covered with a layer of ether and cooled to 0°C. The cooled solution was then made alkaline with concentrated sodium

hydroxide and the ester extracted into the ether layer. The ether extract was dried overnight over anhydrous sodium sulfate. After drying, the solution of ester was transferred to a dry three necked flask. The flask was fitted with a dropping funnel, an ebulating tube for maintaining an atmosphere of dry nitrogen, and a reflux condenser fitted with a soda lime tube to eliminate atmospheric CO_2 from the reaction vessel. A solution of 1 gm of lithium aluminum hydride in 50 ml of ether was then added dropwise to the ester solution. The reducing solution was added over a period of ten minutes. The reaction mixture was allowed to stand for an additional fifteen minutes with agitation (bubbling nitrogen) before the excess lithium aluminum hydride was decomposed by the cautious addition of 5 ml of water. The resulting precipitate was filtered, washed twice with 50 ml of ether and finally with three 50 ml volumes of hot 95 per cent ethanol. A pale yellow oil was obtained after evaporation of the combined filtrates. The product was dissolved in acid-ethanol and stored in the refrigerator. Identifications were made by comparing the relative rates of movement in paper chromatography with the relative R_f's reported by Fromageot.

Results of the Reduction of Pepsin Figure 5 illustrates the pattern observed when the amino alcohol isolated from reduced pepsin was chromatographed on paper. Included in this chromatogram are a number of authentic amino alcohols for comparative purposes. The alcohol moved with an R_f intermediate between those



F IGURE 5

IDENTIFICATION OF THE AMINO ALCOHOL FROM REDUCED PEPSIN

of alaninol and glutamidiol. None of the amino alcohols corresponding to the common amino acids move with this R_f. Therefore, it was concluded that the material isolated from reduced pepsin was either alaninol or glutamidiol. Fromageot (47) had reported that the presence of lithium and aluminum salts would slow up the chromatographic movement of amino alcohols. These salts can be carried over from the reduction step during the ether extraction. Therefore, the tentative conclusion was drawn that the amino alcohol isolated from reduced pepsin was alaninol.

CHAPTER VI

EXPERIMENTS ON THE DIGESTION OF PEPSIN WITH CARBOXYPEPTIDASE

The use of carboxypeptidase for the determination of C-terminal residues in proteins does not give unequivocal results. Due to the specificity requirements of the enzyme (111), certain amino acids cannot be detected, e.g., proline and hydroxyproline. When basic or acidic amino acids (lysine, glutamic acid, etc.) occupy the terminal position, the cleavage is slow and requires relatively large amounts of enzyme. Aromatic amino acids are most easily cleaved by this exopeptidase, followed by the long chain aliphatic amino acids and finally by the shorter chain acids. Considering this relative specificity, it is possible to visualize a number of anomalous results. For example, if the C-terminal residue is hydrolyzed slowly and the second amino acid from the carboxyl end is hydrolyzed rapidly, it is possible that approximately equivalent amounts of both amino acids may be cleaved in the same time interval. The results would indicate two chains whereas in reality there is only one. Also, if a

protein consisted of two open chains one of which terminated in proline or hydroxyproline, the carboxypeptidase method would indicate only a single chain. These difficulties were not encountered in the present study since it had already been determined that there was only one species at the C-terminal end of pepsin. The method was therefore used only to corroborate the tentative conclusion drawn from the reduction experiments and to obtain a semi-quantitative estimation of the number of C-terminal residues.

Carboxypeptidase Digestion 200 mgm of pepsin were dissolved in 5 ml of 0.1 M phopphate buffer at pH 7.6 0.3 ml of a carboxypeptidase suspension containing 8 mg of enzyme was then added. The temperature during digestion was 28-29°C. At 1,2,4, and 5 hours, 1 ml aliquots were withdrawn and added to 1 volume of 0.6 N HCl in 95 per cent ethanol. The precipitate was centrifuged and aliquots of the order of 20-30λ were applied to a paper chromatogram. Known quantities of amino acids corresponding to 1, 2 and 4 end groups per mole of pepsin were applied to the same chromatogram. The chromatogram was developed with 77 per cent ethanol. The spots were developed with ninhydrin and the intensities compared in order to estimate the number of end groups.

The minimal distance, direction and time of solvent flow, and the type paper were the same as described for the

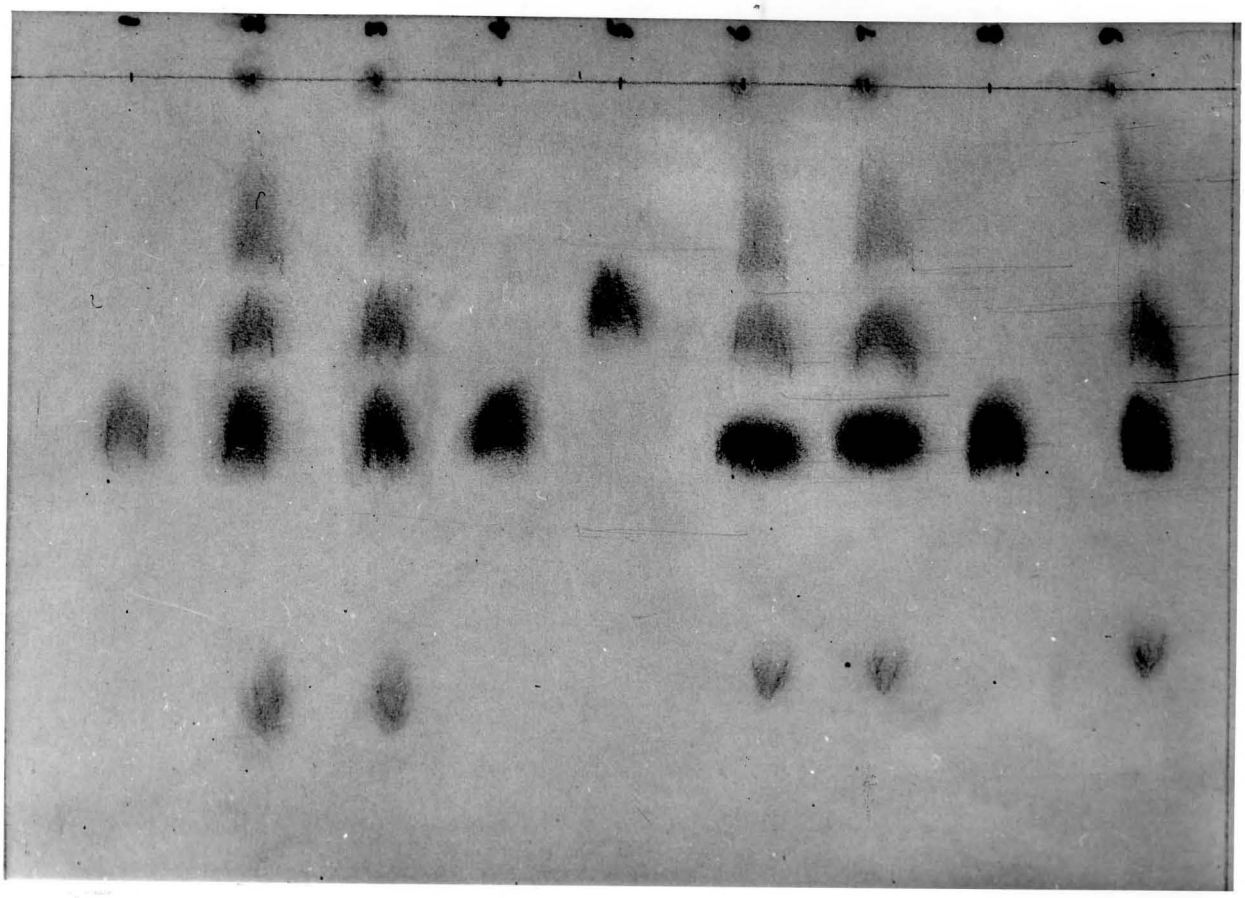


FIGURE 6

SPOT DILUTION CHROMATOGRAM OF AMINO ACIDS LIBERATED FROM PEPSIN BY THE ACTION OF CARBOXYPEPTIDASE

(1) 1.5μg alanine (2) 30λ 4 hr. digest (3) 30λ 5 hr. digest (4) 3μg alanine (5) 2.4μg glutamic acid (6) 3μg alanine+30λ 4 hr. digest (7) 3μg alanine + 30λ 5 hr. digest (8) 6μg alanine (9) 60λ 5 hr. digest. 1.5μg alanine=1 end group, 2.4 μg glutamic acid=1 end group.

chromatography of the amino alcohols. The color developing reagent was the same ninhydrin solution used in the chromatography of the amino alcohols.

Results Several amino acids and some slow moving peptide material appeared on the chromatograms of the carboxypeptidase digests of pepsin. The most intense spot corresponded in movement to alanine. A spot dilution chromatogram of the alanine spot is shown in Figure 6. The chromatograms faded before a photograph could be made and therefore a drawing was prepared to illustrate the results.

If there are two C-terminal alanine residues in pepsin, the amount of alanine to be expected in the 30λ aliquot of the supernatant taken for chromatography is about 3μg. Approximately 2μg of glutamic acid correspond to one end group of this amino acid. From the figure it can be seen that the intensity of the alanine spot corresponds most closely to two end groups. The estimations were made by visual comparison. The intensity of the glutamic acid spot is less than one end group. This amino acid was also estimated since it was the next most intense spot on the chromatogram. The patterns on the right of the figure illustrate the results obtained when double amounts of all the components of interest were chromatographed. Patters 6 and 7 are mixtures of alanine with the four and five hour digests respectively. These patterns also serve as mixed chromatograms further strengthening the identification of the C-terminal residue as

alanine.

In addition to the glutamic acid and alanine spots, there were spots corresponding in movement to leucine (or isoleucine), phenylalanine, methionine, and an unidentified amino acid with an R_f smaller than glutamic acid. The movement of this spot in 77 per cent ethanol corresponded to the group, serine, glycine, aspartic acid. The faster moving spots are not shown in the figure. The substances shown in the figure are, from top to bottom, peptide material, unidentified amino acid, glutamic acid, alanine and methionine.

CHAPTER VII

INVESTIGATION OF THE C-TERMINAL GROUP OF PEPSIN AS THE THIOHYDANTOIN

In the methods published by Waley and Watson and by Baptist and Bull the recoveries of the thiohydantoins were measured by spot dilution chromatography of the amino acid after hydrolysis of the thiohydantoin. No attempt to determine the losses that might be incurred in this step were reported. In the present studies, the thiohydantoin itself was measured spectrophotometrically in an attempt to improve the quantitative aspects by eliminating the hydrolysis of the hydantoin. The method was used in the hope of obtaining better quantitative data than that yielded by the carboxypeptidase experiments.

Thiohydantoin Technique In the chromatographic studies, 50 mg of pepsin were dissolved in 5 ml of acetic anhydride containing 10 per cent acetic acid and 250 mg of ammonium thiocyanate. The mixture was kept at 40-50°C for four hours with mechanical stirring. At the end of this time the protein was precipitated with 10 volumes of acetone. The precipitate was washed three times with 5 ml volumes of acetone. The dried protein was

suspended in 2 ml of 10 per cent HCl and heated over a boiling water bath for one hour. The protein was again precipitated with 10 volumes of acetone and washed as before. The supernatant and washings were taken to dryness and the residue was dissolved in about 0.5 ml of ethanol. This solution was used in the chromatography.

In the spectrophotometric studies, the residue, after evaporation of acetone, was taken up in 10 ml of 0.25 M phosphate buffer at pH 6.45. This solution was then extracted three times with 10 ml volumes of ethyl acetate and the extract was diluted to 50 ml. The measurements were made at 265 mμ in a Beckman Model DU spectrophotometer.

Blank determinations were made by subjecting the protein to the same procedure except that the ammonium thiocyanate was omitted in the initial reaction.

Recovery experiments were carried out by adding 1 ml of a solution containing 3.49×10^{-5} moles of 5-methyl-2-thiohydantoin and 1 ml of 20 per cent HCl to a protein preparation produced in the same manner as in the blank determination or as in the experimental determination. The hydrolysis and subsequent procedure were the same as described for the spectrophotometric determination.

Chromatography The paper, minimal distance of solvent flow and direction of solvent flow were the same as described for the chromatography of the amino alcohols. The time required for the

minimal distance of a solvent flow was about 3-4 hours. The solvents used were those proposed by Sjoquist (110) with the substitution of isooctane for heptane. For initial separations the solvent used was isooctane-n-butanol-85 per cent formic acid, 40:20:40 (Sj-B). The components were shaken in a separatory funnel and the lower layer was used to saturate the atmosphere of the chamber. For secondary separations two other solvents were used, isooctane-n-butanol-85 per cent formic acid, 20:40:40 (Sj-c) and isooctane-anhydrous pyridine, 70:30 (Sj-A). These two solvents are single phase.

The thiohydantoins were located on a paper by spraying with an iodine-potassium iodide-sodium azide reagent which gives a fading brown background with white spots (18). The background can be stabilized by pretreating the paper with starch. The background is then blue. The reagent is prepared by mixing, immediately before use, equal volumes of M/100 I_2 in M/2 KI and M/2 NaN_3. The sulfur in the thiohydantoins is said to catalyze the reaction

$$2\ NaN_3 + I_2 = 2NaI + 3N_2.$$

The chromatograms were dried five minutes at 95°C before spraying with the reagent.

Preparation of Thiohydantoins The thiohydantoins corresponding to the amino acids leucine and alanine were prepared according to the method of Johnson and Nicolet. The thiohydantoins were

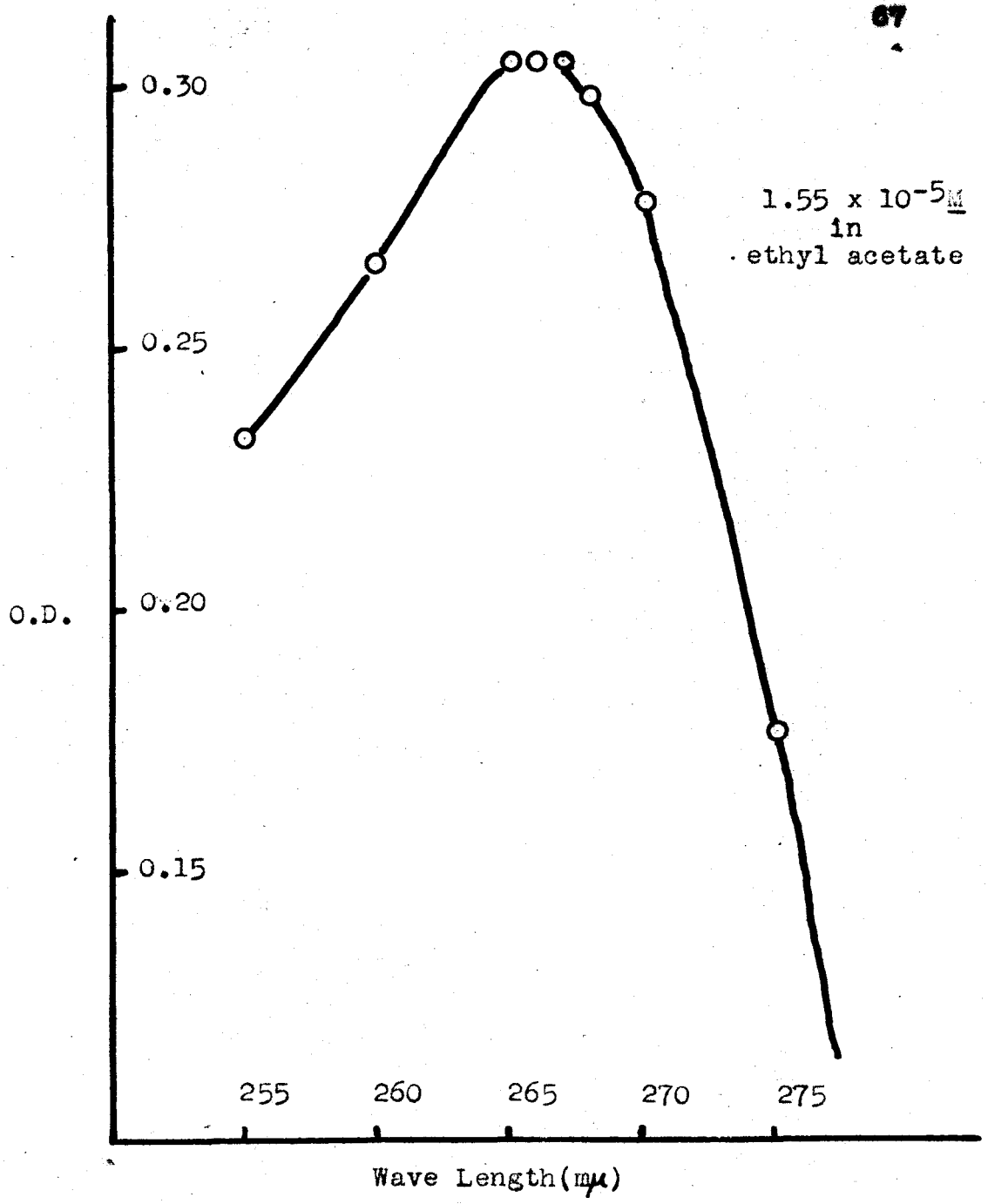


FIGURE 7

ABSORPTION SPECTRUM OF 5-METHYL-2-THIOHYDANTOIN

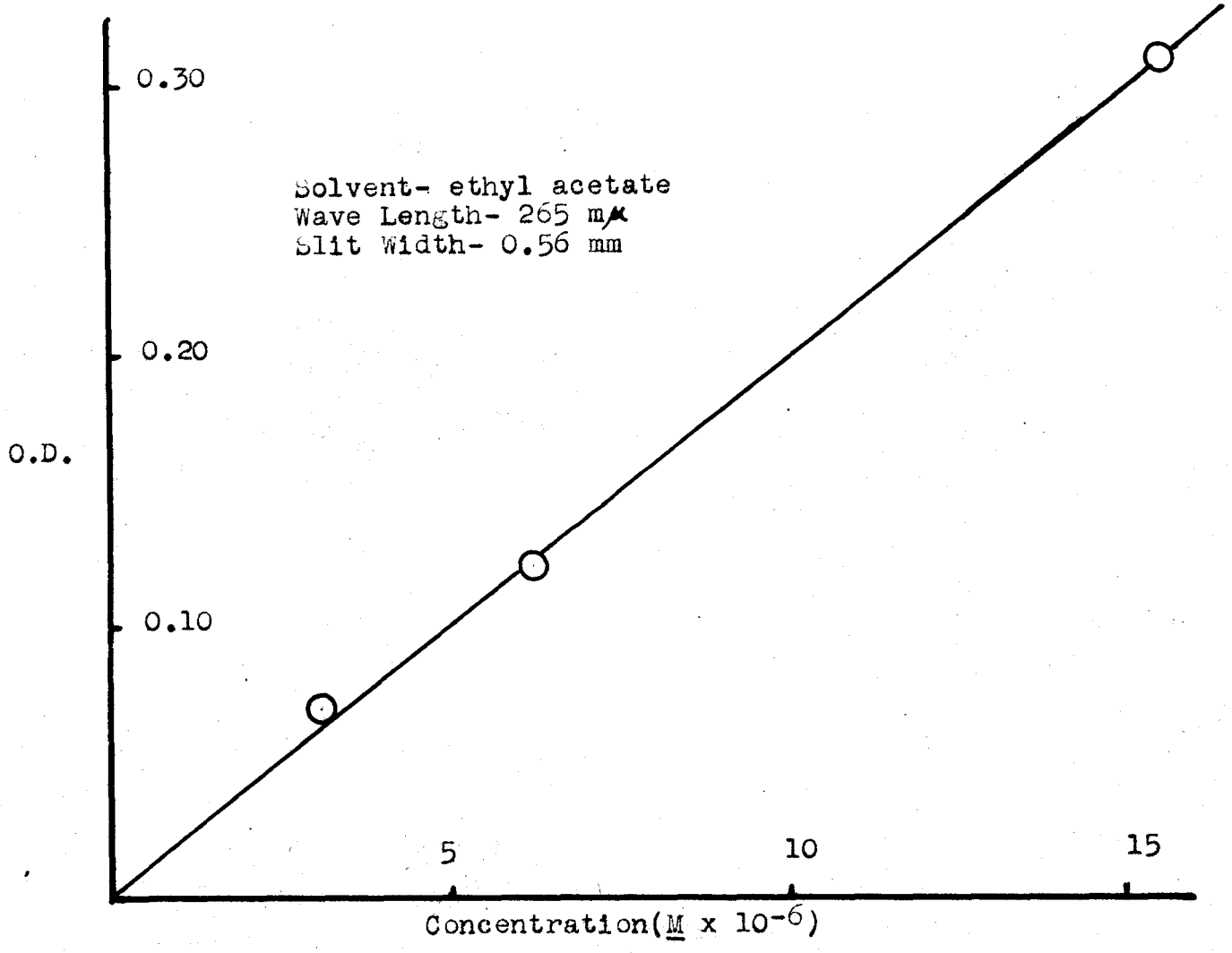


FIGURE 8

STANDARD CONCENTRATION CURVE OF 5-METHYL-2-THIOHYDANTOIN

obtained in crystalline form and the melting points corresponded to those in the literature. 5-Methyl-2-thiohydantoin: M.P. Lit. 163-165°C, Obs. 162-163°C. 5-isobutyl-2-thiohydantoin: M.P. Lit. 174-176°C, Obs. 175-177°C.

The preparation of the alanine thiohydantoin (5-methyl-2-thiohydantoin) is a typical procedure. 2 gm of alanine and 2.6 gm of potassium thiocyanate were ground together in a mortar. The mixture was suspended in 10 ml of acetic anhydride and placed over a boiling water bath for thirty-five minutes. The solution was then cooled to room temperature and 60 ml of cold water were added and the solution placed in a refrigerator for about one hour. The orange-yellow crystals were filtered, dried and recrystallized from aqueous ethanol. The crystals of N-acetyl-5-methyl-2-thiohydantoin were a light tan in color. M.P. Lit. 171°C, Obs. 169-171°C. 0.5 gm of this N-acetyl-5-methyl-2-thiohydantoin was suspended in 5 ml of 10 per cent HCl and refluxed for forty minutes. The solution was then cooled to room temperature and placed in an ice bath. A large crop of white crystals was obtained. The 5-methyl-2-thiohydantoin was recrystallized from 95 per cent ethanol. M.P. Lit. 163-165°C, Obs. 162-163°C.

Results The chromatographic studies on pepsin using the thiohydantoin technique gave negative results, i.e., there was no sulfur on the chromatograms of the acetone extract. These re-

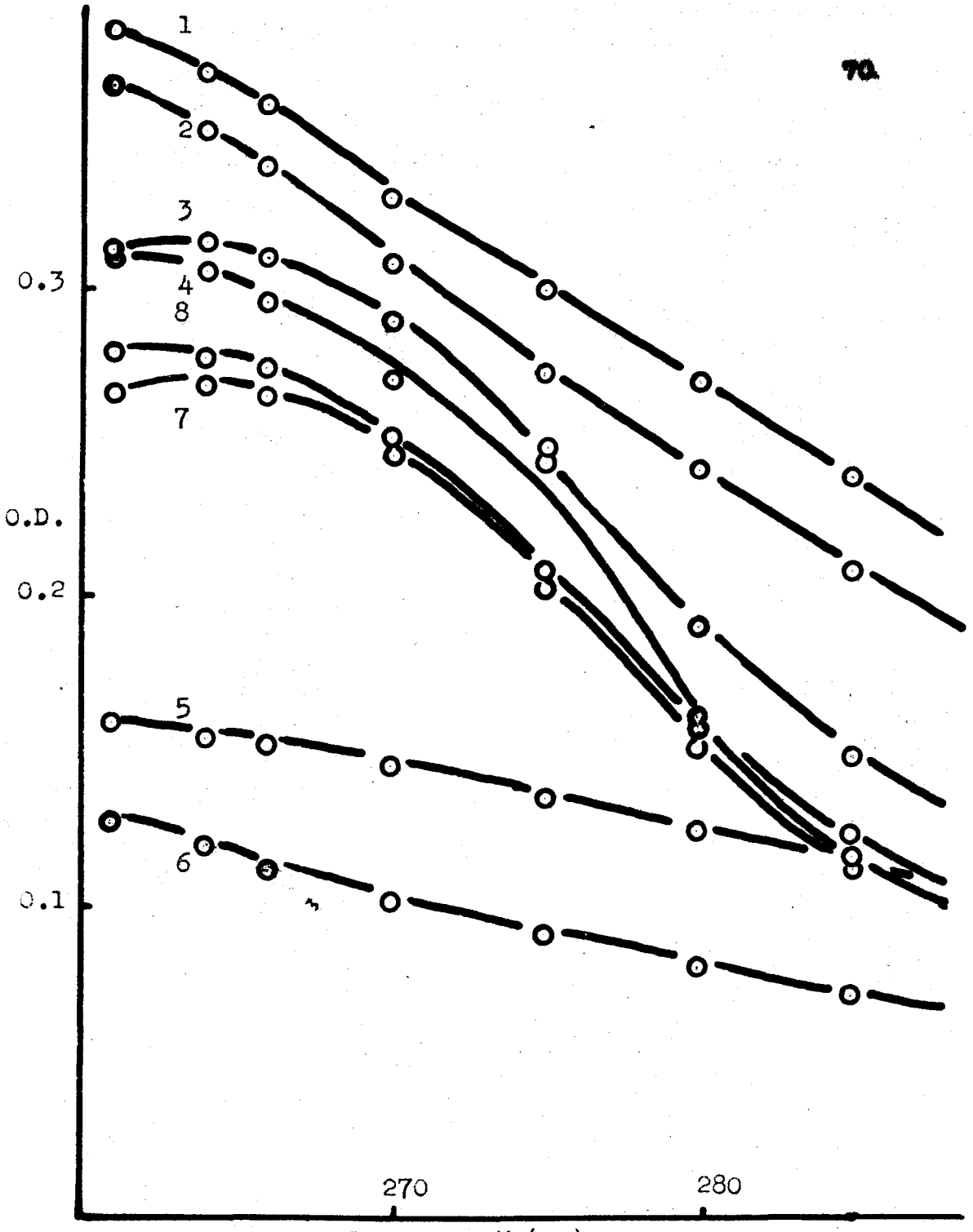


FIGURE 9

SPECTRA OF ETHYL ACETATE EXTRACTS FROM THIOHYDANTOIN EXPERIMENTS

sults could be interpreted to mean that the method was unsuccessful at one of the steps or that there was no C-terminal residues, in pepsin. That the first conclusion and not the last was most probably correct was demonstrated by the fact that a protein known to have a C-terminal group also gave negative results. The technique was applied to lysozyme which has a single C-terminal group of leucine. No evidence of 5-isobutyl-2-thiohydantoin could be found on the chromatograms in this experiment.

Figure 9 summarizes the results obtained in the spectrophotometric studies. Curves 1 and 2 were made at one-half the dilution used to make the other curves and therefore cannot be directly compared with them. A comparison of these curves with the absorption spectrum of 5-methyl-2-thiohydantoin shows a similarity only for curves 3 and 4 and curves 7 and 8. These are the preparations to which 5-methyl-2-thiohydantoin had been added. The curves are flatter than the spectrum of the pure compound indicating the presence of interfering substances. Curves 1 and 2 indicate that there was some ultraviolet absorbing material extracted even in the blank determinations. Calculated at 266 mμ this amounts to 0.0125 moles of 5-methyl-2-thiohydantoin per mole of pepsin used. This value was substracted from the values obtained from calculations made with the other curves. In the recovery experiments, curves 3 and 4, a corrected value of 22.5 per cent was obtained. This would indicate that the losses

reported by the other authors were not incurred during the hydrolysis of the thiohydantoin but rather may be due to adsorption on the protein or incomplete cleavage from the protein. In the experiments represented by curves 7 and 8, after application of the recovery factor practically all (3.1×10^{-6} moles) of the added 5-methyl-2-thiohydantoin was recovered. There was, however, no additional hydantoin from the reaction with ammonium thiocyanate. Curves 5 and 6 indicate, from the shape, that there is probably no thiohydantoin present. If there were one C-terminal group in pepsin, the optical density at 266 m should be at least as high as that in curves 3 and 4. The actual calculated value was 0.12 end groups and when corrected for recovery was increased only to 0.51 end groups.

CHAPTER VIII

DETERMINATION OF THE C-TERMINAL GROUPS OF PEPSIN AS THE DNP-AMINO ALCOHOLS

Since the results of the thiohydantoin technique indicated no improvement over the results obtained by Waley and Watson and by Baptist and Bull, another method of spectrophotometric determination of the C-terminal group was developed.

Dinitrophenyl compounds are relatively stable and have an adequate absorbancy coefficient at 350 mμ for use in spectrophotometric measurements of small amounts of materials. The use of DNP-amino acids has met with much success in the determination of N-terminal residues and therefore it was decided to convert the C-terminal residue, in the form of the amino alcohol, to the DNP-derivative.

DNP Treatment of Reduced Pepsin The reduction procedure was the same as described in the preceding section. After the hydrolysis the solution was taken to dryness once under reduced pressure, made to 50 ml and neutralized with solid sodium bicarbonate. To this mixture was added 0.25 gm of solid $NaHCO_3$

and 45 ml of acetone containing 0.21 gm of DNFB. The mixture was stirred for three hours after which time 0.2 gm of glutamic acid were added to react with the excess DNFB. The mixture was stirred for an additional two hours. After this time, the mixture was centrifuged and the precipitate was washed twice with 10 ml volumes of water and three times with 10 ml volumes of ethanol. The acetone and ethanol were evaporated off under reduced pressure. The resulting solution was extracted three times with 50 ml volumes of ether. The ether extract was back-extracted twice with 50 ml volumes of 0.05 M NaOH and once with 50 ml of water.

After drying over anhydrous sodium sulfate, the ether was removed from the extract and the residue was dissolved in 10 ml of 95 per cent ethanol. Two ml aliquots of this solution were used in the identifications and measurements. The aliquot was first passed through an alumina column to remove dinitrophenol. The eluate was taken to dryness, dissolved in a minimal amount of the solvent to be used on the silicone treated column and then chromatographed. The proper fraction was collected, the solvent evaporated under reduced pressure, and the residue made to volume in 95 per cent ethanol. Measurements were made in the Beckman spectrophotometer at 350 mμ. The concentration of the eluate was determined from a standard curve. Identifications were made by running mixed chromatograms with authentic

DNP-amino alcohols.

In recovery experiments, 0.1 ml of a solution of alanine containing 7.7 mg/ml was added to the ether suspension of pepsin. Sufficient excess lithium aluminum hydride was added to react with the water from the alanine solution. The procedure was then carried out as described in the previous section.

Chromatography

DNP-amino alcohols were separated on columns of Hyflo Supercel impregnated with silicone. This treatment rendered the adsorbent hydrophobic. The process of chromatography on columns of hydrophobic adsorbents is called inverse phase chromatography. Known mixtures separated on these columns contained about 100 g of each component.

The silicone treated Hyflo Super-cel was prepared as follows: 50 gm of Hyflo Super-cel were suspended in 125 ml of 10 per cent acetic acid. The suspension was filtered and washed twice with 125 ml volumes of 10 per cent acetic acid. The material was then washed with water until neutral. 30 ml of ethyl acetate then were added and the resulting slurry was filtered on a Buchner funnel. The filter cake was powdered and dried overnight at 100°C. The dried material was ground in a mortar. The powder was transferred to a glass stoppered round bottom flask. 75 ml of a 5 per cent solution of Rhodorsil 240

(Societe des Usines Chimiques, Rhone-Poulenc, Paris) in chloroform were added and the mixture was shaken vigorously for one minute. Rhodorsil 240 is a partially hydrolysed polymer of chlorosiloxane. The mixture was filtered on a Buchner, transferred to filter paper and allowed to dry in air. The last traces of chloroform were removed by washing twice with 75 ml volumes of ethanol. HCl from the Rhodorsil was removed by washing with distilled water until the filtrate gave no test with silver nitrate. The final product was then dried in an air oven at 100°C for forty hours.

The column was prepared by grinding, in a mortar, 6 gm of the Hyflo Super-cel and 1.2 ml of the organic phase of the solvent system to be used. The Super-cel was then suspended in 30 ml of the aqueous phase of the solvent and poured carefully down the side of the column. The column was rotated between the hands between each addition to remove any trapped air bubbles. The column was then allowed to settle by gravity. It was found that columns packed by positive pressure or by vacuum did not give satisfactory results.

When the meniscus of the solvent just touched the surface of the column, the mixture to be separated was added in about 1-2 ml of the aqueous phase of the solvent and allowed to settle into the column. The developing solvent was then added carefully so as not to disturb the surface of the column.

It was found during the course of the studies that separations on these columns could only be effected below 18°C and should be carried out preferably at about 12-15°C. For this purpose the columns were jacketed and tap water in this temperature range was circulated through them during the run.

In order to remove dinitrophenol from mixtures before applying them to the Super-cel columns, the mixtures were first passed through an activated alumina column about 3 cm. long. The dinitrophenol remains in the upper portion of the column while remaining components are washed through. The solvent used on the alumina column was 95 per cent ethanol.

The solvents used on the silicon treated columns were those employed by Jutisz (67). These were ternary systems of acetic acid, butyl acetate and water. The components were mixed in either the ratio 50:10:100 or 20:10:100 and shaken in a separatory funnel. The two layers were allowed to equilibrate under running tap water. The temperature at which the aqueous layer was drawn off was 12°C.

A standard concentration curve for DNP-alaninol is shown in Figure 10. The curve was prepared from solutions of the compound in 95 per cent ethanol. The measurements were made in a Beckman Model DU spectrophotometer with a photomultiplier attachment. The sensitivity control of the photomultiplier unit was set at "4". The optical densities were read at 350mμ with a slit width of 0.03 mm. A tungsten light source was used.

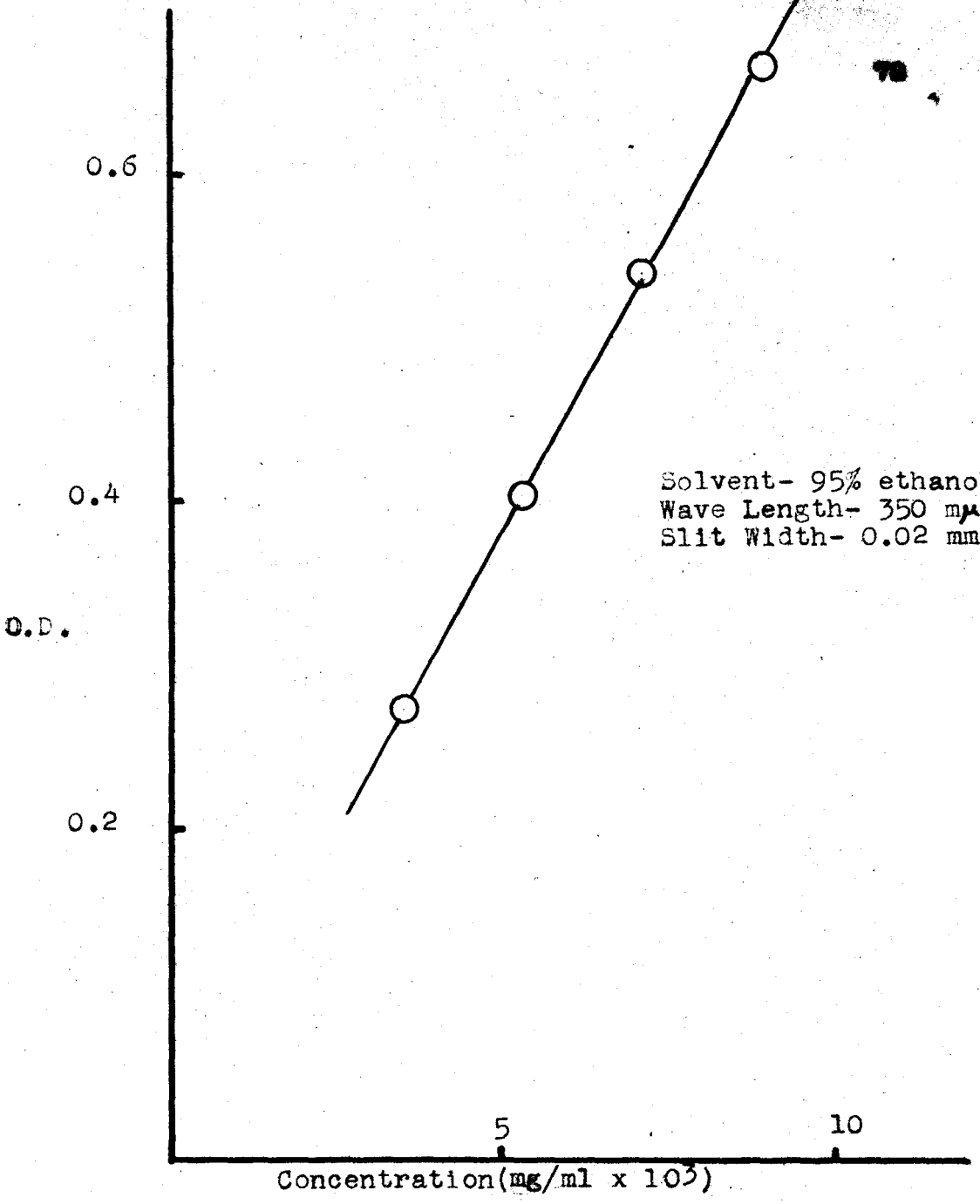


FIGURE 10

STANDARD CONCENTRATION CURVE OF DNP-ALANINOL

Ionography In early attempts to separate the DNP-amino alcohols in the DNP-reduction technique, ionography was employed and the DNP-amino alcohols were measured on the paper strips according to the method described by Marbach (77) and Urbin (120). For this purpose a Bausch and Lomb diffraction grating monochrometer with a mercury lamp source was employed. The strip was moved past the light source (350mμ) and the optical density was recorded automatically on a Brown recorder. The area under the recorded curve was measured with a planimeter and plotted against the amount of material applied. The strips were prepared by adding known volumes (λ quantities) of a solution of known concentration to strips wetted with a phosphate-citrate buffer (M 0.25, Γ/2 = 0.01. ph 7.8). The strips were then subjected to a current at 10 v/cm. The distance of migration usually required to effect good separation of dinitrophenol, dinitroaniline, and the DNP-amino alcohols from the pepsin preparations was about 6 cm. This distance of migration was also used in preparing the standard curves. The standard concentration curve for DNP-alaninol is shown in Figure 11.

Preparation of DNP-amino Alcohols The DNP-derivatives of ethanolamine, alaninol, leucinol, glutamidiol, and aspartidiol were prepared. DNP-ethanol amine and DNP-alaninol were obtained in crystalline form. The other derivatives were obtained as oils. The melting points of DNP-ethanolamine (88-89°C) and DNP-alaninol

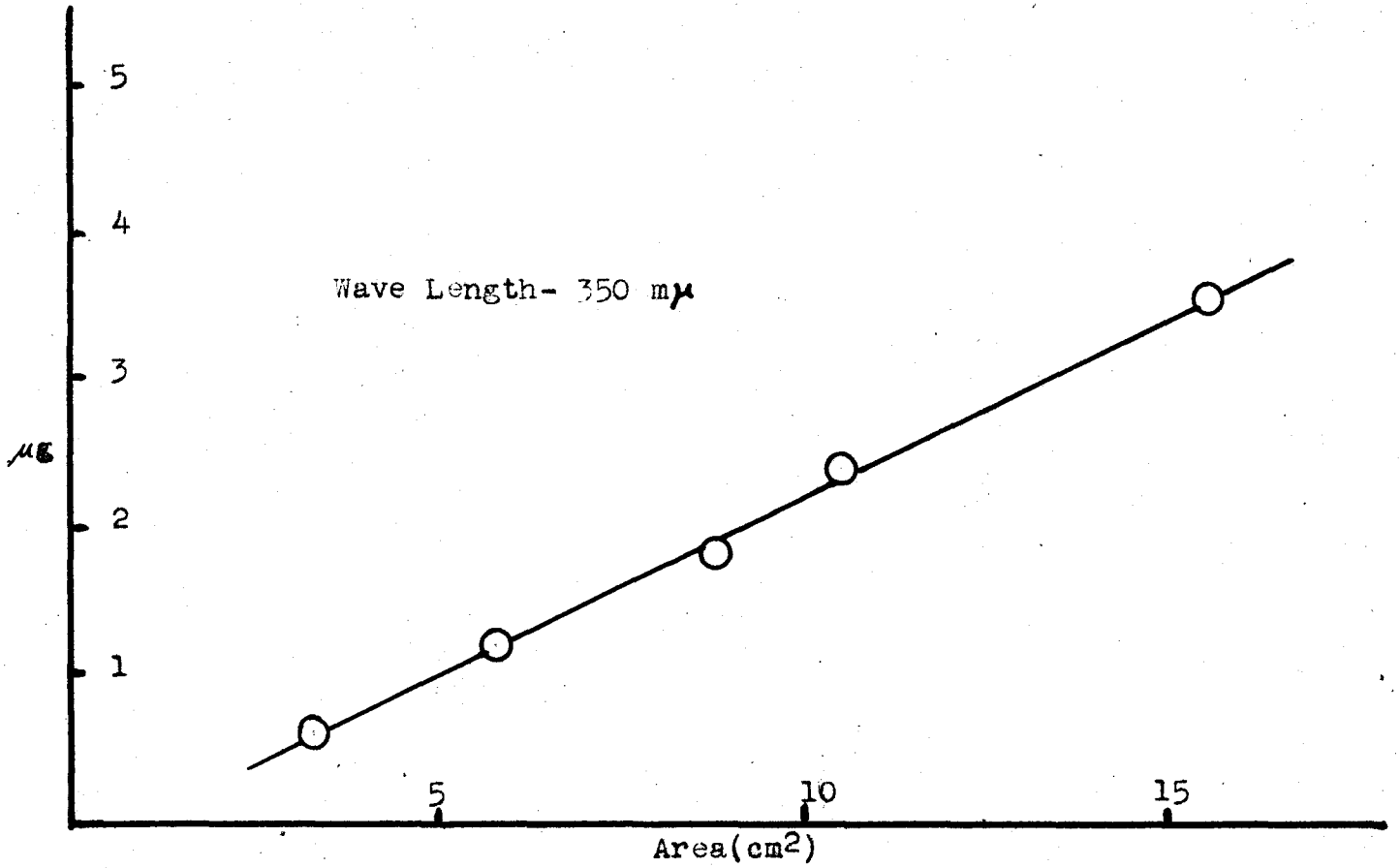


FIGURE 11

STANDARD CURVE OF DNP-ALANINOL ON PAPER

(92-94°C) agree well with those reported by Jutisz et al. (67). The preparations of Jutisz et al. had not yet been published at the time the compounds used in this study were prepared. These are the first reports in the literature of the preparation of these compounds. A typical preparation is that of DNP-ethanolamine (DNP-glycinol).

0.5 gm of ethanolamine was dissolved in 50 ml of 2 per cent $NaHCO_3$. To this solution was added 10 ml of 95 per cent ethanol containing 1.54 gm of 2,4-dinitrofluorobenzene. The reaction mixture was shaken for three hours after which, sufficient glycine was added to react with the excess DNFB. The mixture was shaken for an additional two hours. The mixture was then extracted with ether until the extracting solvent was colorless. After drying the ether extract over anhydrous sodium sulfate, the ether was evaporated and the DNP-ethanol amine was recrystallized from ethanol-isooctane.

Results In the ionographic studies it was not possible to separate DNP-amino alcohols. However, since the reduction and carboxypeptidase experiments indicated only the presence of alanine as an end group separation was not essential. At the pH chosen for the runs on the ionograph, dinitrophenol moved to the positive electrode, dinitroaniline remained essentially at the point of application, and the DNP-amino alcohols moved to the negative electrode. While the separation of these compounds

was good, there was a yellow background along the length of migration which interferred with the measurements. The average value obtained with duplicate preparations was 3.4 end groups.

When Dr. Jutisz made available the details of his inverse phase method of chromatography of DNP-amino alcohols, the same samples used in ionography were put through Super-cel columns. There was some material which remained absorbed at the surface of the column. The remainder of the colored material passed through the column as a single diffuse band. The average value obtained from the eluates was 2.5 end groups. It was later found that dinitrophenol could be removed by putting the mixture through an alumina column. These first columns employed, it was found later, did not separate DNP-alaninol and dinitroaniline and consequently the value of 2.5 end groups was too high. In the final recovery and experimental determinations, the precautions regarding temperature and method of packing the column etc. as described under "Chromatography" were employed. In the recovery experiments, 86 per cent of the added alanine, measured as DNP-alaninol, was recovered. The uncorrected value for the C-terminal residue was 1.5 end groups. When corrected for recovery, the value was 1.7 C-terminal residues.

Mixed chromatrograms were run for identification purposes. Only two bands were present on the columns. The fast fraction from the preparation was collected and mixed in one

case with DNP-alaninol and in another case with dinitroaniline. The mixture with authentic DNP-alaninol gave only a single band on the column whereas the mixture with dinitroaniline gave two bands. These results confirm the two previous identifications of the C-terminal residue as alanine. The slower moving band was mixed with dinitroaniline and chromatographed. Only a single band was obtained.

CHAPTER IX

DETERMINATION OF THE N-TERMINAL SEQUENCE OF PEPSIN

In most of the modifications of Edman's degradation in use at present, the hydantoin after cleavage is extracted from the solution in which the cleavage occurred. The phenylthiohydantoins are soluble in acetone and therefore it was decided to eliminate the extraction step by precipitating the protein with acetone and recovering the hydantoin from the supernatant. In addition, the removal of excess phenylisothiocyanate and N-methylmorpholine from the reaction mixture by benzene extraction usually leads to emulsions. Both of these compounds are soluble in acetone and therefore an acetone precipitation of the PTC-protein was substituted for the extraction. The protein is washed with additional amounts of acetone. Although there is some loss of protein in the precipitation, the method worked sufficiently well to allow the determination of the first four residues in the sequence.

Phenylthiohydantoin Degradation 50 mg of pepsin were dissolved in 1.5 ml of water to which was added 0.25 ml of N NaOH and 1.5

ml of N-methylmorpholine. To this solution was then added 0.25 ml of phenylisothiocyanate, (PITC), and the mixture was stirred vigorously for one and one half hours. At the end of this time the protein was precipitated with 10 volumes of acetone, centrifuged and washed with three 10 ml volumes of acetone. The dry protein was dissolved in 3 ml of 70 per cent acetic acid which is 1.5 M with respect to HCl. The solution was stirred and kept at 40-45°C for thirty minutes to effect cleavage of the phenylthiohydantoin. The protein was then precipitated with 10 volumes of acetone, centrifuged, and washed three times with 10 ml volumes of acetone. The protein was then ready for the second treatment with PITC. The combined supernatant and washings were evaporated to dryness and the residue was dissolved in 0.5 ml of 95 per cent ethanol. This solution was used for chromatography.

As will be shown later in the chapter, the degradation yielded no phenylthiohydantoins beyond the fourth step. In order to determine the nature of the fifth residue, the protein from the fourth step was treated by Sanger's technique as follows: It was suspended in 2 ml of water. 0.15 gm of $NaHCO_3$ and 1 ml of 95 per cent ethanol containing 0.14 gm of DNFB were added. The mixture was stirred at room temperature for three hours. The mixture was then made acid to congo red and 30 ml of acetone were added. The protein was washed three times with 10 ml volumes of acetone. The dry protein was then transferred to a pyrex tube,

1 ml of 12 N HCl was added and the tube was sealed. The sealed tube was heated in a boiling toluene bath (114°) for sixteen hours. The contents were cooled, diluted to 5 ml with water and extracted three times with 5 ml volumes of ether. The ether was dried over anhydrous sodium sulfate, evaporated off, and the residue was taken up in a few drops of 95 per cent ethanol. This solution was used for chromatography of the ether soluble DNP-compounds.

Chromatography The chromatography of the phenylthiohydantoins was the same in all respects as the chromatography of the thiohydantoins.

The solvents used to chromatograph DNP-amino acids were those used in the DNP studies on the N-terminal groups (126). The solvent used for preliminary separations was toluene-pyridine-ethylenechlorohydrin-NH_3 ("toluene" solvent). The solvent was prepared as follows: The organic components were mixed in the volume ratio 5:1:3 and shaken in a separatory funnel. Aqueous 0.8N NH_3 was then poured carefully down the side of the funnel and the two layers allowed to equilibrate for one hour. The aqueous layer was then drawn off and used to saturate the atmosphere of the chromatographic chamber. The organic layer was filtered to remove droplets of water. Fresh solvent must be prepared for each run. A second solvent used in the chromatography of DNP-amino acids was phenol-isoamyl alcohol-water in the ratio 1:1:1 ("Phenol" solvent). The components were shaken in

a separatory funnel and the organic layer was used in the solvent trough.

Preparation of Phenylthiohydantoins Phenylthiohydantoins corresponding to glycine, aspartic acid, glutamic acid, leucine, phenylalaine, lysine, and valine were prepared according to the method of Edman (35). The compounds were obtained in crystalline form except for the glutamic acid hydantoin which was an amorphous white powder. The observed melting points of the phenylthiohydantoins and those from the literature are listed in Table VII.

TABLE VII

MELTING POINTS OF SOME PHENYLTHIOHYDANTOINS

Amino Acid	Observed (uncorr.)	Literature (uncorr.)
Glycine	242°C	245-248°C
Alanine	184-185.5°C	185°C
Valine	212°C	206-208°C
Leucine	171-173°C	178°C
Phenylalanine	189-190°C	187°C
Lysine	165-166°C	162-164°C
Aspartic Acid	226°C	229°C

The general method of preparation was as follows: 0.01 moles of the amino acid were dissolved in 50 ml of 50 per

cent pyridine-water. The pH was adjusted to 9 with N NaOH and the solution was heated in a bath to 40°C. This temperature was maintained throughout the reaction. 2.4 ml of phenylisothiocyanate were added and the mixture was stirred vigorously. The pH was readjusted every five minutes to 9.0 with N NaOH as the reaction proceeded. When the pH remained constant at 9.0 the reaction was complete. This usually required about thirty minutes. The reaction mixture was extracted three times with 50 ml volumes of benzene. An amount of N HCl equivalent to the amount of N NaOH used in the reaction was added to the extracted solution. A mixture of white amorphous solid and a pale yellow syrup, which solidified on standing in the refrigerator overnight, was obtained. This phenylthiocarbamylamino acid was not purified but was suspended in 30 ml of N HCl and refluxed for two hours. On cooling, the hydantoin solidified. The suspension was evaporated three times under reduced pressure to remove HCl. Finally the solution was filtered and the hydantoin recrystallized. Most of the compounds were recrystallized from acetic acid. The hydantoins of lysine, leucine, and aspartic acid were recrystallized from ethanol.

Spectrophotometry The spectra of a few phenylthiohydantoins were determined and are included here (Figures 12, 13, 14), although they were not used in quantitative studies as originally

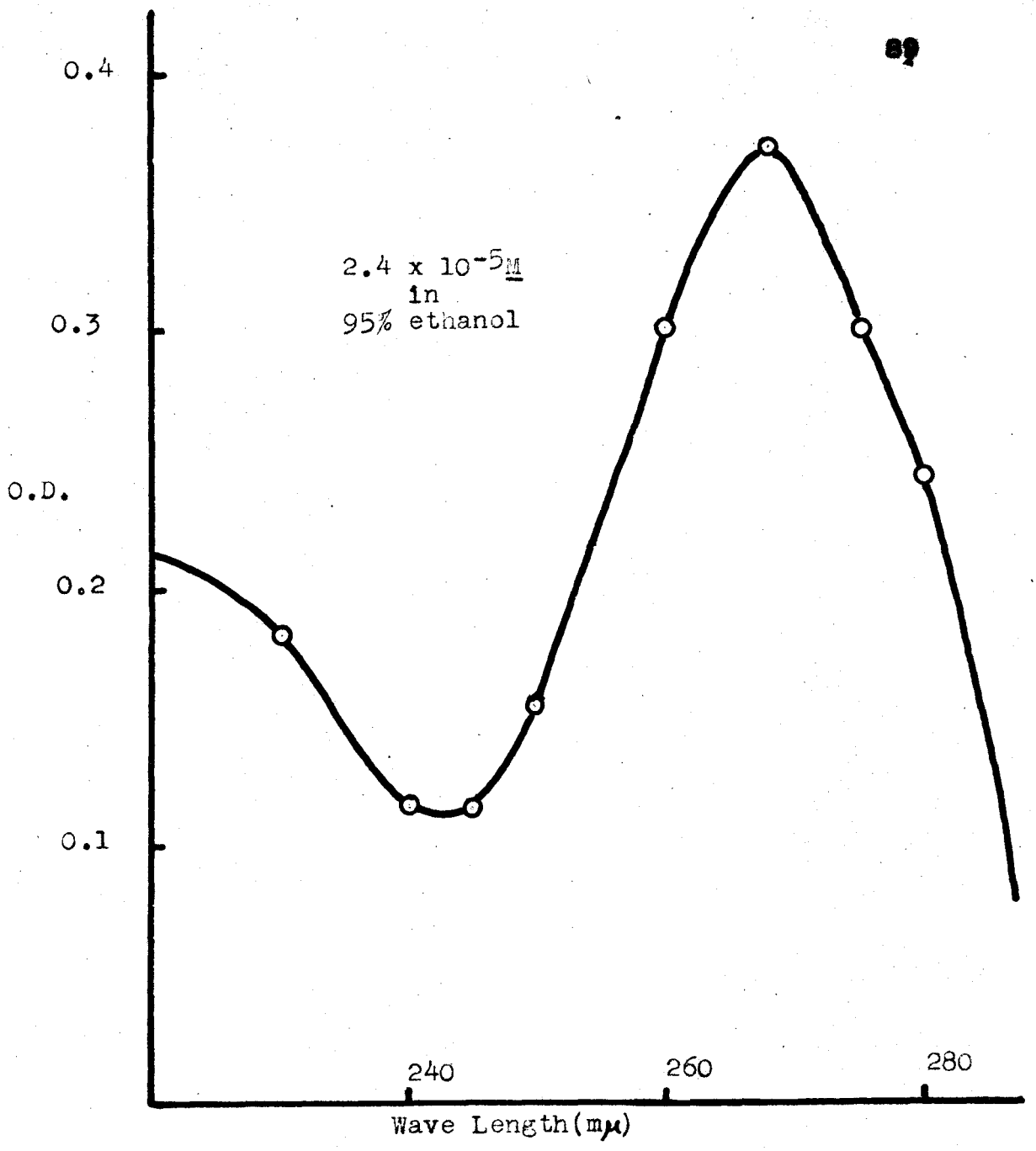


FIGURE 12

ABSORPTION SPECTRUM OF 5-METHYL-3-PHENYL-2-THIOHYDANTOIN

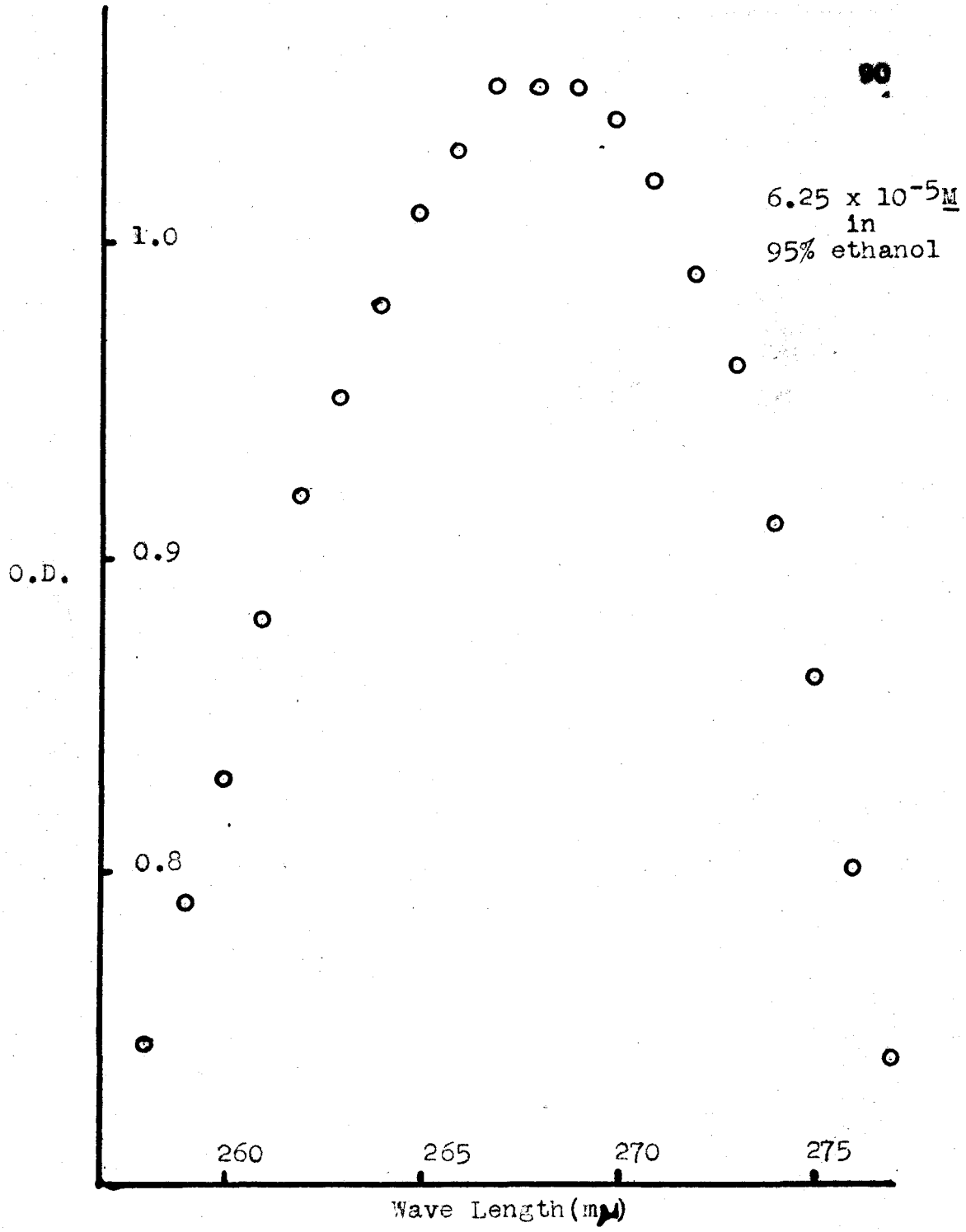


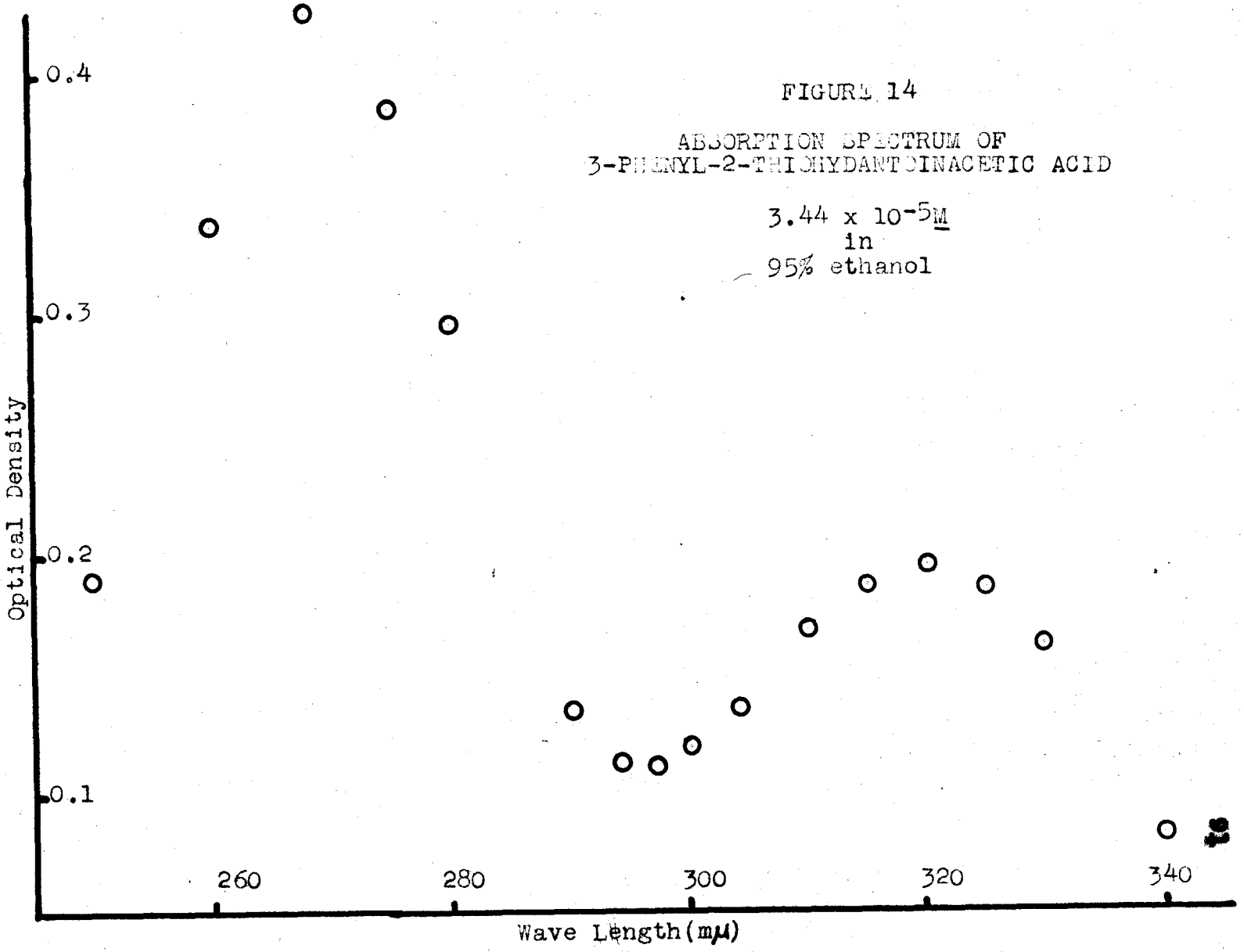
FIGURE 13

ABSORPTION SPECTRUM OF 5-ISOBUTYL-3-PHENYL-2-THIOHYDANTOIN

FIGURE 14

ABSORPTION SPECTRUM OF
3-PHENYL-2-THIOHYDANTOINACETIC ACID

3.44×10^{-5} M
in
95% ethanol



intended.

Results Figures 15, 16, and 17 illustrate the results obtained with this technique. Figure 15 is a chromatogram of the hydantoins obtained at steps 1, 2, 3, and 4 and a mixture of the hydantoins of aspartic acid, glycine and leucine. The material obtained at steps 3 and 4 moved like the phenylthiohydantoin of aspartic acid, however, the sulfur containing amino acids also move slowly. In order to check the identity of the substances obtained at the third and fourth steps, a solvent was employed in which the aspartic acid hydantoin moves much faster than the sulfur amino acids. The results of this chromatogram are shown in Figure 16. The results establish that the compounds in question are hydantoins and not sulfur amino acids. With the solvent employed on the chromatogram shown in Figure 15, the thiohydantoins of aspartic and glutamic acids cannot be separated. To establish which of these compounds was obtained at steps 3 and 4 solvent Sj-A was used. The results are shown in Figure 17. The chromatogram identifies the third and fourth residues as aspartic acid. The identity of the first two residues is evident from the first chromatogram (Figure 15). The identity of the first residue agrees with the identification obtained with the DNP-method (125).

In order to check the reliability of the degradation, it was applied to lysozyme. The sequence of the first five amino acids in lysozyme is known. The sequence is Lys.Val.Phe.Gly.Ser. Figure 18 is a chromatogram of the phenylthiohydantoins isolated

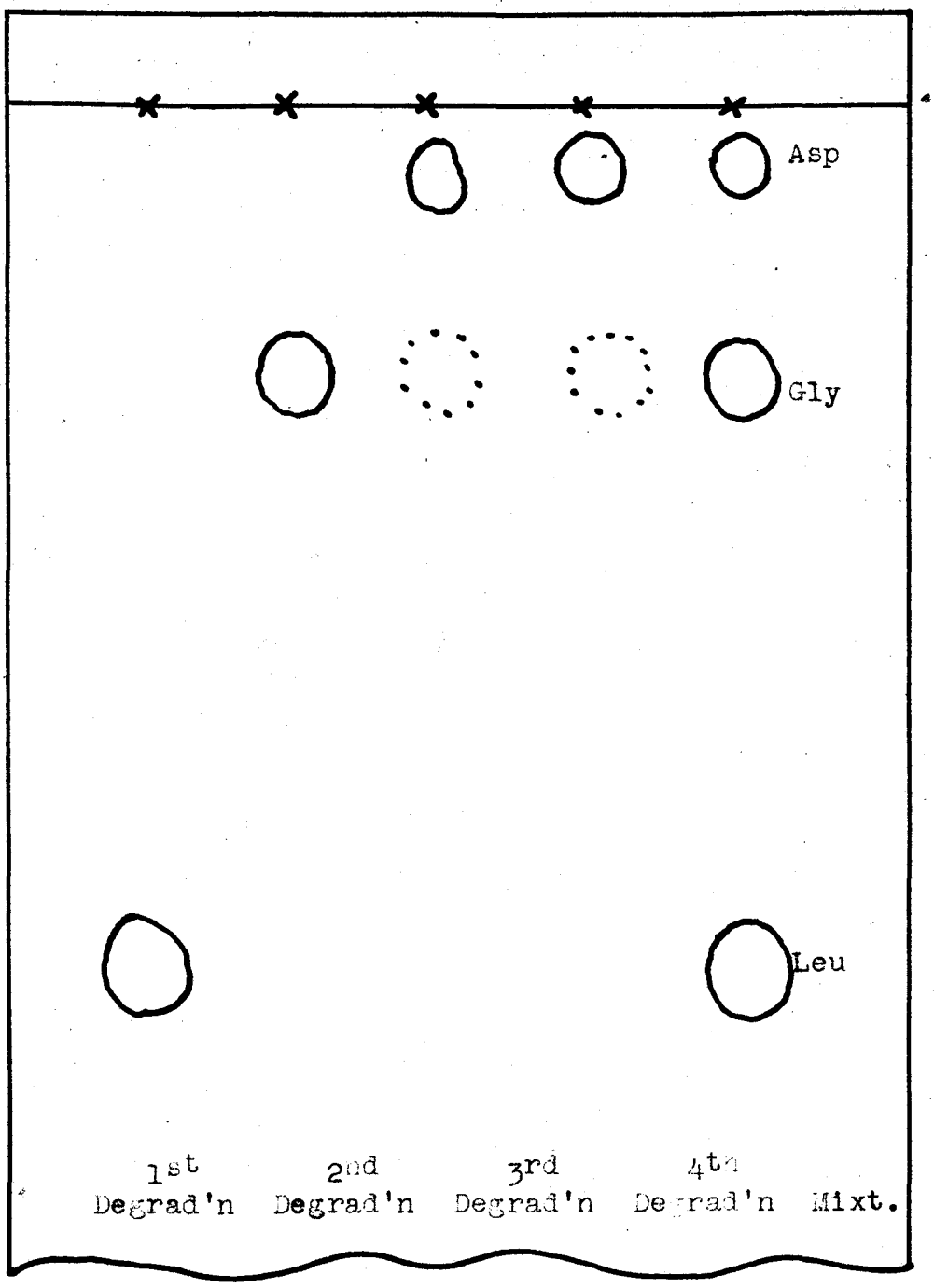


FIGURE 15

THIOHYDANTOINS OF THE N-TERMINAL SEQUENCE OF PEPSIN

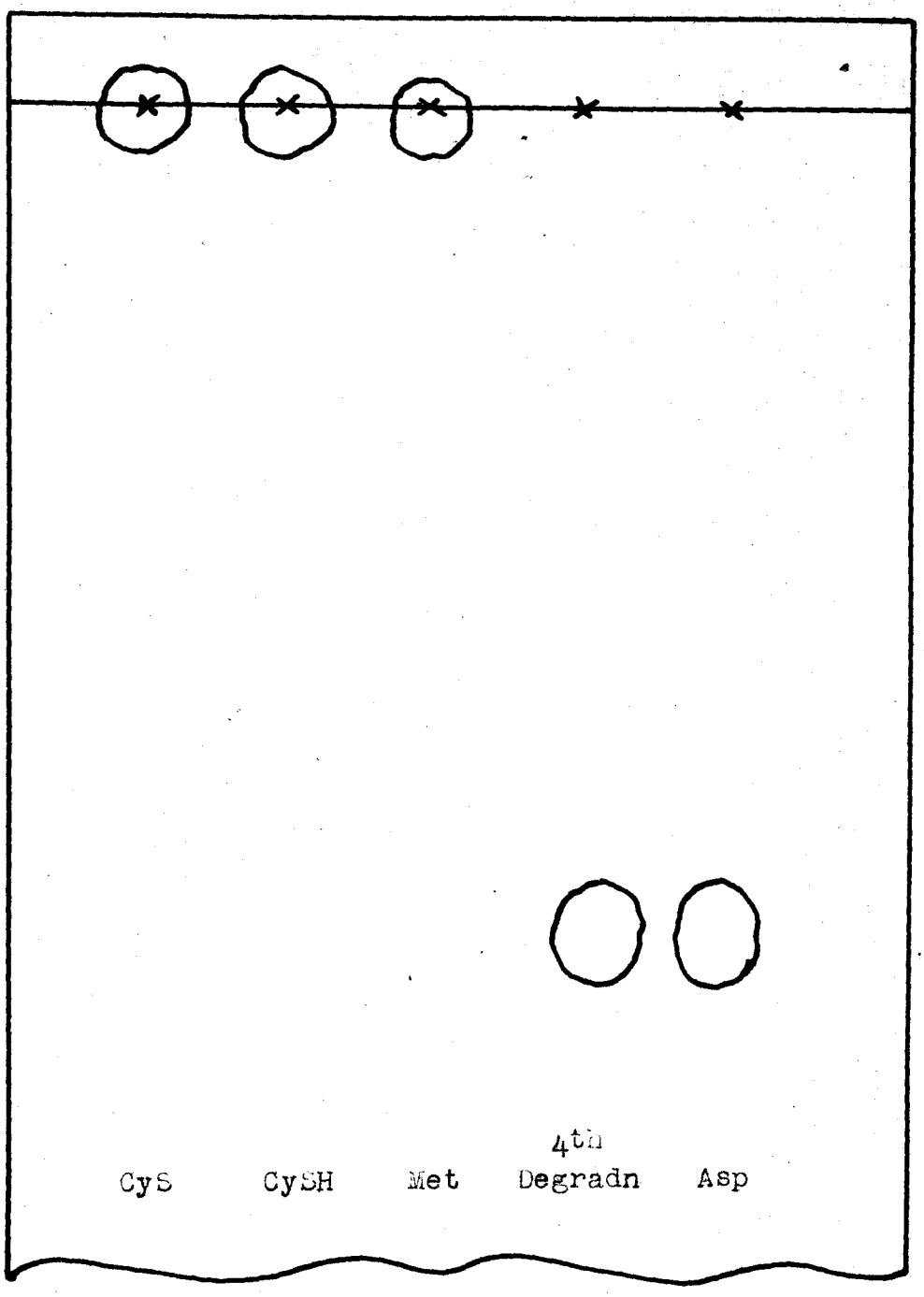


FIGURE 16

CHROMATOGRAPHIC SEPARATION OF THE SULFUR AMINO ACIDS AND 3-PHENYL-2-THIOHYDANTOINACETIC ACID

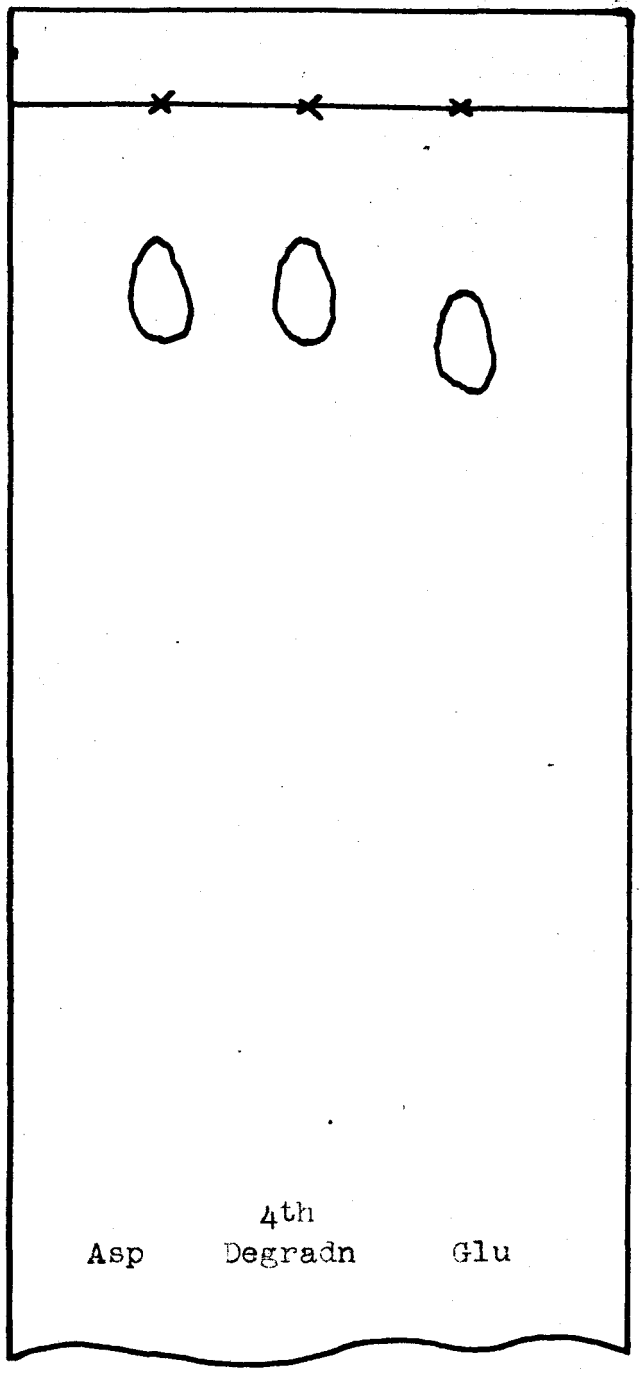


FIGURE 17

CHROMATOGRAPHIC SEPARATION

OF THE THIOHYDANTOINS OF ASPARTIC AND GLUTAMIC ACIDS

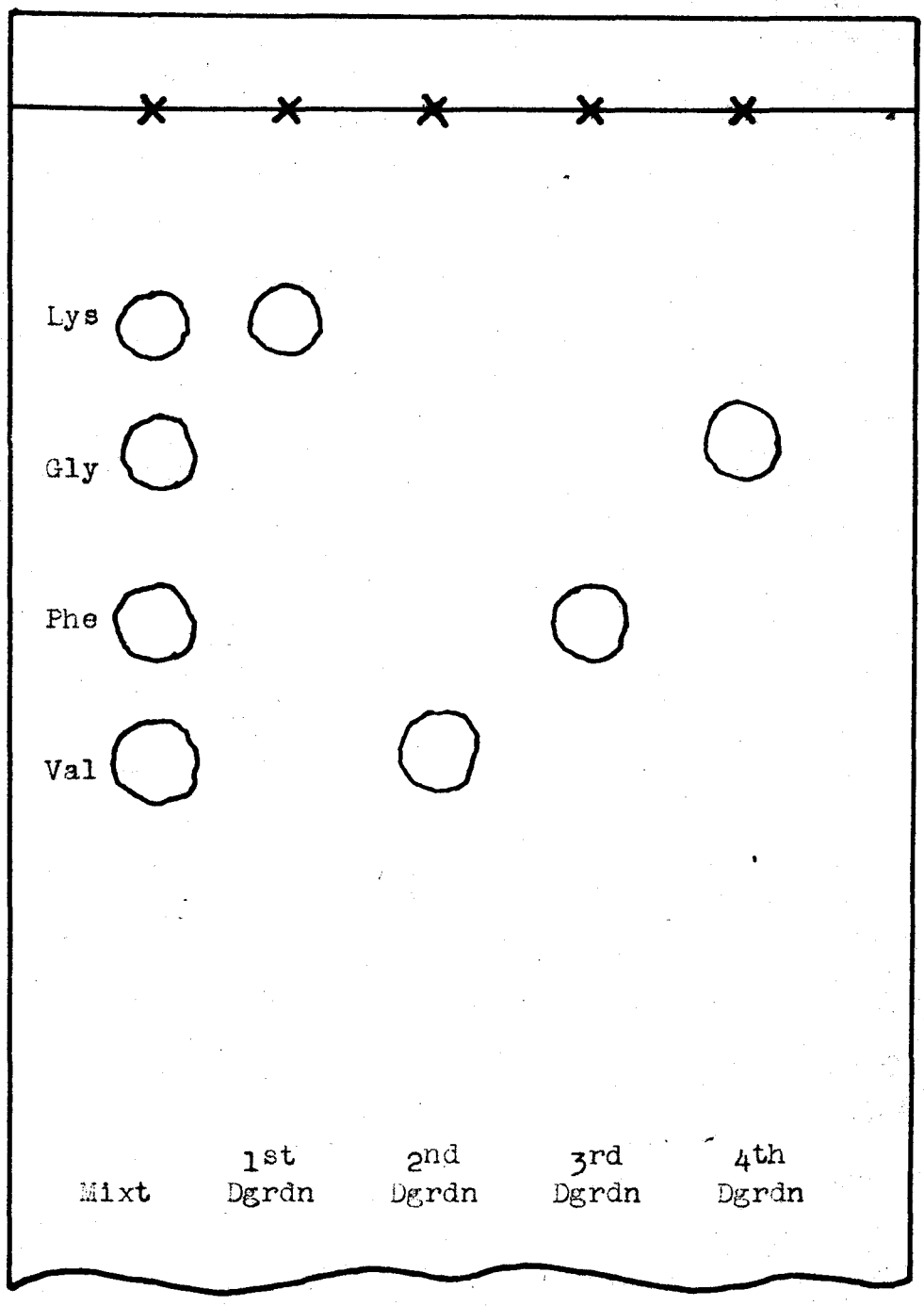


FIGURE 18

THIOHYDANTOINS OF THE N-TERMINAL SEQUENCE OF LYSOZYME

from lysozyme at the first four steps of the degradation. From the position of the authentic hydantoins on the chromatogram it can be seen that the degradation yielded the same results as those reported in the literature.

As mentioned before, the degradation gave negative results at the fifth residue. Attempts to identify the fifth residue, by Sanger's technique have not as yet been successful. Figure 19 shows a chromatogram of the ether soluble DNP material resulting from the application of Sanger's method to the protein after the fourth residue had been removed. The slow moving spot moves like dinitrophenol. DNP-proline and DNP-alanine also have the same R_f in the "toluene" solvent. The slow spot was cut out, eluted and rechromatographed using the "phenol" solvent. It was anticipated that the fifth residue might be proline, however, the results of the chromatogram (Figure 20) demonstrated that the slow spot was dinitrophenol. Identification of the fast spot from the "toluene" chromatogram has not been made. It may be stated that compounds moving faster than dinitrophenol in the "toluene" solvent include phenylalanine, tryptophane, tyrosine and lysine.

The results of the degradation studies therefore demonstrate the N-terminal sequence:

leu.gly.asp.asp_.

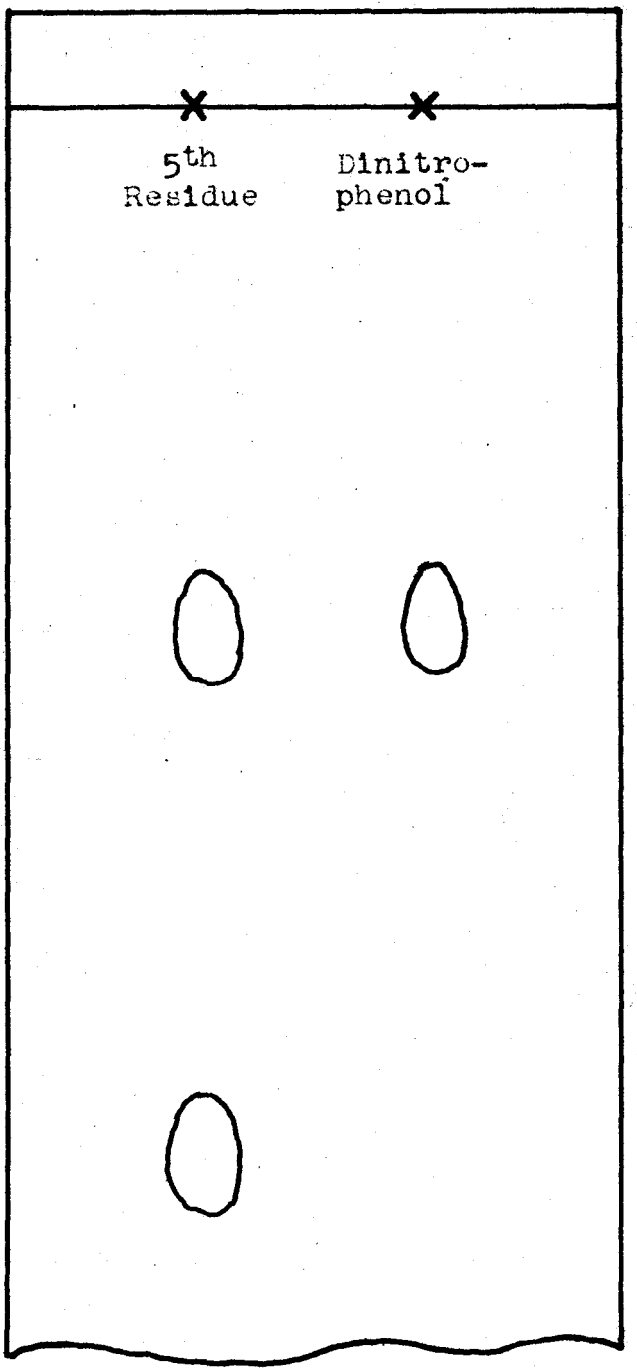


FIGURE 19

ETHER SOLUBLE DNP-COMPOUNDS FROM THE

DINITROPHENYLATION OF THE FIFTH RESIDUE OF PEPSIN

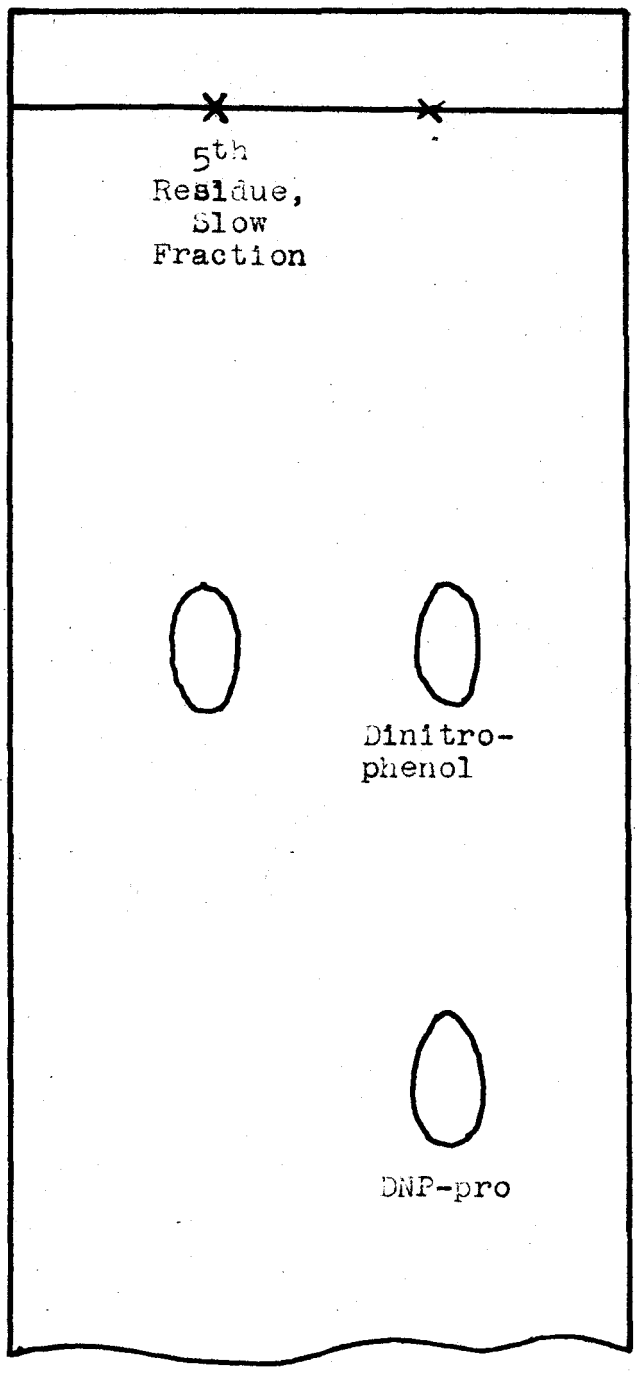


FIGURE 20

"PHENOL" CHROMATOGRAM OF THE SLOW FRACTION

FROM THE "TOLUENE" CHROMATOGRAM

CHAPTER X

DISCUSSION AND SUMMARY

Four methods have been applied to the determination of the terminal free carboxyl groups of pepsin. The first method, reduction with lithium aluminum hydride, established that there was only one species of amino acid at the C-terminus but gave only a tentative identification of this amino acid as alanine.

A subsequent method, carboxypeptidase digestion, corroborated the identification of the C-terminus as alanine and in addition gave an estimate of two end groups per mole of pepsin. Also the enzyme cleaved several other amino acids in small amounts, the one appearing in largest amount after alanine being glutamic acid. The other identified amino acids were leucine (or isoleucine), phenylalanine, and methionine. These amino acids may occupy positions near the C-terminal groups. Glutamic acid may occupy the second position at one or both C-terminal ends.

The attempts to make the thiohydantoin methods of Waley

and Watson and Baptist and Bull quantitative did not meet with success. It may be that this method is applicable only to small peptides.

Finally, the reduction technique of Fromageot was modified by applying the DNP-method to the liberated amino alcohols. The results of this method gave essentially the same quantitative results as the carboxypeptidase experiment and also substantiated the two previous identifications of the C-terminal residue as alanine.

The application of a modification of Edman's degradation to the determination of the N-terminal sequence in pepsin established the sequence

Leu.Gly.Asp.Asp_.

The fifth residue in the sequence appears to be unreactive towards phenyisothiocyanate. An attempt to identify this residue as the DNP-amino acid was unsuccessful although it was established that the residue is not proline or any of the amino acids which move slower than DNP-proline on chromatographing with the "toluene" solvent. Among the amino acids which have not been ruled are phenylalanine, tyrosine and tryptophane.

In the chapter on the characteristics of pepsin, it was noted that Flavin has established that the phosphorus in pepsin is present in the form of O-phosphoserine and that the serine appears in the sequence

Thr.Ser. Glu_.
|
PO_3H_2

A summary of the findings of the present investigations lus the data from the literature can be presented in the ollowing schematic structure of the pepsin molecule:

u.Gly.Asp.Asp.....Thr.Ser.Glu.....(Met,Phe.Leu.Glu).Ala
|
PO_3H_2 ⋮....(Met,Phe.Leu.Glu.).Ala

BIBLIOGRAPHY

1. Abderhalden, E. and Brockmann, H. (1930) "Beitrag Zur Konstionsermittlung Von Proteinen Bzw. Polypeptiden", Biochem. Z., 225, 386-408.

2. Abderhalden, E. and Stix, W. (1923) "Weiterer Beitrag Zur Der Konstitution Der Proteine. Untersuchungen Über Den Aufbau Des Seidenfibroins", Z. Physiol.Chem., 129 143-156.

3. Acher, R., Chauvet, J. and Fromageot, C. (1952) "L'Ordre D'Enchainement De Quelques Acides Amines Dans La Vasopressine De Boeuf", Biochim. Biophys. Acta, 9, 471-472.

4. Acher, R., and Chauvet, J. (1953) "La Structure De La Vasopressine De Boeuf", Biochim, Biophys. Acta. 12, 487-488.

5. Akabori, S., Ohno, K.and Narita K. (1952) "On The Hydrazinolysis of Proteins and Peptides: A Method for the Characterization of Carboxyl-terminal Amino Acids in Proteins", Bull. Chem. Soc. Jap., 25, 214

6. Akabori, S., Sakota, N. and Ohno, K. (1952) "The Mechanism of the Activation of α-Chymotrypsinogen to α-Chymotrypsin" J. Biochem., 39, 29

7. Anfinsen, C., Flavin, M. and Farnsworth, J. (1952) "Preliminary Studies on Ribonuclease Structure", Biochim. Biophys. Acta, 9, 468.

8. Bailey, K. (1951) "End Group Assay in Some Proteins of the Keratinmyosin Group", Biochem. J., 49, 23-27.

9. Baptist, V. and Bull, H. (1953) "Determination of the Terminal Carboxyl Residues of Peptides and Proteins", J. Am. Chem. Soc., 75. 1727.

10. Bergmann, M. (1934) "Synthesis and Degradation of Proteins in the Laboratory and in Metabolism", Science, 79, 439-445.

11. Bergmann, M. and Zervas, L. (1936) "A Method for the Stepwise Degradation of Polypeptides", J. Biol. Chem. 113, 341-357.

12. Bergmann, M. (1935) "Complex Salts of Amino Acids and Peptides. II Determination of l-Proline With The Aid of Rhodantic Acid. The Structure of Gelatin", J. Biol. Chem., 110, 471-479.

13. Bergmann, M. and Fruton, J. (1941) "The Specificity of Proteinases", Advan. Enzym., 1, 63-98.

14. Bettzeiche, F. (1926) "Uber Die Einwirkung Von Grignardreagens Zur Aminosauren. Die Bestimmung Der Freien Carboxylgruppen In Peptiden", Z. Physiol. Chem., 161, 178-190.

15. Boissonnas, R. (1952) "Degradation Des Peptides Neutres A Partir De Leur Extremite Carboxylique", Helv. Chim. Acta, 35, 2226-2228.

16. Brand, E., Saidel, L.J., Goldwater, W. H., Kassell, B., and Ryan, F. J. (1945) "The Empirical Formula of β-Lactoglobulin", J. Am. Chem. Soc., 67, 1524-1532.

17. Bull, H., Hahn, W. and Baptist, V. (1949) "Filter Paper Chromatography", J. Am. Chem. Soc., 71, 550-553.

18. Chargaff, E., Levine, C. and Green, C. (1948) "Techniques for the Demonstration by Chromatography of Nitrogenous Lipide Constituents, Sulfur-containing Amino Acids, and Reducing Sugars", J. Biol. Chem., 175- 67-71.

19. Chibnall, A. and Rees, M. (1951) "The Amide and Free Carboxyl Groups of Insulin", Biochem. J., 48, xlvii.

20. Chibnall, A. and Rees, M. (1952) "Further Observations on the Amide and Free Carboxyl Groups of Insulin", Biochem. J., 52, iii.

21. Christensen, J. (1945) "Free Chemical Groups of Tyrocidin" J. Biol. Chem., 160, 75-82.

22. Christensen, H. (1952) "Attempted Successive Applications of the Edman Degradation to Insulin", Acta Chem. Scand. 6, 1555-1556.

23. Dahlerup-Petersen, B., Linderstrom-Lang, K. and Ottesen, N. (1952) "Stepwise Degradation of Peptides", Acta Chem. Scand., 6, 1135.

24. Davie, E. and Neurath, H. (1952) "C-terminal Groups of Trypsinogen, DFP-trypsin and Carboxypeptidase", J. Am. Chem. Soc., 74, 6305-6306.

25. Davie, E. and Neurath, H. (1953) "Identification of the Peptide Split from Trypsinogen During Autocatalytic Activation", Biochim. Biophys. Acta, 11, 442.

26. Davoll, H., Turner, R., Pierce, J. and duVigneaud, V. (1951) "An Investigation of the Free Amino Groups in Oxytocin and Desulfurized Oxytocin Preparations", J. Biol. Chem. 193, 363-370.

27. Desnuelle, P., Rovery, M. and Bonjour, G. (1950) "Quelques Nouvelle Observations Sur Le Mode D'Attaque De La Globine De Cheval Et De L'Albumine D'Oeuf Par La Pepsine" Biochim. Biophys. Acta, 5, 116-1137.

28. Desnuelle, P., Rovery, M. and Fabre, C. (1951a) "Etude Des Restes N-terminaux Dans Les Serumalbumines De Diverses Speces (Suive D'Une Remarque Sur La Stabilite Des Dinitrophenylaminoacides Pendant L'Hydrolyse)", Compt. Rend., 233, 987-989.

29. Desnuelle, P., Rovery, M. and Fabre, C. (1951b) "Sur Les Derives Dinitrophenyles Du Chymotrypsinogene Et De L'α-Chymotrypsine Cristallisee", Compt. Rend. 233, 1496-1498.

30. Desnuelle, P., Rovery, M. And Fabre, C. (1952) "Extremites N-terminaux De La Proteine De L'α-Chymotrypsine", Biochim. Biophys. Acta, 9, 109.

31. Desnuelle, P., and Fabre, C. (1954) "Etude Des Sequences N-terminales De La Trypsine Et Du Trypsinogene", Bull Ste. Chim. Biol., 36, 181-183.

32. Dieu, H. and Bull, H. (1949) "Monolayers of Pepsin and of Insulin", J. Am. Chem. Soc., 71, 450-452.

33. duVigneaud, V., Lawler, H. and Popenoe, E. (1953) "Enzymatic Cleavage of Glycinamide from Vasopressin and a Proposed Structure for this Pressor-antidiuretic Hormone of the Posterior Pituitary", J. Am. Chem. Soc., 75, 4880.

34. DuVigneaud, V., Ressler, C., Swan, J., Roberts, C.,Katsoyannis, P. and Gordon, S. (1935) "The Syntheses of an Octapeptide with the Hormonal Activity of Oxytocin", J. Am. Chem. Soc., 75, 4879.

35. Edman, P. (1950a) "The Preparation of Phenylthiohydantoins", Acta Chem. Scand., 4, 277-283.

36. Edman, P. (1950b) "Method for Determination of the Amino Acid Sequence in Peptides", Acta Chem. Scand., 4, 283-293.

37. Edman, P. (1953) "Note on the Stepwise Degradation of Peptides Via Phenylthiopydantoins", Acta Chem. Scand., 7, 700.

38. Evans, G. and Reith, W. (1954) "Studies on the Determination of the Sequences of Amino Acids in Peptides and Proteins 3. The Syntheses of 3-(4'Dimethylamino-3'5'-dinitrophenyl) hydantoin Derivatives of Various Amino Acids, and Their Use for the Determination of N-terminal Amino Acids", Biochem. J.,56, 111-116.

39. Fischer, E. (1902) "Ueber Einige Derivate Des Glykocolls, Alanins, Und Leucins", Ber., 35, 1095-1115.

40. Fischer, E. and Abderhalden, E. (1907) "Bildung Von Polypeptiden Bei Der Hydrolyse Der Proteins", Ber., 40, 3544-3562.

41. Flavin, M. (1954) "Threonylphosphoseryl Glutamic Acid Isolated from Pepsin", Fed. Proc. 13, 208.

42. Fletcher, C., Lowther, A. and Reith, W. (1954) "Studies on the Determination of the Sequences of Amino Acids in Peptides and Proteins 2. The Separation of Methyl Esters of N-2, 4-dinitrophenyl Derivatives of Amino Acids by Adsorption Chromatography and the Free Amino Groups of Insulin", Biochem. J., 56, 106-110.

43. Flowers, H. and Reith, W. (1953) "Studies on the Determination of the Sequences of Amino Acids in Peptides and Proteins 1. The Preparation, Properties and Chromatographic Adsorption of the Azobenzene-p-sulphonyl Drivatives of Various Amino Acids, Biochem.J., 53, 657-662.

44. Fraenkel-Conrat, H. and Fraenkel-Conrat, J. (1951) "A Method for Determination of the Amino Acid Sequence of Proteins" Acta Chem. Scand., 5, 1409-1410.

45. Fraenkel-Conrat, J. and Porter, R. (1952) "The Terminal Amino Groups of Conalbumin, Ovomucoid, and Avidin", Biochim. Biophys. Acta, 9, 557-562.

46. Fraenkel-Conrat, H. and Harris, J. (1953) "The Peptide Chain of Tobacco Mosaic Virus", Abstracts, 124th Meeting, Am. Chem. Soc.

47. Fromageot, C., Jutisz, M., Meyer, D. and Penasse, L. (1950), "Methods Pour La Caracterisation Des Groupes Carboxyliques Terminaux Dans Les Proteines. Application A L'Insulin", Biochim, Biophys. Acta, 6, 283-289.

48. Fromageot, C. and Jolles, P. (1952) "Caracterisation Du Residueβ-Aspartique (-NH.CH(COOH)CH_2.CO-) Dans L'Insulin" Biochim. Biophys. Acta, 9, 416-418

49. Gladner, J. and Neurath, H. (1952) "C-terminal Groups in Chymotrypsinogen and DFP-α-chymotrypsin in Relation to the Activation Process", Biochim. Biophys. Acta, 9, 335-336.

50. Gladner, J. and Neurath, H. (1953) "Carboxyl Terminal Groups of Proteolytic Enzymes I. The Activation of Chymotrypsinogen toα-Chymotrypsin", J. Biol. Chem., 205, 345-360.

51. Gurin, S. and Clarke, H. (1934) "Allocation of the Free Amino Groups in Proteins and Peptides", J. Biol. Chem., 107, 395-419.

52. Harrington, C. and Pitt Rivers, R. (1944) "The Syntheses of Cysteine-(Cystine)-tyrosine Peptides and the Action Thereon of Crystalline Pepsin", Biochem. J. 38, 417-422.

53. Harris, J. (1952) "The Use of Carboxypeptidase for the Identification of Terminal Carboxyl Groups in Polypeptides and Proteins. Asparagine as a C-terminal Residue in Insulin", J. Am. Chem. Soc., 74, 2944-2945.

54. Harris, J. and Knight, C. (1952) "Action of Carboxypeptidase on Tobacco Mosaic Virus", Nature, 170, 613.

55. Havinga, E. (1953) "Comparison of the Phosphorus Content, Optical Rotation, Separation of Hemes and Globin, and Terminal Amino Acid Residues of Norman Adult Human Hemoglobin and Sickle Cell Anemia Hemoglobin", Proc. Nat. Acad. Sci., 39, 59-64.

56. Herriot, R. (1935) "Acetylation of Tyrosine in Pepsin", J. Gen. Physiol., 19, 283-299.

57. Herriot, R. (1937) "Inactivation of Pepsin by Iodine and the Isolation of Diiodotyrosine from Iodinated Pepsin", J. Gen. Physiol., 20, 335-352.

58. Herriot, R. (1954) The Mechansim of Enzyme Action, edited by McElroy, W.D. and Glass B., The Johns Hopkins Press, Baltimore, p. 24.

59. Herriot, R. and Northrop, J. (1934) "Crystalline Acetyl Derivatives of Pepsin", J. Gen. Physiol.,18, 35-37.

60. Hofmeister, F. (1902) "Uber Bau Und Gruppierung Der Eiweiss-Korper", Ergeb. Physiol., 1, 758-802.

61. Holley, R. and Holley, A. (1952) "A New Stepwise Degradation of Peptides", J. Am. Chem. Soc., 74,5445-5448.

62. Ingram, V. (1950) "Micro End Group Analysis for Peptides", Nature 166, 1038.

63. Ingram, V. (1951) "Autolysis Products of Pepsin", Nature 167, 83

64. Ingram, V. (1952) "The Identification of Peptide End Groups as Dimethylamino Acids", J. Biol. Chem., 193-201.

65. Jensen, H. and Evans, E. (1935) "Studies on Crystalline Insulin XVIII. The Nature of the Free Amino Groups in Insulin and the Isolation of Phenylalanine and Proline from Crysalline Insulin", J. Biol. Chem., 108, 1-9.

66. Jutisz, M. (1954) "Identification Des Residues C-terminaux Des Peptides Et Des Proteines", Bull, Ste. Chim. Biol., 36, 109-116.

67. Jutisz, M., Privat de Garilhe, M. and Suquet, M. (1954) "Micromethod De Separation Des N-dinitrophenylaminoalcools Par Chromatographie "Phase Inverse", Bull. Ste. Chim. Biol., 36, 117-126.

68. Karrer, P., Portmann, P., Suter, M. (1948) "Die Reduktion Von α-Aminocarbonsauren Zu Aminoalkoholen Mittels Lithiumaluminumhydrid", Helv. Chim. Acta, 31, 1617-1623.

69. Landmann, W., Drake, M. and White, W. (1953) "Studies on Pituitary Adrenocorticotropin. VI. An N-terminal Sequence of Corticotropin A", J. Am. Chem. Soc., 75, 4370.

70. Landmann, W., Drake, M. and Dillaha, J. (1953) "Paper Chromatography of the 3-Phenyl-2-thiohydantoin Derivatives of Amino Acids with Application to End Group and Sequence Studies", J. Am. Chem. Soc., 75, 3638.

71. Lens, J. (1949) "The Terminal Carboxyl Groups of Insulin", Biochim. Biophys. Acta, 3, 367-379.

72. Leonis, J. and Levy, A. (1951) "Und Methode Permettant La Degradation Controlee Des Peptides, Au Moyen Du Sulfure De Carbone", Bull. Ste. Chim. Biol., 33, 779-783.

73. Levy, A. (1950) "The Removal of Terminal Groups from Peptides" J. Chem. Soc., 404-405.

74. Li, C. H. and Ash, L. (1953) "The Nitrogen Terminal End Groups of Hypophyseal Growth Hormone", J. Biol. Chem., 203, 419-424.

75. Lichtenstein, N. (1938) "The Behavior of Peptides When Heated In β-Naphthol", J. Am. Chem. Soc., 60, 560-563.

76. Livermore, A. and duVigneaud, V. (1949) "Preparation of High Potency Oxytocic Material by the Use of Countercurrent Distribution, J. Biol. Chem., 180, 365-373.

77. Marbach, E. (1954) Dissertation, Loyola University, Chicago.

78. Mueller, J., Pierce, J., Davoll, H. and duVigneaud, V. (1951) "The Oxidation of Oxytocin with Performic Acid", J. Biol. Chem., 191, 309-313.

79. Northrop, J. (1925) "A Test for Diffusible Ions. II. The Ionic Nature of Pepsin", J. Gen. Physiol., 7, 603-614.

80. Northrop, J. (1930) "Crystalline Pepsin. I. Isolation and Tests of Purity. II. General Properties and Experimental Methods", J. Gen. Physiol., 13, 739-780.

81. Northrop, J., Kunitz, M. and Herriot, R. (1948) Crystalline Enzymes, Columbia University Press, New York.

82. Ohno, K. (1953) "On the Structure of Lysozyme 1. Quantitative Estimation of Carboxyl-terminal Amino Acid by Improved Hydrazinolysis Method", J. Biochem., 40, 621-627

83. Nystrom, R. and Brown, W. (1947) "Reduction of Organic Compounds by Lithium Aluminum Hydride. II Carboxylic Acids" J. Am. Chem. Soc., 69, 2548-2549.

84. Ottesen, M. and Wollenberger, A. (1952) "Stepwise Degradation of the Peptides Liberated in the Transformation of Ovalbumin to Plakalbumin", Nature, 170, 801.

85. Pauling, L., Corey, R. and Branson, H. (1951) "The Structure of Proteins: Two Hydrogen-bonded Helical Configurations of the Polypeptide Chain", Proc. Nat. Acad. Sci., 37, 205-211.

86. Philipot, J. and Eriksson-Quensel, I. (1933) "An Ultracentrifugal Study of Pepsin and Insulin", Nature, 132 932.

87. Pierce, J. and duVigneaud, V. (1950a) "Preliminary Studies on the Amino Acid Content of a High Potency Preparation of the Oxytocic Hormone of the Posterior Lobe of the Pituitary Gland", J. Biol. Chem., 182, 359-366.

88. Pierce, J. and duVigneaud, V. (1950b) "Studies on High Potency Oxytocic Material from Beef Posterior Pituitary Lobes", J. Biol. Chem. 186, 77-84.

89. Pierce, J., Gordon, S. and duVigneaud, V. (1952) "Further Distribution Studies on the Oxytocic Hormone of the Posterior Lobe of the Pituitary Gland and the Preparation of an Active Crystalline Flavinate", J. Biol. Chem. 199, 929-940.

90. Popenoe, E. and duVigneaud, V. (1953) "Degradative Studies on Vasopressin and Performic Acid-oxidized Vasopressin", J. Biol. Chem., 205, 133-43.

91. Popenoe, E. and duVigneaud, V. (1954) "A Partial Sequence of Amino Acids in Performic Acid-oxidized Vasopressin", J. Biol. Chem., 206, 353-360.

92. Porter, R. and Sanger, F. (1948) "The Free Amino Groups of Hemoglobin", Biochem. J., 42, 287-294.

93. Privat De Garilhe, M., Maier-Huser, H. and Formageot, C. (1951) "Absence De Groupements Amines Libres Dans L'Ocytocine", Biochim. Biophys. Acta, 7, 471-476.

94. Reid, E. (1951) "Relative Importance of Free α-and ε-Amino Groups for the Biological Activity of the Growth Hormone", Nature, 168-955.

95. Reith, W. and Waldron, N. (1953) "On the Determination of the Sequence of Amino-acids in Peptides", Biochem. J., 53 XXXV.

96. Reith, W. and Waldron, N. (1954) "Studies on the Determination of the Sequence of Amino Acids in Peptides and Proteins 4. The Synthesis of 3-(4'-dimethylamino-3',5' dinitrophenyl)-2-thiohydantoin Derivatives of Various Amino Acids and Their Use for Amino Acid Sequence Determinations", Biochem. J., 56, 547-549.

97. Roche, J. and Schiller, J. (1944) "Sur La Presence Dans Diverses Proteines De Petit Quantities De Groupments Amines Appartenant A La Phenylalanine", Compt. Rend., 219, 38-49.

98. Rovery, M. and Desnuelle, P. (1954) "Application Comparee De La Technique De Sanger Et De La Technique D'Edman-Fraenkel-Conrat A La Determination Des Residue N-terminaux Des Proteines", Bull. Ste. Chim. Biol., 36, 95-103.

99. Rovery, M., Fabre, C. and Desnuelles, P. (1952) "Etude Des Extremities N-terminales Du Trypsinogene Et De La Trypsine De Boeuf", Biochim. Biophys. Acta, 9, 702.

100. Rovery, M., Fabre, C. and Desnuelle, P. (1953) "Etude De L'Activation Du Chymotrypsinogene Et Du Trypsinogene De Boeuf Par Determination Des Residus N-terminaux dans Les Proteines Et Les Enzymes Correspondants", Biochim. Biophys. Acta, 12, 547-559.

101. Sanger, F. (1945) "The Terminal Amino Acid Residue of Insulin", Biochem. J., 39, 507-515.

102. Sanger, F., (1946) "The Free Amino Group of Gramicidin S", Biochem. J., 40, 261-262.

103. Sanger, F. (1947) "Oxidation of Insulin by Performic Acid", Nature, 160, 295-296.

104. Sanger, F. (1948) "Some Peptides from Insulin", Nature, 162, 491-492.

105. Sanger, F. (1949) "The Terminal Peptides of Insulin", Biochem. J., 45, 563-574.

106. Sanger, F. and Tuppy, H. (1951) "The Amino Acid Sequence in the Phenylalanyl Chain of Insulin", Biochem. J., 49, 463-490.

107. Sanger, F. and Thompson, E. (1952) "The Amino Sequence In the Glycyl Chain of Insulin", Biochem. J.,

108. Schlack, P. and Kumpf, W. (1926) "Uber Eine Neue Methode Zur Ermittlung Der Konstitution Von Peptiden", Z. Physiol. Chem.,154,125-170.

109. Schroeder, W. (1952) "Sequence of Four Amino Acids at the Amino End of the Single Polypeptide Chain of Lysozyme", J. Am. Chem. Soc., 74, 281-282.

110. Sjoquist, J. (1953) "Paper Strip Identification of Phenyl Thiohydantoins", Acta Chem. Scand., 7, 447-448.

111. Smith, E. (1953) "Specificity of Certain Peptidases and Their Use in the Study of Peptide and Protein Structure", The Chemical Structure of Proteins, Little, Brown and Company Boston.

112. Staub, A., Sinn, L. and Behrens, O. (1954) "Characterization of Glucagon (Hyperglycemic-glycogenolytic Factor)", Fed. Proc., 13, 303-304.

113. Thompson, A. (1952) "The C-terminal Residue of Lysozyme", Nature, 169, 495-496.

114. Thompson, E.O.P. (1954) "Crystalline Papain IV Free Amino Groups and N-terminal Sequence", J. Biol. Chem., 207, 563-574.

115. Tibbs, J. (1951) "Identification of the Free Carboxyl Groups in Peptides", Nature, 168, 910.

116. Tiselius, A., Henschen, G. and Svensson, H. (1948) "Electrophoresis of Pepsin", Biochem. J., 32, 1814-1818.

117. Turner, R., Pierce, J. and duVigneaud, V. (1951) "The Desulfurization of Oxytocin", J. Biol. Chem., 193, 359-361.

118. Turner, R. and Schmerzler, G. (1953) "C-terminal Residues of Ovomucoid and Ovalbumin", Biochim. Biophys. Acta, 11, 586.

119. Udenfriend, S. and Velick, S. (1951) "The Isotope Derivative Method of Protein Amino End Group Analysis", J. Biol. Chem., 190, 733-740.

120. Urbin, M. (1954) Dissertation, Loyola University, Chicago

121. Velick, S. and Udenfriend, S. (1951) "The Amino End Group and the Amino Acid Analysis of Salmine", J. Biol. Chem., 191, 233-238.

122. Velick, S. and Udenfriend, S. (1953) "The Composition and Amino Acid End-groups of Glyceraldehyde-3-phosphate Dehydrogenase", J. Biol. Chem., 203, 575-581.

123. Waley, S. and Watson, J. (1951) "The Stepwise Degradation of Peptides", J. Chem. Soc., 2394-2397.

124. White, W. (1953) "Studies on Pituitary Adrenocorticotropin. VII, A C-terminal Sequence of Corticotropin A", J. Am. Chem. Soc., 75, 4877.

125. Williamson, M. and Passmann, J. (1952) "The Terminal Amino Group of Pepsin", J. Biol. Chem., 199, 121-216.

126. Williamson, M. Unpublished data.

127. Wolstenholme, G. and Cameron, M. (1953) The Chemical Structure of Proteins, Little, Brown and Company, Boston.

ADDENDUM

After the completion of this thesis, a report of studies on the structure of pepsin by Herriot(58) appeared in print. The results of these studies are summarized and compared with the results of this thesis in the following table.

	This thesis	Herriot
N-terminus	1 Leu	1 Ileu
N-terminal sequence	Leu.Gly.Asp.Asp-	Ileu.Gly-
C-terminus	2 Ala	Ala
C-terminal sequence	-(Phe,Asp,Val,Leu, Glu)Ala	-Val.Leu(Ileu). Ala

Herriot "favors" isoleucine as the single N-terminal group. The differentiation between leucine and isoleucine was made using t-amyl alcohol as the solvent. A paper chromatogram using this solvent requires 2-3 weeks development to obtain a separation. The choice of leucine made by Williamson and Passmann(125) was based on a two dimensional paper chromatogram using collidine-water and phenol-NH_4OH as the solvents.

Herriot identified the C-terminal group by means of carboxypeptidase but gives no indication as to the number of alanine residues detected. The text can be interpreted as proposing a single C-terminal group.