INFORMATION TO USERS

This dissertation was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again – beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

University Microfilms

300 North Zeeb Road
Ann Arbor, Michigan 48106

A Xerox Education Company

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/3382

73-4938

CHEN, Kirk C. S., 1941-
STRUCTURE AND FUNCTION OF PORCINE PEPSIN.

The University of Oklahoma, Ph.D., 1972
Biochemistry

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

STRUCTURE AND FUNCTION OF PORCINE PEPSIN

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requriements for the

degree of

DOCTOR OF PHILOSOPHY

BY

Kirk C. S. Chen

Oklahoma City, Oklahoma

1972

STRUCTURE AND FUNCTION OF PORCINE PEPSIN

APPROVED BY

Jordan J. N. Lang

A. Chadwick Cox

Robert Delaney

Jan A. Hartsuck

John R. Tobitt

DISSERTATION COMMITTEE

PLEASE NOTE:

Some pages may have

indistinct print.

Filmed as received.

University Microfilms, A Xerox Education Company

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to Dr. Jordan J. N. Tang, Head of Laboratory of Protein Studies, Oklahoma Medical Research Foundation, for his advice, guidance, and friendship.

I am indebted to Dr. Jean A. Hartsuck, of Laboratory of Protein Studies, Oklahoma Medical Research Foundation, for discussions and suggestions.

I wish to thank the members of my dissertation committee, Dr. Jordan J. N. Tang, Dr. Chadwick A. Cox, Dr. Robert Delaney, Dr. Jean A. Hartsuck, and Dr. John R. Sokatch for their advice in the preparation of this dissertation.

I am also indebted to Mrs. Mary Rebecca Enochs for her friendly assistance throughout this work.

I want to express my gratitude to Mrs. Donnie Freeman and Mrs. Ruth A. Reed who performed the tedious work of preparing glassware for me.

I also wish to express my gratitude to the Oklahoma Medical Research Foundation for the provision of facilities, equipment, and material and for financial support while a predoctoral fellow.

TABLE OF CONTENTS

	Page
LIST OF TABLES	v
LIST OF ILLUSTRATIONS	vi
Chapter	
I. INTRODUCTION	1
II. MATERIALS AND METHODS	9
III. RESULTS	18
IV. DISCUSSION	71
V. SUMMARY	81
BIBLIOGRAPHY	83

LIST OF TABLES

Table		Page
1.	Amino Acid Compositions of Porcine Pepsinogen and Porcine Pepsin..	4
2.	Summary of Electrophoretic Mobility, Amino Acid Composition and Sequence Determination of Peptides.....	28
3.	Recovery of EPNP-Aspartyl Peptides from 10°-Reacted Pepsin in Column Chromatography (Sephadex G-25).....	41
4.	Recoveries of Cyanogen Fragments of 800 mg of Aminoethylated Porcine Pepsin from Sephadex G-75 Column.....	49
5.	Amino Acid Analysis of Purified CB5.....	52
6.	Summary of Electrophoretic Mobility, Amino Acid Composition and Sequence Determination of Peptides from CB5.....	59

LIST OF ILLUSTRATIONS

Figure — Page

1. Chromatographic patterns of peptides from peptic digest EPNP-pepsin inactivated at 30°(A) and at 10°(B) on a column of Sephadex G-25 19

2. Two-dimensional high-voltage electrophoretic pattern of peptides recovered under peak P_1 in Fig. 1 (left), and peak P_2 in Fig. 1 (right) 21

3. Chromatographic patterns of peptides from subtilisin digest of EPNP-pepsin inactivated at 30°(A) and at 10°(B) on a column of Sephadex G-25 23

4. Chromatographic patterns of peptides from papain digest of EPNP-pepsin inactivated at 30°(A) and at 10°(B) on a column of Sephadex G-25 24

5. Chromatographic patterns of peptides from ficin digest of 30°-reacted EPNP-pepsin 25

6. Chromatographic patterns of peptides from thermolysin digest of EPNP-pepsin inactivated at 30°(A) and at 10°(B) on a column of Sephadex G-25 26

7. Chromatographic patterns of peptides from α-chymotryptic digest of 30°-reacted EPNP-pepsin 27

8. Sequence of EPNP-peptides from pepsin 30

9. Summary of the possible chemical reactions of EPNP-alkylated methionyl peptide (P_1) 32

10. The absorbancy at 315 nm of peptide P_1 during the Edman sequential degradation 34

11. Mobility of EPNP-peptides P_1 and P_2 in high-voltage electrophoresis at pH 6.0 before and after treatment with alkaline solution 43

12. Chromatogram of cyanogen bromide fragments of 800 mg of aminoethylated porcine pepsin on Sephadex G-75 48

LIST OF ILLUSTRATIONS--Continued

Figure		Page
13.	Chromatographic patterns of crude CB5 on a column of DEAE-Sephadex A-25	51
14.	Chromatographic patterns of peptides from tryptic digest of CB5 on a column of Sephadex G-25	54
15.	Chromatographic patterns of peptides from thermolysin digest of CB5 on a column of Sephadex G-25	55
16.	Chromatographic patterns of peptides from α-chymotryptic digest of CB5 on a column of Sephadex G-25	56
17.	Chromatographic patterns of peptides from papain digest of CB5 on a column of Sephadex G-25	57
18.	Electrophoretic patterns of tryptic digested peptides at pH 2.0	58
19.	Amino Acid Sequence of CB5	70
20.	Carboxyl Terminal Sequence of Porcine Pepsin (Residue 241 - 321)	80

STRUCTURE AND FUNCTION OF PORCINE PEPSIN

CHAPTER I

INTRODUCTION

Many important discoveries have been made in the study of the structure and function of enzymes in the last two decades. As a result, we have reached a point of understanding that, for at least a few enzymes, the detailed mechanism of catalysis can be attributed to the functional groups in the three-dimensional model of the enzymes. There is a group of pancreatic proteolytic enzymes such as trypsin, chymotrypsin, elastase, carboxypeptidase A which has been well studied. These investigations have contributed significantly to our understanding of the structure-function relationship in enzymes. The experiences resulting from such studies have shown that information from several major areas of investigation are necessary. First, the primary structure provides a rational chemical frame work of the enzyme molecule. Second, a three dimensional model resulting from crystallographic studies gives the topological relationships of the functional groups. Third, chemical modifications are instrumental in locating the active sites. And finally, kinetic and other studies suggest a possible mechanism which can

pinpoint the roles of the functional groups. Each of these studies contribute to the detailed mechanisms of enzymic catalysis.

However a group of acidic proteases from the stomach has not been well studied. There are several reasons for such deficiency: First, pure acidic proteases are not readily available. The problem of purity of porcine pepsin was a serious problem and was not solved until quite recently (1), gastricsin was isolated only about ten years ago (2), and the best method of preparation of rennin yields a product which still contains two components (3). Second, porcine pepsin has a molecular weight of 34,000, consists of 321 amino acid residues and is larger than the pancreatic proteolytic enzymes. It also contains a large number of acidic residues which usually make sequence study difficult. Third, since acidic proteases are labile in neutral and alkaline solution, there are considerable restrictions on application of available methods for their structure-function studies.

In spite of these facts, acidic proteases present a challenging problem. We have developed an interest specifically for this problem as it relates to the structure and function of porcine pepsin for several reasons. First, most proteolytic enzymes whose structure and function have been examined have neutral or slightly alkaline pH. Included among this group of enzymes are chymotrypsin, trypsin, elastase, subtilisin, carboxypeptidase A and papain. Pepsin on the other hand, has an optimum at pH 2 which suggests a very different mechanism in its catalytic action from that of alkaline proteases. It would be interesting to understand the relationship between structure and

function in an acidic protease. Second, just as the pancreatic proteases, chymotrypsin, trypsin and elastase, are structurally homologous, two gastic enzymes, pepsin and gastricsin, were also found to be homologous in their sequences (4). They are also found to be homologous to rennin, a milk clotting enzyme (4, 5). The study of the structure and function of pepsin would be an extension of such a comparison. More specifically, the structure of porcine pepsin is particularly interesting. Most kinetic studies in acidic proteases have been carried out in porcine pepsin. The sequences around 4 methionine and 6 half-cystine residues in porcine pepsin are known (5). These sequences provide the overlaps for ordering fragments from specific chemical and enzymic cleavages at these residues, thus encouraging the study of the complete structure of the pepsin molecule. Additionally, a specific, substrate-like pepsin inhibitor, 1,2-epoxy-3-(p-nitrophenoxy) propane (abbreviated as EPNP), which reacts at the active site of pepsin, has been discovered recently (6). This provided an opportunity for the study of the active site structure of pepsin. The goal of this dissertational research was to elucidate the partial structure of porcine pepsin, by specifically taking advantage of the facts mentioned above.

A great deal of ground work on the structure of pepsin had been done. Porcine pepsin and its precursor pepsinogen had been analyzed for amino acid composition by several investigators. The most recent data are in Table I (1). The total number of amino acid residues per molecule in pepsinogen and pepsin are about 362 and 320, respectively. Porcine pepsin and its precursor, porcine pepsinogen each consists of a single peptide chain containing 3 disulfide bridges.

Table I

Amino Acid Compositions of Porcine Pepsinogen and Porcine Pepsin [a]

Amino Acid	Porcine Pepsinogen	Porcine Pepsin
Lysine	10	1
Histidine	3	1
Arginine	4	2
Aspartic Acid	44	40
Threonine	26	25
Serine	46	43
Glutamic Acid	28	26
Proline	19	16
Glycine	35	34
Alanine	19	16
Half Cystine	6	6
Valine	23	20
Methionine	4	4
Isoleucine	25	23
Leucine	33	28
Tyrosine	17	16
Phenylalanine	15	14
Tryptophan	(5)	(5)
Total	362	320
Amide	27	27

[a] From Rajagopalan et al. (1).

Their amino-terminal amino acid residues are isoleucine and leucine respectively and the carboxyl terminal amino acid residue is alanine in both proteins (7). It is well established that pepsin is produced by removal of a peptide portion from the amino-terminal sequence of pepsinogen. The difference between the amino acid composition of pepsinogen and pepsin (41 residues) is relatively high in their content of basic amino acids (9 Lys + 2 His + 2 Arg). The sequence of this 41 amino acid residue peptide which was removed on activation of the zymogen was elucidated by Ong and Perlmann (8). However, in a recent paper by Stepanov et al (9), the peptide sequence of Ong and Perlmann was shown to be incomplete. According to the survey made by Keil in 1970 (10), the data already published on the primary structure of pepsin contained about 63% of its residues in known sequences. Amino acid sequences around three disulfide bridges and four methionine residues of porcine pepsin had been determined by Tang and Hartley using diagonal electrophoretic methods (5). The number of tryptophan residues was not certain, and was reported to be four (11). five (12), and six (13). No tryptophan residues appeared to be removed from pepsinogen upon its activation (8). Amino acid sequences around 4 tryptophan residues had been determined by Dopheide and Jones (11). There is one phosphoserine residue in porcine pepsin. The amino acid sequence around this phosphoserine was revealed by several investigators (5, 14, 15). The sequence of a peptide containing 9 residues from the amino terminus of porcine pepsin was elucidated by Tang (16). The sequence of a 37 amino acid fragment which occupies the carboxyl terminal position of porcine pepsin, and is produced

upon cleavage of pepsin with cyanogen bromide was determined by Kostka et al (17). Earlier work had established the sequence of shorter fragments (27 residues) of the carboxyl-terminal sequence of pepsin (18). Two arginines and the single lysine of the entire porcine pepsin molecule are located in the carboxyl-terminal 20 amino acid unit.

The approach we took for the elucidation of the primary structure of porcine pepsin was as follows. Porcine pepsin was reduced with dithiothreitol in the presence of 6 M guanidine-hydrochloride, and S-aminoethylated with ethyleneimine. These reactions convert 6 cysteines into 6 aminoethylcysteine residues which are susceptible to tryptic digest (19). After dialysis and lyophilization, the reduced and aminoethylated pepsin was subjected to cyanogen bromide cleavage in 70% formic acid. The fragments were separated by column chromatography with Sephadex G-75. Five peaks were obtained. The peptides in each peak were further purified by DEAE-Sephadex A-25. The purified fragments were subjected to amino acid and terminal amino acid analyses. The complete sequences of the porcine pepsin will be assembled from the sequences of the fragments, using the sequences around 4 methionines provided by Tang and Hartley (5). A 44 amino acid residue peptide fragment was obtained, and the structure of this fragment was determined in this work. This sequence is the continuation of the known carboxyl-terminal sequence of 37 amino acid residues. Therefore, a 81 residue sequence at the carboxyl terminal end of the enzyme has been ascertained from this current work.

As to the catalytic function of porcine pepsin, it is important to

understand the nature of catalytic groups. The pH optimum for peptic cleavage of acetyl dipeptides is near pH 2.0. The activity of pepsin depends on the ionization of two groups on the enzyme, one having an apparent pKa of 1.0 and the other of 4.7 (20,21). Possible candidates for these two groups are the carboxyl groups and the imidazole group of the histidyl residue. Photooxidation of histidine does not affect the activity of the enzyme (22). Therefore carboxyl groups are probably involved in peptic catalysis. In order to identify the active carboxyl groups of pepsin, several specific inhibitors have been found for pepsin. In 1965, diphenyldiazomethane was found to inactivate pepsin (23). Since then, a number of diazo compounds have been reported to inactivate pepsin (24,25), such as diazoacetyl-DL-norleucine methyl ester and diazoacetyl-L-phenylalanine methyl ester. It is clear that all diazo inactivators esterify a unique aspartyl residue which is contained in the sequence: Ile-Val-Asp-Thr-Gly-Thr-Ser-Leu (25,28). Due to the fact that these diazo compounds can completely inhibit peptic activity, this reactive aspartyl residue is likely to be involved in enzymic catalysis. Partial inactivation of pepsin can be achieved by means of p-bromophenacyl bromide, which also esterifies the carboxyl group of a different aspartic acid residue (26). This aspartic acid residue probably is not directly involved in peptic catalysis, as only partial inactivation resulted from the reaction. Several substrate-like epoxides were found to act as inactivators of pepsin (6). The most potent among these, 1,2-epoxy-3-(p-nitrophenoxy) propane, caused the specific and irreversible inactivation of all enzymic activity. During this process, 2 molecules of epoxide were covalently

bound to each molecule of enzyme. This study also demonstrated that one of the two reactive sites was located in the active center of pepsin. Two carboxyl groups of the enzyme were proposed to be esterified each by one hydroxyl group of the glycol formed from the epoxide. Both of the sites reactive to 1,2-epoxy-3-(p-nitrophenoxy) propane differed from the carboxyl group known to react with previously reported inactivators, diazoacetyl-DL-norleucine methyl ester (24) and p-bromophenacylbromide (26). Thus the structure of the epoxide-reactive site had not been studied previously. Therefore, studies were undertaken in this work to identify the epoxide reactive residues in pepsin, to determine the amino acid sequence near the modified residues, and to elucidate the chemical linkages involved in the modification. Our results showed that one of the epoxide-modified residues is an aspartic acid esterified on the β-carboxyl group. The other modified residue, however, is a methionyl residue which is located 38 residues from the COOH terminus of the molecule. These current results have recently been reported in the Journal of Biological Chemistry (27).

CHAPTER II

MATERIALS AND METHODS

Materials

Proteolytic Enzymes

Porcine pepsin (2 x crystalline 2800 u/mg lot PMIBA), α-chymotrypsin (3 x crystalline), papain and diisopropylfluorophosphate-treated carboxypeptidase A were purchased from Worthington Corp. Crystalline pepsin obtained from Pentex was also used. Thermolysin was obtained from Calbiochem. Pronase, ficin and trypsin (I x crystalline, DCC treated, type XI) were purchased from Sigma. Subtilisin was a product of Novo Terapeutisk Laboratorium, Copenhagen, Denmark. Aminopeptidase M was obtained from GMBH Chemische FABRIK.

1,2-Epoxy-3-(p-nitrophenoxy) Propane

1,2-epoxy-3-(p-nitrophenoxy) propane (EPNP) was obtained from Eastman Kodak Co.

Dansyl Chloride

5-Dimethyl amino-1-naphthalene sulfonyl chloride (dansyl chloride) was obtained from K and K Laboratories, Inc., Plainview, New York.

Ninhydrin

Ninhydrin was obtained from Pierce Chemical Company, Rockford, Illinois.

Phenylisothiocyanate

Phenylisothiocyanate was obtained from Fisher Scientific Company, Fair Lawn, New Jersey, and was distilled before use.

Ethyleneimine

Ethyleneimine was obtained from K and K Laboratories, Inc., and was distilled over barium oxide before use.

Dithiothreitol

Dithiothreitol was purchased from Calbiochem.

Cyanogen Bromide

Cyanogen bromide was obtained from Eastman Kodak Co.

Other chemicals were of the highest purity, available from commercial sources and were used without further purification.

II

Methods

Preparation of EPNP-Pepsin

EPNP-pepsin modified at 30° was prepared a s described by Tang (6). The enzyme concentration was 4 mg/ml in 0.1 M acetic acid. Excess EPNP powder was added to create a suspension in the incubation solution which was stirred constantly for 48 hours. The incubation mixture was then filtered, adjusted to pH 5.0 with NH_4OH solution, dialyzed at 4° against several changes of water, and lyophilized. The product had about 90% inactivation and about 2 residues of EPNP were incorporated per molecule of inactivated enzyme. The inactivation was carried out in a similar manner at 10°, except that 120 hours of incubation was requried (29).

Proteolytic Digestion of EPNP-Pepsin

a) Peptic digest: Lyophilized EPNP-pepsin (200 mg) was suspended in 25 ml of water. To this solution, concentrated formic acid was added drop by drop until the pH reached 2.0. Crystalline pepsin (20 mg) was added to this insoluble EPNP-pepsin suspension, which was kept at 37° for 72 hr with rigorous stirring. A small amount of trubidity appeared during the acidification, but it usually disappeared in the first few hours of incubation. The digested mixture was directly applied to a Sephadex G-25 column for peptide separation.

b) α-Chymotrypsin digest: Lyophilized EPNP-pepsin (300 mg) from the 30° modification reaction was digested by 15 mg of the enzyme in 15 ml

of 0.2 M N-ethylmorpholine acetate buffer, pH 8.0, for 2 hr at 37°. When EPNP from the 10° modification reaction was used, twice the amount of starting material and digestion volume was used. At the end of incubation, 1 ml of glacial acetic acid was added and the digest was directly applied to a Sephadex column for peptide separation.

c) Thermolysin and subtilisin digest: The conditions for digestion by these two enzymes were identical to those in the α-chymotryptic digestion, except that in the case of thermolysin, $CaCl_2$ (0.001 M) was added.

d) Ficin digest: The conditions were identical to those in the α-chymotryptic digestion except that the pH of the buffer was 7.0, the incubation time was 6 hr, and the solution contained 6% mercaptoethanol.

e) Papain digest: The conditions were identical to those in the α-chymotryptic digestion except that 0.2 M pyridine-acetate buffer, pH 6, containing 6% mercaptoethanol was used. Incubation time was 5 hr.

f) Trypsin digest: The conditions were identical to the α-chymotryptic digestion, except the incubation time was 5 hr.

g) Aminopeptidase M digest: The conditions were identical to α-chymotryptic digestion except the enzyme to substrate ratio was 1:3, and incubation time was 24 hr.

Gel Filtration on Sephadex Columns

a) Separation of EPNP-peptides: A 3 x 150 cm column of Sephadex G-25 was equilibrated and eluted with 0.01 M acetic acid. Fractions of 10-ml

eluents were collected. The optical density of the fractions at 280 and 315 nm was determined spectrophotometrically.

b) Separation of proteotytic digests of CB5: The separation was performed the same as that for the EPNP-peptides except the tryptic peptides were eluted with 0.01 M NH_4OH, but only the optical density at 280 nm was determined.

c) Fractionation of cyanogen bromide fragments: The mixture containing approximately 800 mg of cyanogen bromide fragments from reduced and aminoethylated pepsin was dissolved in 50 ml of 0.1 M ammonium acetate pH 9.0, (a few drops of concentrated ammonium hydroxide was helpful) and applied directly to a 4 x 200 cm column of Sephadex G-75, which had been equilibrated with 0.1 M ammonium acetate pH 9.0. The eluting buffer consisted of 0.1 M ammonium acetate pH 9.0. The optical density of the effluent fractions was read at 280 nm spectrophotometrically.

High-Voltage Paper Electrophoresis

a) EPNP-peptides: The apparatus and conditions for high-voltage paper electrophoresis were the same as previously described (4). In the two-dimensional peptide mapping, the peptide sample was applied as a spot on Whatman No. 1 paper. The first-dimensional electrophoresis was performed at pH 3.5 for 1 hr at approximately 60v/cm. After drying, the paper strip was stitched onto a full sheet of Whatman No. 1 paper and the direction of the second-dimensional electrophoresis was performed at right angles to the

first, at pH 2.0 for 45 min, using the same voltage. In the preparative purification of peptides, peptide fractions (obtained from 200 mg of epoxide-modified pepsin in the Sephadex G-25 column chromatography) were applied as bands on Whatman 3 MM paper. The two-dimensional electrophoresis was performed as just described, except that the time allowed for migration was increased 50% to achieve better separation. The peptide bands were located by dipping a guide strip of the paper into a solution of cadmium-ninhydrin reagent (30). Whenever necessary, high-voltage electrophoresis was also performed at pH 6.0 for some of the peptides.

b) Peptides from CB5: In the preliminary peptide separation, the first dimensional electrophoresis was carried out at pH 2.0 for 30 min at a voltage of 60v/cm. The second-dimensional electrophoresis was performed at pH 3.5 for 1 hr at a voltage of 60v/cm. In this short duration two-dimensional electrophoresis, no peptides moved out of the paper. In the preparative purification of peptides, the two-dimensional electrophoresis was performed as just described, except that the time allowed for migration was increased 50% to 100% to achieve better separation. When necessary, peptides were also purified with pH 6.0 electrophoresis.

Mobilities of peptides after high-voltage electrophoresis were determined relative to aspartic acid at pH 6.0 or to serine at pH 2, so that calculation of the number of acidic groups could be made (31).

Identification of EPNP-Peptides on Paper

The positions of the EPNP-peptides on dried filter paper could be directly visualized after high-voltage electrophoresis and before staining with cadmium-ninhydrin reagent. This was done by holding the paper against an ultraviolet lamp. The EPNP-peptide spots or bands appeared as dark-brown shades. However, this procedure was dependable only in the absence of nearby fluorescent peptides. When confirmation was deemed necessary, the content of EPNP was quantitatively determined by measuring the optical density of the purified peptide at 315 nm.

Determination of EPNP in Peptides

The EPNP content of all peptides was quantitated spectrophotometrically (6). The optical density of an EPNP-peptide solution was determined at 315 nm. A molar extinction coefficient of 11,000 for EPNP was used in the calculation.

Removal of EPNP from EPNP-Peptides

The purified EPNP-peptides were incubated in 1 N NH_4OH at 45° for 12 hr. When the resulting mixtures were subjected to high-voltage electrophoresis, the further purified peptides were not visible as dark-brown spots on the filter paper when observed under ultraviolet light. Nor did these further purified peptides display absorption at 315 nm.

Amino Acid Compositions

The amino acid analyses were performed by the method of Spackman

(32) in a Spinco 120 B analyzer. Peptides were hydrolyzed in 5.7 N HCl for 18 hr at 108° in evacuated tubes. For the samples containing homoserine, the dried HCl-hydrolysate was dissolved in 100 µl of 2 M NH_4OH and incubated at 37° for 3 hr. After each sample was dried under reduced pressure in a desiccator, the amino acid analysis was performed. For the analysis of homoserine, the first buffer used in the amino acid analyzer had a pH value of 3.10 instead of the normal 3.28. Homoserine emerged before glutamic acid.

Amino-Terminal Residue Determination

The NH_2-terminal residues of the purified peptides were determined after reaction with dansyl chloride according to the method of Gray and Hartley (33). The dansyl amino acids were identified by thin-layer chromatography on polyamide sheets (34).

Carboxyl-Terminal Residue Determination

A diisopropylfluorophosphate-treated carboxypeptidase A suspension was washed and dissolved according to the method of Ambler (35). Digestion of peptides was carried out for 5 hr at 37° in 0.2 M N-ethylmorpholine acetate buffer, pH 8.0. The incubation mixtures were analyzed for free amino acids in an amino acid analyzer.

Hydrazinolysis was carried out under conditions essentially the same as those described by Fraenkel-Conrat and Tsung (36). The amount of COOH-terminal residue remaining after hydrazinolysis was measured with an amino

acid analyzer.

Edman Degradation

Dansyl-Edman degradation was accomplished according to the method of Gray (37).

Lithium Borohydride Reduction

The reduction of peptides by $LiBH_4$ was carried out according to the method of Wilcox (38). The peptide (0.02 μ moles) was suspended in 0.2 ml of freshly distilled anhydrous tetrahydrofuran in a small test tube to which 2 mg of repurified $LiBH_4$ was added. The material in the tube was frozen in a dry ice bath, evacuated with a vacuum pump, and sealed over a flame. The sample was heated in a 65° oven for 6 hr. After the tube was opened, the contents were dried in a vacuum desiccator over concentrated H_2SO_4 and hydrolyzed with HCl for amino acid analysis.

CHAPTER III

RESULTS

Part I: Amino Acid Sequence Around the Epoxide-reactive Residues in Pepsin

Purification of EPNP-Peptides (from EPNP-Pepsin Reacted at 30°)

The chromatographic pattern of EPNP-peptides from a peptic digest on a column of Sephadex G-25 is shown in Fig. 1. The front peak appeared at 400 ml of the eluent, thus occupying approximately the same elution position as pepsin. It was assumed that the materials contained in this peak were large-molecular weight polypeptides resulting from incomplete digestion of EPNP-pepsin. This was confirmed by a preliminary experiment using high-voltage electrophoresis at pH 2. The material from this peak had very slow mobility. Moreover, marked "tailing" of the peptide spots occurred. consequently, no further attempt was made to purify the peptides from this peak. The last peak in the chromatogram (Fig. 1) was EPNP-OH, which was identified by thin-layer chromatography on polyamide sheets (6). Two EPNP-peptide peaks, P_1 and P_2, were found at 700 and 900 ml in the effluent, respectively. The fractions under these two peaks were pooled, evaporated

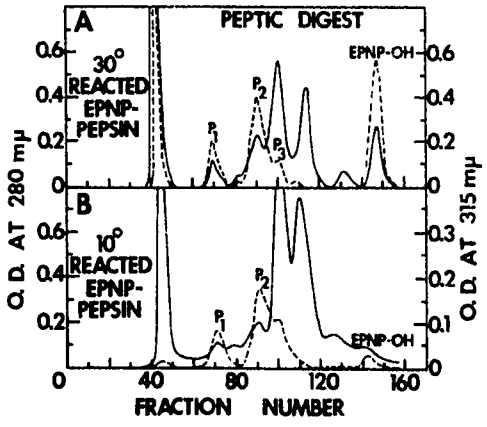


Fig. 1 - Chromatographic patterns of peptides from peptic digest EPNP-pepsin inactivated at 30°(A) and at 10°(B) on a column of Sephadex G-25. The solid line represents optical density at 280 nm and the broken line represents optical density at 315 nm.

to dryness in a rotating evaporator, and further purified by two-dimensional paper electrophoresis. The electrophoretic patterns of these pooled fractions are shown in Fig. 2. The EPNP-peptides, P_1 and P_2, were clearly separated from the other peptides. A minor peak, P_3, could not be observed on the paper and was not purified. The preparative electrophoretic separations were performed and the peptides P_1 and P_2 were recovered from the paper by elution.

Purification of EPNP-peptides from other enzymic digests proceeded similarly. The chromatographic separation of the subtilisin digest of EPNP-pepsin is shown in Fig. 3A. The peptides in the fractions under the 315 nm absorption peaks S_1, S_2, S_{2A}, and S_3 were further purified in high-voltage electrophoresis from which EPNP-peptides S_1, S_2, and S_3, were obtained. Fraction S_{2A} gave rise to an EPNP-peptide, S_{2A}, the sequence of which was revealed, in further studies to be identical to that of S_2. This peptide probably resulted from the cross-contamination by peak S_2, however. Peak S_{2A} may have contained a number of other EPNP-peptides in small quantities. Data from the peptide S_{2A} experiments are thus not discussed in the paragraphs below.

The chromatographic pattern of the papain digest of EPNP-pepsin on a column of Sephadex G-25 is shown in Fig. 4A. Three EPNP-peptide peaks, Pa_1, Pa_2 and Pa_3 appeared. The peptides giving rise to these peaks were purified on high-voltage electrophoresis.

The chromatographic separation of ficin digest of EPNP-pepsin

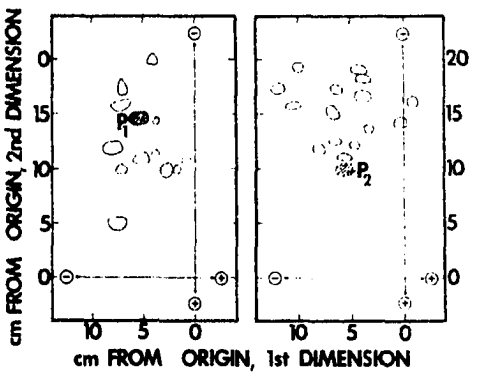


Fig. 2 - Two-dimensional high-voltage electrophoretic pattern of peptides recovered under peak P_1 in Fig. 1 (left), and peak P_2 in Fig. 1 (right). The first dimension was obtained at pH 3.5 and the second dimension at pH 2.0. The positions of peptide P_1 and P_2 were observed as dark brown spots against ultraviolet light. The peptides were then revealed using cadmium-ninhydrin reagent.

produced three EPNP-peptide peaks, F_1, F_2 and F_3 (Fig. 5A). Peptides F_2 and F_3 were purified from high-voltage electrophoresis. Further purification of material recovered from the fractions under peak F_1, however, failed to produce an EPNP-peptide. Since this peak emerged early in the chromatography, it probably represented a number of large EPNP-containing peptides resulting from incomplete digestion of EPNP-pepsin.

The thermolysin digest of EPNP-pepsin produced several EPNP-peptide peaks, Th_0, Th_1, and Th_2 on Sephadex column chromatography (Fig. 6A). The relatively large peak of EPNP-OH probably resulted from the hydrolysis of the EPNP groups from pepsin by the esterase activity of thermolysin. Peptides Th_1 and Th_2 were purified from the material recovered from the fractions under the peak. Peak Th_0, however, contained large molecular-weight material. Attempts to purify EPNP-peptide from this peak were unsuccessful.

α-Chymotryptic digests of EPNP-pepsin produced only one peak, C_1 (Fig. 7). An EPNP-peptide C_1 was obtained on high-voltage electrophoresis. The electrophoretic mobility and characteristics of the purified peptides are listed in Table 2.

Amino Acid Sequences of EPNP-Peptides

The amino acid compositions and other characteristics of the purified EPNP-peptides are shown in Table 2. Further studies on the sequence of these peptides revealed that each possessed one of two unique sequences, and the

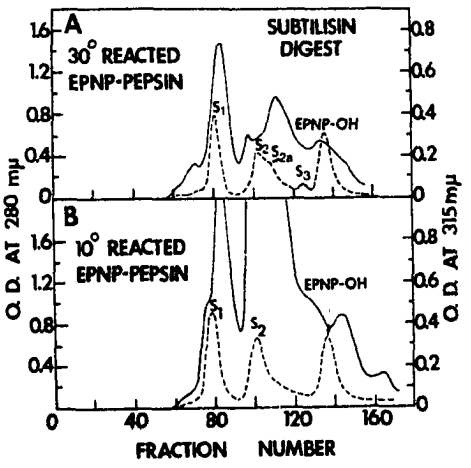


Fig. 3 - Chromatographic patterns of peptides from subtilisin digest of EPNP-pepsin inactivated at 30°(A) and at 10°(B) on a column of Sephadex G-25. The solid line represents optical density at 280 nm and the broken line represents optical density at 315 nm.

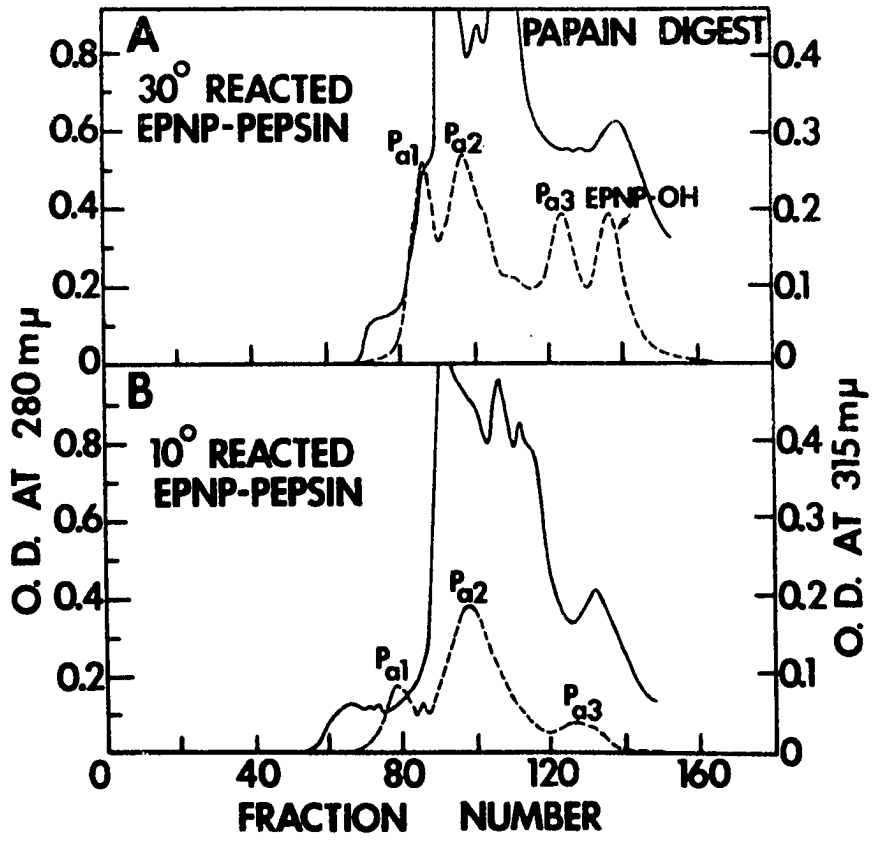


Fig. 4 - Chromatographic patterns of peptides from papain digest of EPNP-pepsin inactivated at 30°(A) and at 10°(B) on a column of Sephadex G-25. The solid line represents optical density at 280 nm and the broken line represents optical density at 315 nm.

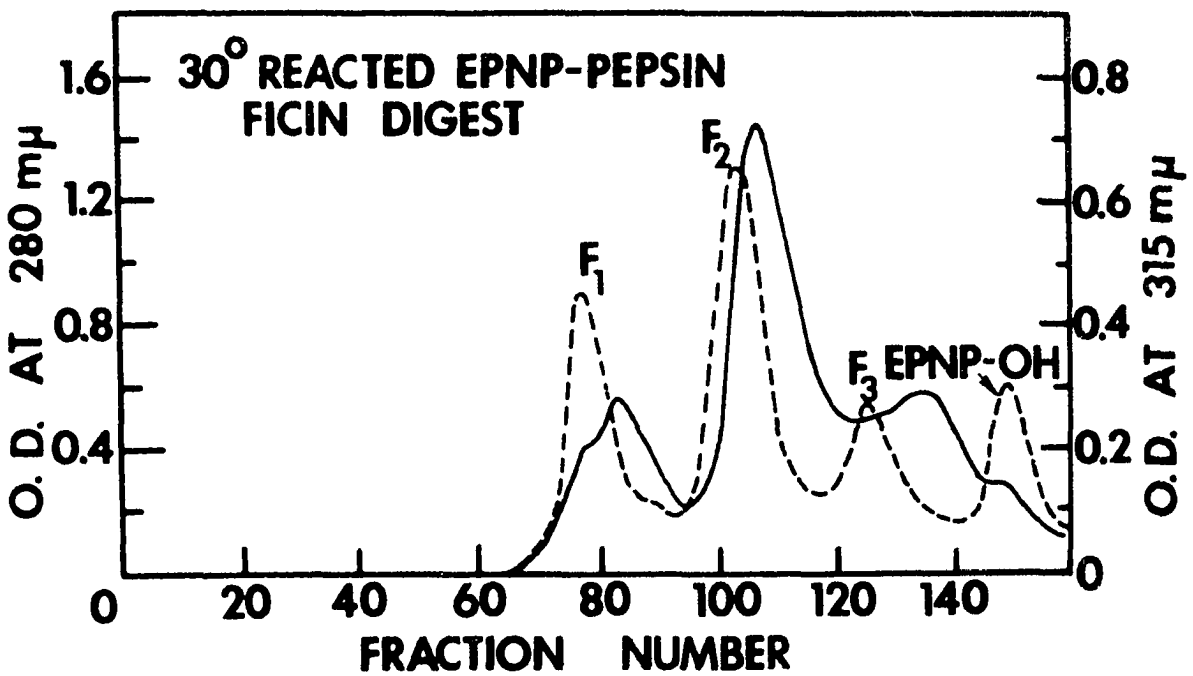


Fig. 5 - Chromatographic patterns of peptides from ficin digest of 30°-reacted EPNP-pepsin. The solid line represents optical density at 280 nm and the broken line represents the optical density at 315 nm.

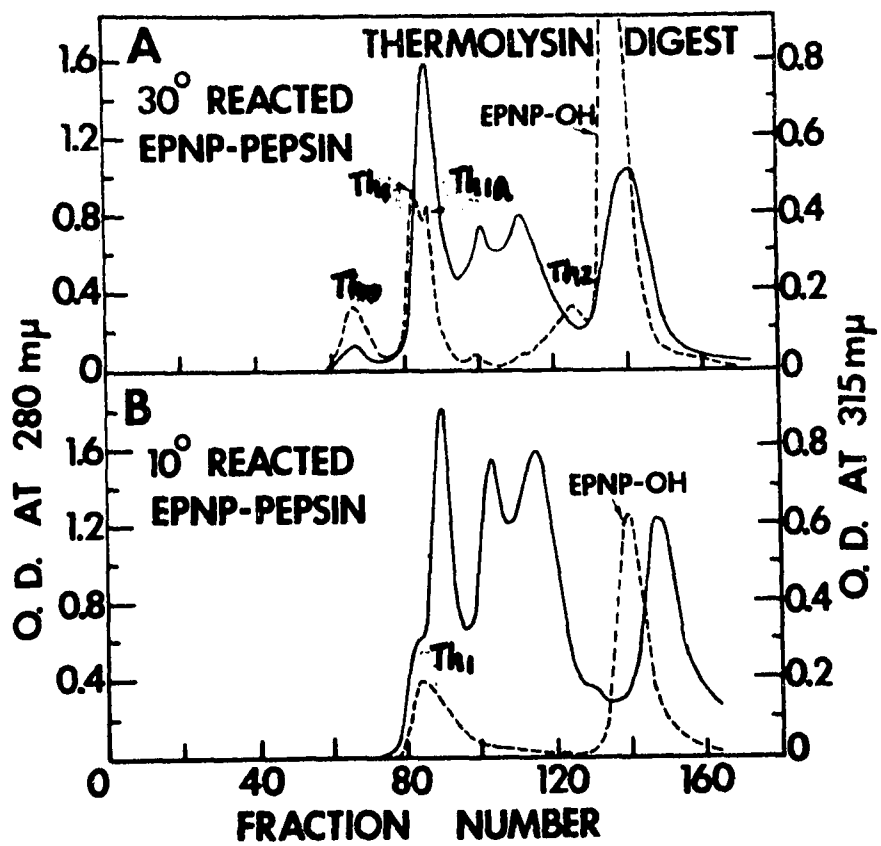


Fig. 6 - Chromatographic patterns of peptides from thermolysin digest of EPNP-pepsin inactivated at 30°(A) and at 10°(B) on a column of Sephadex G-25. The solid line represents optical density at 280 nm and the broken line represents optical density at 315 nm.

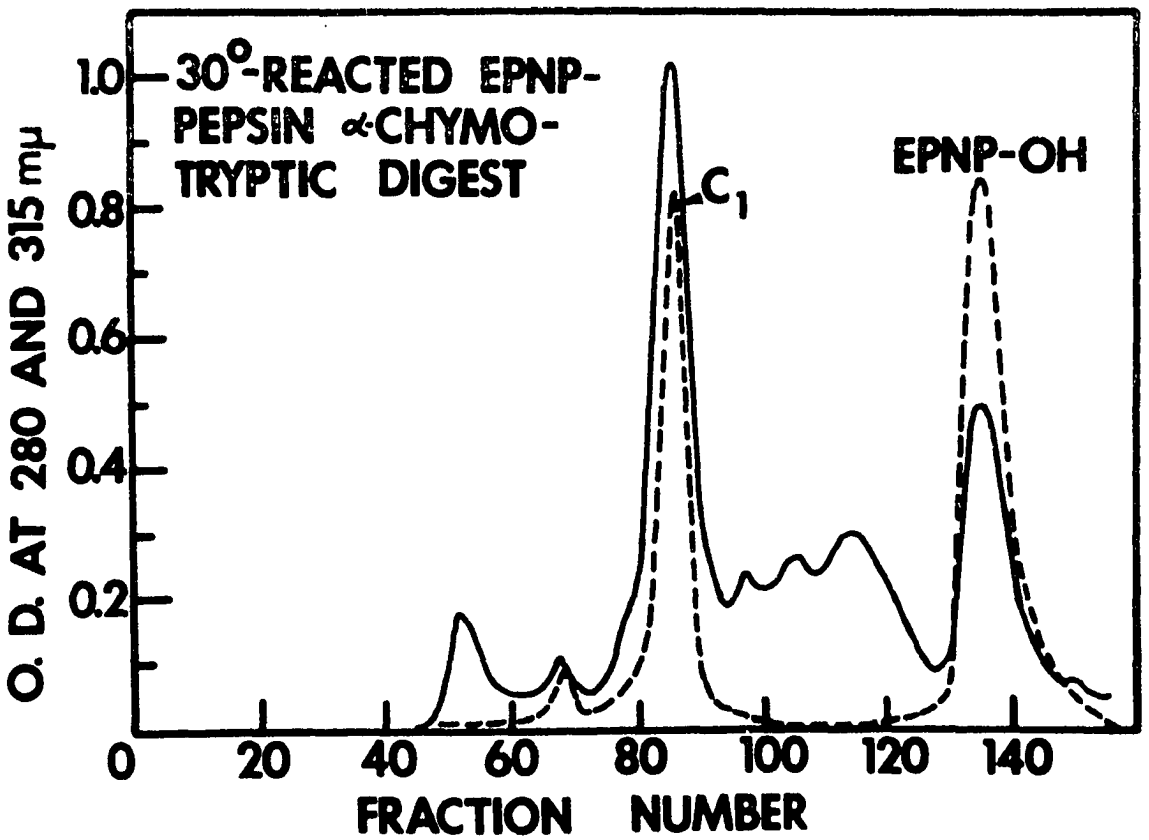


Fig. 7 - Chromatographic patterns of peptides from α-chymotryptic digest of 30°-reacted EPNP-pepsin. The solid line represents optical density at 280 nm and the broken line represents the optical density at 315 nm.

TABLE 2

Summary of Electrophoretic Mobility, Amino Acid Composition and Sequence Determination of Peptides

Peptide No.	Electrophoretic Mobility			Cadmium-Ninhydrin color	Residue of EPNP[b]	Total Residues of amino acid	Amino Acid Compositions[c]	Sequence Determination[d]
	pH 2.0	pH 6.0	pH 3.5					
P_1	0.74	-0.41	16.7	Red	1.0	11	$Asp_{1.0}Thr_{0.9}Ser_{1.5}Glu_{1.8}Pro_{0.7}Gly_{1.6}Val_{0.6}Hsr_{0.2}$	Glu-Gly-Met(EPNP)-Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu
P_1S_1	0.75	-0.30	—	Red	+	9	$Asp_{1.0}Thr_{1.1}Ser_{1.8}Glu_{1.0}Pro_{0.9}Gly_{1.3}Val_{0.6}Hsr_{0.2}$	Glu-Gly-Met(EPNP)-Asp-Val-Pro-Thr-Ser-Ser
P_1S_2	0.96	-0.69	—	Yellow	-	2	$Glu_{1.1}Gly_{1.0}$	Gly-Glu
P_1N_1	0.87	-0.31	—	Red	+	7	$Asp_{1.0}Thr_{0.9}Glu_{1.0}Pro_{0.8}Gly_{1.1}Val_{0.7}Hsr_{0.2}$	Glu-Gly-Met(EPNP)-Asp-Val-Pro-Thr (H) Ser-Ser-Gly-Glu
P_1N_2	0.71	-0.50	—	Yellow-Red	-	4	$Ser_{1.9}Glu_{1.0}Gly_{1.0}$	Ser-Gly-Glu
P_1N_3	0.85	-0.55	—	Red	+	3	$Ser_{0.9}Glu_{1.0}Gly_{1.1}$	Ser
P_1N_4	1.0	-0.07	—	Red	-	1	Ser	Gly-Met(EPNP)-Asp-Val-Pro-Thr
$P_1N_1E_1$	—	-0.06(N)	—	Yellow	+	6	N.D.	Glu-Gly-Met-Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu
P_1A_1	0.37	-0.67	—	Red	-	11	$Asp_{1.0}Thr_{0.9}Ser_{1.9}Glu_{1.9}Pro_{1.0}Gly_{2.0}Val_{0.9}Met_{0.6}$	Glu-Gly-MetO$_2$-Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu
P_1A_1O	—	—	—	—	-	11	$Asp_{1.0}Thr_{1.0}Ser_{1.7}Glu_{1.9}Pro_{0.7}Gly_{1.8}Val_{0.8}MetO_{0.6}$	Glu-Gly-Met-Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu
P_1A_2	1.1	-0.33	—	Red	-	11	$Asp_{1.0}Thr_{0.9}Ser_{1.9}Glu_{1.9}Pro_{0.9}Gly_{1.8}Val_{0.8}Hsr_{0.3}$	Glu-Gly-Met(EPNP)-Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu-Leu
C_1	0.64	-0.41	15.1	Red	1.0	12	$Asp_{1.0}Thr_{0.9}Ser_{1.8}Glu_{2.1}Pro_{1.0}Gly_{1.9}Val_{0.8}Leu_{1.0}Hsr_{0.2}$	Phe-Glu-Gly-Met(EPNP)-Asp-Val-Pro-Thr
T_1	0.74	-0.30	19.5	Red	1.0	8	$Asp_{1.0}Thr_{1.1}Glu_{1.0}Pro_{1.2}Gly_{1.0}Val_{1.2}Phe_{1.0}Hsr_{0.2}$	Glu-Gly-Met(EPNP)-Asp-Val-Pro-Thr-Ser-Ser
S_1	0.76	-0.31	19.0	Red	1.2	9	$Asp_{1.0}Thr_{0.9}Ser_{1.9}Glu_{1.0}Pro_{0.8}Gly_{1.1}Val_{0.8}Hsr_{0.22}$	Asp(EPNP)-Thr-Gly-Ser-Ser-Asn
P_3	0.47	-0.08(N)	15.2	brownish yell	1.1	6	$Asx_{1.9}Thr_{1.0}Ser_{1.9}Gly_{1.0}$	Asp-Thr-Gly-Ser-Ser-Asn
P_3-EPNP	0.36	-0.45	—	Red	-	6	$Asp_{2.0}Thr_{1.0}Ser_{1.9}Gly_{1.1}$	Asp-Thr-Gly-Ser
P_3H_1	1.1	-0.65	—	Red	-	4	$Asp_{1.0}Thr_{1.0}Ser_{0.6}Gly_{1.0}$	Thr-Gly
P_3H_2	1.1	-0.1(N)	—	Orange	-	2	$Thr_{1.0}Gly_{1.0}$	Thr-Gly-Ser-Ser-Asn
P_3E_1	0.67	N	—	Yellow	-	5	$Asp_{1.0}Thr_{1.0}Ser_{1.8}Gly_{0.9}$	Ile-Val-Asp(EPNP)
P_2	0.60	N	17.8	Red	1.0	3	$Asp_{1.0}Val_{0.7}Ile_{0.6}$	Val-Ile-Phe-Asp(EPNP)
T_2	0.51	-0.06(N)	14.8	Red	1.1	4	$Asp_{1.1}Val_{0.5}Ile_{0.5}Phe_{1.0}$	Val-Ile-Phe-Asp
T_2-EPNP	0.64	-0.34	—	Red	-	4	N.D.	Ile-Val-Asp(EPNP)
S_2	0.60	-0.08(N)	17.8	Red	1.0	3	$Asp_{1.0}Val_{0.9}Ile_{0.9}$	Asp(EPNP)
S_2	0.43	-0.07(N)	7.0	brownish yell	1.0	1	Asp	Ile-Val-Asp(EPNP)
Po_2	0.60	N	17.8	Red	1.0	3	$Asp_{1.0}Val_{0.9}Ile_{0.6}$	Ile-Val-Asp
Po_2-EPNP	0.77	-0.42	—	Red	-	3	N.D.	Val-Ile-Phe-Asp(EPNP)
Po_3	0.51	N	14.8	Red	+	4	$Asp_{1.0}Val_{0.7}Ile_{0.5}Phe_{1.0}$	Val-Ile-Phe-Asp
Po_3-EPNP	0.62	-0.34	—	Red	-	4	N.D.	

a. Electrophoretic mobilities are relative to serine (+1.0) at pH 2.0 and to aspartic acid at pH 6.0. At pH 3.5, the mobility is expressed as centimeters from origins after electrophoresis for 120 min at 36 v/cm. Under these conditions, serine migrated 9.5 cm. (N) indicates the peptides which are considered neutral in electrophoresis at pH 6.0.

b. The figures are residues of EPNP per peptide molecule. The + and - signs indicate the presence or absence of ultraviolet-quenching property of the peptide spots on filter paper, which is characteristic of the presence of EPNP.

c. The amino acid analyses are molar ratios. Serine was corrected for 15% of loss during the hydrolysis. N.D. indicates the peptides whose composition was not determined. Hsr and $MetO_2$ are the abbreviations for homoserine and methionine sulfone, respectively.

d. Sequences determined by the dansyl-Edman procedure are indicated as arrows toward the right (→) under the residues. Amino acids released by carboxypeptidase A are indicated as arrows toward the left (←) under the residues. The residues determined by hydrazinolysis are indicated by arrows containing the letter H ($\overset{H}{\rightarrow}$).

sequence data can be categorized accordingly (Fig. 8). For convenience of comparison, peptides that arise from the same sequence will be discussed together hereafter.

EPNP-Peptide Sequence I and the Presence of EPNP Alkylated Methionine

Peptide P_1 contained 1 residue of EPNP and about 10 amino acid residues. The mobility of P_1 on high-voltage electrophoresis at pH 2.0, however, indicated that this peptide contained 2 positive charges. This was surprising, since peptide P_1 contained no basic amino acid. Therefore, the possibility was considered that EPNP might have alkylated a methionyl residue in this peptide, which had then become a sulfonium salt, thus providing an additional positive charge. It is known that the sulfonium salt of methionine can be partially converted to homoserine during hydrolysis in 6 N HCl (39). An analysis for homoserine in peptide P_1 revealed about 0.15 residue after acid hydrolysis. Therefore, the presence of an EPNP-alkylated methionyl residue in P_1 was considered likely.

The dansyl-Edman procedure showed the third and fifth residues to be valine. Since peptide P_1 contained only one valine, the third residue or fifth residue possibly represented alkylated methionine and the observed dansyl-valine may have come from contamination by another peptide. Consequently, peptide P_1 was incubated in 1 N NH_4OH to remove EPNP from the peptide. When the resulting sample was subjected to high-voltage electrophoresis at pH 2.0, 2 new peptides were obtained. Peptide P_1A_1 had a slower mobility and

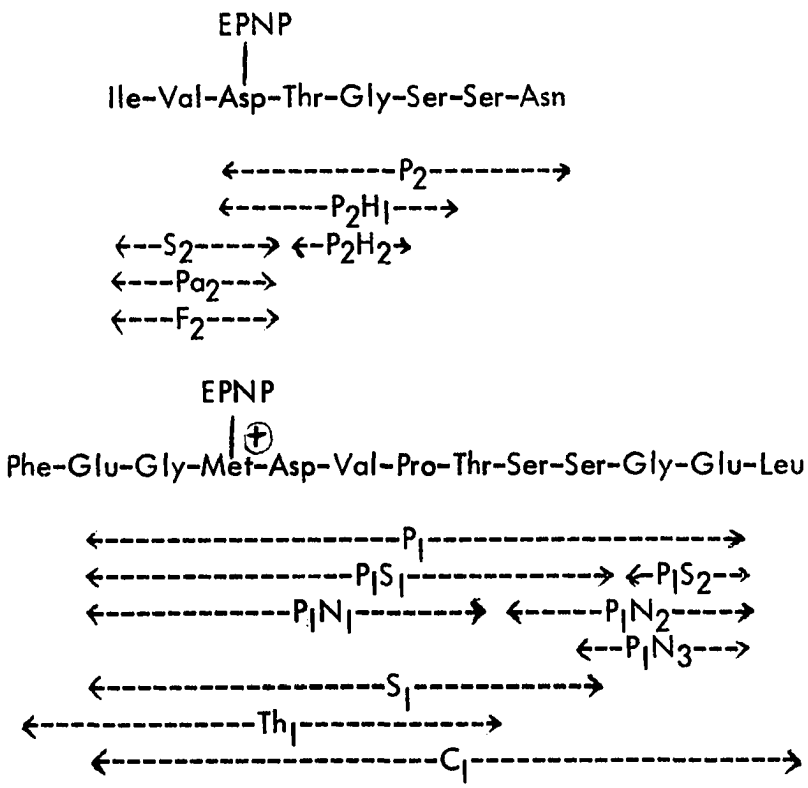


Fig. 8 - Sequences of EPNP-peptides from pepsin. The double-headed arrows indicate the part of the sequence covered by a given peptide. The perfix on the peptide number indicates the enzyme digest, where P=pepsin, Pa=papain, S=subtilisin, F=ficin, N=pronase, Th=thermolysin, and C=α-chymotrypsin.

peptide P_1A_2 a slightly faster mobility than P_1. Neither peptide contained EPNP. The amino acid compositions of the peptides were identical except that P_1A_1 now contained 0.6 residue of methionine.

When the peptide P_1A_1 was oxidized with performic acid, the resulting peptide, P_1A_1O, showed 0.6 residue of methionine sulfone (Table 2). The sequence of the first three residues of peptide P_1A_1O, using the dansyl-Edman procedure, was established to be $\underrightarrow{Glx\text{-}Gly\text{-}MetO_2}$-. This confirmed the previous assumption that the third residue might prove to be methionine. The chemical reactions involved, including those discussed below, are summarized in Fig. 9.

Peptide P_1A_2 contained two positive charges on high-voltage electrophoresis at pH 2.0. Therefore, the methionyl residue in this peptide must still have been a sulfonium salt. Because the incubation of peptide P_1 in 1 N NH_4OH produced a bright yellow color in solution, we suspected that the phenoxy bond in EPNP was hydrolyzed to produce p-nitrophenol (Fig. 9). This was confirmed by the identification of p-nitrophenol, using thin-layer chromatography on a polyamide sheet and an authentic sample as the standard (solvents: benzene: glacial acetic acid, 9:1, and n-heptane: n-butanol: glacial acetic acid, 3:1:1). No EPNP-OH was observed in this thin-layer chromatography. It seemed likely, therefore, that peptide P_1A_2 contained a methionyl residue in the form of the sulfonium salt of 1,2-dihydroxy-propane (Fig. 9). This was confirmed by the amino acid analysis of peptide P_1A_2 which contained 0.3 residue of homoserine (Table 2). The formation of homoserine from the

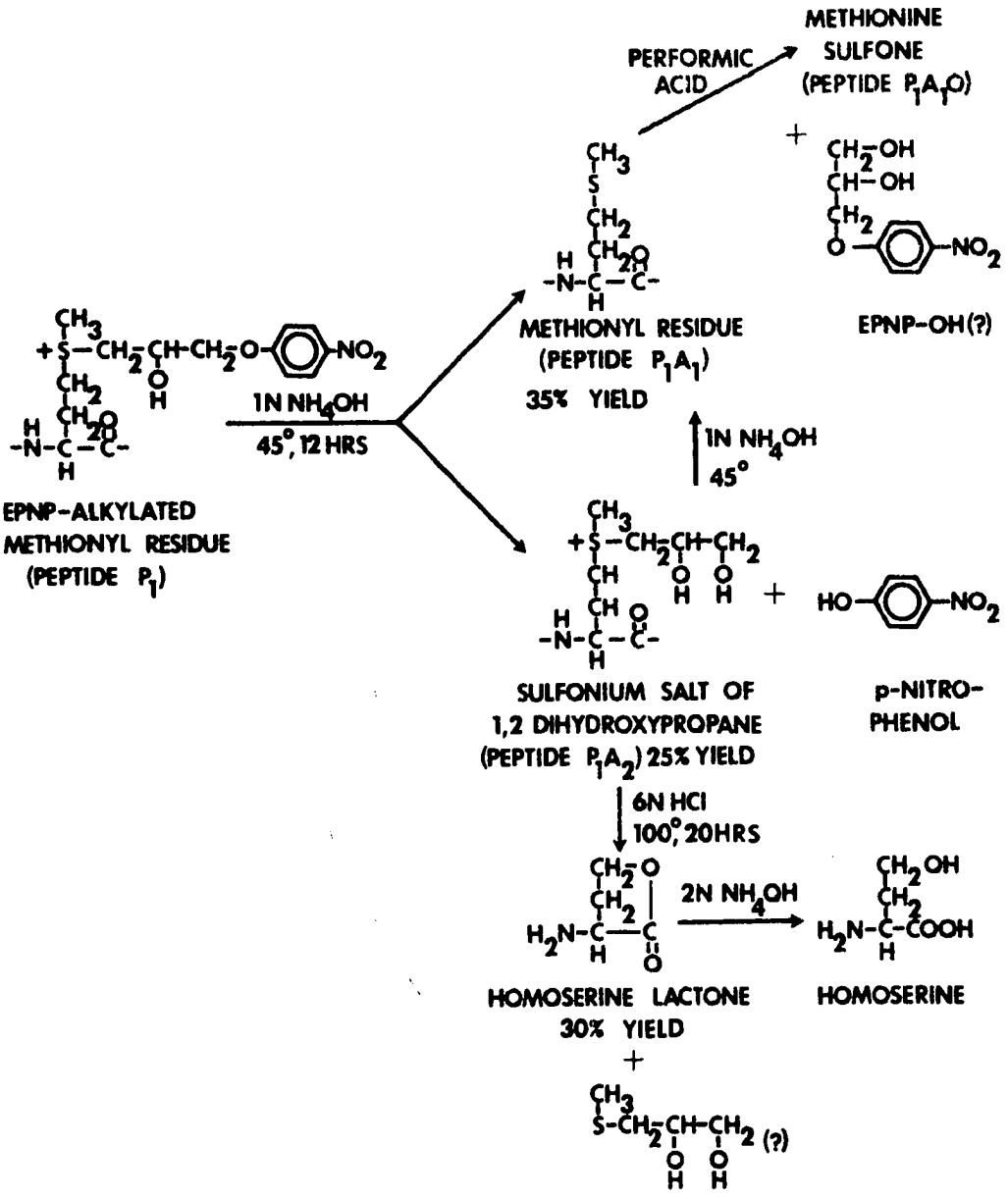


Fig. 9 - Summary of the possible chemical reactions of EPNP-alkylated methionyl peptide (P_1).

sulfonium salts of methionine is known to occur in the cyanogen bromide treatment of methionyl residues (40) and also in methionine alkylated by iodoacetamide (39,41). The lower yield of homoserine in peptide P_1 was probably due to the large alkylating group (EPNP) which hindered the cyclization of homoserine lactone. Additional confirmation of the position of the EPNP group in peptide P_1 was obtained by measuring optical density at 315 nm after each step in the Edman degradation. As indicated in Fig. 10, the removal of methionine from the peptide caused a large drop in optical density (O.D.) at 315 nm, in contrast to other Edman degradation steps, which produced only small changes.

Peptide P_1 was digested for 10 hr with subtilisin (substrate: enzyme= 20:1) and subjected to high-voltage electrophoresis at pH 2.0. Two peptides were obtained. Peptide P_1S_1 contained 9 residues from the amino terminus of P_1. The dansyl-Edman procedure revealed the sequence of the first 5 residues to be identical to that of P_1. The carboxyl-terminal residue was serine. P_1S_2 was a dipeptide, Gly-Glu. Peptide P_1 was digested for 6 hr with pronase (substrate: enzyme=20:1). Separation of the digestion mixture in high-voltage electrophoresis at pH 2.0 resulted in 4 peptides. Peptide P_1N_1 contained 7 residues from the amino terminus of P_1. The dansyl-Edman procedure yielded 6 residues: $\underrightarrow{\text{Glx}}$-$\underrightarrow{\text{Gly}}$-$\underrightarrow{(\text{Met})}$-$\underrightarrow{\text{Asx}}$-$\underrightarrow{\text{Val}}$-$\underrightarrow{\text{Pro}}$-. The carboxyl-terminal residue was threonine. Peptide P_1N_2 was $\underrightarrow{\text{Ser}}$-$\underrightarrow{\text{Ser}}$-$\underrightarrow{\text{Gly}}$-$\underrightarrow{\text{Glu}}$. Peptide P_1N_3 was $\underrightarrow{\text{Ser}}$-$\underrightarrow{\text{Gly}}$-$\underrightarrow{\text{Glu}}$, and P_1N_4 was free serine.

The information described above was sufficient to construct the

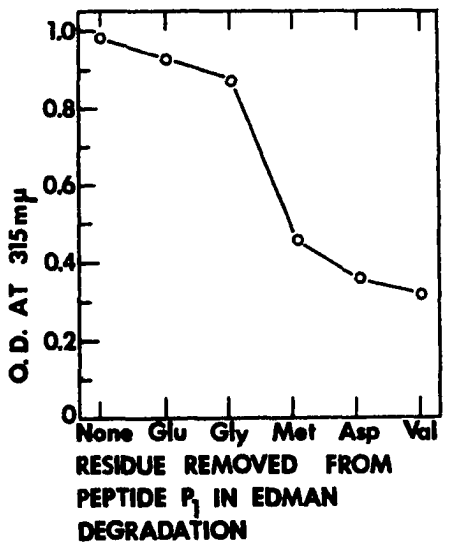


Fig 10 - The optical density at 315 nm of peptide P_1 during the Edman sequential degradation. The removal of EPNP-alkylated methionine, the third residue, produced a large drop in optical density at 315 nm.

sequence of P_1 to be Glu-Gly-Met(EPNP)-Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu (see Table 2). The amino-terminal glutamic acid was assigned from the electrophoretic mobility. After removal of the amino-terminal residue of peptide P_1, the resulting peptide had a negative charge on paper electrophoresis at pH 6.0 (Table 2). The aspartyl residue was assigned as the fourth residue based on two findings. First, peptide P_1 contained 2 negative charges on high-voltage electrophoresis at pH 6.0. After one-step Edman degradation, the peptide P_1E_1 contained 1 net negative charge. The results are consistent with both glutamic acids and aspartic acid occurring in the acidic form since the modified methionyl residue contains a positive charge. Second, peptide P_1N_1 was neutral on high-voltage electrophoresis at pH 6.0 after removal of glutamic acid from the amino terminus in Edman degradation. Since the sulfonium salt of methionine contains a positive charge, again the aspartyl residue must have been in acid form.

Peptide C_1 was apparently related to P_1. The amino acid composition showed that it contained leucine in addition to the amino acids in P_1. The amino-terminal sequence was Glu-Gly- and the carboxyl-terminal residue was leucine. Logically, then, the sequence of C_1 must be Glu-Gly-Met(EPNP)-Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu-Leu.

Peptide Th_1 had the same amino acid composition as P_1N_1 but it contained an additional phenylalanine and was related to peptide P_1N_1. The dansyl-Edman procedure showed the three amino-terminal residues to be Phe-Glu-Gly-. The carboxyl-terminal residue was threonine. Therefore, the

sequence of peptide Th_1 had to be Phe-Glu-Gly-Met(EPNP)-Asp-Val-Pro-Thr.

From the sequences of the peptides described above, a unique sequence I can be constructed: Phe-Glu-Gly-Met(EPNP)-Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu-Leu (see the summary of Fig. 8).

EPNP-Peptide Sequence II

Peptide P_2 contained 6 amino acid residues and 1 EPNP group. The dansyl procedure revealed the amino-terminal residue to be aspartic acid. The carboxyl-terminal amino acid was asparagine with a minor contaminant of serine; these were released by carboxypeptidase A and were identified as dansyl derivatives. The following evidence made it apparent that the modifying EPNP group was located on the amino-terminal aspartyl residue: a) Removal of the amino-terminal residue by Edman degradation reduced the original optical density of EPNP at 315 nm by about 60%. b) The peptide, which resulted from one-step Edman degradation and was purified using high-voltage electrophoresis at pH 2.0, contained no EPNP. c) When EPNP was removed from peptide P_2, the resulting peptide changed color on filter paper after ninhydrin-cadmium treatment from its original brownish red (P_2) to red, indicating a change of side chain in the amino-terminal residue.

The first residue of peptide P_2 could be only partially removed by Edman degradation, probably because of a large modifying group. This peptide was therefore incubated in 1 M NH_4OH to remove EPNP (see Chemical Linkage,

below) and repurified by high-voltage electrophoresis at pH 2. This peptide was then subjected to the Edman procedure to remove the amino-terminal aspartyl residue (but the yield still below 50%). The resulting peptide, P_2E_1, after repurification by high-voltage electrophoresis at pH 2.0, had an amino acid composition of $Asp_{1.0}Thr_{1.0}Ser_{1.8}Gly_{0.9}$. The dansyl-Edman procedure revealed the sequence of P_2E_1 to be: Thr-Gly-Ser-Ser-Asn. The asparagine was again identified as a dansyl derivative after removal of 4 residues of P_2E_1. Since peptide P_2E_1 showed a minor NH_2-terminal glycine in addition to threonine, the amino acid composition was determined after removal of the NH_2-terminal residue by Edman degradation. The results showed a composition of $Asp_{1.0}$ $Thr_{0.2}Ser_{2.0}Gly_{1.1}$. The disappearance of 0.8 residue of threonine confirmed that it was located at the NH_2-terminus of P_2E_1. In order to confirm the results of Edman degradation, a partial hydrolysis of peptide P_2 was carried out (in 5.7 N HCl at 100° for 25 min). Two peptides were purified in high-voltage electrophoresis at pH 3.5 and 2.0. Peptide P_2H_1 contained 4 residues (Table 2) and amino-terminal aspartic acid. This peptide corresponded to Asp-Thr-Gly-Ser. The dipeptide P_2H_2 was found to be Thr-Gly by dansyl-Edman procedures. The sequence of P_2 was thus Asp(EPNP)-Thr-Gly-Ser-Ser-Asn.

Peptides Pa_2, F_2, and S_2 were apparently identical tripeptides. The sequence of these peptides was determined by the dansyl-Edman procedure to be: Ile-Val-Asp(EPNP). Since the carboxyl-terminal aspartic acid was the only residue capable of reacting with EPNP, we concluded that the EPNP had modified this aspartic acid. This conclusion was supported by the fact

that these peptides were neutral in high-voltage electrophoresis at pH 6.0. Peptide S_3 contained 1 mole of aspartic acid for each mole of EPNP. Thus it had to be EPNP-modified free aspartic acid.

Two additional peptides, Th_2 and Pa_3, were recovered from minor peaks in chromatography (Figs. 4A, and 6A). The dansyl-Edman procedure showed to have the same sequence Val-Ile-Phe-Asp(EPNP).

In the above sequences, the EPNP-aspartyl residue was found either at the NH_2- or COOH-terminus of the peptides (Fig. 8). These two types of EPNP-aspartyl peptides were mutually exclusive in all the digests used. We previously found that only two sites on pepsin reacted with EPNP(6), one of which (as demonstrated above) was a methionine residue. Therefore, these two types of EPNP-aspartyl peptides must have belonged to the same continuing sequence. In attempts to obtain an overlapping sequence, other proteolytic digestions of EPNP-pepsin were carried out (with elastase, streptococcal proteinase, and pronase). These experiments were unsuccessful, however, partly because the EPNP-peptides appeared to lose their EPNP group during incubation--apparently due to the esterase activity of the proteases. Short-duration, partial hydrolysis of EPNP-pepsin with various proteases also failed to yield an overlapping peptide. However, evidence supporting the overlapping EPNP-aspartyl sequences was obtained from the quantification of the EPNP peptides recovered from 10°-reacted EPNP-pepsin. These results are presented in the following section.

Overlap of EPNP-Aspartyl Peptides in Sequence II

In a separate study (29), it was found that when pepsin was inactivated by EPNP at 4° or 10°, for up to 50 hr, the ratio in moles of EPNP incorporation to each mole of inactivated pepsin molecule approached a value of 1.2. This observation enabled us to quantify the recovery of EPNP-aspartyl peptides from 10°-reacted EPNP-pepsin. The digests of 10°-reacted EPNP-pepsin (having a ratio of EPNP/molecule inactive pepsin=1.40) which had been incubated with pepsin, subtilisin, papain, or thermolysin were subjected to column chromatography on Sephadex G-25. The chromatographic patterns are shown in parts B of Figs. 1, 3, 4, and 6. In general, the chromatographic patterns from 10°-reacted EPNP-pepsin contained fewer minor peaks. The absence of peaks S_{2A} and S_3 in Fig. 3, and that of peak Pa_3 in Fig. 4, was most noticeable. Since most of these peptides represented a minor sequence Val-Ile-Phe-Asp(EPNP), this sequence must have been derived from a minor, nonspecific reaction of EPNP with a site at 30°. This reaction did not occur at 10°. Therefore, this minor sequence must not have overlapped with peptide P_2.

In the thermolysin digest of 10°-reacted EPNP-pepsin (Fig. 6B). Peak Th_1 represented an EPNP-methionyl peptide (see Fig. 8). The other peak was EPNP-OH. This indicated that the aspartyl-bonded EPNP had completely hydrolyzed to EPNP-OH. The total amount of EPNP in both peaks was quantified by using the total optical density at 315 nm; this showed a Th_1 to EPNP-OH ratio of 0.4 to 1.0. In this chromatogram, 76% of the original

EPNP in pepsin was recovered. Since the thermolysin digest did not produce insoluble material, a nearly uniform loss of EPNP from all peaks during chromatography could be assumed. This indicated that of the 1.4 residues of EPNP in 10°-reacted EPNP-pepsin, 0.4 residue had reacted with the methionyl sequence I and 1.0 residue had reacted with an aspartyl residue. Quantitative estimation of recovery in the chromatogram of subtilisin digest confirmed this (Fig. 3B). The recovery of S_1, an EPNP-methionyl peptide, was 0.45 residue. The combined recovery of the EPNP-aspartyl peptide (0.45 residue) and EPNP-OH (0.41 residue) was 0.86 residue--again suggesting an initial 0.4 residue of EPNP-methionine and 1 residue of EPNP-aspartic acid in the starting EPNP-pepsin. The amounts of recovery of peptide Pa_2, S_2, and P_2 are listed in Table 3. Assuming a uniform loss of EPNP, the recovery of the Ile-Val-Asp(EPNP) sequence in Pa_2 and S_2 was 1.0 and 0.51 residues, respectively. The actual recovery of S_2 was 0.55 residue, since in this chromatography almost all of the original EPNP was recovered. It should be noted that the EPNP-aspartyl peptides spontaneously hydrolyzed to EPNP-OH and peptides in the subtilisin digest due to alkaline conditions and also possibly to the esterase activity of the enzyme. Therefore, the value of 0.51 residue for S_2 should be a minimum, since no correction for hydrolysis loss was applied. The papain digest produced considerable amounts of insoluble material which accounted for only 0.3 residues of peptide Pa_2 which was actually recovered. In this case, it is uncertain whether uniform loss in the various fractions was applicable. Therefore, the value of 1 residue based on ratio may not be

Table 3

Recovery of EPNP-Aspartyl Peptides from 10°-Reacted Pepsin in Column Chromatography (Sephadex G-25)

Peptide	Sequence	No. of Residues Recovered[a]	No. of Residues by Ratio[b]
S_2[c]	Ile-Val-Asp(EPNP)	0.55	0.51
Pa_2[d]	Ile-Val-Asp(EPNP)	0.30	(1.00)[f]
P_2[e]	Asp(EPNP)-Thr-Gly-Ser-Ser-Asn	0.88	1.08

a. Number of residues recovered was calculated based on the moles of peptide recovered from 1 mole of EPNP-aspartic acid in the starting EPNP-pepsin.

b. Number of residues by ratio =(amount of an EPNP-peptide recovered x 1.4)/ total amount of EPNP-containing material recovered. This calculation gives the number of residues of an EPNP-peptide assumming that the total recovered EPNP material corresponds to 1.4 residues initially in pepsin.

c. Calculated from the chromatogram shown in Fig. 3B.

d. Calculated from the chromatogram shown in Fig. 4B.

e. Calculated from the chromatogram shown in Fig. 1B.

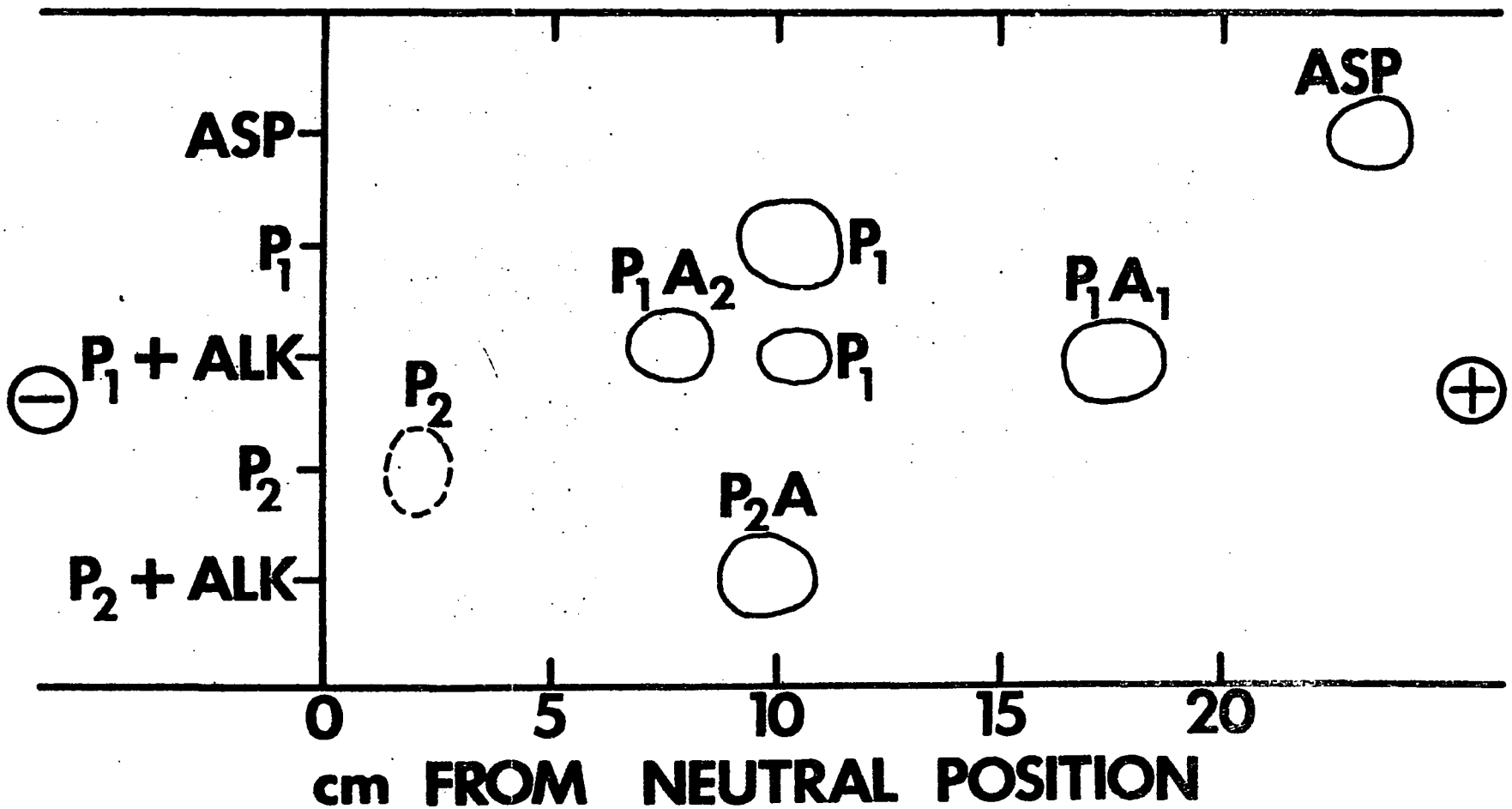
f. Papain digest produced insoluble material. This figure may not be reliable.

justifiable. The actual recovery of P_2 was 0.88 residue; this represented 1.08 residue if uniform loss was assumed. It is apparent that any addition of the number of residues of NH_2-terminal EPNP-aspartyl peptide to the number of residues of COOH-terminal EPNP-aspartyl peptide produces a value exceeding 1. For example, 0.55 residue of S_2 plus 0.88 residue of P_2 would equal 1.43 residue of EPNP-aspartyl peptide. The values of this sum ranged from 1.18 (Pa_2 + P_2 in actual recovered residue) to 1.59 (S_2 + P_2 in residue by ratio). Since the content of EPNP-aspartic acid was only 1 residue in EPNP-pepsin, these two sequences must have belonged to the same EPNP-aspartyl sequence. Therefore, sequence II is: Ile-Val-Asp(EPNP)-Thr-Gly-Ser-Ser-Asn.

Chemical Linkage of Aspartyl Residue and EPNP

Since the linkage of EPNP to the β-carboxyl group of aspartyl residue in Sequence II was expected to be an ester bond, the electrophoretic mobility of peptide P_2 at pH 6.0 was compared before and after the removal of EPNP in 1 M NH_4OH. Fig. 11 shows that peptide P_2 was essentially neutral at pH 6.0. After EPNP was removed, the peptide migrated toward the anode with a mobility corresponding to the gaining of 1 negative charge. Similar experiments were carried out for peptides with COOH-terminal EPNP-aspartic acid (Th_2 and Pa_2). The mobility change on high-voltage electrophoresis at pH 6.0 is listed in Table 2. Additionally, after the incubation of peptide P_2 in 1 M NH_4OH to remove EPNP, the incubation mixture was subjected to thin-layer chromatography on polyamide layers (6). EPNP-OH

Fig. II - Mobility of EPNP-peptides P_1 and P_2 in high-voltage electrophoresis at pH 6.0 before and after treatment with alkaline solution. The peptides were dissolved in 1 M NH_4OH and incubated at 45° for 12 hr


ASP
P1
P1 + ALK
P2
P2 + ALK
ASP
P1A2
P1
P1
P1A1
P2
P2A
⊖
⊕
0
5
10
15
20
cm FROM NEUTRAL POSITION

was found, and no p-nitrophenol which is seen in case of peptide P_1 was observed.

To further test the hypothesis of the ester linkage, peptide P_2 was reduced with $LiBH_4$ according to the method of Wilcox (38). In this method, which is specific for an ester linkage, the β-carboxyl ester of an aspartyl residue is reduced to homoserine. The amino acid composition of reduced P_2 showed a loss of 0.50 residue of aspartic acid and a gain of 0.33 residue of homoserine. When peptide P_2 was incubated in 1 M NH_4OH to remove EPNP before being subjected to the reduction by $LiBH_4$, no homoserine was found in the resulting peptide. Both of the above experiments indicated that EPNP was linked to the β-carboxyl group of aspartyl residue in Sequence II by means of an ester bond.

Part 2: Amino Acid Sequence of A Cyanogen Bromide Fragment-CB5

Reduction and Aminoethylation of Pepsin

The procedure used was a modified form of that described by Slobin and Singer (42). To a 2% solution of porcine pepsin (1 gm. in 50 ml solution) in 6 M guanidine hydrochloride-2 M Tris-0.01 M EDTA, pH 8.2, dithiothreitol was added so that the resulting solution 0.1 M. After flushing with nitrogen gas, the flask was closed with a rubber stopper and stirred for 3 hr at 37°. One milliliter of ethyleneimine (redistilled over BaO) was then added, and the flask was flushed with nitrogen gas. This process was repeated after 30 min. After stirring for 30 more minutes, the content of the flask was dialyzed against distilled water (thoroughly bubbled with nitrogen gas) for 36 hr with several changes of water, then lyophilized. The pH of the reaction mixture after the first ethyleneimine addition was 8.6 and after the second addition was 8.8. Amino acid analysis of the modified pepsin gave 5.8 residues of aminoethylcysteine (based on lysyl content of 1). The color value of aminoethylcysteine in the amino acid analysis was assumed to be 91% of that for lysine (43). The residual cystine in the modified protein was negligible in the amino acid analysis.

Cyanogen Bromide Cleavage of Aminoethylated Pepsin

The procedure used was that described by Steers *et al* (44). A 1% solution of modified protein was prepared by solubilizing the lyophilized

aminoethylated pepsin in 90% formic acid and then diluting this solution to 70% formic acid with distilled water. A 50-fold molar excess of CNBr over methionine residues was used and reaction proceeded at room temperature for 16-20 hr. After cyanogen bromide cleavage, the reaction mixture was diluted approximately 10-fold with distilled water and lyophilized to complete dryness. The methionine content of the sample as shown in the amino acid analysis was less than 5% after 10 hr of reaction.

Gel Filtration of CNBr Fragments

The lyophilized powder of CNBr treated sample was fractionated in a gel filtration column. The chromatogram of cyanogen bromide fragments on Sephadex G-75 column was shown in Fig. 12. The successive peaks were designated as CB1, CB2, CB3, CB4 and CB5, respectively. The fraction contained under each peak was pooled and lyophilized separately. The recoveries of these cyanogen bromide fragments from Sephadex G-75 column were shown in Table 4. The peak, CB5, which emerged alone late in the chromatography, suggested that this is a relatively smaller and pure fraction. Work was thus undertaken to further purify this fraction and study its sequence.

Purification of CB5

About 350 mg of lyophilized CB5, which was obtained from Sephadex G-75 column chromatography, was dissolved in 15 ml of 0.05 M Tris buffer - pH 8.0 and applied directly to a 2 x 60 cm column of DEAE-Sephadex A-25, equilibrated with the same buffer. The chromatogram of CB5 on DEAE-

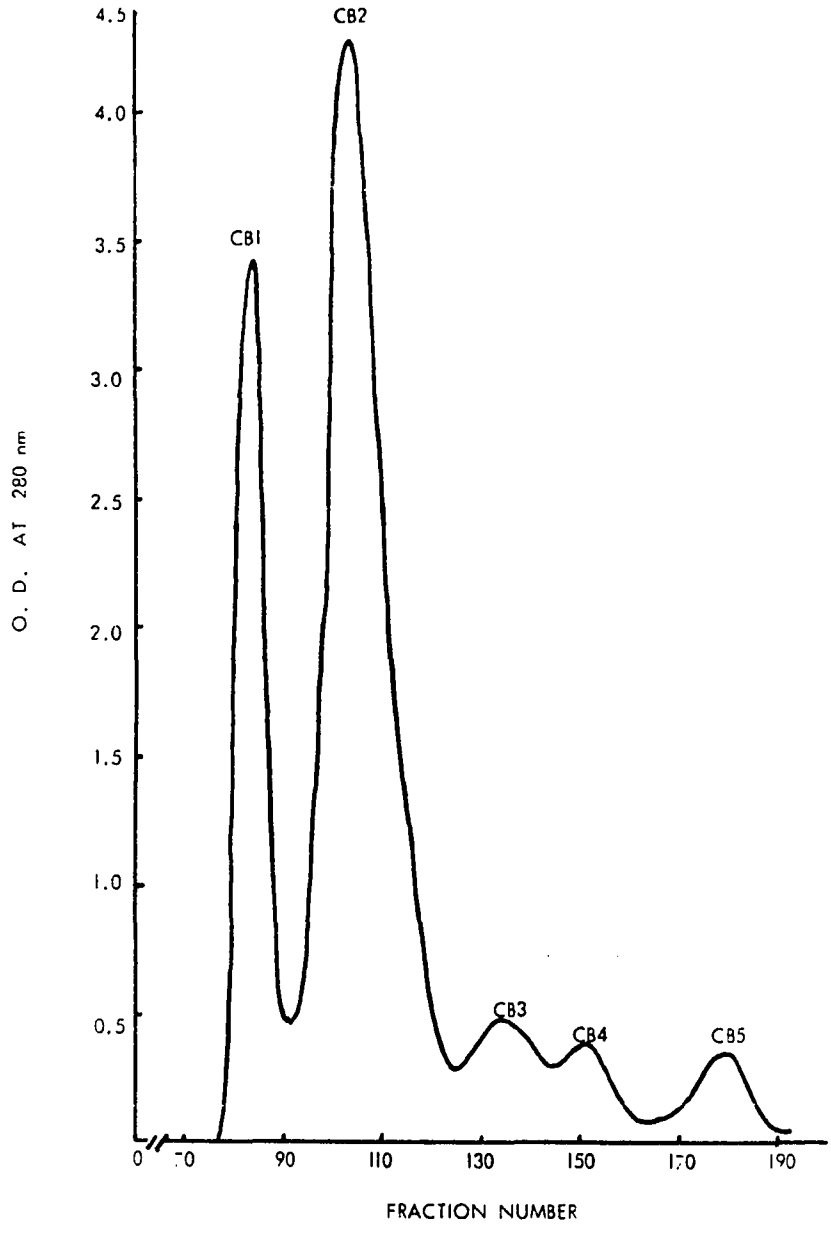


Fig. 12 - Chromatogram of cyanogen bromide fragments of 800 mg of aminoethylated porcine pepsin on Sephadex G-75.

Table 4

Recoveries of Cyanogen Bromide Fragments of 800 mg of Aminoethylated Porcine Pepsin from Sephadex G-75 Column

Cyanogen Bromide Fragments	Yield (mg)	Recoveries (percentage of total put on column)
CB1	75	9.4
CB2	325	40.6
CB3	85	10.6
CB4	60	7.5
CB5	105	13.1
Total	650	81.2

Sephadex A-25 shown in Fig. 13, contains two slightly overlapping peaks. The fraction contained under each peak was pooled, lyophilized, desalted on a 3 x 150 cm column of Sephadex G-25, and then lyophilized again. 210 mg of peptide was recovered from the first peak and 70 mg from the second peak. It is obvious that the first peak represents the major products of the CB5 fraction. Therefore, only this fraction was studied during further structure determination. Amino-terminal analysis of the purified CB5 (from the first peak) showed valine as NH_2-terminus. An additional fluorescent spot which moved very close to the phenylalanine position on the polyamide layer and which was assumed to be due to unhydrolyzed DNS-val dipeptide. COOH-terminal analysis of purified CB5 with carboxypeptidase A (35) showed homoserine as the COOH-terminus, with trace amounts of glycine, glutamic acid, phenylalanine, serine and threonine. The penultimate amino-terminal residue of purified CB5 determined by the dansyl-Edman method was isoleucine. Amino acid compostion of purified CB5 was shown in Table 5. From the composition, the size of CB5 was calculated to be about 40 residues.

Isolation of Peptides from Proteolytic Digest of CB5

In order to elucidate the amino acid sequence of CB5, 60 mg of purified CB5 in 6 ml buffer were digested separately with trypsin, thermolysin, α-chymotrypsin and papain. The mixtures were acidified by adding 1 ml glacial acetic acid, and applied to a 3 x 150 cm column of Sephadex G-25 which was equilibrated and eluted with 0.01 M acetic acid. In tryptic digest,

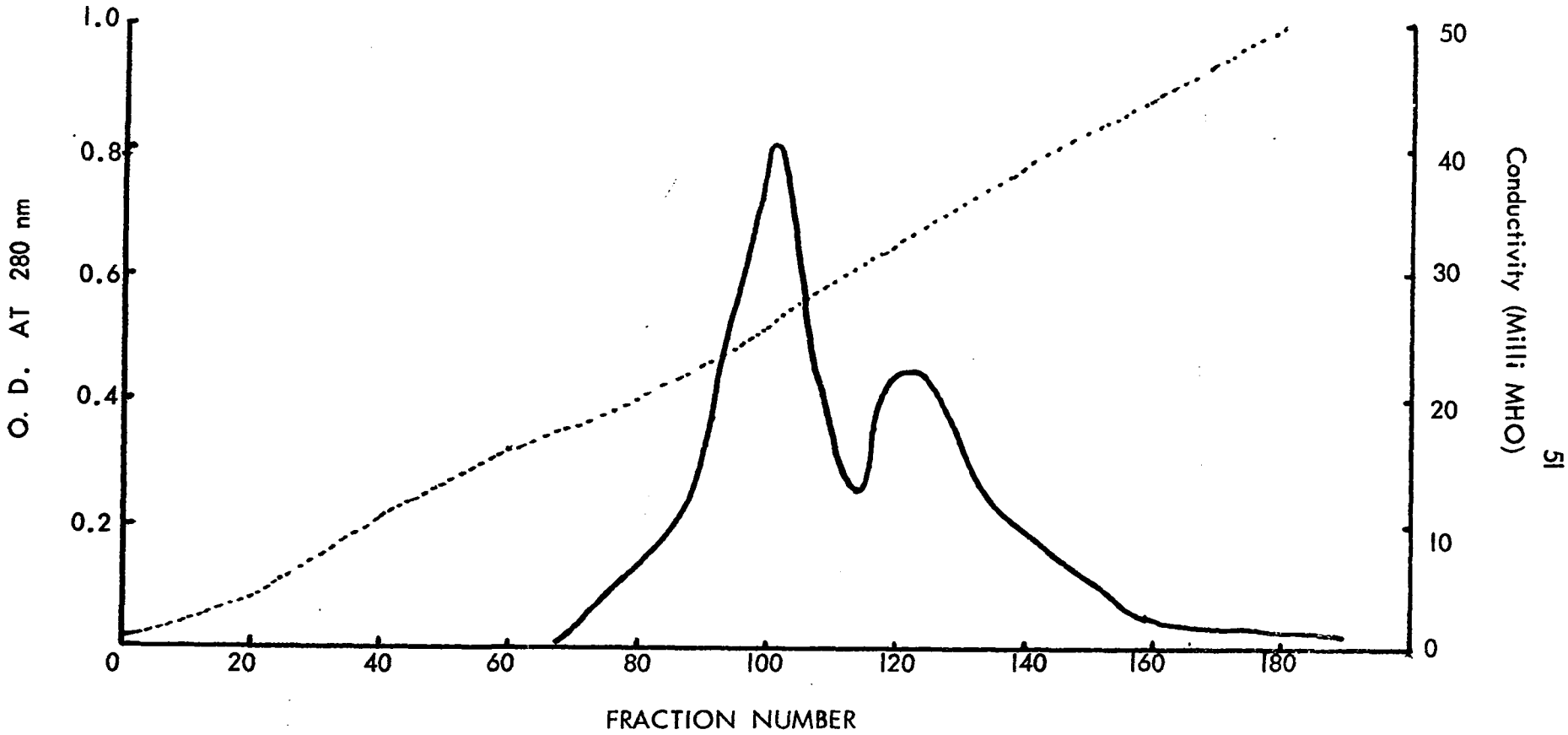


Fig. 13 - Chromatographic patterns of crude CB5 on a column of DEAE-Sephadex A-25. The column was eluted by a NaCl gradient started from 1 liter of 0.05 M Tris pH8.0 to 1 liter of 0.05 M Tris, 1 M NaCl pH 8.0. Each fraction contained 10 ml.

Table 5

Amino Acid Analysis of Purified CB5 [a]

Amino Acid [b]	No. of Residues	Nearest Integer	Found in Peptide [c]
Aminoethylcysteine	1.20	1	2
Aspartic Acid	6.17	6	6
Threonine	2.14	2	2
Serine	7.55	8	8
Homoserine	0.74	1	1
Glutamic Acid	3.25	3	3
Proline	2.20	2	3
Glycine	2.88	3	3
Alanine	1.07	1	1
Valine	2.26	2	3
Isoleucine	3.60	4	5
Leucine	2.86	3	3
Tyrosine	1.90	2	2
Phenylalanine	1.98	2	2
NH_3			(2)
Total		40	44

[a] Acid hydrolysis was conducted with 5.7 N HCl , 108° for 20 hr.

[b] Serine residue was corrected for 15% loss during acid hydrolysis.

[c] Actual amino acid composition determined in sequence studies.

the mixture was applied to the same column directly and eluted with 0.01 M ammonium hydroxide. The chromatograms of trypsin, thermolysin, α-chymotrypsin and papain digests of CB5 were shown in Figs. 14, 15, 16, and 17 respectively. From each chromatogram, 0.2 ml of sample was taken from each tube which had UV absorption. The samples were dried first in a vacuum desiccator, redissolved in a small volume of water, applied to Whatman No. 1 paper and subjected to separation in high voltage electrophoresis at pH 2.0. Fig. 18 shows an example of such an electrophoretic pattern of a tryptic digest. The peptide bands revealed the distribution of peptides in these tubes from which the pooling of tubes were planned. With this type of preliminary information, the fractions from each digest were subdivided into three to four parts; each part was pooled and dried in rotary evaporator. Further peptide separation on high voltage electrophoresis was then performed at pH 2 and where necessary, subsequently at pH 3.5 and 6.0 in order to obtain pure peptides. The peptides which were obtained were listed in Table 6. The amino acid composition as well as sequence determination were also listed.

Amino Acid Sequence Determination

a) Tryptic digest:

Sequence of peptide T_1: Amino acid analysis (Table 5) showed that peptide T_1 contained 4 residues. The results of dansyl-Edman analysis and hydrazinolysis revealed the sequence to be Val-Ile-Ser-Cys(Ae).

Sequence of peptide T_2: Amino acid analysis (Table 5) showed that peptide T_2 contained 11 residues. From dansyl-Edman procedures, the first 4 residues

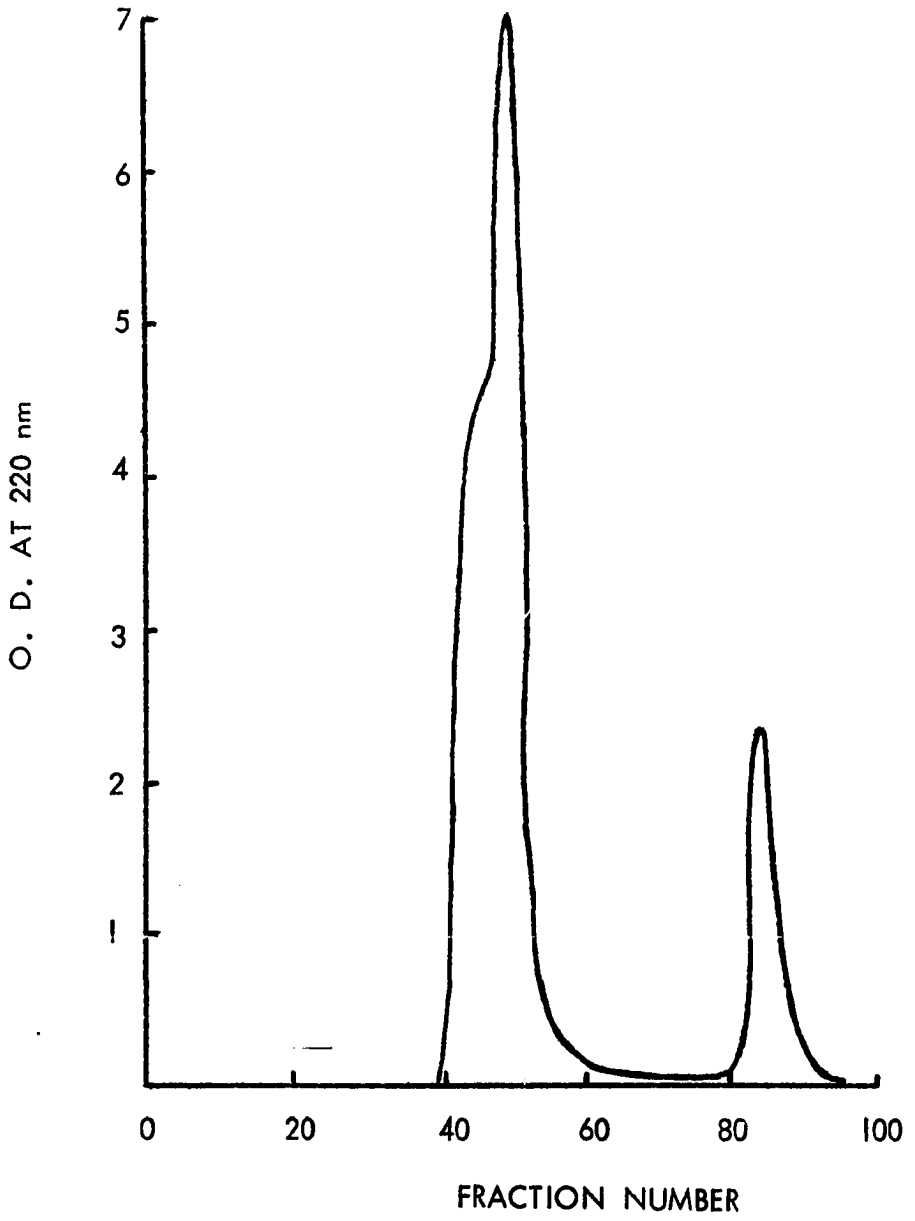


Fig. 14 - Chromatographic patterns of peptides from tryptic digest of CB5 on a column of Sephadex G-25. Each fraction contained 10 ml.

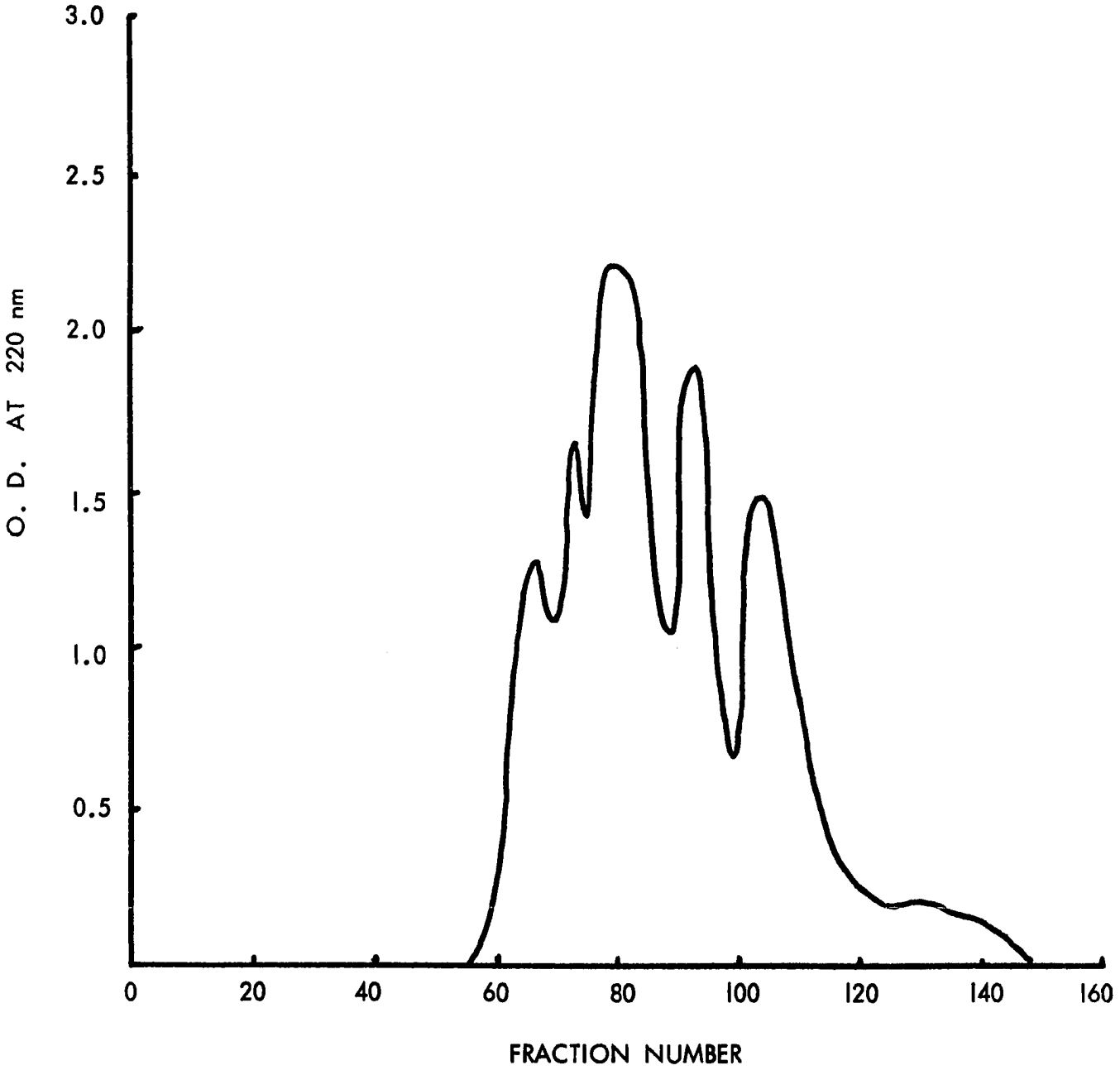


Fig. 15 - Chromatographic patterns of peptides from thermolysin digest of CB5 on a column of Sephadex G-25. Each fraction contained 10 ml.

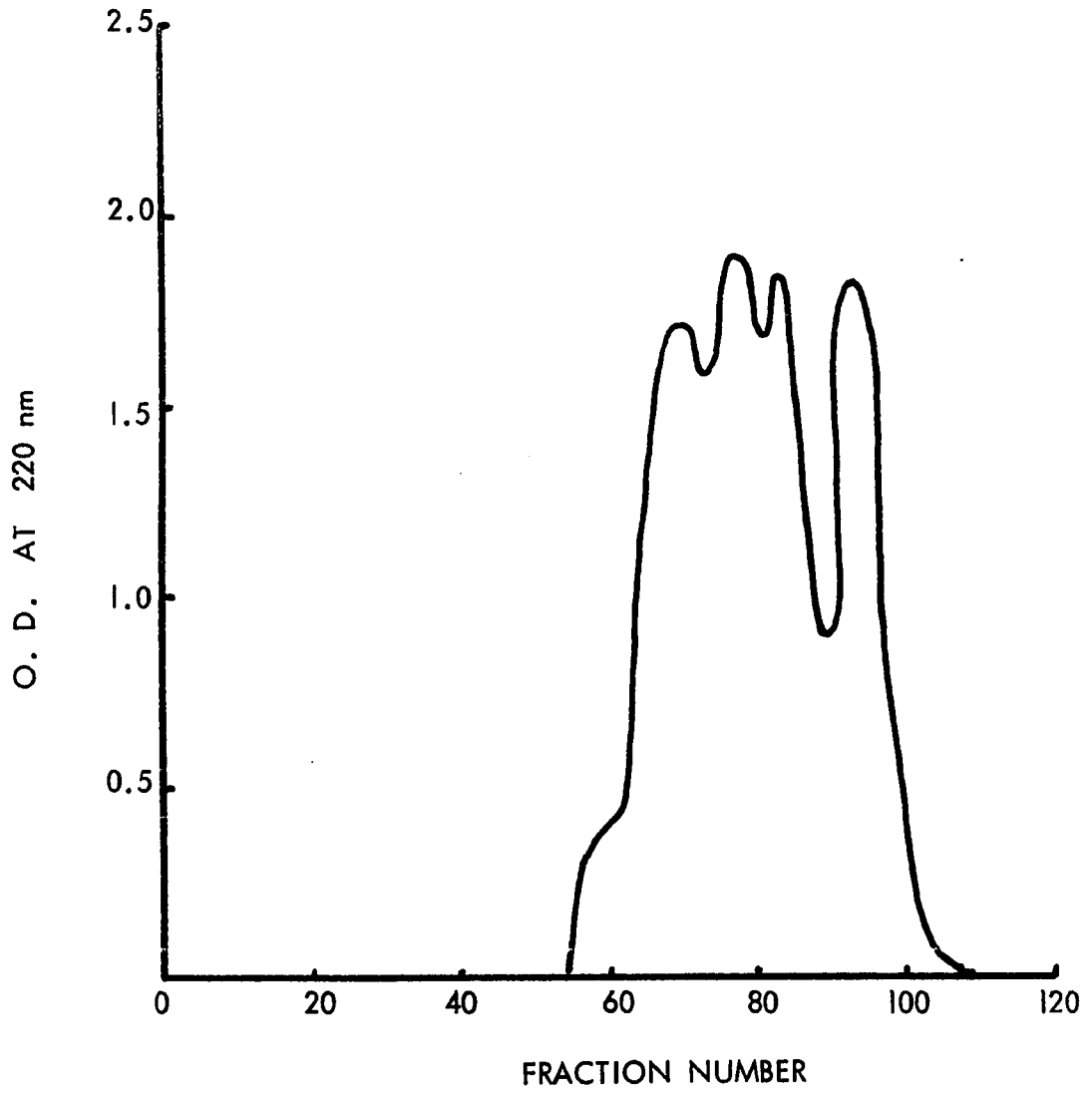


Fig. 16 - Chromatographic patterns of peptides from α-chymotryptic digest of CB5 on a column of Sephadex G-25. each fraction contained 10 ml.

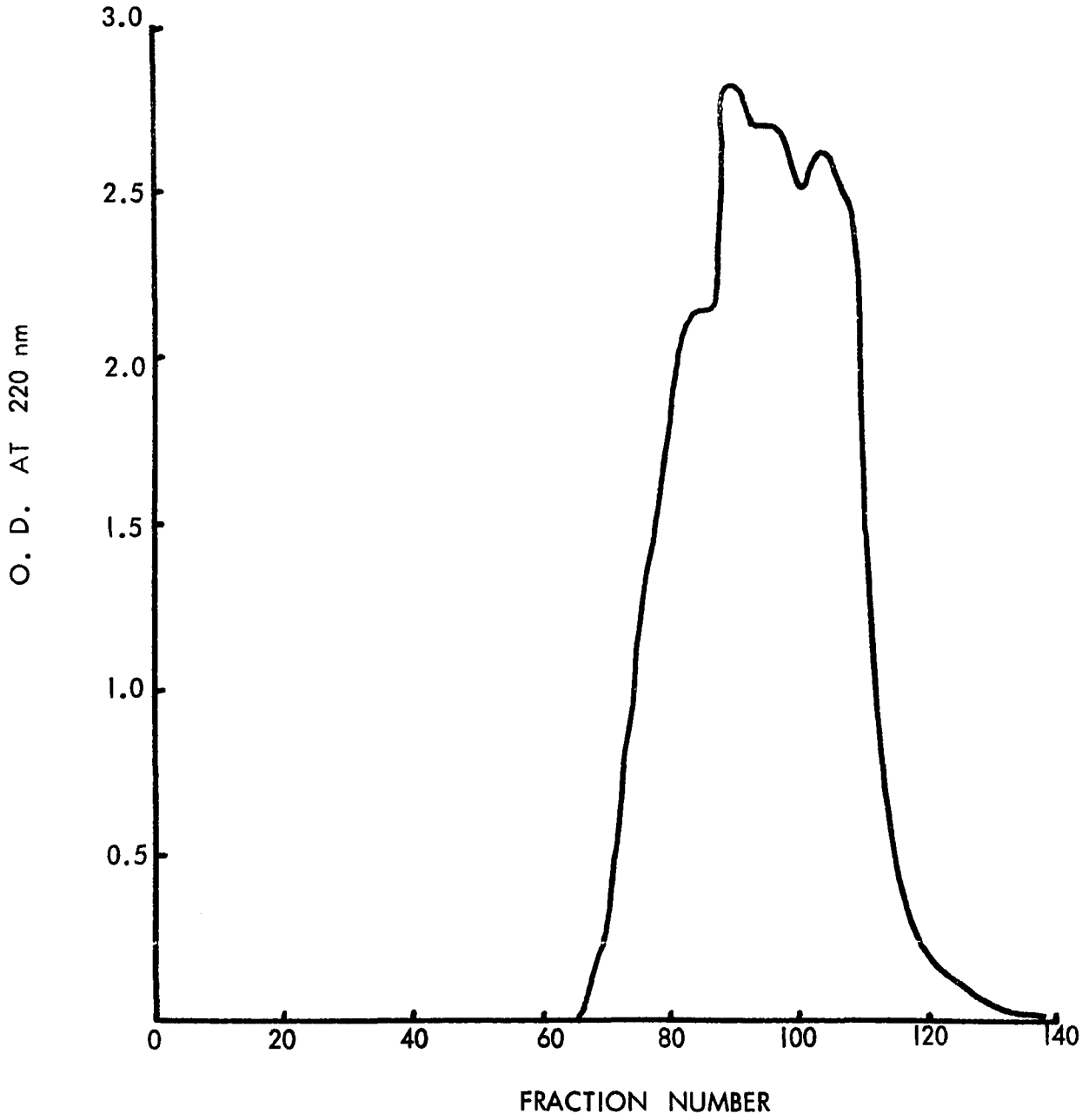


Fig. 17 - Chromatographic patterns of peptides from papain digest of CB5 on a column of Sephadex G-25. Each fraction contained 10 ml.

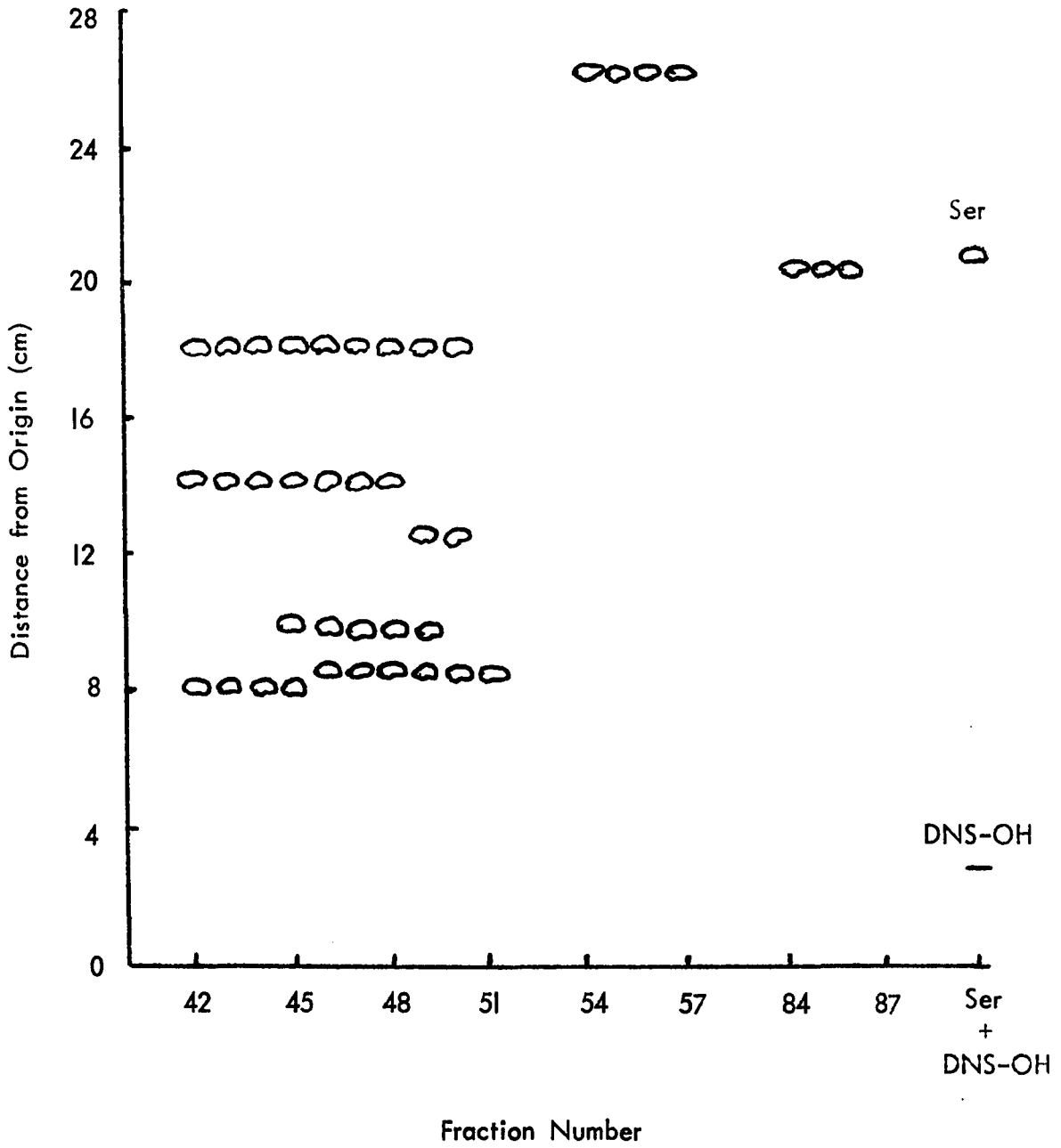


Fig. 18 - Electrophoretic patterns of tryptic digested peptides at pH 2.0. Serine and dansyl sulfone used as markers.

Table 6

Summary of Electrophoretic Mobility, Amino Acid Composition and Sequence Determination of Peptides from CB5

Peptide No.	Mobility [a]		Cd-Nin color	No. of Residues	Amino Acid Compositions [b]	Sequence Determination [c]
	pH 2.0	pH 6.0				
T_1	1.20	+0.28	red	4	$Ser_{1.3}Val_{1.0}Ile_{0.9}Cys(Ae)_{0.5}$	Val-Ile-Ser-Cys(Ae)
T_2	0.38	-0.44	brown →red	11	$Asp_{2.3}Ser_{3.3}Pro_{1.0}Val_{0.6}Ile_{1.7}$ $Leu_{1.1}Phe_{1.0}$	Ser-Ser-Ile-Asp(H)-Ser-Leu-Pro-Asp-Ile-Val-Phe
T_3	0.32	-0.22	yel →red	14	$Asp_{0.9}Thr_{0.7}Ser_{2.0}Glu_{0.7}Pro_{1.8}$ $Gly_{0.8}Ala_{0.8}Val_{1.0}Ile_{1.1}Leu_{1.1}$ $Tyr_{1.5}$	Thr-Ile-Asp-Gly-Val-Gln-Tyr-Pro-Leu-Ser-Pro-Ser-Ala-Tyr
T_{3Thl}	0.46	-0.32	yel →red	8	$Asp_{1.1}Thr_{0.8}Glu_{1.0}Pro_{1.1}Gly_{1.0}$ $Val_{0.8}Ile_{0.8}Tyr_{0.7}$	Thr-Ile-Asp-Gly-Val-Gln-Tyr-Pro
T_{3ThlS1}	0.65	-0.47	yel →red	5	$Asp_{1.0}Thr_{0.9}Gly_{1.0}Val_{1.0}Ile_{0.8}$	Thr-Ile-Asp-Gly-Val
T_{3ThlS2}	0.75	N	red	3	$Glu_{1.0}Pro_{1.1}Tyr_{0.9}$	Gln-Tyr-Pro

Table 6-Continued.

Peptide No.	Mobility [a] pH 2.0	Mobility [a] pH 6.0	Cd-Nin color	No. of Residues	Amino Acid Compositions [b]	Sequence Determination [c]
T_{3Th2}	0.62	N	red	6	$Ser_{2.2}Pro_{1.1}Ala_{1.0}Leu_{0.9}Tyr_{0.8}$	Leu-Ser-Pro-Ser-Ala-Tyr
T_{3Th3}	0.76	N	red	4	$Ser_{2.2}Pro_{1.0}Leu_{1.0}$	Leu-Ser-Pro-Ser
T_{3Th4}	0.93	N	red	2	$Ala_{1.0}Tyr_{1.0}$	Ala-Tyr
T_4	0.75	-0.48	red	8	$Asp_{2.5}Ser_{0.8}Glu_{1.0}Ile_{1.0}Leu_{1.0}$ $Cys(Ae)_{0.5}$	Ile-Leu-Gln-Asp-Asp-Asp-Ser- Cys(Ae) (H)
T_{4E3HI}	1.65	+0.51	brown →red	2	$Ser_{1.0}Cys(Ae)_{0.5}$	Ser-Cys(Ae)
T_5	0.53	-0.31	yel →red	7	$Thr_{0.8}Ser_{1.0}Glu_{1.0}Gly_{1.8}Phe_{0.9}$ $Hse_{0.8}$	Thr-Ser-Gly-Phe-Glu-Gly-Hse
Th_1	0.55	+0.23	red	6	$Ser_{3.3}Val_{1.0}Ile_{1.0}Cys(Ae)_{0.5}$	Val-Ile-Ser-Cys(Ae)-Ser-Ser
Th_2	0.54	-0.60	red	6	$Asp_{2.0}Ser_{0.9}Pro_{1.0}Ile_{1.0}Leu_{1.0}$	Ile-Asp-Ser-Leu-Pro-Asp

Table 6-Continued.

Peptide No.	Mobility[a] pH 2.0	Mobility[a] pH 6.0	Cd-Nin color	No. of Residues	Amino Acid Compositions[b]	Sequence Determination[c]
Th_{2E1}	-	-0.69	red	5		Asp-Ser-Leu-Pro-Asp
Th_{2E2}	-	-0.39	brown →red	4		Ser-Leu-Pro-Asp
Th_3	0.64	N	red	4	$Thr_{1.1}Val_{0.9}Ile_{0.8}Phe_{1.0}$	Ile-Val-Phe-Thr
Th_4	0.48	-0.30	red	7	$Asp_{1.2}Glu_{1.0}Pro_{0.9}Gly_{1.1}$ $Val_{1.0}Ile_{1.0}Tyr_{0.6}$	Ile-Asp-Gly-Val-Gln-Tyr-Pro
Th_{4E2}	-	N	yel	5		Gly-Val-Gln-Tyr-Pro
Th_5	0.76	N	red	4	$Ser_{1.9}Pro_{1.1}Leu_{1.0}$	Leu-Ser-Pro-Ser
Th_6	0.97	N	red	2	$Ala_{1.0}Tyr_{0.9}$	Ala-Tyr
Th_7	0.77	-0.43	red	11	$Asp_{3.0}Thr_{0.8}Ser_{1.7}Glu_{1.0}Gly_{1.2}$ $Ile_{1.0}Leu_{1.0}Cys(Ae)_{0.5}$	Ile-Leu-Gln-Asp-Asp-Asp-Ser- Cys(Ae)-Thr-Ser-Gly
Th_8	0.66	-0.43	red	4	$Glu_{1.0}Gly_{1.2}Phe_{0.9}Hse_{0.8}$	Phe-Glu-Gly-Hse

Table 6-Continued

Peptide No.	Mobility[a] pH 2.0	Mobility[a] pH 6.0	Cd-Nin color	No. of Residues	Amino Acid Compositions[b]	Sequence Determination[c]
Th_{8E2}	-	N	yel	2		Gly-Hse
C_1	0.66	-0.17	red	15	$Asp_{2.0}Ser_{4.3}Pro_{0.9}Val_{1.6}Ile_{2.6}$ $Leu_{1.0}Phe_{1.0}Cys(Ae)_{0.5}$	Val-Ile-Ser-Cys(Ae)-Ser-Ser-Ile- Asp-Ser-Leu-Pro-Asp-Ile-Val-Phe
C_2	0.41	-0.21	yel →red	9	$Asp_{1.1}Thr_{0.8}Glu_{0.9}Pro_{1.2}Gly_{0.8}$ $Val_{1.0}Ile_{1.0}Leu_{1.0}Tyr_{0.7}$	Thr-Ile-Asp-Gly-Val-Gln-Tyr- Pro-Leu
C_3	0.50	N	brown →red	5	$Ser_{1.9}Ala_{1.0}Pro_{1.1}Tyr_{1.0}$	Ser-Pro-Ser-Ala-Tyr
C_4	0.74	-0.36	red	12	$Asp_{2.8}Thr_{0.9}Ser_{1.7}Glu_{1.0}Gly_{1.0}$ $Ile_{1.0}Leu_{1.0}Phe_{0.8}Cys(Ae)_{0.5}$	Ile-Leu-Gln-Asp-Asp-Asp-Ser- Cys(Ae)-Thr-Ser-Gly-Phe
C_5	0.77	-0.56	red	3	$Glu_{1.0}Gly_{1.0}Hse_{0.7}$	Glu-Gly-Hse
Pa_1	0.72	-0.50	brown red	4	$Asp_{1.0}Ser_{2.0}Ile_{1.0}$	Ser-Ser-Ile-Asp

Table 6-Continued.

Peptide No.	Mobility [a] pH 2.0	Mobility [a] pH 6.0	Cd-Nin color	No. of Residues	Amino Acid Compositions [b]	Sequence Determination [c]
Pa_2	0.52	-0.30	red	6	$Asp_{1.0}Pro_{1.1}Val_{0.7}Ile_{0.7}Leu_{1.0}Phe_{0.9}$	Leu-Pro-Asp-Ile-Val-Phe
Pa_3	0.54	-0.36	yel →red	6	$Asp_{1.0}Thr_{0.7}Glu_{0.9}Gly_{0.7}Val_{1.1}Ile_{1.0}$	Thr-Ile-Asp-Gly-Val-Gln
Pa_4	0.75	N	red	3	$Pro_{1.1}Leu_{1.0}Tyr_{0.9}$	Tyr-Pro-Leu
Pa_5	0.55	N	brown →red	7	$Ser_{2.1}Pro_{1.9}Ala_{1.0}Leu_{1.0}Tyr_{0.6}$	Tyr-Pro-Leu-Ser-Pro-Ser-Ala
Pa_6	0.61	N	red	4	$Glu_{1.0}Ile_{0.9}Leu_{0.9}Tyr_{0.7}$	Tyr-Ile-Leu-Gln
Pa_7	0.90	-0.66	red	7	$Asp_{3.0}Thr_{0.7}Ser_{2.0}Cys(Ae)_{0.5}$	Asp-Asp-Asp-Ser-Cys(Ae)-Thr-Ser
Pa_8	0.40	-0.60	red	11	$Asp_{2.6}Thr_{1.0}Ser_{1.7}Glu_{1.0}Ile_{0.7}Leu_{0.7}Tyr_{0.6}Cys(Ae)_{0.5}$	Tyr-Ile-Leu-Gln-Asp-Asp-Asp-Ser-Cys(Ae)-Thr-Ser
Pa_9	0.66	-0.48	red	4	$Glu_{1.0}Gly_{1.0}Phe_{1.0}Hse_{0.9}$	Phe-Glu-Gly-Hse

a. Electrophoretic mobilities are relative to serine (=1.0) at pH 2.0 and to aspartic acid at pH 6.0. (N)—

(Continued on next page)

Table 6 - Continued - Legend

indicates the peptides which are considered neutral in electrophoresis at pH 6.0.

b. The amino acid analyses are molar ratios. Serine was corrected for 15% of loss during the hydrolysis. Hse and Cys(Ae) are the abbreviations for homoserine and aminoethylcysteine, respectively.

c. Sequences determined by the dansyl-Edman procedure are indicated as arrows toward the right(→) under the residues. Amino acids released by carboxypeptidase A are indicated as arrows toward the left (←) under the residues. The residues determined by hydrazinolysis are indicated by arrows containing the letter H ($\overset{H}{\leftarrow}$).

from the NH_2-terminus were Ser-Ser-Ile-Asp-. From a short time carboxypeptidase A digestion, 3 residues from COOH-terminus were Ile-Val-Phe. The sequence of the other 4 residues was not determined in this peptide but it was revealed by a thermolysin peptide (see thermolysin digest below).

Sequence of peptide T_3: Peptide T_3 contained 14 residues. From dansyl-Edman procedures, the first 3 residues from the NH_2-terminus were found to be Thr-Ile-Asp. Carboxypeptidase A digestion revealed the 2 residues from the COOH-terminus to be Ala-Tyr. A digest of peptide T_3 with thermolysin and further purification by paper electrophoresis at pH 2.0, gave peptides T_{3Th1}, T_{3Th2}, T_{3Th3}, and T_{3Th4}. T_{3Th1}, which contained 8 residues as shown in Table 6, was further digested with subtilisin. After purification with paper electrophoresis at pH 2.0, two peptides, T_{3Th1S1} and T_{3Th1S2} were obtained. The amino acid sequences of the peptides were determined to be T_{3Th1S1}: Thr-Ile-Asp-Gly-Val and T_{3Th1S2}: Gln-Tyr-Pro. Peptide T_{3Th1S1} had one negative charge as demonstrated by paper electrophoresis at pH 6.0; thus the acidic form was assigned to Asp. T_{3Th1S2} was neutral by pH 6.0 paper electrophoresis, thus amide was assigned to Gln. The sequence of T_{3Th1}, based on the above information, was established to be Thr-Ile-Asp-Gly-Val-Gln-Tyr-Pro. By the combination of dansyl-Edman and carboxypeptidase A digestion, sequences of peptide T_{3Th2}, T_{3Th3} and T_{3Th4} were determined as listed in Table 6. From this information, the sequence of T_3 was thus established to be Thr-Ile-Asp-

Gly-Val-Gln-Tyr-Pro-Leu-Ser-Pro-Ser-Ala-Tyr.

Sequence of peptide T_4: The sequence of peptide T_4 was revealed from dansyl-Edman analysis and hydrazinolysis to be Ile-Leu-Gln-Asp-Asp(Asp, Ser)-Cys(Ae). Direct evidence of glutamine in this sequences was provided in a peptide from a papain digest which will be described below. The mobility of T_4 on high-voltage electrophoresis at pH 6.0, indicated that this peptide contained 2 negative charges. This meant all the three aspartyl residues were in the acidic forms. This finding was further supported by the amino acid analysis of a digest of T_4 with aminopeptidase M which showed the aspartyl residues to be in the acidic forms. After removal of the first three residues from the NH_2-terminus of T_4 by Edman degradation, partial acid hydrolysis was performed (5.7 N HCl, 100°, 15 min.). A dipeptide, Ser-Cys(Ae) was obtained. The aminoethylcysteine was identified as dansyl derivative without HCl hydrolysis, after the removal of serine from this dipeptide. Thus the sequence of T_4 was assigned showed in Table 6.

Sequence of peptide T_5: Peptide T_5 contained 7 residues, apparently derived from the COOH-terminus of CB5. The sequence of T_5 was determined to be Thr-Ser-Gly-Phe-Glu-Gly-Hse. At pH 6.0, paper electrophoresis indicated T_5 had one negative charge; thus the acidic form was assigned to Glu.

b) Thermolysin digest:

Sequence of peptide Th_1: Peptide Th_1 contained 6 residues. From dansyl-Edman procedures and carboxypeptidase A digestion, this peptide was

determined to be Val-Ile-Ser-Cys(Ae)-Ser-Ser.

Sequence of peptide Th_2: Peptide Th_2 contained 6 residues. The sequence was shown to be Ile-Asp-Ser-Leu-Pro-Asp. Th_2 showed 2 negative charges by paper electrophoresis at pH 6.0. After removal of the second residue, Asp, only one negative charge remained as was shown in Table 6 (Th_{2EI}). Thus both aspartyl residues were assigned to be in the acidic forms.

Sequence of peptide Th_3: Peptide Th_3 is a tetrapeptide with the sequence of Ile-Val-Phe-Thr.

Sequence of peptide Th_4: By using dansyl-Edman procedure, the sequence of peptide Th_4 was established to be Ile-Asp-Gly-Val-Gln-Tyr-Pro. It contained one negative charge as shown by paper electrophoresis at pH 6.0. After two steps of Edman degradation, this peptide became neutral in mobility at pH 6.0 electrophoresis. Therefore acid form was assigned to the aspartyl residue, while the amide form was assigned to the glutamine.

Sequence of peptide Th_5: This is a tetrapeptide with the sequence of Leu-Ser-Pro-Ser.

Sequence of peptide Th_6: Th_6 is a dipeptide Ala-Tyr.

Sequence of peptide Th_7: This peptide contained 10 residues. The partial sequence was determined to be Ile-Leu-Gln-Asp-(Asp, Asp, Ser, Cys(Ae))-Thr-Ser-Gly. The identity of three residues from COOH-terminus were revealed by carboxypeptidase A digestion at pH 5.7 (45). However, it was clear that this peptide is related to peptide T_4 whose sequence was

known. Therefore, the determination of complete sequence of peptide Th_7 was not pursued.

c) α-Chymotryptic digest:

Sequence of peptide C_1: C_1 contained 15 residues. From dansyl-Edman analysis and carboxypeptidase A digestion, it was obvious that this peptide was derived from NH_2-terminus of CB5 and would provide the overlap for T_1 and T_2. Therefore, only the partial sequence of this peptide was established as shown in Table 6.

Sequence of peptide C_2: Peptide C_2 contained 9 residues. The partial sequence of this peptide (Table 6) revealed that it is related to peptide T_3 which had been sequenced.

Sequence of peptide C_3: The sequence of peptide C_3 was found to be Ser-Pro-Ser-Ala-Tyr.

Sequence of peptide C_4: From partial sequence determination (Table 6), this peptide was obviously an extention of that of Th_7. The complete amino acid sequence was therefore not determined.

Sequence of peptide C_5: This is a tripeptide with the sequence of Glu-Gly-Hse.

d) Papain digest:

Sequence of peptide Pa_1: This is a tetrapeptide with the sequence of Ser-Ser-Ile-Asp.

Sequence of peptide Pa_2: By using the dansyl-Edman procedure, the sequence of peptide Pa_2 was established to be Leu-Pro-Asp-Ile-Val-Phe.

known. Therefore, the determination of complete sequence of peptide Th_7 was not pursued.

c) α-Chymotryptic digest:

Sequence of peptide C_1: C_1 contained 15 residues. From dansyl-Edman analysis and carboxypeptidase A digestion, it was obvious that this peptide was derived from NH_2-terminus of CB5 and would provide the overlap for T_1 and T_2. Therefore, only the partial sequence of this peptide was established as shown in Table 6.

Sequence of peptide C_2: Peptide C_2 contained 9 residues. The partial sequence of this peptide (Table 6) revealed that it is related to peptide T_3 which had been sequenced.

Sequence of peptide C_3: The sequence of peptide C_3 was found to be Ser-Pro-Ser-Ala-Tyr.

Sequence of peptide C_4: From partial sequence determination (Table 6), this peptide was obviously an extention of that of Th_7. The complete amino acid sequence was therefore not determined.

Sequence of peptide C_5: This is a tripeptide with the sequence of Glu-Gly-Hse.

d) Papain digest:

Sequence of peptide Pa_1: This is a tetrapeptide with the sequence of Ser-Ser-Ile-Asp.

Sequence of peptide Pa_2: By using the dansyl-Edman procedure, the sequence of peptide Pa_2 was established to be Leu-Pro-Asp-Ile-Val-Phe.

Sequence of peptide Pa_3: From dansyl-Edman procedures and carboxypeptidase A digestion, this peptide was determined to be Thr-Ile-Asp-Gly-Val-Gln.

Sequence of peptide Pa_4: This is a tripeptide with the sequence of Tyr-Pro-Leu.

Sequence of peptide Pa_5: This sequence is obviously a part of sequence of T_3. Its partial sequence was determined as shown in Table 6.

Sequence of peptide Pa_6: This is a tetrapeptide with the sequence of Tyr-Ile-Leu-Gln. It remained at the origin during electrophoresis at pH 6.0; thus the amide form was assigned to Gln.

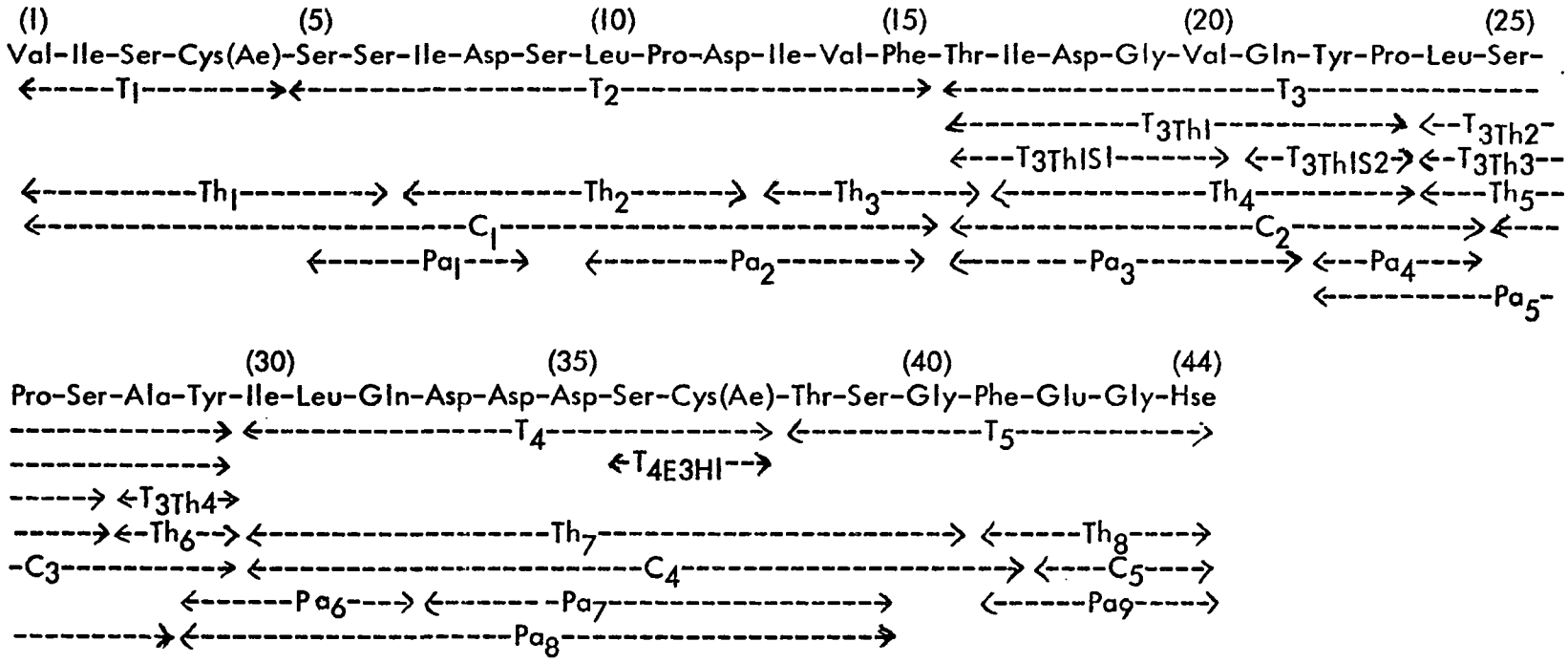
Sequence of peptide Pa_7: This peptide contained 7 residues. It was a part of the sequence of peptide Th_7. It contained two negative charges at pH 6.0 electrophoresis, thus acid form for all three aspartyl residues was assigned.

Sequence of peptide Pa_8: This peptide contained 10 residues. It provided the overlap for Pa_6 and Pa_7. Therefore, only the partial sequence of this peptide was established as shown in Table 6.

Sequence of peptide Pa_9: This peptide is exactly the same as peptide Th_8. Its sequence was determined to be Phe-Glu-Gly-Hse.

Amino Acid Sequence of CB5

From the amino acid sequence of the peptides described above, the complete sequence of 44 residues was assembled as shown in Fig. 19.



Keys of digests: T: Tryptic digest, Th: Thermolysin digest, C:α-Chymotryptic digest, Pa: Papain digest,

S: Subtilisin digest, H: Partial acid hydrolysis.

Fig. 19 - Amino Acid Sequence of CB5

CHAPTER IV

DISCUSSION

EPNP reacts with pepsin at room temperature (or 30°) at a methionyl residue and an aspartyl residue in the following sequences:

Sequence I: Phe-Glu-Gly-Met-Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu-Leu

Sequence II: Ile-Val-Asp-Thr-Gly-Ser-Ser-Asn

The modification of methionine is due to alkylation by EPNP while the modification of aspartic acid is accomplished by the esterification of a β-carboxyl group. In acidic solutions, the protonated intermediate of an epoxide are well known to react with various nucleophiles (46). The thioether of methionine and the β-carboxylate of aspartic acid apparently serve as nucleophiles near the EPNP binding sites of the enzyme, and thus are reactive with EPNP.

The methionine sequence given above agrees with the known sequence of a methionine peptide: Met-Asp-Val-Pro-Thr-Ser (5). It also agrees with the NH_2-terminal sequence of a cyanogen bromide fragment (17) which presumably was derived from the same methionyl sequence. Therefore,

the modified methionine is located 38 residues from the carboxyl terminus of the enzyme (17). However, this current sequence does not agree with the overlap sequence of this methionine (5). The left side of this methionine sequence was previously thought to be Tyr-Gly-Tyr-Gly-Ser-Met. However, the correspoinding sequence found in this study is Phe-Glu-Gly-Met. The reason for the discrepancy is not entirely clear. A critical evaluation of data from the previous work of Tang and Hartley (5) showed that the overlap was based on one peptide, the sequence of which was not completely established. Thus, their evidence was relatively weaker than that of the present study in which the sequences of several methionine peptides were fully determined. Therefore, we suggest that the methionine overlap shown in Sequence I is the correct one.

Although an overlapping peptide for EPNP-aspartyl residue was not found, evidence for the overlap appears sufficient from quantitative estimation of recoveries. Two peptides with the sequence Ile-Val-Asp(EPNP) and Asp (EPNP)-Thr-Gly-Ser-Ser-Asn were obtained separately in yields of 0.55 and 0.88 residues from 1 residue of EPNP-aspartic acid in pepsin molecule. Also, in the various digest of EPNP-pepsin, an EPNP-aspartyl residue was found either at the NH_2-terminus or the COOH-terminus of the peptides. In no case, were these two types of peptides found in the same digest, as one would expect if two EPNP-aspartyl sites were present in modified pepsin. A search of the literature indicated that this EPNP-aspartyl sequence has not been previously reported. Therefore, the location of this site in the primary structure

of the enzyme remains uncertain.

It is interesting that a striking homology exists between the EPNP-reactive aspartyl sequence (Sequence II) and the previously known active-center aspartyl which is reactive to diazo inactivators (25,28).

The EPNP-reactive aspartyl sequence is Ile-Val-Asp-Thr-Gly-Ser-Ser-Asn.

The diazo-reactive aspartyl sequence is Ile-Val-Asp-Thr-Gly-Thr-Ser-Leu. Since only 2 among 8 residues differed, we evaluated the possibility that they might indeed be the same sequence. This seemed unlikely, however, as discussed below:

The Possibility of Differences due to a Sequencing Mistake

We obtained peptide P_2 (sequence: Asp(EPNP)-Thr-Gly-Ser-Ser-Asn) from 20 different preparations, and its composition was always the same. In no case were 2 threonine residues or 1 leucine residue found. Previous data on the diazo-reactive aspartyl sequence studies also appeared to be convincing. Bayliss et al. (25) isolated a tripeptide Thr-Gly-Thr from partial acid hydrolysis of a diazo-reactive peptide, which agreed with the content of 2 threonine residues in the diazo-reactive sequence. The position of leucine was reported by Fry et al. (28) in an octapeptide containing 1 residue of leucine. After one-step Edman degradation, the content of leucine remained essentially unchanged. Carboxypeptidase A released 1 residue of leucine from the octapeptide. Thus, a sequencing error appeared improbable.

The Possibility of Differences due to Different Pepsins

Previous studies have shown that diazoacetyl-DL-norleucine methyl ester and EPNP completely inactivated the same pepsin preparation, resulting in a stoichiometry of modification of 1 in the former case and 2 in the latter (6). This excludes the possibility that the two types of inactivators each react specifically with one type of pepsin.

The Possibility of Differences due to Genetic Variants of Pepsin

We have used 4 different batches of pepsin from Worthington and 2 batches from Pentex. In no case, was an EPNP-peptide with a diazo-reactive aspartyl sequence found. Moreover, the diazo-reactive sequence reported by Bayliss et al. (25) and Fry et al., (28) who used pepsin from different sources (Koch-Light Labs. Ltd., Colnbrook, Buckinghamshire, and Mann Research Lab., N. Y.), agreed in their results. Additionally, a large number of laboratories have engaged in determination of the partial sequence of pepsin. Keil (10) reviewed most of these results; he reported that about 15 separate studies were done on the sequence near the carboxyl-terminal region. A number of separate studies also performed on the phosphoserinyl and cystyl sequences. To the best of our knowledge, no genetic variants in porcine pepsin have been reported. The involvement of the completely different genetic varients appeared to be unlikely.

On the other hand, a previous study (6) indicated that EPNP-modified pepsin still reacts with 1 residue of diazoacetyl-DL-norleucine methyl ester. The

reversed order of reaction also produced the same stoichiometry of incorporation of the two inactivators. These results suggested that pepsin contains two different aspartyl sites which each react specifically with either EPNP or diazo inactivators. In a separate study (29), it was found that the pKa of the EPNP-reactive carboxyl group (2.85) was distinctly lower than that of the diazo-reactive carboxyl group (4.7). These findings support the conclusion that the differences in the two sequences are genuine and that both are present in pepsin.

The EPNP-reactive aspartyl sequence apparently differs from that of the p-bromophenacyl bromide reactive site found by Erlanger et al. (26), which is in a peptide composed of Gly_2, Asp_1, Ser_1, Glu_1. This confirms the previous findings which showed that all three inactivators react at different sites in the pepsin molecule.

The modification of the β-carboxyl group of the aspartyl residue apparently produces an ester linkage to one of the hydroxy groups of EPNP-OH, as ascertained from the following evidence: a) the lability of EPNP-aspartyl peptides in basic solution; b) the removal of EPNP from EPNP-aspartyl peptides, resulting in a gain of a negative charge; and c) the specific reduction of the β-carboxyl ester of the aspartyl residue to homoserine.

A minor sequence Val-Ile-Phe-Asp(EPNP) resulted from modification at higher temperatures (20°-30°). The absence of this peptide in the EPNP-pepsin modified at 10° confirms data from a separate study (29); i.e., EPNP modification is more specific at a lower reaction temperature.

After careful examination of the electrophoretic mobility data in Table 2, an unusual pKa of the EPNP reactive aspartyl residue was found. It is known that the charge and molecular weight of peptides can be related to their electrophoretic mobilities. Based on Stoke's law, Offord demonstrated the interrelationships of relative mobility, charge, and mass of peptides during electrophoresis (31): a) Relative mobility of a peptide is proportional to the charge which the peptide contains when the mass of the peptide is constant; b) The logarithm of the relative mobility is inversely proportional to the logarithm of the mass of the peptide when the charge which the peptide contains is constant. He collected a number of peptides and plotted the logarithm of the relative eletrophoretic mobility at pH 2.0 and pH 6.5 against the logarithm of the molecular weight. Straight lines were obtained and each represented a given number of charges. The pKa of the β-carboxyl group in proteins ranges from 3.0 to 4.7 (47). If the pKa's of the carboxyl groups in peptide P_2, after the removal of the EPNP group, (peptide P_2-EPNP), were normal in pH 2.0, this peptide should contain only one positive charge from the NH_2-terminal amino group. With a molecular weight of 579 it should have a relative mobility to serine equal to 0.71. However, the actual observed value was 0.36. This means that partial ionization of the carboxyl groups of P_2-EPNP occurred during the electrophoresis at pH 2.0. Based on Offord's first generalization, the percentage of ionization of carboxyl groups in peptide P_2-EPNP can be calculated to give a value of 50% ionization of a carboxyl group. In order to determine which carboxyl group had an abnor-

mally low pKa, Edman degradation procedure was performed to remove the NH_2-terminal Asp and the mobility in pH 2.0 was again determined. The relative mobility was found to be 0.67, while the theoretical value was 0.80. The percentage of ionization of the COOH-terminal carboxyl group, now the only carboxyl group remaining is 16% based on the above calculation. If the removal of the first NH_2-terminal aspartyl residue does not affect the ionization of the COOH-terminal asparaginyl residue, then the percentage of ionization of the β-carboxyl group of the NH_2-terminal aspartyl residue is 35%. Using this information and the Henderson-Hasselbalch equation, a pKa of 2.27 was found for the β-carboxyl group of the aspartyl residue in peptide P_2-EPNP. In order to know whether the NH_2-terminal amino group induced the lowering of the pKa on the β-carboxyl of the NH_2-terminal aspartyl residue, P_2-EPNP was dansylated, then subjected to electrophoresis at pH 2.0 on a flat plate apparatus to determine the mobility. This information was calculated in the same manner as described above and a value of 50% ionization of carboxyl groups was also found. The inducing effect of the amino group appeared to be unlikely. Since this epoxide reactive aspartyl residue has to be charged in the enzyme molecule when it reacts with EPNP, a low pKa is expected for this residue. It seems possible that a low pKa of this β-carboxyl group might have been a unique result of the sequence of this peptide, which may produce a specific interaction of side chains in the peptide. For example, hydroxyl groups of threonine or serines might form a hydrogen bond to the β-carboxyl group, thus lowering the pKa of the aspartyl residue. It is not clear, however that the low pKa of the aspartyl residue

of this peptide has the same cause as the low pKa of this group on the native enzyme, which has been determined to be 2.85 (29). More information is needed to reach a firm conclusion on this.

The active site of pepsin was suggested to contain two functional groups, one of which has a slightly high pKa of about 4.7 and another one has an abnormal pKa of 1.0 (20,21). A mechanism for pepsin action was proposed by Knowles (22) in which a tetrahedral intermediate was formed by an acid-nucleophilic catalysis, (the carboxyl group with a pKa of 1.0, serving as the nucleophile, that with a pKa of 4.7, as the acid). Amino-enzyme was postulated to be formed from this tetrahedral intermediate. Using this hypothetical amino-enzyme, the hydrolysis and transpeptidation, which pepsin catalyzes, can be explained. In the recent epoxide inactivator studies conducted by Hartsuck and Tang (29), the EPNP-reactive aspartyl residue, in the Sequence II of this current work, was shown to be in the active site of pepsin, and its β-carboxyl group had a low pKa which would be at least partially charged during peptic digestion. They proposed a new mechanism for peptic catalysis. In this proposal, the epoxide-reactive aspartic acid in the active site was considered to be a nucleophile, while the diazo-reactive aspartic acid served as a proton donor to the amide nitrogen of the substrate. They also postulated that the lower pKa of the epoxide reactive carboxyl could be due to the hydrogen bonding of another near-by, protonated, carboxyl group. The involvement of an arginyl side chain, in polarization of the carbonyl group of the substrate, was also suggested. This was further

supported by the recent finding that the biacetyl modification of arginyl residue, twelfth residue from the COOH-terminus, caused a loss of 85% of peptic activity. In addition, this arginine was thought to be located near the active center of the enzyme (48). Formation of an amino-enzyme intermediate would not be expected from this mechanism. Transpeptidation catalyzed by pepsin can be explained in the following way. There is a strong interaction with the side chain on the imino side of the scissile bond, so that this product remains in the active center, and is available for subsequent transfer. Although the final mechanism of catalysis of pepsin must wait for other information (i.e., X-ray crystallographic studies), the involvement of the aspartyl side chain in Sequence II, determined in this study, appeared to be certain (6).

The Sequence I near the EPNP-alkylated methionine, actually provided the overlap for CB5 and the carboxyl terminal sequence of 38 residues previously completed (17). Thus the sequence of CB5 is located in residue numbers 241 to 284 of porcine pepsin. This gives a combined amino acid sequence of 81 residues (residue number 241 to 321) in the COOH-terminal end of porcine pepsin, as shown in Fig. 20. It is obvious that neither of the active center sequences is present in this portion of the molecule. The role of this structure in pepsin function must wait for further studies.

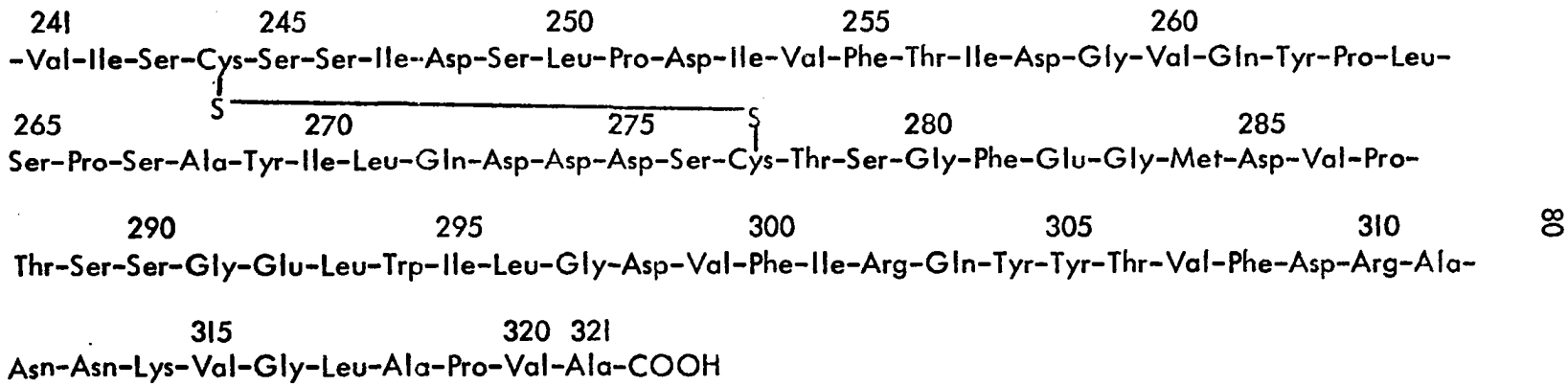


Fig. 20 - Carboxyl Terminal Sequence of Porcine Pepsin (Residue 241 - 321)

CHAPTER V

SUMMARY

The amino acid sequences around two sites reactive with the specific pepsin inactivator, 1,2-epoxy-3-(p-nitrophenoxy) propane (EPNP), were determined to be:

Sequence I: Phe-Glu-Gly-Met(EPNP)-Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu-Leu.

Sequence II: Ile-Val-Asp(EPNP)-Thr-Gly-Ser-Ser-Asn.

EPNP-pepsin was digested separately by proteases and EPNP-peptides were purified by Sephadex column chromatography and high-voltage electrophoresis. When the sequences of the peptides were determined, they fell into one of the two groups shown above--Sequence I or Sequence II. The modification of methionine in Sequence I represented an alkylation by EPNP to form a sulfonium salt. This methionine residue is located 38 residues from the carboxyl terminus of pepsin. The modification of the β-carboxyl group of the aspartyl residue in Sequence II was an esterification reaction. The location of this aspartyl residue in the primary structure of pepsin is uncertain. Extensive homology exists between EPNP-aspartyl Sequence II and the previously reported

diazo-reactive aspartyl sequence; however we believe them to be unique. The sequence confirmed data from a previous study (Tang, J. 1971 J. Biol. Chem. 246:4510) which indicated that EPNP reacts with two main sites of pepsin and that these sites differ from those reactive with diazo and p-bromophenacyl bromide inactivators of pepsin.

Porcine pepsin was reduced, S-aminoethylated, and then subjected to cyanogen bromide cleavage. Reaction mixture was separated on a column of Sephadex-G-75. Five peaks were obtained. One of the cyanogen bromide fragments, CB5 was digested separately by proteases and the resulting peptides were purified by Sephadex column chromatography and high voltage electrophoresis. The sequence of CB5 was determined to be: Val-Ile-Ser-Cys-Ser-Ser-Ile-Asp-Ser-Leu-Pro-Asp-Ile-Val-Phe-Thr-Ile-Asp-Gly-Val-Gln-Tyr-Pro-Leu-Ser-Pro-Ser-Ala-Tyr-Ile-Leu-Gln-Asp-Asp-Asp-Ser-Cys-Thr-Ser-Gly-Phe-Glu-Gly-Met. The carboxyl-terminal methionine of this fragment is the 38th residue from the carboxyl-terminus of porcine pepsin.

BIBLIOGRAPHY

1. Rajagopalan, T. G., Moore, S. and Stein, W. H. 1966 Pepsin from Pepsinogen, J. Biol. Chem. 241:4940.

2. Tang, J., Wolf, S., Caputto, R. and Trucco, R. E. 1959 Isolation and Crystallization of Gastricsin from Human Gastric Juice, J. Biol. Chem. 234:1174.

3. Foltmann, B. 1966 A Review on Prorennin and Rennin, Comptes Rendus Des Travaux Du Laboratoire Carlsberg. 35:192.

4. Huang, W. Y. and Tang, J. 1970 Carboxyl-terminal Sequence of Human Gastricsin and Pepsin, J. Biol. Chem. 245:2189.

5. Tang, J. and Hartley B. S. 1970 Amino Acid Sequence Around the Disulphide Bridges and Methionine Residues of Porcine Pepsin, Biochem. J. 118:611.

6. Tang, J. 1971 Specific and Irreversible Inactivation of Pepsin by Substrate-like Epoxides, J. Biol. Chem. 246:4510.

7. Arnon, R. and Perlmann, G. E. 1963 The Amino Acid Composition of Chromatographically Purified Pepsinogen, J. Biol. Chem. 238:653.

8. Ong, E. B. and Perlmann, G. E. 1968 The Amino-terminal Sequence of Porcine Pepsinogen, J. Biol. Chem. 243:6104.

9. Stepanov, V. M., Timokhina, E. A., Baratova, L. A., Belyanova, L. P., Korzhenko, V. P. and Zhukova, I. G. 1971 On the Activation of Pepsinogen, Biochem. Biophys. Res. Comm. 45:1482.

10. Keil, B. 1970 Acid Proteinases (Chemical Studies of Porcine Pepsin), Phil. Trans. Roy. Soc. Lond. B257:125.

11. Dopheide, T. A. A. and Jones, W. M. 1968 Studies on the Tryptophan Residues in Porcine Pepsin, J. Biol. Chem. 243:3906.

12. Kostka, V., Moravek, L., Kluh, I. and Keil, B. 1969 Studies on S-sulfo-pepsin and its Cyanogen Bromide Hydrolysate, Biochim. Biophys. Acta. 175:459.

13. Blumenfeld, O. O. and Perlmann, G. E. 1959 The Amino Acid Composition of Crystalline Pepsin, J. Gen. Physiol. 42:553.

14. Flavin, M. 1954 Linkage of Phosphoserine to Protein in Pepsin and Ovalbumin, J. Biol. Chem. 210:771.

15. Stepanov, V. M., Vachitova, E. A., Jegorov, C. A. and Avajeva, S. M. 1965 Sequence of Amino Acids Surrounding the Phosphoserine Residue in Hog Pepsin, Biochim. Biophys. Acta. 110:632.

16. Tang, J. 1970 Amino Acid Sequence Near the Amino Terminus of Porcine Pepsin, Biochem. Biophys. Res. Comm. 41:697.

17. Kostka, V., Moravek, L. and Sorm, F. 1970 Amino Acid Sequence of C-terminal Fragment of Hog Pepsin, Eur. J. Biochem. 13:447.

18. Dopheide, T. A. A., Moore, S. and Stein, W. H. 1967 The Carboxyl-terminal Sequence of Porcine Pepsin, J. Biol. Chem. 242:1833.

19. Raftery, M. A. and David Cole, R. 1963 Tryptic Cleavage at Cysteinyl Peptide Bonds, Biochem. Biophys. Res. Comm. 10:467.

20. Jackson, W. T., Schlamowitz, M., Shaw, A. and Trujillo, R. 1969 The Effect of pH on the Kinetic Constants of Peptic Substrates and Inhibitors, Arch. Biochem. Biophys. 131:374.

21. Cornish-Bowden, A. J. and Knowles, J. R. 1969 The pH-Dependence of Pepsin-Catalysed Reactions, Biochem. J. 113:353.

22. Knowles, J. R. 1970 On the Mechanism of Action of Pepsin, Phil. Trans. Roy. Soc. Lond. B257:135.

23. Delpierre, G. R. and Fruton, J. S. 1965 Inactivation of Pepsin by Diphenyl Diazomethane, Proc. Nat. Acad. Sci. U.S.A. 54:1161.

24. Rajagopalan, T. G., Moore, S. and Stein, W. H. 1966 The Inactivation of Pepsin by Diazoacetylnorleucine Methyl Ester, J. Biol. Chem. 241:4295.

25. Bayliss, R. S., Knowles, J. R. and Wybrandt, G. B. 1969 Aspartic Acid Residue at the Active Site of Pepsin. The Isolation and Sequence of the Heptapeptide, Biochem. J. 113:377.

26. Erlanger, B. F., Vratsanos, S. M., Wassermann, N. and Cooper, A. G. 1966 A Chemical Investigation of the Active Center of Pepsin, Biochem. Biophys. Res. Comm. 23:243.

27. Chen, Kirk C. S. and Tang, J. 1972 Amino Acid Sequence around the Epoxide-reactive Residues in Pepsin, J. Biol. Chem. 247:2566.

28. Fry, K. T., Kim, O. K., Spona, J. and Harmilton, G. A. 1970 Site of Reaction of a Specific Diazo Inactivator of Pepsin, Biochemistry 9:4624.

29. Hartsuck, J. A. and Tang, J. 1972 The Carboxylate Ion in the Active Center of Pepsin, J. Biol. Chem. 247:2575.

30. Heilmann, J., Barollire, J. and Watzke, E. 1957 Determination of Amino Acids on Paper Chromatogram, Z. Physiol. Chem. 309:219.

31. Offord, R. E. 1966 Electrophoretic Mobilities of Peptides on Paper and Their Use in The Determination of Amide Groups, Nature 211:591.

32. Spackman, D. H. 1963 Accelerated System for the Automatic Analysis of Amino Acids, Fed. Proc. 22:244.

33. Gray, W. R. and Hartley, B. S. 1963 Fluorescent End-group Reagent for Proteins and Peptides, Biochem. J. 89:59P.

34. Woods, K. R. and Wang, K. P. 1967 Separation of Dansyl-amino Acids by Polyamide Layer Chromatography, Biochem. Biophys. Acta. 133:369.

35. Ambler, R. P. 1967 in C. H. W. Hirts (Editor). Methods in Enzymology, Vol. XI, P. 155, Academic Press, New York.

36. Fraenkel-Conrat, H. and Tsung, C. M. 1967 in C. H. W. Hirs (Editor). Methods in Enzymology, Vol. XI, P. 151, Academic Press, New York.

37. Gray, W. R. 1967 in C. H. W. Hirs (Editor). Methods in Enzymology, Vol. XI, P. 469, Academic Press, New York.

38. Wilcox, P. E. 1967 in C. H. W. Hirs (Editor). Methods in Enzymology, Vol. XI, P. 68, Academic Press, New York.

39. Tang, J. and Hartley, B. S. 1967 A diagonal Electrophoretic Method for Selective Purification of Methionine Peptides, Biochem. J. 102:593.

40. Gross, E. and Witkop, B. 1961 Selective Cleavage of the Methionyl Peptide Bonds in Ribonuclease with Cyanogen Bromide, J. Amer. Chem. Soc. 83:1510.

41. Lawson, W. B., Gross, E., Foltz, C. M. and Witkop, B. 1962 Alkylation and Cleavage of Methionine Peptides, J. Amer. Chem. Soc. 84:1715.

42. Slobin, L. I. and Singer, S. J. 1968 The specific Cleavage of Immunoglobulin Polypeptide Chains at Cysteinyl Residues, J. Biol. Chem. 243:1777.

43. Raftery, M. A. and Cole, R. D. 1966 On the Amino-ethylation of Porteins, J. Biol. Chem. 241:3457.

44. Steers, E. Jr., Craven, G. R., Anfinsen, C. B. and Bethune, J. L. 1965 Evidence for Nonidentical Chains in the β-galactosidase of Escherichia Coli K12, J. Biol. Chem. 240:2478.

45. Titani, K., Narita K. and Okunuki, K. 1962 N-terminal Sequence in Beef- and Horse-Heart Cytochrome c, J. Biochem. 51:350.

46. Parker, R. E. and Isaacs, N. S. 1959 Mechanisms of Epoxide Reactions, Chem. Rev. 59:737.

47. Cohn, E. J. and Edsall, J. T. 1942 Proteins, Amino Acid, and Peptides as Ions and Dipolar Ions (Reinhold Publishing Corporation, New York).

48. Huang, W. Y. and Tang, J. 1972 Modification of an Arginyl Residue in Pepsin by 2.3-Butanedione, J. Biol. Chem. 247:2704.

49. Fruton, J. S. 1971 in P. D. Boyer (Editor). The Enzymes Vol. III, P. 152, Academic Press, New York.