THE UNIVERSITY OF MANITOBA

THE AMINO ACID SEQUENCE OF LIMA BEAN INHIBITOR FRACTION IV

by

Celine Gaik Lin Tan

B.Sc.(Hons) University of Singapore 1966
M. Sc. University of Manitoba 1968

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOCHEMISTRY

WINNIPEG, MANITOBA

FALL, 1970

To Mum

and to Christopher

ACKNOWLEDGEMENTS

I am most grateful to my thesis supervisor, Dr. Frits C. Stevens for his guidance and helpful criticisms during the course of this work and also in the preparation of the manuscript. His encouragement, patience and great enthusiasm was a constant source of inspiration. I also wish to express my gratitude to Dr. M. Blanchaer, Chairman of the Department of Biochemistry.

I acknowledge the technical assistance of Miss Caroline Taylor and Mr. Norbert Grift in performing the amino acid analyses.

I thank the Medical Research Council of Canada for their award of a studentship from 1968-1970.

I also wish to thank Miss Maureen Scanlan and Miss Sandra Kotak for typing the manuscript, and Mrs. Sylvia Jirousek for preparing the figures.

ABSTRACT

Commercial lima bean inhibitor (LBI) was separated into four apparently homogenous fractions. All four fractions were equally active and 1 mg of LBI inhibited ~ 2.5 mg of trypsin and ~ 1.5 mg of chymotrypsin. As judged from their amino acid composition and from peptide maps obtained by trypsin hydrolysis of the reduced and alkylated inhibitors, they are structurally related but not identical.

As a first step in the elucidation of structure-function relationships in LBI it was decided to determine the amino acid sequence of LBI Fraction IV (LBI IV). From a study of the overlapping peptides obtained by tryptic and chymotryptic digests of the reduced and alkylated protein and also a tryptic digest of the reduced, alkylated and guanidinated protein, the entire sequence of 84 amino acid residues could be derived. The peptides obtained by enzymic hydrolysis were purified using combinations of ion exchange chromatography, gel filtration, paper chromatography and paper electrophoresis. Their sequences were obtained by classical methods, using exopeptidases and the Edman sequential degradation procedure; large peptides were first fragmented into smaller overlapping peptides by endopeptidases of wide specificity.

During the course of the investigation it became apparent that the LBI IV preparation used consisted of at least two variant forms. The amino acid residues in positions 37 and 39 respectively are either threonine and leucine or serine

and phenylalanine. Position 26 can be occupied by either serine or alanine. The origin of this heterogeneity is still subject to speculation.

Examination of the complete amino acid sequence of LBI IV revealed the occurrence of a repetitive sequence. When the sequences of residues 23 through 34 and 50 through 61 were compared it could be seen that these two regions were identical in 9 out of 12 positions. Of added interest is the fact that the anti-chymotrypsin site of LBI (Krahn and Stevens, 1970) is located in one of these regions.

TABLE OF CONTENTS

ABBREVIATIONS

APM	Aminopeptidase M
BPTI	Basic Pancreatic trypsin inhibitor
BTEE	Benzoyl tyrosine ethyl ester
Co A	Carboxypeptidase A
Co B	Carboxypeptidase B
DEAE	Diethyl amino ethyl
EDTA	Ethylene diamine tetraacetic acid
LAP	Leucine aminopeptidase
LBI	Lima bean inhibitor
RCM	Reduced carboxymethylated
RCAM	Reduced carboxamidomethylated
GRCM	Reduced, carboxymethylated, guanidinated
PTH	Phenylthiohydantoin
TAMe	p-toluene sulfonyl arginine methyl ester
SDS	Sodium dodecyl sulfate
TPCK	L-(1-Tosylamido-2-phenyl) ethyl chloromethyl ketone

NOMENCLATURE OF THE PEPTIDES

Roman numerals refer to the fractions separated by column chromatography. Arabic numerals refer to the purified peptides. Purified peptides are designated to indicate the agent responsible for their formation: T, tryptic peptides; C, chymotryptic peptides; Tg, tryptic peptides from GRCM-LBI IV; P, pronase peptides; PA, peptides from partial acid hydrolysis; Pn, papain peptides.

The T-, C-, and Tg- peptides are numbered (e.g. T-1) starting with 1 for the peptide from the NH_2-terminus of the protein and proceeding to the highest number for the peptide from the COOH-terminus.

The other subpeptides of T- and C- are numbered arbitrarily.

SYMBOLS

⟶ Half arrows to the right below the amino acid residues indicate results of Edman degradation.

⇢ Half arrows of broken lines to the right below the amino acid residues indicate the results of LAP or APM digestion.

⟵ Half arrows to the left indicate results of degradation by Co A and Co B.

⟶ Arrows indicate the direction of the amino acid sequence, from the amino to the carboxyterminal.

MATERIALS

Lima bean trypsin inhibitor (LBI) was obtained from Worthington (Freehold, N.J.).

LBI-7HA was used to prepare RCM-LBI IV tryptic peptides and RCM-LBI IV chymotryptic peptides.

LBI-9CA was used to prepare RCAM-LBI IV.

LBI-8KB was used to prepare GRCM-LBI IV.

TPCK-Trypsin (TRTPCK 8HA), and chymotrypsin (CDI 7CG), papain (PAP 7IE), leucine aminopeptidase (LAPC 7LA and LAPC 9EB), carboxypeptidase A (COADFP 7KB), and carboxypeptidase B (COBDFP 7GA) were all obtained from Worthington (Freehold, N.J.).

Aminopeptidase M (Rohm and Haas) was purchased from Henley and Co., Inc., (New York, N.Y.).

Pronase (B grade) was purchased from Calbiochem (Los Angeles, Calif.).

TAMe and BTEE were purchased from Mann Research Laboratories (New York, N.Y.).

The Sephadex series were purchased from Pharmacia Fine Chemicals (Montreal, P.Q.).

The Bio-Gel P series and Dowex 50-X2 (200-400 mesh) were purchased from Bio-Rad (Richmond, Calif.).

DEAE-Cellulose DE-11 was purchased from Whatman (England).

Guanidine hydrochloride, β-mercaptoethanol, iodoacetic acid and iodoacetamide were obtained from Sigma Chemical Co. (St. Louis, Mo.). O-methyl isourea was obtained from Nutritional Biochemicals (Cleveland, Ohio).

Phenylisothiocyanate (PITC) was purchased from Eastman Organic Chemicals (Rochester, N.Y.) and the sequanal grade PITC was purchased from Pierce (Rockford, Ill.). Chromar Sheet 500 was obtained from Mallinckrodt (St. Louis, Mo.). For acrylamide gel electrophoresis, the chemicals were obtained from Eastman Organic Chemicals (Rochester, N. Y.).

All other chemicals were reagent grade or better.

INTRODUCTION

A group of structurally dissimilar proteins that inhibit the activity or proteases occur in animal and plant tissues. These naturally occurring inhibitors are capable of combining with certain proteases to form an enzyme-inhibitor complex which is devoid of proteolytic activity. Bearing in mind that one partner in the complex is an enzyme, whereas the other has enzyme inhibiting activity, this unique system offers a chance for investigating certain aspects of protein-protein interactions.

A general survey of their molecular properties enables one to classify these inhibitors into two groups, one with relatively low molecular weight (6,000 - 12,000) and the other with a high molecular weight (20,000 - 60,000). The low molecular weight inhibitors, unlike the high molecular weight inhibitors, are stable to acids and heat. They remain active after exposure of several minutes to 2-3% trichloroacetic acid at temperatures up to 95°C.

The physiological importance of these inhibitors is poorly understood. It is believed that in the tissues in which they exist, they limit the area affected by and the duration of proteolytic processes. In this manner, the inhibitors supplement the other means available to the organism for controlling the action of proteolytic enzymes.

The spectrum of proteases inhibited by any one of these proteins vary considerably. Most reports have centered on the

inhibition of two bovine pancreatic enzymes, trypsin and chymotrypsin. Some inhibitors (chicken ovomucoid, Kazal's bovine pancreatic inhibitor) are specific for trypsin, some (golden pheasant ovomucoid, an inhibitor from potatoes) inhibit only chymotrypsin, while others can inhibit both trypsin and chymotrypsin simultaneously (turkey ovomucoid, lima bean inhibitor and the Bowman-Birk soybean inhibitor). In the case of the lima bean and Bowman-Birk inhibitors, the sites for the inhibition of trypsin and chymotrypsin are distinct and non-overlapping. These have been referred to as "double-headed" inhibitors.

Inactivation of proteases by these inhibitors is the result of a specific interaction between enzyme and inhibitor to form a stable complex. There is considerable speculation in the literature concerning the nature of this interaction, but at the moment, no firm statement can be made concerning the bonds and forces involved in complex formation. Undoubtedly, this remains an object of considerable research.

It seemed of interest to us to contribute to the knowledge of protease inhibitors by studying the inhibitor from lima beans. This inhibitor is a protein of low molecular weight ($< 10,000$), is commercially available in partially purified form, and a simple procedure has been developed for its further purification. Of added interest is the fact that it inhibits both trypsin and chymotrypsin and it has been shown that the inhibition of these two enzymes is independent and takes place

on two different and independent sites on the inhibitor.

Knowledge of the structure of the enzyme, as well as of the inhibitor, is a pre-requisite for a complete understanding of their mode of interaction. The amino acid sequences of trypsin and chymotrypsin are known; the three dimensional structure of chymotrypsin is also known and that of trypsin is presently under study. Because of its small size and its availability, the determination of the amino acid sequence of lima bean trypsin inhibitor seemed a feasible project. It was decided that the amino acid sequence be determined in order to open the way to a systematic study of the molecule and to understand its interactions with trypsin and chymotrypsin. Knowledge of the sequence will also allow comparison with other naturally occurring inhibitors for which sequences have already been determined.

LITERATURE REVIEW

There are several good review articles concerning the nutritional, pharmacological and general aspects of proteinase inhibitors (Green and Neurath, 1954; Laskowski and Laskowski, 1954; Desnuelle, 1960; Pusztai, 1967; Vogel et al., 1968; Weyer, 1968; Feeney and Allison, 1969; Liener and Kakade, 1969). Therefore, this review is limited to a discussion of those inhibitors which have been studied in greatest detail, with particular emphasis on structure and structure-function relationships. The discussion of protease inhibitors is divided into three sections:

I. Inhibitors of Plant Origin

II. Inhibitors of Animal Origin

III. Mechanism of Inhibition - A General Scheme

In the Appendix section of this thesis, there will be a brief discussion on the rationale of amino acid sequence determination of proteins and an evaluation of the methodology used.

I. INHIBITORS OF PLANT ORIGIN

The first proteinase inhibitor from plants was found in soybeans (Read and Haas, 1938). Since then, inhibitors have been found in all of the Leguminosae examined and they are also found to be present in Solanaceae, Gramineae and Chenopodiaceae (Vogel et al., 1968). Most extensively studied are the inhibitors from soybean and lima bean and these are discussed below.

A. Isolation and specificity.

1. Soybean Inhibitors.

a) Kunitz Soybean Inhibitor

Kunitz (1945,1946,1947) was the first to isolate, crystallize, and characterize the soybean inhibitor. A chromatographic procedure yielding preparations of higher purity was described by Rackis et al. (1962) for purification of the Kunitz inhibitor from soybean whey solutions. This inhibitor is specific for trypsin and at neutral pH forms a one-to-one molar complex with the enzyme (Laskowski and Laskowski, 1954). The complex is devoid of proteolytic activity.

b) Bowman-Birk Soybean Inhibitor

A second inhibitor, distinct from the Kunitz inhibitor, was first isolated from soybeans by Bowman (1946). Interest in this second inhibitor remained dormant until Birk (1961) resumed the investigation. She isolated two inhibitors from soybeans. One, specific for trypsin, was identical to the Kunitz inhibitor, and the other, found in the acetone-insoluble fraction of ether extracts of soybean flour, inhibited chymotrypsin, as well as trypsin. The same group of workers (Birk et al., 1963) succeeded in extensively purifying this acetone-insoluble inhibitor, which is now referred to as the "Bowman-Birk" inhibitor. There are other cumbersome descriptive terms used to describe this protein: acetone-insoluble factor, inhibitor AA and soybean trypsin and chymotrypsin inhibitor.

Frattali (1969) using the CM-cellulose

chromatographic procedure of Birk et al. (1963) isolated identical acetone-insoluble inhibitor preparations from two strains of soybean, the Lee and Hawkeye variety.

The Bowman-Birk inhibitor is "double headed" and can inhibit trypsin and chymotrypsin simultaneously (Birk et al., 1967; Frattali and Steiner, 1969). There have been conflicting reports on the stoichiometry of trypsin inhibition. Birk et al. (1963) reported that inhibition of trypsin was stoichiometric. In contrast to that, Frattali (1969) reported that inhibition of trypsin was non-stoichiometric. The difference was attributed to the method of assay. Birk et al. (1963) used the casein digestion method of assay, while Frattali (1969) had used synthetic substrates. Kakade (1970) measured the inhibition of chymotrypsin as a function of the ratio (by weight) of the Bowman-Birk inhibitor to chymotrypsin. Their results would lend support to the view of Birk et al. (1963) that inhibitor and chymotrypsin combine in a one-to-one molar ratio.

Yamamoto and Ikenaka (1967) isolated an inhibitor from soybeans specific for chymotrypsin and trypsin. Because of its sedimentation behaviour, it was referred to as the 1.9 S inhibitor. The similarity of this 1.9 S inhibitor to the Bowman-Birk inhibitor in terms of amino acid composition (Table I), electrophoretic mobility on acrylamide gel and specificity, led Frattali (1969) to propose that the two were identical. The only discrepancy is in the reported molecular

weight and this can be explained by the fact that the Bowman-Birk inhibitor is known to aggregate (see below).

2. Lima Bean Inhibitor

Tauber et al. (1949) first described the isolation of a crystalline trypsin inhibitor from lima beans. Fraenkel-Conrat et al. (1952) obtained an amorphous preparation of the inhibitor which was twice as active in inhibiting trypsin as the crystalline preparation of Tauber. More recently ion exchange chromatography and gel filtration have been used for the further purification of lima bean inhibitor. Jirgensons et al. (1960) observed that the inhibitor from lima beans could be fractionated into several components on DEAE-cellulose. A fractionation procedure combining gel filtration on Sephadex G-75 and chromatography on DEAE-cellulose was described by Jones et al. (1963). By this procedure, four homogenous fractions of the inhibitor were isolated from both bean extracts and commercial preparations of the inhibitor. Thus, there seems to be extensive heterogeneity in the inhibitors from lima beans. The origin of this heterogeneity is not known. Haynes and Feeney (1967) employed a similar fractionation method using DEAE-Sephadex as the ion-exchanger. They observed as many as six chromatographically distinct inhibitor fractions and also found that the relative proportions of these six fractions differed with the variety of bean used. This would suggest that the heterogeneity may be of genetic origin. The several active components could also arise by enzymatic

degradation in the beans, but the general resistance of the inhibitors to enzymic action would make this unlikely.

The four components prepared by Jones et al. (1963) had essentially the same inhibitory activity towards trypsin and each combined with the enzyme in equimolar amounts. Ryan and Clary (1964) first reported that the lima bean inhibitor preparation also inhibited chicken chymotrypsin. They referred to the inhibitor preparation as being "double-headed" since it inhibited turkey trypsin and chicken chymotrypsin. However, their results did not eliminate the possibility that the inhibitor might be a mixture of a trypsin inhibitor and a chymotrypsin inhibitor. Haynes and Feeney (1967) assayed their purified inhibitors for trypsin-inhibitory as well as chymotrypsin-inhibitory activity. They noted that the four components, corresponding to the four components of Jones et al. (1963) inhibited trypsin to the same extent. All four fractions also inhibited chymotrypsin, but to a lesser extent than the corresponding trypsin-inhibitory activity. In addition, the chymotrypsin-inhibitory activity varied among the fractions. Two fractions had two to three times as much chymotrypsin-inhibitory activity as the other two. They also observed that the inhibitor could inhibit trypsin and chymotrypsin simultaneously. This indicated that the two activities were independent of each other, and that these two enzymes were not competing for the same site on the inhibitor.

Plasmin, which resembles trypsin in specificity, was

reported to be inhibited by the lima bean inhibitor (Robbins and Summaria, 1966). However, the reaction between inhibitor and human plasmin is complex. In contrast to several other protease inhibitors which are good inhibitors of bovine trypsin, but do not inhibit human trypsin, lima bean inhibitor has strong inhibitory activity against human trypsin (Feeney et al., 1969). Recently, Trop and Birk (1970) demonstrated that a subfraction of pronase with trypsin-like activity was inhibited 80-90% by the lima bean inhibitor as well as by the two soybean inhibitors.

B. Physiocochemical Properties

1. Soybean Inhibitors

a) Kunitz Soybean Inhibitor

The molecular weight is 21,500±800g/mole (Wu and Scheraga, 1962a). The molecule consists of a single polypeptide chain crosslinked by two disulphide bridges. From measurements of intrinsic viscosity, fluorescence polarization and frictional ratio, the inhibitor appears to be a typical globular protein of low molecular asymmetry and high internal rigidity (Laskowski and Laskowski, 1954; Steiner and Edelhoch, 1963). The α helical content, as estimated from optical rotatory dispersion measurements, is small (Jirgensons, 1961). The physical parameters of aqueous solutions of this inhibitor including optical rotation, ultraviolet absorbance, fluorescent polarization and ultra violet fluorescence intensity are found to be constant between pH 6.0 and 9.0 indicating that no

significant change in molecular conformation occurs in this range (Steiner and Edelhoch, 1963; Wu and Scheraga, 1962b). At acid pH difference spectra, ultra violet fluorescence intensity and optical rotation measurements suggested that a conformational change might occur (Steiner and Edelhoch, 1963).

b) Bowman-Birk Soybean Inhibitor

Based on sedimentation velocity studies, the molecular weight of the Bowman-Birk inhibitor was reported as 24,000g/mole (Birk et al., 1963). Millar et al. (1969) studied the molecular weight of the Bowman-Birk inhibitor by sedimentation equilibrium techniques. They observed a variation of molecular weight with protein concentration and their studies indicated that the inhibitor self associates in solution. The results would be compatible with a mechanism involving a monomer-dimer-trimer equilibrium. The molecular weight at infinite dilution was 9,000g/mole. The phenomenon of self association observed by Millar et al. (1969) would explain the discrepancies in molecular weights reported for this protein in the literature. Kakade et al. (1970) determined the molecular weight of the inhibitor by gel-filtration. They obtained a value of 8250g/mole, which is in close agreement with the minimal molecular weight of 7975 calculated from amino acid analysis for the monomeric species (Frattali, 1969), and the value of 9000 reported by Millar et al. (1969).

The far ultra violet spectrum of the inhibitor was that expected for a random coil conformation (Steiner and Frattali,

1969). However, they consider it unwise to regard the inhibitor as a structureless entity since the large number of disulfide bonds must provide a good deal of molecular restraint. The only conclusion which could be made is that the inhibitor does not possess much helical character.

2. Lima Bean Inhibitor

From osmotic pressure measurements, Fraenkel-Conrat et al. (1952) estimated the molecular weight of lima bean inhibitor to be 9,400g/mole. In good agreement with this value are those calculated on the basis of the amino acid compositions and the stoichiometry of trypsin inhibition. Jones et al. (1963) reported molecular weights of 8,408, 8,291, 9,892 and 9,423g/mole for their four homogeneous components. However, by ultracentrifugation using the Ehrenberg method, Haynes and Feeney (1967) determined a molecular weight of 16,200g/mole and an S_{20w} value of 1.8 S. This suggests the possibility of a particle size represented by weights nearly twice the values obtained by Jones et al. (1963). The possibility of a larger particle size was also indicated by an S_{20w} value of 1.5 S reported by Fraenkel-Conrat et al. (1952). This apparent discrepancy between molecular weight determined from amino acid analysis and inhibitory activity on one hand and ultracentrifugation on the other hand is probably due to aggregation phenomena. This aggregation phenomenon has recently been shown to exist in many trypsin inhibitors of plant origin (Hochstrasser et al., 1969a).

C. Chemical Composition

The amino acid compositions of the plant inhibitors are listed in Table I.

1. Soybean Inhibitors

a) Kunitz Soybean Inhibitor

The amino acid composition of the Kunitz inhibitor was first determined by Wu and Scheraga (1962a) (Table I) and their results were later confirmed by Yamamoto and Ikenaka (1967). The molecule consists of 194 amino acids (Wu and Scheraga, 1962a). The amino acid composition has no unusual features. The amino acid sequence of this inhibitor is being determined by Ikenaka. Partial sequences around the two disulphide bonds were established by Brown et al. (1966) who employed the diagonal electrophoretic technique.

b) Bowman-Birk Inhibitor

The molecule consists of 78 amino acids (Frattali, 1969) (Table I). A striking feature in the amino acid composition of the Bowman-Birk inhibitor is the high content of half cystine (20%), proline (7%), acidic amino acids (20%) and absence of glycine and tryptophan (Frattali, 1969). So far there have been no reports on the primary structure of the Bowman-Birk inhibitor.

2. Lima Bean Inhibitor

Using chemical and microbiological techniques, Fraenkel-Conrat et al. (1952) performed the first amino acid analysis of lima bean inhibitor and they noted the very high half cystine content (16%) and the absence of tryptophan and

methionine. The inhibitors from lima beans and the Bowman-Birk soybean inhibitor have the highest cysteine content of any proteins so far studied. Jones et al. (1963) performed amino acid analyses of the four components which were shown to contain 77, 76, 93 and 86 amino acids respectively (Table I). They noted the striking similarity in the amino acid composition of the four components. Ferdinand et al. (1965) repeated some of the work of Jones et al. (1963) and the amino acid compositions of the components obtained by ion-exchange chromatography of different commercial preparations were found to be similar, but not identical. These results again point to the extensive heterogeneity of the lima bean inhibitor. Indications are that the different components are probably structurally related (Jones et al., 1963).

The similarity in amino acid composition between lima bean inhibitor and the Bowman-Birk soybean inhibitor is very striking (Table I). Both inhibitors lack tryptophan and show a very high content of half cystines, acidic amino acids, serine and proline.

D. Chemical Modification

1. Soybean Inhibitors

a) Kunitz Soybean Inhibitor

Chemical modification of the Kunitz inhibitor has revealed the importance of particular groups of amino acids in its inhibitory activity. When the two disulphide bonds were reduced by mercaptoethanol in 9M urea, the molecule

approached the state of a structureless random coil and lost its inhibitory activity (Steiner, 1965). If only one disulphide was reduced by treatment with 0.25M sodium borohydride the protein retained inhibitory activity (Dibella and Liener, 1969). When the fully reduced inhibitor was allowed to reoxidize in air a partial recovery of activity was observed (Steiner, 1965). If the reoxidized material was fractionated by gel filtration on Sephadex G-100, two fractions were isolated. One of these was the native inhibitor and the other fraction consisted of aggregates of the inhibitor. If the fully reduced inhibitor was allowed to reoxidize under denaturing conditions in 9M urea, the product contained inactive species in which "incorrect" pairing of half cystine residues had occurred (Steiner, 1965).

Steiner (1966) observed that 10 of the 11 lysine groups in the inhibitor could be converted to homoarginine by treatment with methyl isourea without a significant loss in activity. When 1,2,-cyclohexanedione was used to modify the arginyl residues, 80% of the inhibitory activity was abolished with modification of 50% of the arginines (Liu et al., 1968). The Kunitz inhibitor is classified as an "arginine inhibitor" in which a specific arginine residue is essential for the inhibitory activity (Ozawa and Laskowski, 1966). Two of the four tyrosines could be iodinated without loss in activity. Iodination of all four tyrosines under denaturing conditions resulted in complete and irreversible loss in activity (Steiner, 1966).

At least one of the three tryptophans could be oxidized by N-bromosuccinimide without major loss in activity, but oxidation of all three tryptophans under denaturing conditions resulted in complete and irreversible inactivation (Steiner, 1966). These findings of Steiner indicated that modification of readily available groups of a particular amino acid permitted retention of activity, while alteration of groups, which reacted only after disruption of the native structure prevented reformation of the active molecule.

b) Bowman-Birk Soybean Inhibitor

No chemical modification studies have been reported.

2. Lima Bean Inhibitor

The inhibitor prepared by Fraenkel-Conrat et al. (1952) was shown to retain activity after exposure to extremes of pH and temperature. However, destruction of the tertiary structure of the protein abolished all inhibitory activity. This was achieved by oxidation of the disulfide bonds with performic acid, or their reduction by mercaptoethanol (Jones et al., 1963). When the reduced inhibitor was allowed to re-oxidize in air, it was found to regain full activity (Jones et al., 1963). Initially it was thought that the oxidized as well as the reduced-carboxamidomethylated inhibitor was still resistant to the action of trypsin. It was later pointed out by Ferdinand et al. (1965), that incompleteness of the reduction could account for this resistance; it only requires

1% of residual inhibitory activity to prevent tryptic action. The inhibitors from lima bean have the highest half-cystine content of any proteins so far studied, resulting in a high degree of crosslinking by -S-S- bonds; some or all of these disulfide bonds would have to remain intact to maintain the molecular conformation for interaction with trypsin.

Techniques of chemical modification have been employed to determine the groups on the protein which are important for the interaction with trypsin. In view of the stability of the inhibitor to extremes of temperature and pH, inactivation produced by chemical reagents was regarded as a consequence of chemical alteration of the molecule rather than to non-specific denaturation. The inhibitor was not inactivated by esterification of its carboxyl groups, by iodination of its phenolic and imidazole groups or by coupling of these two groups with diazobenzene sulfonic acid (Fraenkel-Conrat et al., 1952).

Chemical modification of the amino groups by acetylation (Fraenkel-Conrat et al., 1952), amidination (Haynes and Feeney, 1967) or treatment with trinitrobenzene sulfonic acid (Haynes and Feeney, 1967), destroyed the trypsin inhibitory activity with no effect on the chymotrypsin inhibitory activity.

A kinetic analysis of the trinitrobenzene sulfonic acid reaction led Haynes and Feeney (1967) to conclude that a "fast reacting" lysine residue was probably critical for trypsin inhibitory activity. Chemical modification of all the amino

groups by guanidination did not abolish trypsin inhibitory activity (Haynes and Feeney, 1968). This was unexpected since other chemical modification of the amino groups always resulted in loss of trypsin inhibitory activity. The authors proposed that homoarginine was functioning as the primary binding site on the inhibitor in place of lysine.

In support of this argument (Haynes and Feeney, 1968) chemically modified the guanidinated lima bean inhibitor with 1,2-cyclohexanedione, a relatively specific reagent for the modification of both arginine and homoarginine. The result was a loss of trypsin inhibitory activity. Since modification of arginine residues in native lima bean inhibitor had little effect on the trypsin inhibitory activity (Liu et al., 1968), the lima bean inhibitor is placed in the class of inhibitors that require a lysine residue for trypsin inhibitory activity (see discussion on mechanism of action below), however, homoarginine can substitute for this lysine.

II. INHIBITORS OF ANIMAL ORIGIN

A. General

Inhibitors of proteolytic enzymes occur in various organs of different mammals, but only a few are well characterized. The inhibitors from bovine and porcine organs are the ones that have been studied to the greatest extent. In the bovine species, inhibitors occur in the pancreas, pancreatic juice, lungs, parotid glands and colostrum (Vogel et al., 1968).

In dogs and cats, inhibitors occur in the submandibular glands (Vogel et al., 1968). Inhibitors from seminal vesicles and semen have been reported to occur in all mammals studied including man (Vogel et al., 1968). The serum of humans, sheep and cow also contain inhibitors but little is known concerning their mode of interaction. Serum inhibitors are typically high molecular weight proteins and in humans, have been found in the post-albumin region, the α_1- and α_2-globulin fractions (Vogel et al., 1968).

In addition to the mammalian inhibitors, proteinase inhibitors have so far been found in nematodes (Peanasky and Laskowski, 1960; Portmann and Fraefel, 1967; Fraefel and Acher, 1968; Kucich and Peanasky, 1970), the eggs of certain birds (Vogel et al., 1968) and in the pancreas of the African lung fish (Reeck et al., 1970).

B. Kunitz Pancreatic Inhibitor or Basic Pancreatic Trypsin Inhibitor (BPTI)

1. Isolation and Specificity

Kunitz and Northrop (1936) were the first to isolate the basic trypsin inhibitor from pancreatic tissue of the cow and they crystallized it as well as its one-to-one molar complex with trypsin. However, the inhibitor preparation was not chromatographically pure. Kassell et al. (1963) modified the purification procedure employing ion-exchange chromatography and gel filtration. They obtained BPTI which was homogeneous as judged by chromatography, sedimentation

velocity, amino acid analysis and amino terminal analysis. Besides inhibiting trypsin BPTI also inhibits chymotrypsin (Wu and Laskowski, 1955), kallikrein (Werle and Appel, 1959), thrombin (Ferguson, 1942), and a subfraction of pronase with trypsin-like activity (Trop and Birk, 1970).

Inhibitors similar or identical to BPTI are also found in other organs of cattle. All ruminants investigated also had a Kunitz type inhibitor in pancreas and other organs. It is not found in other mammals. The physiological function of this inhibitor is unknown although it has been suggested that it is concerned with intracellular processes since it is not detectable in secretions or in the blood (Vogel et al., 1968).

2. Chemical Composition and Primary Structure

The amino acid composition of BPTI was first determined by Kassell et al. (1963) and it is listed in Table II. The inhibitor contains 58 amino acids, is devoid of histidine and tryptophan, and three amino acids are present as single residues, methionine, serine and valine. The molecular weight is 6513g/mole. The inhibitor is a highly basic protein with isoelectric point above pH 10 (Green and Work, 1953).

The molecular structure of BPTI was worked out almost simultaneously by three groups (Chauvet et al., 1964, 1966a,b, 1967; Kassell et al., 1965a,b; Dlouha et al., 1965). The protein consists of a single polypeptide chain crosslinked by three disulfide bridges (Kassell and Laskowski Sr., 1965b). The presence of three disulfide bridges in a molecule with a

molecular weight of 6513, and the strong resistance to enzymatic digestion suggested a compact structure for the molecule (Kassell and Laskowski Sr., 1956, 1965).

3. Proteinase Inhibitors Similar to BPTI

Proteinase inhibitors from bovine organs, other than the pancreas, have been shown to be identical to BPTI. A kallikrein inactivator from bovine parotid glands, lymph nodes and spleen was described for the first time by Kraut et al. (1928). It inhibited kallikrein, the proteinase which liberates peptides of hypotensive effect (kallidin, bradykinin) from the serum α_2-globulin fraction. Kraut et al. (1960) succeeded in purifying the kallikrein inactivator of the bovine parotid gland. The molecular weight and sequence studies on the inactivator showed it to be identical to BPTI (Anderer, 1965b; Anderer and Hörnle, 1966). A kallikrein inactivator from bovine lungs which also inhibited trypsin and chymotrypsin (Werle and Kaufmann-Boetsch, 1960) was also found to be identical to BPTI in primary structure (Anderer, 1965a).

A heterogeneous mixture of proteinase inhibitors was obtained from cow colostrum (Laskowski and Laskowski, 1951). Čechová et al. (1969) resolved cow colostrum inhibitor into three electrophoretically homogeneous components. The one present in highest amount was investigated in detail. The complete sequence of its 67 amino acids was elucidated and a comparison with the sequence of BPTI revealed considerable homology between the two, with at least twenty-one amino

acids in identical positions.

4. Chemical Modification

One or more of the disulfide bonds of BPTI are necessary for its inhibitory activity. Complete reduction of the three disulfide bonds results in loss of activity while air oxidation of the reduced inhibitor restores the activity (Chauvet and Acher, 1966; Avineri-Goldman et al., 1967). Kress and Laskowski (1967) succeeded in selectively reducing a single disulfide bond (Cys 14 - Cys 38) with sodium borohydride. The product was active, but lost its activity upon alkylation with iodoacetic acid. On the other hand, when partially reduced, BPTI was reacted with iodoacetamide or ethyleneimine, the derivatives retained trypsin inhibitory activity, but were inactive towards chymotrypsin (Kress et al., 1968). However, inhibition towards trypsin was temporary, enzyme activity was regained and the derivatives were gradually digested by the enzyme (Kress et al., 1968). This phenomena has been described as "temporary" inhibition and has also been observed in two other inhibitors of trypsin, namely, Kazal type inhibitors from pancreas (Laskowski and Wu, 1953; Greene et al., 1966; Burck et al., 1967), and chicken ovomucoid (Gorini et al., 1952; 1953).

The presence of a single methionine in BPTI stimulated interest in its possible role in the activity of the inhibitor. Kassell (1964), oxidized the single methionine to the sulfoxide and showed that this had no effect on trypsin inhibitory activity.

A spectophotometric titration of BPTI showed that three out of four tyrosines were exposed (Sherman and Kassell, 1968). Treatment of the inhibitor with tetranitromethane (Meloun et al., 1968) or iodine (Sherman and Kassell, 1968) resulted in modification of two and three tyrosines respectively, without any loss in trypsin inhibitory activity. Optical rotatory dispersion measurements revealed that nitration caused no appreciable change in the conformation of the peptide backbone (Meloun et al., 1968). Thus tyrosine residues of BPTI are probably not involved in interaction with trypsin.

When the free carboxyl groups of BPTI were esterified, the inhibitor lost 75% of its inhibitory activity (Avineri-Goldman et al., 1967). Chauvet and Acher (1968) attempted to chemically modify the free carboxyl groups with carbodiimide and glycine ethyl ester. Three out of five carboxyl groups were modified including the carboxyl-terminal residue. The inhibitor retained its activity so the authors concluded that the carboxyl-terminal was not involved in the binding with trypsin.

Since trypsin hydrolyzes at bonds involving lysine and arginine residues, these residues were suspected to be points at which the inhibitor might attach itself to trypsin. The four ε-amino groups of lysine were chemically modified to homoarginine residues without any effect on inhibitory activity (Kassell and Chow, 1966; Chauvet and Acher, 1967a). However, when free amino groups of the protein were acetylated (Avineri-

Goldman et al., 1967; Chauvet and Acher, 1967a) the inhibitor lost 75 - 93% of its trypsin inhibitory activity.

When free BPTI and the BPTI-trypsin complex were allowed to react with N-carboxy-DL-alanine, it was found that lysine 15 was shielded in the complex with trypsin (Chauvet and Acher, 1967b). Lysine 15 was thus implicated as the residue involved in the interaction between inhibitor and trypsin. Kress and Laskowski (1968) presented further evidence that lysine 15 was indeed the "reactive site" residue. Their experimental results indicated that complex formation between trypsin and BPTI, which contained carboxamidomethyl groups in half cystine residues 14 and 38, resulted in cleavage of a single lysyl bond in the inhibitor. The "reactive site" was identified as lys 15 - ala 16.

Removal of three amino acids from the N-terminus (Kassell and Chow, 1966) or lengthening of the N-terminus by addition of an ala-ala peptide (Chauvet and Acher, 1967a) had no appreciable effect on the inhibiting capacity of BPTI.

C. Kazal Inhibitor or Pancreatic Secretory Trypsin Inhibitor

1. Isolation and Specificity

Kazal et al. (1948) were the first to isolate an acidic trypsin inhibitor as a byproduct of commercial insulin preparations from bovine pancreas. They did not characterize it to any extent because of the limiting quantities that could be prepared. Their preparation was also reported as heterogeneous. It is referred to as the Kazal inhibitor,

other designations include - acidic secrtory inhibitor, specific inhibitor APTI; SPTI; PSTI.

Since that time at least three groups of workers (Greene et al., 1966; Cerwinsky et al., 1967; Fritz et al., 1966, 1967a) have independently isolated this acidic trypsin inhibitor from bovine pancreas as well as bovine pancreatic juice. This acidic pancreatic trypsin inhibitor is distinct from the Kunitz inhibitor.

Greene et al. (1966) using gel filtration and ion-exchange chromatography techniques, succeeded in preparing a homogenous fraction of the acidic trypsin inhibitor from bovine pancreatic juice in yields sufficient for chemical characterization. Cerwinsky et al. (1967) applied a slightly different purification procedure and isolated the acidic pancreatic tissue inhibitor from bovine pancreas. At about the same time a polypeptide of the same amino acid composition as the Kazal inhibitor was isolated from bovine pancreas by formation of the trypsin-inhibitor complex with soluble trypsin (Fritz et al., 1966) and water-insoluble trypsin resins (Fritz et al., 1967a).

The Kazal inhibitor is the only inhibitor present in bovine pancreatic juice and is a constituent of the exocrine secretion (Greene et al., 1966; Fritz et al., 1966). It inhibits trypsin and forms a one-to-one molar complex with the enzyme (Greene et al., 1966; Burck et al., 1967). However, the trypsin-inhibitor complex is unstable and in the presence of an excess trypsin in the solution, the inhibitor becomes the substrate

for the unbound trypsin. This phenomena has been referred to as "temporary inhibition" and was also observed by Laskowski and Wu (1953) and Fritz et al. (1967b).

The Kazal inhibitor does not inhibit chymotrypsin or kallikrein (Greene et al., 1966; Burck et al., 1967; Fritz et al., 1966). However, it was shown to be a potent inhibitor of thrombin clotting activity (Burck et al., 1967).

2. Chemical Composition and Primary Structure

The molecular weight of the Kazal inhibitor is 6155g/mole from amino acid analysis (Greene et al., 1966) and 6500g/mole from sedimentation equilibrium (Cerwinsky et al., 1967). The molecule contains 56 amino acids and these are listed in Table II. Histidine, phenylalanine, and tryptophan are absent and there is a preponderance of acidic amino acids. The amino acid sequence of the 56 residues was first reported by Greene and Giordano (1969) and Greene and Bartelt (1969). When the amino acid sequence of the Kazal inhibitor is compared with that of the Kunitz inhibitor, one notes an absence of any similarity between the two.

3. Kazal Type Inhibitor from Other Mammals

The Kazal type inhibitor is also present in the pancreas of other mammals. Two trypsin inhibitors of the Kazal type were isolated from porcine pancreatic juice (Greene et al., 1968) as well as porcine pancreas (Tschesche et al., 1969). Inhibitor I has a molecular weight of 6024g/mole and contains 56 amino acids. Inhibitor II has 52 amino acids and

has a molecular weight of 5609g/mole (see Table II). The NH_2-terminal sequence determined by Greene et al. (1968) as well as the amino acid compositions of I and II, indicates that Inhibitor II may be equivalent to residues 5-56 of Inhibitor I. Tscheche et al. (1969) reported the sequence of Inhibitor I which shows considerable homology with the Kazal inhibitor of bovine pancreas and bovine pancreatic juice.

Hochstrasser et al. (1969b) compared the tryptic peptides of reduced and carboxymethylated trypsin inhibitor from sheep pancreas with those of the Kazal inhibitor of bovine pancreas. They concluded that the inhibitor from sheep pancreas was identical to the Kazal inhibitor with the exception of one alanine in the sheep inhibitor replacing a serine in the bovine inhibitor.

III. MECHANISM OF INHIBITION

A. Physical Evidence for the Formation of Enzyme-Inhibitor Complex

1. Ultracentrifugation

The complex of trypsin and ovomucoid was demonstrated as early as 1954 in the ultracentrifugation studies of Ram et al. (1954). Further evidence for the enzyme inhibitor complexes was furnished by Rhodes et al. (1960). The sedimentation patterns for various enzyme-ovomucoid mixtures indicated single peaks of the complexes with values of sedimentation constants higher than that of the ovomucoid preparations

alone.

2. Gel Filtration

The physical presence of a complex between inhibitor and enzyme has been aptly demonstrated by gel filtration studies of inhibitors from plants and animals (Haynes and Feeney, 1967; Dlouhá and Keil, 1969). On gel filtration of a mixture of trypsin and BPTI, a peak corresponding to a high molecular weight component, was eluted. This high molecular weight component was devoid of inhibitory activity, and on the basis of its elution volume as well as the amino acid composition, would represent the one-to-one molar complex between enzyme and inhibitor (Dlouhá and Keil, 1969).

That lima bean inhibitor is "double headed" was also demonstrated by gel filtration studies (Haynes and Feeney, 1967). The binary complex of trypsin and lima bean inhibitor eluted from a Sephadex G-100 column, had chymotrypsin inhibitory activity. So it appeared that the inhibitory sites for trypsin and chymotrypsin were present on the same molecule.

3. Electrophoresis

Polyacrylamide gel electrophoresis has been instrumental in identifying the complex of chymotrypsin-Bowman-Birk soybean inhibitor (Frattali, 1969). When the combining weights were calculated, it was concluded that inhibitor and chymotrypsin combined in a one-to-one molar ratio.

4. Spectral Changes Following Formation of Enzyme-Inhibitor Complexes

Interaction of either trypsin or chymotrypsin with the Bowman-Birk soybean inhibitor produced characteristic differences in the ultraviolet spectra. Examination of the circular dichroism spectra of enzyme-inhibitor complexes yielded interesting results (Steiner and Frattali, 1969). When the inhibitor bound to chymotrypsin, there was a positive difference in the near ultraviolet region where tyrosine residues contribute to the circular dichroism spectrum. This did not occur on the interaction of the inhibitor with trypsin. This result could be indicative of the involvement of one of the inhibitor's tyrosine residues in the binding site for chymotrypsin. In the far ultraviolet region, where the amide bond absorbs, there was a negative difference with the trypsin-inhibitor complex, whereas the chymotrypsin inhibitor complex showed little change. The effect with trypsin was interpreted as an actual structural change when enzyme and inhibitor combined.

B. Amino Acid Residues Involved in the Reactive Site of Inhibitors

1. By Chemical Modification

Chemical modification studies on inhibitors from widely different sources have shown that either one or more lysine or arginine residues are essential for inhibition of trypsin. A survey of 19 different inhibitors (Fritz et al., 1969)

showed that they were inactivated either by reaction with maleic anhydride (modification of lysine residues) or by reaction with 2-3 butanedione (modification of arginine residues). Since trypsin is specific for hydrolysis of peptide bonds involving the carboxyl group of lysine or arginine, it is reasonable to put forward the hypothesis that the reactive site of trypsin inhibitors resembles a tryptic substrate. Direct evidence for the involvement of a trypsin sensitive peptide bond in their active site was obtained by the partial proteolysis studies discussed below.

In those inhibitors which are known to have independent sites for trypsin and chymotrypsin modification of lysine residues resulting in loss of trypsin inhibitory activity does not affect chymotrypsin inhibitory activity. Examples of this are turkey ovomucoid (Stevens and Feeney, 1963), and lima bean inhibitor (Haynes and Feeney, 1967). Amino acids involved in the chymotrypsin inhibitory activities of these inhibitors have not as yet been directly identified by chemical modification.

2. By Partial Proteolysis

a) Anti-trypsin Sites

Incubation of chicken ovomucoid and Kunitz soybean trypsin inhibitor (Ozawa and Laskowski, 1966) with catalytic amounts of trypsin at acid pH resulted in the hydrolysis of a single arginyl-alanyl bond in the former and an arginyl-isoleucyl peptide bond in the latter. The modified inhibitor

was fully active but combined with trypsin about fifty times more slowly than the virgin one. When the modified inhibitors were treated with carboxypeptidase B, the newly formed COOH-terminal arginine was released and the des-arginine modified inhibitor was totally inactive (Finkenstadt and Laskowski, 1965; Ozawa and Laskowski, 1966).

On the basis of these observations, the arginyl-isoleucyl bond of the Kunitz soybean trypsin inhibitor and the arginyl-alanyl peptide bond of chicken ovomucoid has been referred to as the "reactive site" which interacts with trypsin. Since that time, partial proteolysis of several other inhibitors have confirmed the existence of arginyl-X and lysyl-X active sites. Table III lists the reactive sites of several inhibitors which have been determined by partial proteolysis.

Incubation of the Bowman-Birk soybean inhibitor (Birk et al., 1967; Frattali and Steiner, 1969) or lima bean inhibitor (Krahn and Stevens, 1970) with catalytic quantities of trypsin at low pH resulted in conversion of the inhibitors to the trypsin modified form. This trypsin modified inhibitor retained inhibitory activity towards chymotrypsin and trypsin. Treatment with carboxypeptidase B resulted in loss of trypsin inhibitory activity, but not the chymotrypsin inhibitory activity.

Thus the chemical modification studies and partial proteolysis studies allow a general conclusion be made that the reactive site of trypsin inhibition is a trypsin susceptible bond, but appropriate conformation around this site would also

be a requirement, since denatured inhibitors do not inhibit trypsin (Jones et al., 1963; Steiner, 1965).

b) Anti-chymotrypsin Site

By analogy to the anti-trypsin reactive site it is reasonable to propose the hypothesis that the anti-chymotrypsin sites contain a chymotrypsin sensitive bond. Incubation of the Bowman-Birk soybean inhibitor (Frattali and Steiner, 1969) and the lima bean inhibitor (Krahn and Stevens, 1970) with catalytic quantities of chymotrypsin at low pH resulted in conversion of the inhibitors to chymotrypsin modified inhibitors which retained inhibitory activity towards trypsin, but became less active against chymotrypsin. Recently Krahn and Stevens (1970) have identified the chymotrypsin sensitive bond of LBI, thus presumably the anti-chymotrypsin site of LBI is a leucyl-seryl bond located 29 residues from the COOH-terminal end of the molecule.

C. A General Scheme for the Mechanism of Inhibition

The simplest equation describing the interaction of trypsin and inhibitor and accounts for the partial proteolysis experiments described above is (Finkenstadt and Laskowski, 1965; Ozawa and Laskowski, 1966):

$$T + I \rightleftarrows C \rightleftarrows T + I^* \qquad (1)$$

Where T is trypsin, I and I* are native (virgin) and trypsin modified inhibitors respectively and C, the trypsin inhibitor complex. This equation implies that virgin and modified inhibitors are in equilibrium with one another. For

Kunitz soybean trypsin inhibitor the equilibrium constant $Khyd = \frac{(I^*)}{(I)}$ has been shown to be 6± 1 at pH 4.0 (Niekamp et al., 1969). Equilibrium concentrations of 70%-80% modified, 20%-30% virgin and 30% modified, 70% virgin have been obtained in the tryptic cleavage of the active sites of chicken ovomucoid (Haynes and Feeney, 1968a) and bovine pancreatic secretory inhibitor (Rigbi and Greene, 1968) respectively.

Another consequence of equation (1) is the reversibility of the reaction. Experimentally this was demonstrated as follows: a complex formed from equimolar amounts of trypsin modified soybean inhibitor and trypsin could be rapidly dissociated by suddenly dropping the pH to 2.0. The inhibitor was isolated and found to be essentially 100% virgin (Finkenstadt and Laskowski, 1967). This experiment allowed for virtually complete resynthesis of the cleaved reactive site bond and offers further proof that the Arg_{64}-Ileu bond was indeed the reactive site of the inhibitor.

Laskowski and Sealock (1969) even succeeded in replacing the active site $arginine_{64}$ by a lysine residue. Kunitz soybean $desarginine_{64}$ inhibitor was incubated with trypsin, carboxypeptidase B and lysine at pH 6.7. Within 10 days they had made a complex of trypsin and inhibitor in which a lysine had replaced $arginine_{64}$. This complex could be dissociated and and free inhibitor was isolated. Various experiments showed that indeed a virgin inhibitor had been synthesized in which lysine had replaced $arginine_{64}$.

On the basis of these and other similar experiments, Laskowski proposed (Finkenstadt and Laskowski, 1966) that the trypsin inhibitor reaction consists of the cleavage of one especially sensitive bond in the inhibitor by trypsin and of subsequent formation of a covalent bond between trypsin and inhibitor (possibly an ester bond between the active seryl of trypsin and the newly formed COOH-terminal of the inhibitor). Unfortunately all attempts to trap the acyl complex have thus far failed and in all cases virgin inhibitor and the enzyme were obtained upon dissociation of the complex.

The model has been subject to criticism. It has been argued (Haynes and Feeney, 1968b) that peptide bond cleavage is not a necessary step in the formation of the enzyme-inhibitor complex. This is mainly based on the observation that in both BPTI and LBI, known to contain a lysine in their active site, lysine can be chemically converted to homoarginine (a residue not believed to be susceptible to tryptic hydrolysis) without loss of activity (Kassell and Chow, 1966; Chauvet and Acher, 1967a; Haynes and Feeney, 1968b).

Therefore it is not clear whether peptide bond cleavage and acyl formation are necessary for complex formation and inhibition or whether these inhibitors merely act as substrate-like substances which bind with the enzyme to form a non-productive enzyme-substrate complex.

At present, not enough information is available on the antichymotrypsin site of chymotrypsin inhibitors although

indications are that they act in a similar fashion in that their reactive site resembles a chymotrypsin substrate (Krahn and Stevens, 1970).

EXPERIMENTAL AND RESULTS

I. PURIFICATION AND CHARACTERIZATION OF INHIBITORS FROM LIMA BEANS

A. Purification

The commercially prepared samples of lima bean trypsin inhibitor were purified according to the procedure described by Jones et al. (1963). All the steps for purification were performed at room temperature. Protein was detected by measurement of absorbance at 280 mμ.

Gel Filtration on Sephadex G-75

Amounts of 500 mg of the commercial inhibitor obtained from Worthington Biochemical Corporation were dissolved in 5-10 ml 0.1M ammonium bicarbonate, pH 7.8. This solution was applied to a Sephadex G-75 column (2.5 x 80 cm) equilibrated with 0.1M ammonium bicarbonate, pH 7.8 and the column was developed with the same buffer at a flow rate of 20-30 ml/hr. 4.5 ml fractions were collected. The elution pattern of gel filtration is shown in Figure 1.

All of the detectable trypsin and chymotrypsin activity was associated with the second peak. The fractions containing inhibitor were pooled and were dialyzed against water and then were lyophilized.

Chromatography on DEAE-cellulose

Inhibitor fractions (500 mg) obtained from the gel filtration step were dissolved in the starting buffer 0.01M sodium phosphate, pH 7.6 and applied to a DEAE-cellulose column

(2.5 x 80 cm) equilibrated with the starting buffer. The column was eluted by an exponential gradient of salt, 0.01M sodium phosphate and 0.4M NaCl, pH 7.6 applied through a 1-litre mixing chamber containing the starting buffer. The column was operated at a flow rate of 20 ml/hr and effluen fractions (4.5 ml) were collected. A typical elution profi is shown in Figure 2a. The four peaks of protein were pooled and were dialyzed against water for 4 hours and were lyophilized. Small amounts (2 mg) of each fraction were redissolved in water and were assayed for trypsin and chymotrypsin inhibitory activity. Four peaks of protein showed inhibitory activity I, II, III, IV. Each of them was then rechromatographed on DEAE-cellulose under the same conditions as described. The result of rechromatography of the material in the four peaks is shown in Figure 2, b, c, d, e.

B. Trypsin and Chymotrypsin Inhibition

The assay system described by Rhodes et al. (1957) was employed. The synthetic substrates p-toluene sulphonyl arginine methyl ester (TAMe) and benzoyl tyrosine ethyl ester (BTEE) are hydrolyzed by trypsin and chymotrypsin respectively, at an optimum pH of 7.9 to 8.4 to release H^+. The hydrolysis of the substrate results in a decrease of pH, which is measured by the change in colour of the indicator (m-nitrophenol) from yellow to colorless. To measure the trypsin or chymotrypsin inhibitory activity of an unknown preparation, a standard curve is first plotted giving the rate of color disappearance

as a function of the amount of turkey ovomucoid added to 30-40 μg of enzyme. The unknown is then located on the curves and its inhibitory activity can be ascertained. The conditions and concentrations used are as described by Rhodes et al. (1957). Figures 3 and 4 give the standard curves for trypsin and chymotrypsin inhibitory activity of a sample of turkey ovomucoid.

Inhibitory activities of Fractions I - IV

All four fractions I, II, III, IV showed approximately equal inhibitory activity against trypsin. The results are given in Table IV. All four fractions inhibited chymotrypsin, but this activity was about half that of the corresponding trypsin-inhibitory activity.

C. Amino Acid Analyses

Samples containing 0.05-0.2 μmoles of protein or peptide were hydrolyzed with 6N HCl at 110°C in sealed, evacuated tubes. The analyses of protein were done in duplicate after 24 and 72 hours of hydrolysis. All other analyses were carried out, after 22 hour hydrolysis, on the Beckman-Spinco 120C automatic amino acid analyzer by the method of Spackman et al. (1958) as outlined in the Spinco manual.

The amino acid compositions of the four inhibitor fractions are given in Table V. Their compositions are very similar, showing a high content of half cystines, a high content of acidic amino acids and absence of methionine and tryptophan.

D. Determination of Free Sulfhydryl Groups

The four inhibitor fractions were assayed for free sulfhydryl residues, with Ellman's reagent according to the method described by Fernandez et al. (1964). All four fractions showed an absence of free sulfhydryl groups. This in turn would suggest that the 14 half cystines are involved in disulphide linkages.

E. Polyacrylamide Gel Electrophoresis

Acrylamide gel electrophoresis of the four inhibitor fractions

This was carried out according to the method described by Ornstein and Davis (1964) on 7½% gels in 0.03M glycine buffer, pH 8.3. Samples of 25 µl to 50 µl containing 200-500 µg protein in 20% sucrose were layered on the surface of the gels. Electrophoresis was carried out at a constant current of 2.5 ma per tube, with the positive electrode in the lower chamber, for 1 hr. At the end of this time, the gels were immersed in the dye solution of amido black (1g/100ml of 7½% acetic acid). The dye was decanted after an hour and the gels were left to destain in 7½% acetic acid.

Figure 5 shows the electrophoresis pattern of the four inhibitor fractions. All four fractions appear to be homogenous.

Polyacrylamide gel electrophoresis of LBI IV and RCAM-LBI IV

The preparation of carboxamidomethylated inhibitor

Fraction IV (RCAM-LBI IV) will be described in Section II, Part A (i). The native inhibitor LBI-IV and the RCAM-LBI IV were subjected to polyacrylamide gel electrophoresis in sodium dodecyl sulphate (SDS), according to the method described by Weber and Osborn (1969). The proteins were dissolved at a concentration of about 400 µg per ml in 0.1% SDS, 0.1% β-mercaptoethanol and 0.01M phosphate buffer, pH 7.0 and then incubated at 37° C for 2 hours before they were applied to the 10% gels. Polyacrylamide gels were prepared by mixing a 20% acrylamide solution with an equal volume of 0.01M phosphate buffer, pH 7.0 containing 0.1% SDS and 0.1% β-mercaptoethanol, and to this solution, 1.5 ml freshly made 1½% ammonium persulphate were added.

Samples of 25 µl to 50 µl of native and RCM - LBI IV were layered on the gel beds. Electrophoresis was carried out at a constant current of 8 ma per tube with the positive electrode in the lower chamber. At the end of 4 hrs, the gels were immersed in the staining solution of 0.25% Coomasie brilliant blue (dissolved in 50% methanol). The dye was decanted after 1 hr and the gels were left to destain in 7½% acetic acid. Figure 6 shows the electrophoresis pattern of native Fraction IV and the denatured RCAM-LBI IV. The native and the denatured inhibitor appear as identical homogenous bands.

F. Peptide Maps

The similarities in the amino acid compositions of

the four inhibitors in lima beans suggested that these components might be structurally related (Jones et al., 1963). So a comparative study of their structure was attempted. The problem was approached by subjecting the four fractions to trypsin digestion followed by an examination of their peptide maps. Amounts of 20 mg of denatured protein (carboxymethylated) were digested with trypsin (2% w/w) at 37°C for 1 hour. The reactions were carried out in a pH-stat with the pH kept constant at pH 8.0. At the end of 1 hour, aliquots were removed for peptide mapping. The experimental details are described in later sections, (Section II, Part B (i) and Section III, Part A (ii)).

The peptide maps of the tryptic digest of Fractions, I II, III and IV were very similar. Figure 7 shows the peptide maps of Fractions I and IV.

Since the structural differences between the four inhibitor fractions were not detectable on the peptide maps of their tryptic digests, it was decided to elucidate the primary structure of one of these inhibitors, which might then serve as the basis of comparison for the other inhibitor fractions. Fraction IV of lima bean inhibitor - LBI IV was thus selected for sequence studies.

II. PREPARATION OF PEPTIDES FOR AMINO ACID SEQUENCE DETERMINATION OF LBI-IV

A. Chemical Modification Reactions on LBI IV

i) Reduction and alkylation

Reduction with mercaptoethanol and alkylation with iodoacetic acid or iodoacetamide, was carried out by a modification of the method of Hirs (1967). 150 - 180 mg inhibitor were dissolved in a solution of 8 ml 5M guanidine hydrochloride containing 0.2% EDTA, previously flushed with N_2 for 20 minutes and adjusted to pH 8.5 with 25% trimethylamine. Reduction was carried out for 16 hours at room temperature using a 200 to 400 fold molar excess (over the molarity of the disulfide bridges) of β-mercaptoethanol (2 additions of 1.5 ml each, 30 minutes apart). During the course of the reaction the pH was kept constant at pH 8.0 by addition of 2N sodium hydroxide using a Radiometer TTT 11 pH-stat equipped with a magnetic valve. Subsequent carboxymethylation or carboxamidomethylation was achieved by the slow addition of 8.1 g of iodoacetic or iodoacetamide (dissolved in 2-3 ml 1N sodium hydroxide) to the reaction mixture. This reaction was carried out in the dark at room temperature and the pH was kept constant at pH 8.0 by addition of 4N sodium hydroxide using the pH-stat. After the base uptake had stopped, the reaction mixture was titrated down to pH 4.0 with glacial acetic acid. The reagents were removed by gel filtration on a column of Sephadex G-25 (2.5 x 80 cm) using 10% acetic acid as the eluant. This treatment of the protein resulted in complete inactivation of the inhibitory activity. Amino acid analysis of a hydrolysate

of the derivative revealed the presence of carboxymethylcysteine. The amino acid composition of the carboxymethylated inhibitor - (RCM - LBI IV) is presented in Table VI. From this, it is evident that the reduction and alkylation reaction was complete. The carboxamidomethylated inhibitor will be referred to as RCAM - LBI IV.

ii) Guanidination

Guanidination of the RCM - LBI IV was carried out according to the method described by Kimmel (1967). O-methyl isourea HCl (0.9 g) was dissolved in 3 ml water and the pH adjusted to 10.5 by the addition of 2.5N NaOH. The volume was brought to 7.5 ml with water. RCM - LBI IV (170 mg) was dissolved in 7.5 ml water and added to the prepared solution of O-methyl isourea. The reaction was allowed to proceed at room temperature for 65 hr during which time the pH was constantly checked and if needed, adjusted to 10.5 by the addition of 2.5N NaOH. The reaction was terminated by lowering the pH to 2.2 using glacial acetic acid. The protein was isolated by gel filtration on a Sephadex G-25 column (2.5 x 80cm) equilibrated in 10% acetic acid which was also used as eluant. The product obtained in this manner will be referred to as carboxymethylated, guanidinated inhibitor - GRCM - LBI IV. Amino acid analysis of this derivative revealed the presence of homoarginine. The amino acid composition of GRCM - LBI IV is presented in Table IX. From the analysis, it is evident that the guanidination reaction

was complete since less than 0.1 mole/mole of lysine was detected.

B. Tryptic Peptides From RCM - LBI IV and RCAM - LBI IV

i) Digestion with trypsin

The reactions were carried out at 37°C in a 20 ml capacity vessel and under a stream of N_2. 150 - 180 mg RCM - or RCAM - LBI IV were dissolved in 9 ml water (i.e. 2% solution) and the solution was adjusted to pH 8.0 by the addition of 40 µl - 60 µl of a 25% solution of trimethylamine. The digestion mixture was maintained at pH 8.0 with the aid of a pH-stat, titrigraph and syringe assembly (Radiometer TTT 11, SBR 2C, SBUI) with 0.05N sodium hydroxide as the titrating agent. The normality of the base was calculated on the basis of the number of trypsin sensitive bonds, a half equivalent of base per bond. At zero time 100 µl of TPCK-trypsin solution (1.8 mg/100 µl water) were added to the digestion mixture. There was rapid uptake of base within 2 minutes. Further additions of 50 µl enzyme solution (1.8 mg/100 µl water) were added after 12 and 20 minutes to give a final enzyme/substrate ratio of 2% (W/W). The trypsin digestion curve of RCAM - LBI IV is shown in Figure 8. After 70 minutes when base uptake had stopped, the reaction was terminated by adjusting the pH to 4.4 with glacial acetic acid. An aliquot was removed for peptide mapping (Section III, Part A (ii)) and the rest of the solution was immediately applied to a column for separation of the tryptic peptides.

ii) Isolation of tryptic peptides

Two different methods were used to isolate the tryptic peptides of RCM - LBI IV and RCAM - LBI IV. The first method involved separation of tryptic peptides of RCM - LBI IV on a cation-exchange column (1.9 x 100 cm) of Dowex 50-X2 at 40°C. The resin was washed and the column was prepared as described by Schroeder (1967). The column was equilibrated with 0.2M pyridinium-acetate buffer at pH 3.1, before application of the digest. Fractions of 4.5 ml were collected. The digest was fractionated by elution with 2 litres of 0.2M pyridinium-acetate buffer at pH 3.1, followed by a linear gradient which was established with 2 litres of 2.0M pyridinium-acetate buffer at pH 5.0. Elution was completed by passing 2 litres of 2.0M pyridinium-acetate buffer at pH 5.0, and, finally 2 litres of 4.0M pyridinium-acetate buffer at pH 5.7. For detection of peptides samples of 200 µl each were removed from alternate tubes and analyzed after alkaline hydrolysis by the ninhydrin method of Hirs et al. (1956).

The elution profile for the separation of tryptic peptides RCM - LBI IV is shown in Figure 9.

Appropriate fractions under each chromatographic peak were pooled, taken to dryness on a rotatory evaporator at 36°C, and then dissolved in deionized water and stored frozen in stoppered vials. Aliquots were taken from each fraction for peptide maps and, where necessary, further purified by paper electrophoresis or paper chromatography. In the instance when

the peptides were shown to be pure by the peptide maps, aliquots were taken for amino acid analysis after acid hydrolysis.

In the second method, tryptic peptides obtained from trypsin digestion of RCAM - LBI IV were separated on a column of Bio-Gel P-6 (2.5 x 90 cm) equilibrated in 10% acetic acid, which was also the eluant. The column was operated at room temperature with a flow rate of 21 ml per hour. Fractions of 3.7 ml were collected. Figure 10 shows the elution profile for the separation on Bio-Gel P-6. Appropriate fractions under each peak were pooled and from then on treated in the same way as described in the previous paragraph.

iii) Tryptic peptides from Dowex 50-X2 (see Figure 9)

Most of the column fractions eluted from the Dowex 50-X2 column were pure. Peaks IV, VII and X were further purified by paper chromatography in Solvent II. Peak IV yielded two peptides - Peptide T-2 and Peptide T-3. Peak VII yielded Peptides T-1 and T-3a. An attempt was made to purify Peak X by paper chromatography in Solvent II. The material was sparingly soluble and consequently most of it was lost. Table VI lists the amino acid compositions of the 6 tryptic peptides. Yields represent actual recoveries, no corrections have been made for losses incurred during purification.

Peptide T-1 was eluted as three peaks, Peak VII, VIII and IX. The presence of 2 lysines in Peptide T-1 indicates that one was resistant to trypsin digestion. Peptide T-2 was eluted in Peaks IV and V. The composition of Peaks I

and II indicated that they were derived from Peptide T-6, so no further work was done on them.

There were three arginine peptides, Peptides T-2, T-3 and T-3a although the analysis of inhibitor Fraction IV shows a content of 2 arginines. Peptides T-3 and T-3a were isolated in yields of 41% and 9.2% respectively. The sum total of the yields do not exceed that of any of the other tryptic peptides to any extent. This discrepancy in arginine content may be explained if Peptides T-3 and T-3a were homologous peptides. This proposal is based on the sequence analyses of these 2 peptides which will be described in the next section. Threonine and leucine of Peptide T-3 seems to be replaced by serine and phenylalanine in Peptide T-3a.

Peptide T-5 was not analyzed since most of it was lost through mishandling. Another attempt was made to obtain Peptide T-5. Trypsin digestion was performed on 150 mg RCAM - LBI IV and the peptides were isolated by gel filtration on Bio-Gel P-6.

Tryptic Peptides from Bio-Gel P-6 (see Figure 10)

Peaks I and II were pure and consisted of Peptide T-1. Peak III was rechromatographed on Bio-Gel P-6, it contained Peptide T-5. Due to its sparing solubility, much of it was again lost through mishandling. There was thus insufficient material for sequence analysis. Peak IV consisted mainly of Peptide T-2 and Peptide T-4. These two peptides - T-2 and T-4, were isolated by paper electrophoresis at pH 1.9 and subsequently paper chromatography in Solvent II.

Peak VI consisted of Peptide T-3 and Peptide T-6 which were isolated by paper chromatography in Solvent I. Peak VII contained Peptide T-3a and Peptide T-6 which were purified by paper chromatography in Solvent II. The amino acid compositions of the seven peptides are listed in Table VII.

It was observed that Peptides T-3 and T-3a were eluted at different elution volumes although both contained the same number of amino acids. The more hydrophobic Peptide T-3a would be retarded since it could interact with the gel. It was also observed that the acidic Peptide T-6 was eluted as two peaks, together with Peptides T-3 and T-3a.

C. Chymotryptic Peptides of RCM - LBI IV

i) Digestion with chymotrypsin

The experimental conditions for chymotrypsin digestion of 180 mg RCM - LBI IV were essentially the same as that described for trypsin digestion. At zero time and 20 minutes 100 μl of a solution of chymotrypsin (1.8 mg/100 μl water) were added to the digestion mixture. After 2 hours when base uptake ceased the reaction was terminated by adjusting the pH to 2.1 with concentrated hydrochloric acid. The chymotrypsin digestion curve is shown in Figure 11. An aliquot of the digested protein was removed for peptide mapping and the rest of the solution was immediately applied to a cation-exchange column (1.9 x 100 cm) of Dowex 50-X2.

ii) Isolation of chymotryptic peptides

The conditions for the fractionation of chymo-

tryptic peptides on Dowex 50-X2 were identical to those described in Section II, Part B (ii). Figure 12 shows the elution pattern of the Dowex 50-X2 separation of the chymotryptic peptides of RCM - LBI IV.

Appropriate fractions under each chromatographic peak were pooled, taken to dryness on a rotatory evaporator at 36°C and then dissolved in deionized water. The fractions were stored in stoppered vials. Aliquots were removed from each fraction for peptide maps and, where necessary, the peptides were further purified. In the instance when the peptides were shown to be pure by the peptide maps, aliquots were taken for amino acid analysis after acid hydrolysis.

Five peaks were eluted from the ion-exchange column. Three of them, Peaks I, II and V were shown to be pure on the peptide map and contained Peptides C-6, C-2a and C-4 respectively.

Peak III was shown to consist of four major components on the peptide map. It was fractionated into 2 components, Peak III-1 and Peak III-2 by gel filtration on a column of Bio-Gel P-4 (see Figure 13).

Peak III-1 was Peptide C-2. Peak III-2 was further fractionated on a column of Bio-Gel P-2 into 2 major components; Peak III-2-A and Peak III-2-B (see Figure 14). Peak III-2-A was purified by paper electrophoresis at pH 4.7 to yield Peptide C-3 and Peptide C-5, which were recovered in low yields as a result of the several steps employed in their purification.

Peak III-2-B should contain low molecular weight material on the basis of its elution volume from the Bio-Gel P-2 column. A peptide map revealed 2 components and an attempt was made to isolate them on paper electrophoresis at pH 4.7. These peptides were obtained in very poor yields due to the many steps in purification. Consequently, amino acid analysis of only one of them was possible. The results of the analysis showed the peptide to be still heterogenous. No further work was done on the material.

Peak IV contained 2 major components which were separated by paper electrophoresis at pH 4.7 to yield Peptide C-3a and Peptide C-7.

The amino acid compositions of the purified chymotryptic peptides are tabulated in Table VIII.

D. Tryptic Peptides From GRCM - LBI IV

i) Digestion with trypsin

Homoarginine is reported to be resistant to trypsin hydrolysis (Shields et al., 1959). The RCM - LBI IV was guanidinated before trypsin hydrolysis in order to obtain specific cleavage at the two arginine residues. Hopefully, the GRCM - LBI IV tryptic peptides (Tg) should provide certain overlaps allowing arrangement of some of the tryptic peptides (T). The experimental conditions for trypsin digestion of GRCM - LBI IV were the same as that described in Section III, Part B (i). The digestion curve of GRCM - LBI IV with trypsin is illustrated in Figure 15.

ii) Isolation of GRCM - LBI IV tryptic peptides

Five GRCM - LBI IV tryptic peptides were isolated from a column (2.5 x 80 cm) of Bio-Gel P-6 with 10% acetic acid as the eluant (see Figure 16). Peak I contained two components. An attempt was made to purify the two components on a column (2.5 x 180 cm) of Sephadex G-25, but the resolution was poor. The elution pattern is illustrated in Figure 17. Peak B was pooled and was shown to be pure on a peptide map. This component was Peptide Tg-1.

The material from Peak A was fractionated at 40°C on a column (1.4 x 60 cm) of Dowex 50-X2 at a flow rate of 30 ml per hour. The resin was washed and the column prepared by the procedures of Schroeder (1967). The column was equilibrated with 0.2M pyridinium-acetate buffer at pH 3.1 which was also the starting buffer. Fractions of 3.4 ml were collected. The column was developed by a gradient which was established with 500 ml of starting buffer and 1200 ml of 2.0M pyridinium-acetate buffer at pH 5.0. The elution pattern is illustrated in Figure 18. Two peptides were isolated in pure form - Peptides Tg-1 and Tg-4.

Peak II - A peptide map showed 1 major and 5 minor components. These were separated on paper chromatography into four fractions. Each of these fractions were purified by paper electrophoresis at pH 1.9. The major component in Peak II was Peptide Tg-2. The minor components were Peptides Tg-1 and Tg-4.

Peak III - A peptide map of this peak showed it to be very heterogenous. It was purified in the same manner as Peak II and yielded mainly Peptide Tg-2.

Peaks IV and V were pure. The former was Peptide Tg-3 and the latter was Peptide Tg-3a.

The amino acid compositions of the GRCM - LBI IV tryptic peptides are listed in Table IX.

From the specificity of trypsin digestion, only the two arginines were expected to be cleaved by the enzyme. However, five peptides were obtained from the digest.

Peptide Tg-1 - This peptide, with homoarginines substituting for lysines, was identical to Peptide T-1 except for three additional residues, glutamic acid, aspartic acid and proline. However, the GRCM - LBI IV was prepared from a different batch of commercial inhibitor than the RCM - LBI IV. The discrepancy might be a reflection of the heterogeneity of the lima bean inhibitor. An overnight digestion with carboxypeptidase B and A released Homoarg, 1.00; Thr, 0.77; Ser, 0.26; Leu, 0.18; establishing the COOH-terminal Leu-Ser-Thr-Homoarg.

Peptide Tg-2 was identical in composition to Peptide T-2.

Peptide Tg-3 and Peptide Tg-3a were identical in compositions to the proposed homologous peptides - Peptide T-3 and Peptide T-3a. Although Peptides Tg-3 and Tg-3a both contained five amino acids, the latter was retarded on gel filtration. The presence of phenylalanine in Peptide Tg-3a might increase the elution volume since hydrophobic amino

acids are known to interact with the gel.

Peptide Tg-4 represents the sum of Peptides T-4, T-5 and T-6.

III. DETERMINATION OF THE AMINO ACID SEQUENCE OF THE PURIFIED PEPTIDES

A. Methods

i) Hydrolysis of peptides

Digestion with trypsin or chymotrypsin

1.0 µmole of peptide dissolved in 500 µl 0.1M NH_4HCO_3, pH 7.8 was treated with 40 µl of TPCK-trypsin or chymotrypsin (16 mg/ml in 0.1M NH_4HCO_3, pH 7.8). The mixture was incubated at 37°C for 4 hours and the reaction was stopped by freezing. An aliquot was used for peptide mapping, and the peptides were isolated by paper electrophoresis or paper chromatography or a combination of the two systems.

Digestion with papain

Hydrolysis with papain was performed with 1 to 2 µmoles of dried peptide and 400-800 µl papain solution (2 mg/ml in 0.005M phosphate buffer at pH 7.5, containing 5 µl of 2,3-dimercaptopropanol/10 ml). The mixture was incubated for 4 hours or overnight at 37°C and digestion was stopped by freezing. An aliquot (0.1 µmole peptide) was used for peptide mapping and the peptides were isolated by paper electrophoresis or paper chromatography or a combination of the two systems.

Digestion with pronase

1-2 µmoles of peptide dissolved in 1.5 ml 0.1M NH_4HCO_3, pH 7.8 were treated with 60 µl pronase solution (1 mg/ml in 0.1M NH_4HCO_3, pH 7.8). The mixture was incubated for 1 hour at 37°C and the reaction was terminated by freezing. An aliquot was used for peptide mapping. The peptides were isolated by paper electrophoresis or paper chromatography or a combination of the two systems.

Partial acid hydrolysis

Partial acid hydrolysis was performed as described by Light (1967). 1-2 µmoles peptide were incubated with 5.7N HCl at 105°C for 15 minutes, in sealed evacuated tubes. The reaction was terminated by immersion of the hydrolysate tubes in ice. The tubes were opened and their contents were taken to dryness at reduced pressure. The contents were redissolved in 100-200 µl water and an aliquot (10-20 µl) was used for peptide mapping. The peptides were isolated by paper electrophoresis or paper chromatography or a combination of the two systems.

ii) Peptide mapping and purification of peptides

Peptide maps were run on Whatman No. 3MM paper. Electrophoresis was performed in Savant electrophoresis tanks with pyridinium-acetate buffer at pH 4.7 (0.025M pyridine-0.035M acetic acid) at 50 volts per cm for 30 minutes or with acetic acid - formic acid buffer at pH 1.9 (25 ml of 98-100% formic acid and 87 ml glacial acetic acid diluted to 1 liter)

at 50 volts per cm for 30 minutes.

Chromatography was performed in Solvent I (1-butanol-acetic acid-water, 200:30:75, v/v) or in Solvent II (1-butanol-pyridine-acetic acid-water, 120:80:24:96, v/v). The same systems or combinations of these were used for purification of peptides. Peptide maps were developed with the ninhydrin-collidine reagent (Margoliash and Smith, 1962) and the Pauly reagent for histidine and tyrosine (Easley, 1965).

In the instance when peptides were purified on paper, they were eluted from paper with water at room temperature in a closed chamber. Concentration of peptide material was effected by rotatory evaporation at 36°C. The purified peptides were redissolved in 1-1.5 ml water and stored at 0°C.

iii) Sequence determination

a. Determination of NH_2-terminal

Leucine aminopeptidase digestion

0.1 µmole of peptide in 100 µl of 0.1M NH_4HCO_3 at pH 7.8 and 25 µl of 0.025M $MgCl_2$ was incubated with 5 µl enzyme (40 µg or 5.2 units). Digestion was performed at 37°C for 4 hours and terminated by freezing. Aliquots of 10 µl were examined by paper chromatography in Solvent system I and when necessary, the rest of the solution was submitted to quantitative amino acid analysis.

Edman degradation

Peptides were degraded sequentially by a modification of the basic Edman procedure as described

by Kasper and Smith (1966). To the peptide residue (0.1-2 µmoles) contained in a 5 ml glass-stoppered, conical centrifuge tube, were added 100 µl of deionized water, 200 µl of 1% (V/V) phenyl isothiocyanate in analytical grade pyridine, and 25 µl of 25% trimethylamine. The contents were mixed and incubated at 37°C for 2 hours, 10 µl of undiluted phenylisothiocyanate was then added with mixing, and the incubation was continued for another 2 hours. Distilled water (0.3 ml) was added, the contents mixed by agitation and the mixture was extracted three times with 0.5 ml portions of cyclohexane and four times with 0.5 ml portions of benzene. Residual organic solvent was removed by bubbling nitrogen through the aqueous phase at 30°C. The solution of the peptide derivative was taken to dryness at reduced pressure over NaOH flakes.

Cyclization was accomplished with anhydrous trifluoracetic acid at 25°C for 1 hour. After removal of trifluoracetic acid in a desiccator over NaOH the residue was dissolved or suspended in 0.5 ml of 0.01N HCl. The PTH-amino acid was extracted three times with 0.5 ml portions of peroxide free ether. An aliquot equivalent to 0.05 µmoles of residual peptide was removed from the aqueous phase for amino acid analysis. The ether solution of the PTH derivative was placed in a stream of nitrogen to remove all solvent. The PTH-amino acid was recyclized in 0.6 ml of an acetic acid:HCl (5:1) solution for 1 hour at 37°C. At the end of this time the solution was taken to dryness in a desiccator over NaOH.

The PTH-amino acid was dissolved in 50 µl of 90% acetic acid and identified by paper chromatography on Chromar Sheet 500 with the solvent system of Edman and Sjöquist (1956), Solvent F. The iodine-azide spray was used to develop the chromatogram.

b. Determination of COOH-terminal Carboxypeptidase A and B

The method was essentially that described by Stevens et al., (1967). To approximately 0.1 µmole peptide in 0.1M NH_4HCO_3, pH 7.8 were added 5 µl diisopropylfluorophosphate-treated carboxypeptidase A or B (24 µg/5 µl of 0.1M NH_4HCO_3, pH 7.8). Digestion was performed at 37°C for 4 hours or overnight and terminated by freezing. Aliquots were removed at 1 and 4 hours and were examined by paper chromatography in Solvent I and, when necessary, were submitted to quantitative amino acid analysis.

c. Amide determination using aminopeptidase M

Peptides were hydrolyzed with aminopeptidase M according to the method described by Evans et al. (1968). 0.1 µmole peptide in 450 µl 0.1M NH_4HCO_3, pH 7.8 and 50 µl 20mM $MgCl_2$ was incubated with 50 units enzyme. Digestion was allowed to proceed for 4 hours or 18 hours at 37°C and was terminated by freezing. Aliquots of 50 µl were examined by paper chromatography in Solvent I and where necessary, the rest of the solution was submitted to quantitative amino acid analysis.

B. Amino Acid Sequences of Tryptic Peptides

PEPTIDE T-1 (Residues 1 through 28): Ser-Gly-His-His-Glu-His-Ser-Thr-Asp-Glx-Pro-Ser-Glx-Ser-Ser-CMCys-Lys-Pro-CMCys-Asn-His-CMCys-CMCys-CMCys-Leu-$\overset{\text{Ser}}{\underset{\text{Ala}}{}}$-Thr-Lys-

Sequences studies were done on material from Peaks I and II obtained from the Bio-Gel P-6 column (Figure 10). Table X summarizes the sequence analysis. The NH_2-terminal Ser-Gly was established by Edman degradation and leucine aminopeptidase digestion. Carboxypeptidase B and A digestion would establish the COOH terminal as (Leu,Ala)-Thr-Lys. From a study of chymotryptic Peptide C-2 (see below) evidence for the existence of a serine, alanine replacement in position 26 was obtained. The carboxyterminal sequence of T-1 is therefore tentatively shown as -Leu-$\overset{\text{Ser}}{\underset{\text{Ala}}{}}$-Thr-Lys. The large Peptide T-1 was further degraded into smaller fragments so that the entire sequence might be established by means of overlapping peptides.

i) Pronase digestion of Peptide T-1

Hydrolysis of 2 µmoles Peptide T-1 with pronase for 1 hour at 37°C yielded 5 major and 5 minor components which were separated first into four fractions by paper electrophoresis at pH 4.7. One of them was pure Peptide T-1-P-2. The three other electrophoretic fractions were further resolved into peptides by paper chromatography in Solvent II. Sequence studies were done on 4 peptides which were reasonably

pure and were obtained in good yields.

Peptide T-1-P-1 (Residues 1 through 6): Ser-Gly-His-His-Glu-His- The sequence of this basic peptide was established by Edman degradation and carboxypeptidase A digestion. This peptide is assigned the NH_2-terminal portion of Peptide T-1 on the basis of the known NH_2-terminal sequence of Peptide T-1 (Ser-Gly) and a consideration that only a single glycine is present in LBI IV.

Peptide T-1-P-2 (Residues 7 through 18): Ser-Thr-Asp-Glx-Pro-Ser-Glx-Ser-Ser-CMCys-Lys-Pro- Aminopeptidase M digestion of this peptide released Ser,1.00; Thr,0.50. This peptide was resistant to carboxypeptidase A and B digestion. The Edman degradation was poorly performed and will not be considered. The peptide was further subjected to a more extensive digestion with pronase (3 hours at 37°C) to yield Peptides T-1-P-2a and T-1-P-2b which were purified by paper electrophoresis at pH 4.7. These Peptides T-1-P-2a and T-1-P-2b account for the total amino acid composition of Peptide T-1-P-2.

Peptide T-1-P-2a (Residues 7 through 14): Ser-Thr-Asp-Glx-Pro-Ser-Glx-Ser- This was a very acidic peptide. The sequence was determined by 6 successive steps of the Edman degradation and carboxypeptidase A digestion. The presence of

aspartic acid at position 9 was confirmed by digestion with leucine aminopeptidase. By comparison of the NH_2-terminal residues of Peptides T-1-P-2a and its parent, Peptide T-1-P-2, Peptide T-1-P-2a is believed to represent the NH_2-terminal portion of its parent peptide. This point will be further substantiated from the results obtained on the chymotryptic Peptide C-2.

Peptide T-1-P-2b (Residues 15 through 18): Ser-CMCys-Lys-Pro- The sequence of this peptide was established by leucine aminopeptidase digestion. It represents the COOH-terminal portion of the parent Peptide T-1-P-2.

It was observed that Peptide T-1-P-2b contained cysteic acid instead of carboxymethylcysteine. The latter has been reported to be unstable under acid conditions (Groskopf et al., 1969). It would appear that the cysteic acid is derived from carboxymethylcysteine during handling of the peptide. During the course of sequence determinations on other peptides, this phenomenon was repeatedly observed when peptides were isolated by paper electrophoresis and the carboxymethyl cysteine values as indicated in some tables by * represent the sum of cysteic acid, carboxy-

methylcysteine as well as half cystine values.

Peptide T-1-P-3 (Residues 19 through 22): CMCys-Asn-His-CMCys- The sequence was established by Edman degradation. The presence of asparagine was confirmed by digestion with aminopeptidase M and identification of free asparagine by paper electrophoresis at pH 4.7.

Peptide T-1-P-6 (Residues 27 and 28): Thr-Lys- This was a dipeptide and would be the COOH-terminal of Peptide T-1 from the results of carboxypeptidase digestion on Peptide T-1 and from the specificity of trypsin hydrolysis.

ii) Chymotrypsin digestion of Peptide T-1

The sequence of Peptide T-1 could not be entirely established by studies of its pronase peptides, therefore a chymotrypsin digest of Peptide T-1 was attempted. The results are also summarized in Table X. 1.5 µmoles peptide were treated with 1% chymotrypsin (W/W) at 37°C for 3½ hours. A peptide map of the digest revealed three major peptides, two of which were isolated in good yield by paper electrophoresis at pH 4.7.

Peptide T-1-C-1 (Residues 1 through 20): Ser-Gly-His-His-Glu-His-Ser-Thr-Asp-Glx-Pro-Ser-Glx-Ser-Ser-CMCys-Lys-Pro-CMCys-Asp- This peptide was assigned the NH_2-terminal portion of Peptide T-1 since it

contained the single glycine which is known to be in position 2. Peptide T-1-C-1 was further digested with pronase for 3 hours at 37°C. Seven peptides were revealed on a peptide map of the digest, three of which were obtained pure and in good yield by paper electrophoresis at pH 4.7.

Peptide T-1-C-1-P-1 (Ser(1), Gly(1), His(2), Glu(1))

Peptide T-1-C-1-P-2 (Ser(3), Thr(1), Asp(1), Glu(2), Pro(1))

Digestion of Peptide T-1-C-1-P-2 with leucine aminopeptidase released Ser,1.00; Thr,0.85; Asp,0.15.

Peptide T-1-C-1-P-3 (Ser(1), CMCys(1), Lys(1), Pro(1))

Digestion of Peptide T-1-C-1-P-3 with leucine aminopeptidase released Ser,1; CMCys,0.54; Lys,0.35

Peptide T-1-C-2 (Residues 23 through 28): CMCys-CMCys-Leu-Ala-Thr-Lys- This peptide was assigned the COOH-terminal portion since its composition indicates that it is derived from the known COOH-terminal of Peptide T-1. One step of the Edman degradation established the NH_2-terminal as carboxymethylcysteine. From considerations of the carboxypeptidase digestion of Peptide T-1 and the results obtained on Peptide C-2 the second CMCys residue must be in position 24. Figure 19 illustrates the sequence of Peptide T-1 derived from the information presented above.

Residues 1 through 22

Peptide T-1-C-1-P-1 was assigned the NH_2-terminal of the parent Peptide T-1-C-1 on consideration of its composition and the known NH_2-terminal of Peptides T-1 and T-1-P-1. The sum of Peptides T-1-C-1-P-2 and T-1-C-1-P-3 represent Peptide T-1-P-2. Thus Peptide T-1-C-1 provides the overlap between the NH_2-terminal Peptide T-1-P-1 and Peptide T-1-P-2. From their respective amino acid compositions, it can be seen that Peptide T-1-C-1 contains the same residues as the sum of Peptides T-1-P-1 and T-1-P-2 with an additional aspartic acid and carboxymethylcysteine. Thus the pronase peptide which follows T-1-P-2 would be one beginning with the dipeptide CMCys-Asp or Asp-CMCys. The condition is satisfied by only one pronase peptide - Peptide T-1-P-3 which begins with CMCys-Asn. In this fashion, the order of the peptides derived from residues 1 through 22 is established. From the known carboxyterminal of T-1 we can assign Peptide T-1-C-2 as carboxyterminal and following residues 1 through 22. Added evidence for this arrangement and the overlap between residues 22 and 23 is obtained from studies on the chymotryptic Peptide C-2 (Table XVII).

PEPTIDE T-2 (Residues 29 through 35): Ser-Ileu-Pro-Pro-Glx-CMCys-Arg-

Table XI summarizes the sequence studies on this peptide, isolated, pure, from Peak V (Figure 9). Digestion with carboxypeptidase B and A for 4 or 18 hours released only arginine. Four steps of the Edman degradation showed the NH_2-terminal sequence to be Ser-Ileu-Pro-Pro. Partial acid hydrolysis of

of 2 μmoles of the peptide in 5.7N HCl at 105°C for 40 minutes yielded 10 peptides of which three acidic peptides were isolated by paper electrophoresis and chromatography in Solvent I:

T-2-PA-1 (Glu(1), Cysteic (1))

T-2-PA-2 (Glu (1))

T-2-PA-3 (Ser(1), Ileu (1), Pro(2), Glu(1))

From these results, the entire sequence of Peptide T-2 may be established. The positions of Glu and CMCys are placed in the order -Glx-CMCys on consideration of the compositions of the 2 peptides - T-2-PA-1 and T-2-PA-3 - which would give the overlap.

PEPTIDE T-3 (Residues 36 through 40): CMCys-Thr-Asp-Leu-Arg-

The sequence studies on this peptide are summarized in Table XII. This was an acidic peptide, purified from Peak IV (Figure 9) by paper chromatography in Solvent II. A 4 hour digestion of the peptide with carboxypeptidase B and A released Arg. An 18 hour digestion with carboxypeptidase B and A released Arg, 1.00; Leu,0.25. Three steps of the Edman degradation established the NH_2-terminal CMCys-Thr-Asp. The presence of aspartic acid was confirmed by digestion of Peptide T-3 with aminopeptidase M and identification of aspartic acid by paper chromatography in Solvent I. The above information would then establish the entire sequence of the Peptide T-3.

PEPTIDE T-3a (Residues 36 through 40): CMCys-Ser-Asp-Phe-Arg-

Table XIII summarizes the sequence studies on this peptide. This was an acidic peptide. It was purified from Peak VII

(Figure 9), by paper chromatography in Solvent II. Digestion of the peptide with carboxypeptidase B released Arg,1.00; Phe,0.73; thus establishing the COOH-terminal as - Phe-Arg. Three steps of the Edman degradation established the NH_2-terminal as CMCys-Ser-Asp. Digestion of the peptide with aminopeptidase M confirmed the presence of aspartic acid which was identified by paper chromatography in Solvent I. This peptide has been designated Peptide T-3a because it seems to be homologous to Peptide T-3 with serine and phenylalanine in Peptide T-3a replacing threonine (37) and leucine (39) in Peptide T-3. Furthermore, there are only two arginines in LBI - IV but three tryptic peptides containing arginine were isolated. Since the sequence of Peptide T-2 is unique, and the sequences of the other two Peptides T-3 and T-3a are very similar, it is proposed that the latter two are homologous. This existence of the 2 homologous peptides will also be evident when the chymotryptic peptides are examined.

PEPTIDE T-4 (Residues 41 through 49): Leu-Asp-Ser-CMCys-His-Ser-CMCys-Ala-Lys- Table XIV summarizes the sequence studies on this peptide obtained from Peak VI (Figure 9). Digestion with carboxypeptidase B for 4 hours released: Lys,1.00; Ala,0.81; CMCys,0.48. An overnight digestion with carboxypeptidase B and A released: Lys,1.00; Ala,1.00; CMCys,1.00; Ser,0.41. This would establish the COOH-terminal to be Ser-CMCys-Ala-Lys. Four steps of the Edman degradation showed the NH_2-terminal to be Leu-Asp-Ser-CMCys. The presence of

aspartic acid in Peptide T-4 was confirmed by a digestion with aminopeptidase M and identifying the free amino acids by paper electrophoresis as well as by amino acid analysis. In the Edman degradation, the values for histidine and lysine were inconsistent. However, lysine is assigned the COOH-terminal residue on consideration of the specificity of trypsin and the results of carboxypeptidase digestion on Peptide T-4. The assignment of histidine as residue 45 was confirmed from results of peptides obtained by papain digestion of Peptide T-4. 2 μmoles of Peptide T-4 were subjected to a 3 hour digestion with papain at 37°C. A peptide map of the digest showed eight major components, five of which were obtained in good yield and in pure form by paper electrophoresis at pH 4.7, and paper chromatography in Solvent I.

Peptide T-4-Pn-1 (Asp(1), Ser(1), Cysteic(1)) overlaps with Peptides T-4-Pn-7 (Leu(1), Asp(1), Ser(1)) as well as T-4-Pn-8 (Leu(1), Asp(1)). This confirms the NH_2-terminal sequence of Peptide T-4 as determined by Edman degradation. Peptides T-4-Pn-5 and T-4-Pn-6 were identical except that cysteic acid in the former had replaced carboxymethylcysteine in the latter. The results of the COOH-terminal determination on Peptide T-4 indicate that the tripeptides (T-4-Pn-5 and T-4-Pn-6) would lie adjacent to the COOH-terminal lysine. The only position that histidine can occupy in Peptide T-4 would be residue 45. This point will be unequivocally established during sequence determination of chymotryptic Peptide C-4 in Section III, Part C.

PEPTIDE T-6 (Residues 75 through 84): Ser-Ser-His-Ser-Asp-Asp-Asp-Asn-Asn-Asn-COOH

Table XV summarizes the sequence studies on this peptide. Amino acid analysis revealed the presence of Asp(6), Ser(3), His(1). Three steps of the Edman degradation showed the NH_2-terminal to be Ser-Ser-His. Carboxypeptidase A digestion for 4 hours showed the presence of Asn. Peptide T-6 was resistant to chymotrypsin and papain digestion. An experiment in which Peptide T-6 was incubated with aminopeptidase M, yielded some interesting results and enabled the entire sequence to be elucidated. Five stoppered tubes, each containing 0.2 µmoles peptide in 450 µl of 0.1M NH_4HCO_3, pH 7.8 and 50 µl 20mM $MgCl_2$ were incubated with 50 units of aminopeptidase M at 37°C. At 2, 10, 24, 30 and 48 hours after incubation, one tube at a time was removed from the incubation bath and the reaction was terminated by freezing. Aliquots of 20 µl from each tube were spotted on Whatman 3MM for paper electrophoresis and the rest of the sample was subjected to amino acid analysis. This time course study on the aminopeptidase M digestion of Peptide T-6 establishes the NH_2-terminal as Ser-Ser-His-Ser (see Table XV). Even after the 48 hours digestion with amino-peptidase M, Peptide T-6 was not completely degraded into amino acids. Ser(3), and His(1), were released after 48 hours and the residual peptide T-6-A-1 was more acidic than aspartic acid as shown by paper electrophoresis.

2 µmoles of Peptide T-6 were subjected to a 48 hour

digestion with aminopeptidase M and the acidic Peptide T-6-A-1 was isolated by paper electrophoresis. Amino acid analysis showed the presence only of aspartic acid. Five steps of the Edman degradation were performed on Peptide T-6-A-1. At each step the PTH-amino acid was identified so that the NH_2-terminal sequence of Asp-Asp-Asp-Asn-Asn was established. After five steps, the residue was identified as asparagine by paper electrophoresis.

Identification of the residues removed at each step of the Edman degradation was also confirmed in the following way. At each step of the Edman degradation, an aliquot of the residual peptide was removed from the aqueous phase and subjected to paper electrophoresis at pH 4.7 along with standards of aspartic acid and asparagine. After the paper strips were developed with the ninhydrin collidine reagent the distances of peptides from the origin were measured. The electrophoretic mobilities at pH 4.7 relative to that of aspartic acid are tabulated in Table XVI. A decreasing mobility indicates the removal of aspartic acid during the Edman degradation. Based on this consideration, the sequence of Asp-Asp-Asp-Asn-Asn was established. Knowledge of the specificity of trypsin digestion would place Peptide T-6 as the COOH-terminal peptide of LBI IV.

C. Amino Acid Sequences of Chymotryptic Peptides

PEPTIDE C-2 (Residues 9 through 28): Asp-(Glx,Pro,Ser,Glx)

(Ser,Ser, CMCys,Lys) (Pro,CMCys,Asn,His,CMCys)-CMCys-CMCys-Leu-Ser/Ala-Thr-Lys-

This peptide was isolated from Peak III (Figure 12). Its purification is illustrated in Figures 13 and 14. Table XVII summarizes the sequence work on this peptide. Only 1 step of the Edman degradation was successful and showed the presence of aspartic acid as the NH_2-terminal residue. A three hour digest with carboxypeptidase B and A released Lys,1.00; Thr,0.55; Ser,0.27; Ala,0.069; Leu,0.29. This together with the compositions of peptides obtained by partial acid hydrolysis (see below) indicates a COOH-terminal sequence of Leu-Ser/Ala-Thr-Lys. In Peptide T-1 residue 26 was identified as alanine. From these results of carboxypeptidase digestion of Peptide C-2 residue 26 seems to be serine or alanine. When the peptides obtained from partial acid hydrolysis of Peptide C-2 are examined below, it also becomes evident that residue 26 could possibly be serine or alanine. This might be another example of the heterogeneity within LBI IV with the conservative replacement at residue 26 of serine, alanine.

To establish the rest of the sequence of Peptide C-2 by the method of overlapping peptides, it was decided on a partial acid hydrolysis. 1.5 μmoles of Peptide C-2 were incubated with 5.7N HCl at 105°C for 15 minutes in a sealed evacuated tube. The reaction was terminated by freezing, the tubes were opened and the contents were taken to dryness at reduced pressure.

A peptide map revealed about 20 components which were separated first into 6 fractions by paper electrophoresis at pH 4.7. Each electrophoretic fraction was then resolved by paper chromatography in Solvent II. Only 9 peptides were isolated in reasonably pure form, but they were recovered in yields sufficient only for amino acid analysis. It was also observed that the carboxymethylcysteine in the original parent peptide had been recovered in the peptides as cysteic acid. The amino acid compositions of the peptides derived from partial acid hydrolysis of Peptide C-2 are listed in Table XVII.

Peptides C-2-PA-5, C-2-PA-6, C-2-PA-7 were obviously derived from the known sequence at the COOH-terminal end of the parent, Peptide C-2. Peptide C-2-PA-6 overlaps with the known COOH-terminal end of Peptide C-2. Peptides C-2-PA-3 and C-2-PA-4 are related to Peptide C-2-PA-5. Peptides C-2-PA-1 and C-2-PA-2 are related. The sum of Peptides C-2-PA-1, C-2-PA-5 and C-2-PA-8 (serine) represent Peptide C-2. Since Peptide C-2-PA-5 represents the COOH-terminal portion of Peptide C-2, the peptide which occupies the NH_2-terminal portion has to begin with aspartic acid which is the known NH_2-terminal of Peptide C-2. This condition is met by Peptide C-2-PA-1 as well as Peptide C-2-PA-2.

Peptide C-2 represents the residues 9 through 28 of Peptide T-1 (see Page 57) when the following points are considered: the composition of Peptide C-2 accounts for residues

9 through 28 of Peptide T-1; the hexapeptide COOH-terminal portion of Peptide C-2 and Peptide T-1 are identical; Peptides C-2-PA-1 and C-2-PA-2 include the residues 9 through 11 and 9 through 13 of Peptide T-1; Peptides C-2-PA-3 and C-2-PA-4 include residues 15 through 17 and 14 through 17 of Peptide T-1; Peptide C-2-PA-5 includes residues 13 through 28 of Peptide T-1; Peptides C-2-PA-6 and C-2-PA-7 include residues 23 through 26 and 26 through 28 of Peptide T-1.

Although the entire sequence of Peptide C-2 was not established, the peptides obtained by the partial acid hydrolysis confirm the sequence of residues 9 through 28 as established in the sequence work on Peptide T-1. Furthermore, Peptide C-2 substantiates the sequence of Peptide T-1-P-2 in this manner Peptide T-1-P-2a $\longrightarrow$ T-1-P-2b with the former in the NH_2-terminal position (see Page 58).

PEPTIDE C-2a (Residues 9 through 25)

The amino acid composition of this peptide is shown in Table VIII. This peptide is related to Peptide C-2 except that it is three amino acid residues shorter. However, the analysis value of carboxymethylcysteine is low. Some half cystine was present in the analysis, but the cysteic acid was not analyzed. Destruction of carboxymethylcysteine to cystine and cysteic acid during acid hydrolysis could result in the low value of carboxymethylcysteine. Since the cysteic acid was not analyzed, the true content of carboxymethylcysteine in

this peptide was not reflected in the analysis.

Carboxypeptidase digestion released 0.8 μmole/μmole of leucine and one step of the Edman established aspartic acid as the NH_2-terminal. No further work was done on this peptide due to insufficient material. However, this subpeptide of Peptide C-2 with leucine at the COOH-terminal indicates that cleavage by chymotrypsin had occurred at this point. Since the yield of this peptide was much lower than that of Peptide C-2, it would seem that this was cleaved at a lower rate.

PEPTIDE C-3 (Residues 29 through 39): Ser-Ileu-Pro-Pro-(Glx, CMCys,Arg)-CMCys-Thr-Asp-Leu-

Table XVIII summarizes the sequence studies on this peptide. This peptide was isolated from Peak III (Figure 12) and its purification is illustrated in Figures 13 and 14. An overnight digestion with carboxypeptidase A and B showed the COOH-terminal to be CMCys-Thr-Asp-Leu. The NH_2-terminal Ser-Ileu-Pro-Pro was established by the Edman degradation. A comparison of these results together with the sequence of Peptide T-2 and Peptide T-3, would provide the overlap linking Peptide T-2 to Peptide T-3 with the former in the NH_2-terminal position to the latter. In order to establish the identity of residue 33 (Glx) an attempt was made to degrade Peptide C-3 by enzymatic digestion with aminopeptidase M. This was unsuccessful as the enzyme preparation could remove only the serine residue.

PEPTIDE C-3a (Residues 29 through 39): (Ser,Ileu,Pro,Pro,Glx, CMCys,Arg,CMCys,Ser,Asp)-Phe-

Carboxypeptidase A digestion released phenylalanine. Due to lack of material no further work was done on this peptide. Peptide C-3a seems to be homologous to Peptide C-3. By comparison with Peptides T-2 and T-3a, this Peptide C-3a would seem to link Peptide T-2 to Peptide T-3a.

PEPTIDE C-4 (Residues 40 through 45): Arg-Leu-Asp-Ser-CMCys-His-

This peptide was isolated in pure form from Peak V (Figure 12). Table XIX summarizes the sequence studies on this peptide. An overnight digestion with carboxypeptidase A and B released His,1.00; CMCys,0.80; Ser,0.65. The COOH-terminal would be Ser-CMCys-His. Three steps of the Edman degradation established the NH_2-terminal as Arg-Leu-Asp. The entire sequence was thus established.

PEPTIDE C-5 (Residues 46 through 55): Ser-CMCys-Ala-Lys-CMCys-CMCys-Ileu-Ser-Thr-Leu-

This peptide was isolated from Peak III. Its purification is illustrated in Figures 13 and 14. Table XX summarizes the sequence studies on Peptide C-5. An overnight digestion of this peptide with carboxypeptidase A and B released Leu,1.00; Thr,0.30; Ser,0.24; indicating the COOH-terminal to be Ser-Thr-Leu. The rest of the peptide (1.5 µmoles) was subjected to tryptic digestion.

Trypsin digestion of Peptide C-5

To 1.5 µmoles peptide in 200 µl 0.1M NH_4HCO_3, pH 7.8

was added 40 µl of TPCK-trypsin solution (64 µg/40 µl water). Digestion was performed at 37°C for 4 hours and was terminated by freezing. A peptide map of an aliquot showed the presence of 2 peptides. These were separated by paper electrophoresis at pH 4.7. From the specificity of trypsin hydrolysis, Peptide C-5-T-1 (Ser(1), CMCys(1), Ala(1), Lys(1)) would be placed at the NH_2-terminal end of the original peptide. The other peptide was Peptide C-5-T-2 (CMCys(2), Thr(1), Ser(1), Ileu(1), Leu(1)).

Peptide C-5-T-1 (Residues 46 through 49): Ser-CMCys-Ala-Lys-

Two steps of the Edman degradation established the NH_2-terminal Ser-CMCys. Lysine is placed at the COOH-terminal on the basis of the specificity of trypsin digestion. By difference the alanine is placed before lysine. So the sequence is as shown above.

Peptide C-5-T-2 (Residues 50 through 55): CMCys-CMCys-Ileu-Ser-Thr-Leu-

There was sufficient peptide material for four steps of the Edman degradation. Three steps gave the NH_2-terminal sequence of CMCys-CMCys-Ileu. The values for serine are inconsistent. However, Peptide C-5-T-2 was derived from the parent Peptide C-5, whose COOH-terminal was established as Ser-Thr-Leu. On this basis, the proposed sequence of Peptide C-5-T-2 is as shown above.

PEPTIDE C-6 (Residues 56 through 68): Ser-Ileu-Pro-Ala-Gln-CMCys-Val-Thr-(Ileu,Asp)-Asx-Asp-Phe-

This peptide was from Peak I of the cation exchanger and would

be the most acidic of the chymotryptic peptides. Table XXI summarizes the sequence studies on this peptide. An overnight digestion with carboxypeptidase A resulted in the release of only phenylalanine. Three steps of the Edman degradation established the NH_2-terminal as Ser-Ileu-Pro. To establish the rest of the sequence by means of overlapping peptides, an attempt was made at the enzymic digestion of the peptide. Papain as well as thermolysin failed to degrade the peptide. Perhaps the proper experimental conditions had not been employed. However, an enzymatic digestion with pronase was successful and allowed much of the sequence to be established. Pronase digestion of 1.5 µmole of this peptide and separation of the resulting peptides by paper chromatography in Solvent II and paper electrophoresis at pH 4.7 yielded 12 peptides, of which 8 were obtained in reasonably pure form. Their amino acid compositions are listed in Table XXI. No further work could be done on these peptides because of the limiting amounts that were recovered.

The complete sequence of Peptide C-6 was elucidated through a different approach. Limited proteolysis of LBI IV by chymotrypsin at low pH's results in the cleavage of one peptide bond Leu_{55}-Ser_{56}. On reduction and alkylation of the modified inhibitor, two peptides were obtained, one comprising the NH_2-terminal portion of 55 amino acids and the other comprising the COOH-terminal portion of 29 amino acids (Krahn and Stevens, 1970) with an NH_2-terminal Ser-Ileu-Pro. This COOH-

terminal portion of the protein comprised the sum total of amino acids in Peptide C-6 and Peptide C-7 and is here referred to as Peptide C-(6+7).

Peptide C-(6+7) (Residues 56 through 84)

4 μmoles were subjected to chymotrypsin digestion for 4 hours at 37°C and the reaction was terminated by freezing. A peptide map of the digest showed 2 major components and several minor components. The components were separated by paper chromatography in Solvent II and paper electrophoresis at pH 4.7. The two major components were Peptide C-6 and Peptide C-7. Two minor components, Peptide C-6a (Ser(1), Ileu(1), Pro(1), Ala(1), Glu(1), Val(0.5), Thr(0.5)) and Peptide C-6b (CMCys(1), Val(1), Thr(1), Ileu(1), Asp(3), Phe(1)) were derived from the parent Peptide C-6. Three steps of the Edman degradation on Peptide C-6b gave the sequence CMCys-Val-Thr. This data, together with the results on the pronase peptides, would establish the sequence of Peptide C-6.

An enzymatic digestion of Peptide C-6 with aminopeptidase M was unsuccessful. The presence of glutamine was confirmed by digestion of Peptide C-6-P-3 with aminopeptidase M and identification of the amide on paper electrophoresis. Peptide C-6 was the most acidic of the chymotryptic peptides and the derived pronase peptides - Peptides C-6-P-6, C-6-P-7 and C-6-P-8 were all acidic. The dipeptides, C-6-P-6 and C-6-P-8, were

acidic so aspartic acids are probably present. However, from these results, residue 66 remains as Asx.

PEPTIDE C-7 (Residues 69 through 83): CMCys-Tyr-Glu-CMCys-Pro-Lys-(Ser,Ser,His,Ser,Asp,Asp,Asp,Asn, Asn-COOH).

This peptide was isolated from Peak IV (Figure 12). Table XXII summarizes the sequence studies on it. Digestion with carboxypeptidase B and A released asparagine. The Edman degradation indicated the NH_2-terminal sequence of CMCys-Tyr-Glu. However, the value for tyrosine was low in step 1, which might be due to destruction during acid hydrolysis. The assigned position for tyrosine was confirmed from a leucine aminopeptidase digest of Peptide C-7, which released CMCys,1.00; Tyr,0.53; Glu,0.33. The presence of tyrosine in position 70 indicates its resistance to chymotrypsin digestion. This could be due to the presence of proline in the environment of the potential tyrosine site. In order to establish the rest of the sequence, Peptide C-7 was hydrolyzed with trypsin which should cleave the lysine-X bond.

Trypsin digestion of Peptide C-7

1 µmole of Peptide C-7 was subjected to an overnight digestion with trypsin at 37°C. Two major components were isolated by paper chromatography in Solvent II.

Peptide C-7-T-1 (Residues 69 through 74): CMCys-Tyr-Glu-CMCys-Pro-Lys-

An overnight digestion with carboxypeptidase B and A released only lysine. Leucine aminopeptidase digestion

released CMCys,1.00; Tyr,0.50; Glu,0.25; establishing the NH_2-terminal sequence of CMCys-Tyr-Glu. Peptide C-7-T-1 is assigned the NH_2-terminal portion of the parent Peptide C-7 on the basis of trypsin specificity and from the known NH_2-terminal sequence of Peptide C-7, CMCys-Tyr-Glu. Proline is known to be resistant to carboxypeptidase digestion. It is also known that leucine aminopeptidase digestion stops at one residue before proline. Based on these considerations and the results of carboxypeptidase digestion on Peptide C-7-T-1 and leucine aminopeptidase digestion on Peptides C-7 and C-7-T-1, the proline in Peptide C-7-T-1 is assigned as residue 73. Of the two carboxylmethylcysteines in Peptide C-7-T-1, one is established as residue 69, the only position left for the other is as residue 72.

Peptide C-7-T-2 (Residues 75 through 83): His,(1); Asp,(5); Ser,(3).

This peptide was identical in composition to tryptic Peptide T-6 except for the value of aspartic acid. No sequence work was done on it. However from results of the sequence analysis of Peptide T-6 it is possible that chymotrypsin had hydrolyzed the Asn_{83}-Asn_{84}. Unfortunately free asparagine was not detected during the purification procedures of the chymotryptic peptides.

IV. AMINO ACID SEQUENCE OF LBI IV

Figure 20 shows the complete amino acid sequence of LBI IV as reconstructed from the sequence information obtained in the tryptic and chymotryptic peptides as well as the amino acid compositions of peptides obtained from a tryptic digest of guanidinated RCM-LBI IV. Attempts to determine the amino-terminal and carboxyterminal residues in the denatured protein were unsuccessful.

Evidence for the Arrangement of Residues 29 through 84

The arrangement of Peptides T-2 through T-6 (residues 29 through 84) is well substantiated by the appropriate overlaps provided by the chymotryptic peptides and in some cases also by the tryptic peptides obtained from the guanidinated protein.

The sequence studies on Peptide C-3 and C-3a indicate that they connect Peptide T-2 $\longrightarrow$ Peptide T-3 or T-2 $\longrightarrow$ T-3a. From the considerations of the carboxylterminal of Peptides T-3 and T-3a the chymotryptic peptide which follows Peptide C-3 should be one with arginine at its aminoterminal. Only one chymotryptic peptide meets this requirement. This is Peptide C-4, so Peptide C-3 $\longrightarrow$ Peptide C-4. In turn Peptide C-4 has a sequence which shows that it connects Peptide T-3 with a tryptic peptide whose aminoterminal sequence is Leu-Asp-Ser and the only tryptic peptide which meets this requirement is Peptide T-4. From the carboxylterminal portion of Peptide T-4, the chymotryptic peptide which follows Peptide C-4 should have

the aminoterminal Ser-CMCys-Ala-Lys. This condition is met only by Peptide C-5.

The results of partial proteolysis of LBI by chymotrypsin (Krahn and Stevens, 1970) and the sequence studies on the resulting small fragment Peptide C-(6+7) establishes the arrangement of Peptides C-6 and C-7 and also places them at the carboxyterminal end of the protein. Furthermore the amino acid composition of Peptide Tg-4 accounts for the amino acids in Peptides C-7, C-6, C-5 and part of C-4, therefore it establishes an overlap between Peptides C-5 and C-6. This arrangement is further substantiated by the fact that the amino acid composition of Peptide T-5 accounts for all of Peptide C-6, the carboxyterminal portion of Peptide C-5 and the aminoterminal portion of Peptide C-7.

Evidence for the Placement of Peptide T-1 as the Amino-Terminal Fragment

Because of the peculiar susceptibility of the Lys_{28}-Ser_{29} peptide bond to hydrolysis by both trypsin and chymotrypsin and by trypsin after conversion of Lys_{28} to homoarginine no overlaps linking residues 1 through 28 to residues 29 through 84 was obtained. However the results of the acrylamide gel electrophoresis in the presence of SDS and mercaptoethanol (Page 39) would indicate that LBI IV is made up of a single polypeptide chain. On the basis of trypsin specificity it can then be argued that the only position that Peptide T-1 (residues 1 through 28) can occupy is at the aminoterminal of

the protein. Added support for this argument can be obtained from results of partial proteolysis by chymotrypsin on LBI III (Krahn and Stevens, 1970). It was shown that partial proteolysis of LBI III (98 residues) resulted in two fragments. The smaller one (29 residues) came from the carboxylterminal end of LBI III and is identical in composition to the carboxylterminal end of LBI IV (Peptide C-(6+7)). The larger fragment (69 residues) accounted for the aminoterminal portion of LBI III. By analogy partial proteolysis by chymotrypsin of LBI IV would result in a small 29 residue fragment and a larger 55 residue fragment. Since the overlaps of residues 29 through 84 are well established Peptide T-1 has to be assigned as the aminoterminal fragment of the original protein. The only argument against this arrangement is the fact that hydrolysis occurred at Lys_{28} by chymotrypsin, and in the guanidinated derivative at $homoarginine_{28}$ by trypsin. The peculiar behaviour of this labile peptide bond will be discussed in the section on Discussion.

DISCUSSION

I. AMINO ACID SEQUENCE DETERMINATION OF LBI IV

Prior to any sequence analysis of naturally occurring inhibitors of proteolytic enzymes, it is necessary to denature them before they can be enzymatically digested. In the case of LBI IV the seven disulfide bonds were reduced and then blocked by alkylation. Under these conditions, the protein was completely denatured as reflected by the loss of inhibitory activity. That the product of reduction and alkylation was a single polypeptide chain was shown by its disc gel electrophoresis pattern in the presence of SDS and β-mercaptoethanol.

LBI IV contains four lysines and two arginines which means that seven peptides are expected in the tryptic digest of this protein. Seven peptides were indeed isolated from a trypsin digest, but an inspection of their amino acid compositions showed that three of them contained arginine and one peptide (Peptide T-1) contained two lysines. All seven peptides were subjected to sequence analyses except Peptide T-5 which was sparingly soluble and was lost through mishandling.

Hydrolysis of peptides with leucine aminopeptidase was valuable in supplementing the results of the Edman degradation as well as establishing the presence of asparagine, aspartic acid, glutamine and glutamic acid. Aminopeptidase M was also used with the hope of a complete enzymatic hydrolysis of the

peptides. However the commercial sample of the enzyme failed to cleave at proline residues so the purpose in employing the enzyme was defeated.

Peptide T-1 was the largest tryptic peptide (28 residues) that had to be sequenced. It was eluted in three peaks from the Dowex 50-X2 column. It may be that deamidation products were eluted in the earlier peaks. It was found that Peptide T-1 contained a Lys-Pro bond which had not been cleaved by trypsin, as can be expected since bonds of the type Arg-Pro and Lys-Pro are known to be resistant to attack by trypsin. Because of its narrow specificity chymotrypsin was chosen as the enzyme for the further fragmentation of this large peptide. It hydrolyzed at the carboxyl side of an asparagine and a carboxymethylcysteine residue. Pronase, which is known to be a mixture of endopeptidases was useful in degrading Peptide T-1 into several small peptides. A combination of digestion with pronase and with chymotrypsin followed by pronase yielded sufficient overlapping peptides to place the 28 residues in sequence.

The results of carboxypeptidase hydrolysis of Peptide T-1 indicated the presence of alanine in position 26. However in the sequence analysis of the corresponding chymotryptic peptide as well as the corresponding peptide obtained from a tryptic digest of guanidinated LBI IV position 26 was occupied by serine. It would seem that there are two variant forms of LBI IV with respect to position 26; these two forms

differ by a serine-alanine replacement. In general, in other examples of homologous proteins, a serine-alanine replacement is not uncommon.

The determination of the amino acid sequence of Peptide T-6 was complicated by the presence of a large number of aspartic acid, asparagine and serine residues. Fortunately, after extensive digestion of Peptide T-6 with aminopeptidase M, the residual peptide was found to contain the six Asx residues originally present in T-6. It was possible to perform five steps of the Edman degradation on this residual peptide and the PTH derivative could be identified after each step. The results were substantiated by examining the electrophoretic mobility of the residual peptide after each step of the Edman degradation.

At first glance it is difficult to reconcile the presence of three unique arginine containing peptides (T-2, T-3 and T-3a) in the tryptic digest of LBI IV with the amino acid analysis of the protein, which shows only two arginine residues per molecule. However, closer examination of the sequences of Peptides T-3 and T-3a showed them to be identical except for a threonine-serine replacement in position 37 and a leucine-phenylalanine replacement in position 39. The corresponding chymotryptic peptides showed the same replacement pattern. The replacements at both positions are considered conservative; in position 37 an aliphatic alcohol group is retained, while in position 39 a bulky aliphatic hydrophobic side chain is

replaced by a bulky aromatic hydrophobic side chain. The sum of the yields of Peptides T-3 and T-3a is about equal to the yield of Peptide T-2 and therefore it is suggested that they are variants of the same peptide and are the result of microheterogeneity in the original protein preparation. This is in agreement with the non integral values obtained for phenylalanine (1.36) and leucine (3.38) in the amino acid analysis of this preparation. The significance of this heterogeneity as observed in positions 26, 37 and 39 will be discussed later.

With one exception (peptide bond $lysine_{28}$-$serine_{29}$) the specificity exhibited by chymotrypsin towards LBI IV was in agreement with that towards other proteins as reported in the literature (Hill, 1965). Hydrolysis appeared to be quantitative in most cases as judged from the absence of overlapping chymotryptic peptides. In the original chymotryptic digest complete cleavage occurred at the carboxyl side of Thr_8, Phe_{39}, Phe_{68}, His_{45}, Leu_{39} and Leu_{55} and partial cleavage also occurred at Leu_{25}. During the further extensive digestion of Peptide T-1, chymotrypsin also cleaved at Asn_{20} and $CMCys_{22}$. The single tyrosine in position 70 was resistant to chymotrypsin digestion, probably due to the presence of a proline residue nearby.

$Lysine_{28}$ was cleaved by chymotrypsin. When it was chemically modified to homoarginine it was also cleaved by trypsin. This seemingly anomalous behaviour of trypsin and

chymotrypsin has been previously documented in sequence studies of other proteins. For example, Bradshaw et al. (1969) in reporting the structure of a cyanogen bromide fragment of carboxypeptidase A noted that an asparagine-proline bond was partially cleaved by both trypsin and chymotrypsin. Also Elzinga et al. (1968) have reported cleavage of a tyrosine-valine peptide bond in carboxypeptidase B by trypsin. Apparently in LBI IV the Lys_{28}-Ser_{29} bond is highly susceptible to proteolysis. It may be that some structural feature, remaining even after denaturation, is responsible for the ease with which this bond is cleaved, regardless of the specificity of the enzyme used.

Peptides resulting from the chymotryptic digest of LBI IV proved very useful in the alignment of the tryptic peptides. Furthermore the sequence studies on Peptides C-5, C-6 and C-7 provided the information needed to establish the sequence of residues 50 through 74 since, because of its insolubility, Peptide T-5 was isolated in amounts insufficient for sequence analysis.

II. THE OCCURRENCE OF A REPETITIVE SEQUENCE IN LBI IV

Genetic, Structural and Functional Implications

As illustrated below, a striking feature of the amino acid sequence of LBI IV is the occurrence of a repetition in sequence in two separate portions of the polypeptide chain.

15 22
Ser-CMCys-Lys-Pro-CMCys-Asn-His-CMCys

43 49
Ser-CMCys-His-Ser-CMCys-Ala-Lys -----

23 34
CMCys-CMCys-Leu-Ser/Ala-Thr-Lys-Ser-Ileu-Pro-Pro-Glx-CMCys

50 61
CMCys-CMCys-Ileu-Ser-Thr-Leu-Ser-Ileu-Pro-Ala-Gln-CMCys

It can be seen that the sequences from residues 23 through 34 and 50 through 61 are clearly homologous. Without assuming any deletions or additions these two regions are identical in 9 out of 12 positions; there is a conservative replacement of Ileu for Leu and the two other replacements involve an alanine-proline and a leucine-lysine substitution respectively. Assuming only one deletion, these homologous sequences can be extended further to include residues 15 through 22 on one hand and 43 through 49 on the other hand.

In the regions of residues 23 through 34 and 50 through 61 it can be seen that in each case 9 amino acids lie in between two cysteine residues and it is therefore tempting to speculate that one could be dealing with two homologous disulfide loops. This striking feature of the amino acid sequence of LBI IV should be considered in terms of genetic origin and structural and functional implications.

The presence of repetitive sequences in proteins has been previously observed (Smith et al., 1968) and is generally

interpreted as involving an extension of shorter peptide chains by a process of duplication of nucleotide sequences within a gene. Some possible mechanisms for achieving this have been reviewed by Dixon (1966). It is generally accepted that the linear arrangement of amino acids in a polypeptide chain predetermines the folding of the protein. In view of this it would not be unlikely that regions containing the repetitive sequences of LBI IV show considerable structural homology.

It was found (Krahn and Stevens, 1970) that the anti-chymotrypsin site of LBI involves a chymotrypsin sensitive Leu-Ser peptide bond located 29 residues in from the carboxy-terminal end of the protein molecule. It can be seen that this bond between residues 55 and 56 is found approximately in the middle of one of the repetitive sequences. The corresponding position in the other repetitive sequence is occupied by Lys_{28}-Ser_{29}. In view of the fact that trypsin inhibitors contain either an Arg-X or a Lys-X peptide bond in their active site against trypsin (Ozawa and Laskowski, 1966) and that LBI has been shown to be of the Lys-X type (Haynes and Feeney, 1968b; Fritz et al., 1969; Krahn and Stevens, 1970) it is tempting to propose the lysyl-seryl peptide bond located in the other repetitive sequence as the anti-trypsin site of LBI IV. At this stage this proposition is purely hypothetical. Support for it can be derived from the following considerations.

There are four lysine residues in LBI IV. Two of these (in positions 49 and 74) are situated in fairly close proximity to the anti-chymotrypsin site, since the inhibitory sites of this double-headed inhibitor are independent and non-overlapping it is not very likely that either of these lysines forms a part of the anti-trypsin site. Of the remaining two lysines, one (position 17) is followed by a proline residue resulting in a Lys-Pro peptide bond which was found to be resistant to trypsin hydrolysis. Therefore $lysine_{28}$, located in the repetitive sequence, is the most likely candidate for the anti-trypsin site of LBI. If $lysine_{28}$ is proved to be the anti-trypsin site one could postulate that the active sites of this double-headed inhibitor are situated at the apexes of homologous disulfide loops.

In terms of its genetic origin one can then consider two possibilities. The ancestral gene was either entirely different from either of the two present day active site region (i.e. it was not a proteolytic inhibitor) or the ancestral gene was in fact very similar to one of the present day active sites (i.e. it was either a trypsin or a chymotrypsin inhibitor). In the first alternative, a process of gene duplication was followed by divergent evolution of the two portions of the duplicated gene. In the second alternative, gene duplication was followed by the divergent evolution of one portion of the gene and the other portion retained the ancestral form.

If lysine$_{28}$ is not proven to be the anti-trypsin site, the argument may be presented in another light. One portion of the duplicated gene conserved the ancestral form (i.e. it was not a proteolytic inhibitor) whilst the other portion evolved, maybe by selective pressure, to become a chymotrypsin inhibitor.

If lysine$_{28}$ is shown to be the anti-trypsin site one could then postulate a very attractive model for this double-headed inhibitor. The two disulfide loops, one containing the site for trypsin and the other for chymotrypsin would be structurally similar and spatially oriented so as not to overlap.

III. HETEROGENEITY OF LBI IV

In the region between residues 35 through 43 there is evidence for two variant forms of LBI IV; these forms are distinguished by a threonine $\longleftrightarrow$ serine replacement at position 37 and leucine $\longleftrightarrow$ phenylalanine replacement at position 39. In the tryptic and chymotryptic digests of denatured LBI IV two homologous peptides were recovered in different proportions. So far the evidence points to the fact that the two replacements are linked; that is each molecule of LBI IV possesses one set of the two replacements. The other position in LBI IV which exhibits a conservative replacement is the serine $\longleftrightarrow$ alanine at position 26. This does not seem to be linked with the other two replacements at positions 37 and 39; that

is either variant form (37, 39) could have serine or alanin at position 26. Since such conservative replacements would probably not alter the biological activity of the protein, the most interesting aspect of this variation is its genetic origin.

The variant forms could either arise from allelomorphism, or from mutational events after the gene duplication process which yielded the double-headed inhibitor. To eliminate one or the other possibility, or both possibilities it would be necessary to survey the distribution of the two forms in a large population of individual plants. If the two forms are allelic, then they should be distributed throughout the population and segregate in accordance with Mendelian laws. The probability of any one plant being homozygous for either type of LBI IV, as opposed to being heterozygous, would be 1:4.

From the present results, it is impossible to state whether the two variant forms are allelic or are the products of divergent evolution after gene duplication.

IV. COMPARISON OF THE AMINO ACID SEQUENCE OF LBI IV WITH THAT OF OTHER INHIBITORS.

The amino acid sequence of LBI IV bears no obvious resemblance to that of the three other inhibitors whose sequences are known - BPTI, Kazal inhibitor and the inhibitor from Ascaris. It also bears no resemblance to the partial sequences

of the Kunitz soybean inhibitor determined by Ikenaka (personal communication). The Bowman-Birk soybean inhibitor is very similar in composition to LBI, however there are no reports on its amino acid sequence.

LITERATURE CITED

Anderer, F. A. (1965a), Z. Naturforsch Pt.b, 20,462.

Anderer, F. A. (1965b), Z. Naturforsch Pt.b, 20,499.

Anderer, F. A., and Hornle, S. (1966), J. Biol. Chem. 241,1569.

Avineri-Goldman, R., Snir, I., Blauer, G., and Rigbi, M. (1967), Arch. Biochem. Biophys. 121,107.

Birk, Y. (1961), Biochim. Biophys. Acta 54,378.

Birk, Y., Gertler, A., and Khalef, S. (1963), Biochem. J. 87,281.

Birk, Y., Gertler, A., and Khalef, S. (1967), Biochim. Biophys. Acta 147, 402.

Bowman, D. E. (1946), Proc. Soc. Exptl. Biol. Med. 63,547.

Bradshaw, R. A., Babin, D. R., Nomoto, M., Srinivasin, N. G., Ericsson, L. H., Walsh, K. A., and Neurath, H. (1969), Biochemistry 8,3859.

Brown, J. R., Lerman, N., and Bohak, Z. (1966), Biochem. Biophys. Res. Commun. 23,561.

Burck, P. J., Hamill, R. L., Cerwinsky, E. W., and Grinnan, E. L. (1967), Biochemistry 6,3180.

Čechová, D., and Muszyńská, G. (1970), FEBS Letters 8,84.

Čechová, D., Švestková, V., Keil, B., and Šorm, F. (1969), FEBS Letters 4,155.

Cerwinsky, E. W., Burck, P. J., and Grinnan, E. L. (1967), Biochemistry 6,3175.

Chauvet, J., and Acher, R. (1966), Bull. Soc. Chim. Biol. 48,1284.

Chauvet, J., and Acher, R. (1967), Bull. Soc. Chim. Biol. 49,985.

Chauvet, J., and Acher, R. (1967a), Biochem. Biophys. Res. Commun. 27,230.

Chauvet, J., and Acher, R. (1967b), J. Biol. Chem. 242,4274.

Chauvet, J., and Acher, R. (1968), FEBS Letters 2,17.

Chauvet, J., Nouvel, G., and Acher, R. (1964), Biochim. Biophys. Acta 92,200.

Chauvet, J., Nouvel, G., and Acher, R. (1966a), Biochim. Biophys. Acta 115,121.

Chauvet, J., Nouvel, G., and Acher, R. (1966b), Biochim. Biophys. Acta 115,130.

Desnuelle, P. (1960), in The Enzymes, Vol. IV, Boyer, P. D., Lardy, H., and Myrback, K., Ed., New York, N. Y., Academic Press, p 128.

Dibella, F. P., and Liener, I. E. (1969), J. Biol. Chem. 244,2824.

Dixon, G. H. (1966), in Essays in Biochemistry, Vol. II, Campbell, P. N., and Greville, G. F., Ed., New York, N. Y., Academic Press, p 128.

Dlouha, V., and Keil, B. (1969), FEBS Letters 3,137.

Dlouha, V., Pospisilova, D., Meloun, B., and Sorm, F. (1965), Coll. Czech. Chem. Commun. 30,1311.

Easley, C. W. (1965), Biochim. Biophys. Acta 107,386.

Edman, P., and Sjoquist, J. (1956), Acta Chem. Scand. 10,1507.

Elzinga, M., Lai, C. Y., and Hirs, C. H. W. (1968), Arch. Biochem. Biophys. 123,353.

Evans, W. H., Landon, M., and Smith, E. L. (1968), J. Biol. Chem. 243,2172.

Feeney, R. E., and Allison, R. G. (1969), The Evolutionary Biochemistry of Proteins, New York, N. Y., Wiley-Interscience Publishers.

Feeney, R. E., Means, G. E., and Bigler, J. C. (1969), J. Biol. Chem. 244,1957.

Ferdinand, W., Moore, S., and Stein, W. H. (1965), Biochim. Biophys. Acta 96,524.

Ferguson, J. H. (1942), Proc. Soc. Exptl. Biol. Med. 51,373.

Fernandez-Diez, M. J., Osuga, D. T., and Feeney, R. E. (1964), Arch. Biochem. Biophys. 107,449.

Finkenstadt, W. R., and Laskowski, M., Jr. (1965), J. Biol. Chem. 240, PC 962.

Finkenstadt, W. R., and Laskowski, M., Jr. (1967), J. Biol. Chem. 242,771.

Fraefel, W., and Acher, R. (1968), Biochim. Biophys. Acta 154,615.

Fraenkel-Conrat, H., Bean, R. C., Ducay, E. D., and Olcott, H. S. (1952), Arch. Biochem. Biophys. 37,393.

Frattali, V. (1969), J. Biol. Chem. 244,274.

Frattali, V., and Steiner, R. F. (1969), Biochem. Biophys. Res. Commun. 34,480.

Fritz, H., Fink, E., Gebhardt, M., Hochstrasser, K., and Werle, E. (1969), Hoppe-Seyler's Z. Physiol. Chem. 350, 933.

Fritz, H., Schult, H., Hutzel, M., Weidemann, M., and Werle, E. (1967a), Hoppe-Seyler's Z. Physiol. Chem. 348,308.

Fritz, H., Hüller, I., Weidemann, M., and Werle, E. (1967b), Hoppe-Seyler's Z. Physiol. Chem. 348,405.

Fritz, H., Woitinas, F., and Werle, E. (1966), Z. Physiol. Chem. 345,168.

Gorini, L., and Audrain, L. (1952), Biochim. Biophys. Acta 8,702.

Gorini, L., and Audrain, L. (1953), ibid. 10,570

Green, N. M., and Neurath, H. (1954), in The Proteins, 1st ed., Vol. IIB, Neurath, H., and Bailey, K., Ed., New York, N. Y., Academic Press, p 1057.

Green, N. M., and Work, E. (1953), Biochem. J. 54,257.

Greene, L. J., and Bartelt, D. C. (1969), J. Biol. Chem. 244,2646.

Greene, L. J., and Giordano, J. S., Jr. (1969), J. Biol. Chem. 244,285.

Greene, L. J., DiCarlo, J. J., Sussman, A. J., Bartelt, D. C., and Roark, D. E. (1968), J. Biol. Chem. 243,1804.

Greene, L. J., Rigbi, M., and Fackre, D. S. (1966), J. Biol. Chem. 241,5610.

Groskopf, W. R., Summaria, L., and Robbins, K. C. (1969), J. Biol. Chem. 244,3590.

Haynes, R., and Feeney, R. E. (1967), J. Biol. Chem. 242,5378.

Haynes, R., and Feeney, R. E. (1968a), Biochim. Biophys. Acta 159,209.

Haynes, R., and Feeney, R. E. (1968b), Biochemistry 7,2879.

Hill, R. L. (1965), Advan. Protein Chem. 20,37.

Hirs, C. H. W. (1967), Methods Enzymol. 11,199.

Hirs, C. H. W., Moore, S., and Stein, W. H. (1956), J. Biol. Chem. 219,623.

Hochstrasser, K., Muss, M., and Werle, E. (1967), Hoppe-Seyler's Z. Physiol. Chem. 348,1337.

Hochstrasser, V. K., Illchmann, K., and Werle, E. (1969a), Hoppe-Seyler's Z. Physiol. Chem. 350,655.

Hochstrasser, V. K., Schramm, W., Fritz, H., Schwarz, S., and Werle, E. (1969b), Hoppe-Seyler's Z. Physiol. Chem. 350,893.

Hochstrasser, V. K., Illchmann, K., and Werle, E. (1969c), Hoppe-Seyler's Z. Physiol. Chem. 350,929.

Hochstrasser, K., and Werle, E. (1969), Hoppe-Seyler's Z. Physiol. Chem. 350,249.

Jirgensons, B. (1961), Arch. Biochem. Biophys. 94,59.

Jirgensons, B., Ikenaka, T., and Gorkuraki, V. (1960), Makromol. Chemie 39,149.

Jones, G., Moore, S., and Stein, W. H. (1963), Biochemistry, 2,66.

Kakade, M. L., Simons, N. R., Liener, I. E. (1970), Biochim. Biophys. Acta 200,168.

Kasper, C. B., and Smith, E. L. (1966), J. Biol. Chem. 241,3754.

Kassell, B. (1964), Biochemistry 3,152.

Kassell, B., and Chow, R. B. (1966), Biochemistry 5,3449.

Kassell, B., and Laskowski, M., Sr. (1956), J. Biol. Chem. 219,203.

Kassell, B., and Laskowski, M., Sr. (1965), Fed. Proc. 24 ,3.

Kassell, B., Radicevic, M., Ansfield, M. J., and Laskowski, M., Sr. (1965a), Biochem. Biophys. Res. Commun. 18,255.

Kassell, B., and Laskowski, M., Sr. (1965b), Biochem. Biophys. Res. Commun. 20,463.

Kassell, B., Radicevic, M., Berlow, S., Peanasky, R. J., and Laskowski, M., Sr. (1963), J. Biol. Chem. 238,3274.

Kazal, L. A., Spicer, D. S., and Brahinsky, R. A. (1948), J. Am. Chem. Soc. 70,3034.

Kimmel, J. R. (1967), Methods Enzymol. 11,584.

Krahn, J., and Stevens, F. C. (1970), Biochemistry 9,

Kraut, H., Frey, E. K., and Bauer, E. (1928), Z. Physiol. Chem. 175,97.

Kraut, H., Korbel, W., Scholtan, W., and Schultz, F. (1960), Z. Physiol. Chem. 321,90.

Kress, L. F., and Laskowski, M., Sr. (1967), J. Biol. Chem. 242,4925.

Kress, L. F., Wilson, K. A., and Laskowski, M., Sr. (1968), J. Biol. Chem. 243,1758.

Kucich, U., and Peanasky, R. J. (1970), Biochim. Biophys. Acta 200,47.

Kunitz, M. (1945), Science 101,668.

Kunitz, M. (1946), J. Gen. Physiol. 29,149.

Kunitz, M. (1947), J. Gen. Physiol. 30,291.

Kunitz, M., and Northrop, J. H. (1936), J. Gen. Physiol. 19,991.

Laskowski, M., Jr., and Laskowski, M., Sr. (1951), J. Biol. Chem. 190,563.

Laskowski, M., and Laskowski, M., Jr. (1954), Adv. Protein Chem. 9,203.

Laskowski, M., Jr., and Sealock, R. W. (1969), Biochemistry 8,3703.

Laskowski, M., and Wu, F. C. (1953), J. Biol. Chem. 204,797.

Liener, I., and Kakade, M. L. (1969), Toxic Constituents of Plant Foodstuffs Liener, I., Ed., New York, N. Y., Academic Press, p 7.

Light, A. (1967), Methods Enzymol. 11,417.

Liu, W. H., Feinstein, G., Osuga, D. T., Haynes, R., and Feeney, R. E. (1968), Biochemistry 7,2886.

Margoliash, E., and Smith, E. L. (1962), J. Biol. Chem. 237,2151.

Meloun, B., Frič, I., and Šorm, F. (1968), Europ. J. Biochem. 4,112.

Millar, D. B. S., Willick, G. E., Steiner, R. F., and Frattali, V. (1969), J. Biol. Chem. 244,281.

Niekamp, C. W., Hixson, H. F., Jr., and Laskowski, M., Jr. (1969), Biochemistry 8,16.

Ornstein, L., and Davis, B. J. (1964), Disc electrophoresis, Distillation Products Industries, Rochester, New York, N. Y.

Ozawa, K., and Laskowski, M., Jr. (1966), J. Biol. Chem. 241,3955.

Peanasky, R. J., and Laskowski, M., Sr. (1960), Biochim. Biophys. Acta 37,167.

Portmann, P., and Fraefel, W. (1967), Helv. Chim. Acta 50,2078.

Pusztai, A. (1967), Nutrition Abstr. Rev. 37,1.

Rackis, J. J., Sasame, H. A., Mann, R. K., Anderson, R. L., and Smith, A. K. (1962), Arch. Biochem. Biophys. 98,471.

Ram, S. J., Terminiello, L., Bier, M., and Nord, F. F. (1954), Arch. Biochem. Biophys. 52,451.

Read, J. W., and Haas, L. W. (1938), Cereal Chem. 15,59.

Reeck, G. R., Winter, W. P., and Neurath, H. (1970), Biochemistry 9,1398.

Rhodes, M. B., Bennett, N., and Feeney, R. E. (1960), J. Biol. Chem. 235, 1686.

Rhodes, M. B., Hill, R. M., and Feeney, R. E. (1957), Anal. Chem. 29,376.

Rigbi, M., and Greene, L. J. (1968), J. Biol. Chem. 243,5457.

Robbins, K. C., and Summaria, L. (1966), Immunochemistry 3,29.

Ryan, C. A., and Clary, J. J. (1964), Arch. Biochem. Biophys. 108,169.

Schroeder, W. A. (1967), Methods Enzymol. 11,351.

Sherman, M. P., and Kassell, B. (1968), Biochemistry 7,3634.

Shields, G. S., Hill, R. L., and Smith, E. L. (1959), J. Biol. Chem. 234,1747.

Smith, E. L., Delange, R. J., Evans, W. H., Landon, M., and Markland, F. S. (1968), J. Biol. Chem. 243,2184.

Spackman, D. H., Stein, W. H., and Moore, S. (1958), Anal. Chem. 30,1190.

Steiner, R. F. (1965), Biochim. Biophys. Acta 100,111.

Steiner, R. F. (1966), Arch. Biochem. Biophys. 115,257.

Steiner, R. F., and Edelhoch, H. (1963), J. Biol. Chem. 238,925.

Steiner, R. F., and Frattali, V. (1969), J. Agr. Food. Chem. 17,513.

Stevens, F. C. (1969), Proc. Can. Fed. Biol. Soc. 12,16.

Stevens, F. C., and Feeney, R. E. (1963), Biochemistry 2,1346.

Stevens, F. C., Glazer, A. N., and Smith, E. L. (1967), J. Biol. Chem. 242,2764.

Tauber, H., Kershaw, B. B., and Wright, R. D. (1949), J. Biol. Chem. 179,1155.

Trop, M., and Birk, Y. (1970), Biochem. J. 116,19.

Tschesche, V. H., Wachter, E., Kupfer, S., and Niedermeier, K. (1969), Hoppe-Seyler's Z. Physiol. Chem. 350,1247.

Vogel, R., Trautschold, I., and Werle, E. (1968), Natural Proteinase Inhibitors, New York and London, Academic Press.

Weber, K., and Osborn, M. (1969), J. Biol. Chem. 244,4406.

Werle, E., and Appel, W. (1959), Z. Naturforsch 14B,648.

Werle, E., Kaufmann-Boetsch, B. (1960), Z. Physiol. Chem. 319,52.

Weyer, E. M. (1968), Ann. N. Y. Acad. Sci. 146,361.

Wu, F. C., and Laskowski, M. (1955), J. Biol. Chem. 213,609.

Wu, Y. V., and Scheraga, H. A. (1962a), Biochemistry 1,698.

Wu, Y. V., and Scheraga, H. A. (1962b), Biochemistry 1,905.

Yamamoto, M., and Ikenaka, T. (1967), J. Biochem. 62,141.

Figure 1 Gel filtration of LBI on Sephadex G-75.

500 mg of crude LBI was applied to a column (2.5 x 80 cm) of Sephadex G-75. The column was equilibrated and developed with 0.1M NH_4HCO_3, pH 7.8 at a flow rate of 20-30 ml/hr. Fractions of 4.5 ml were collected. Details are described in the text.

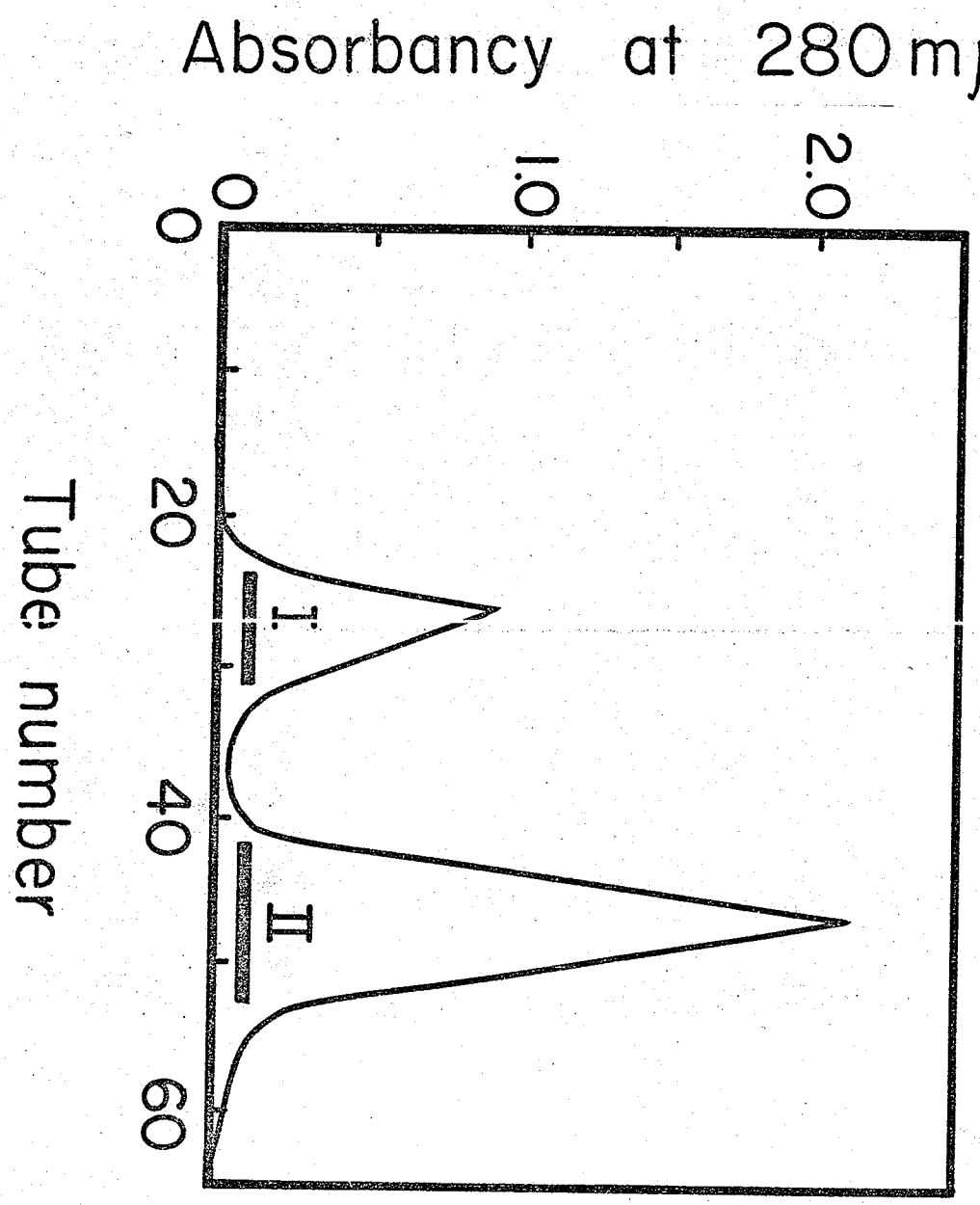

Absorbancy at 280 mμ
0
1.0
2.0
Tube number
0
20
40
60
I
II

Figure 2a Chromatography of LBI on DEAE-cellulose.

500 mg of protein from Peak II (see Figure 1) were chromatographed on a column (2.5 x 80 cm) of DEAE-cellulose. The flow rate was 20 ml/hr and effluent fractions of 4.5 ml were collected. The column was developed by an exponential gradient of salt, 0.01M sodium phosphate and 0.4M sodium chloride, pH 7.6 applied through a 1-liter mixing chamber containing 0.01M sodium phosphate pH 7.6. Details are described in the text.

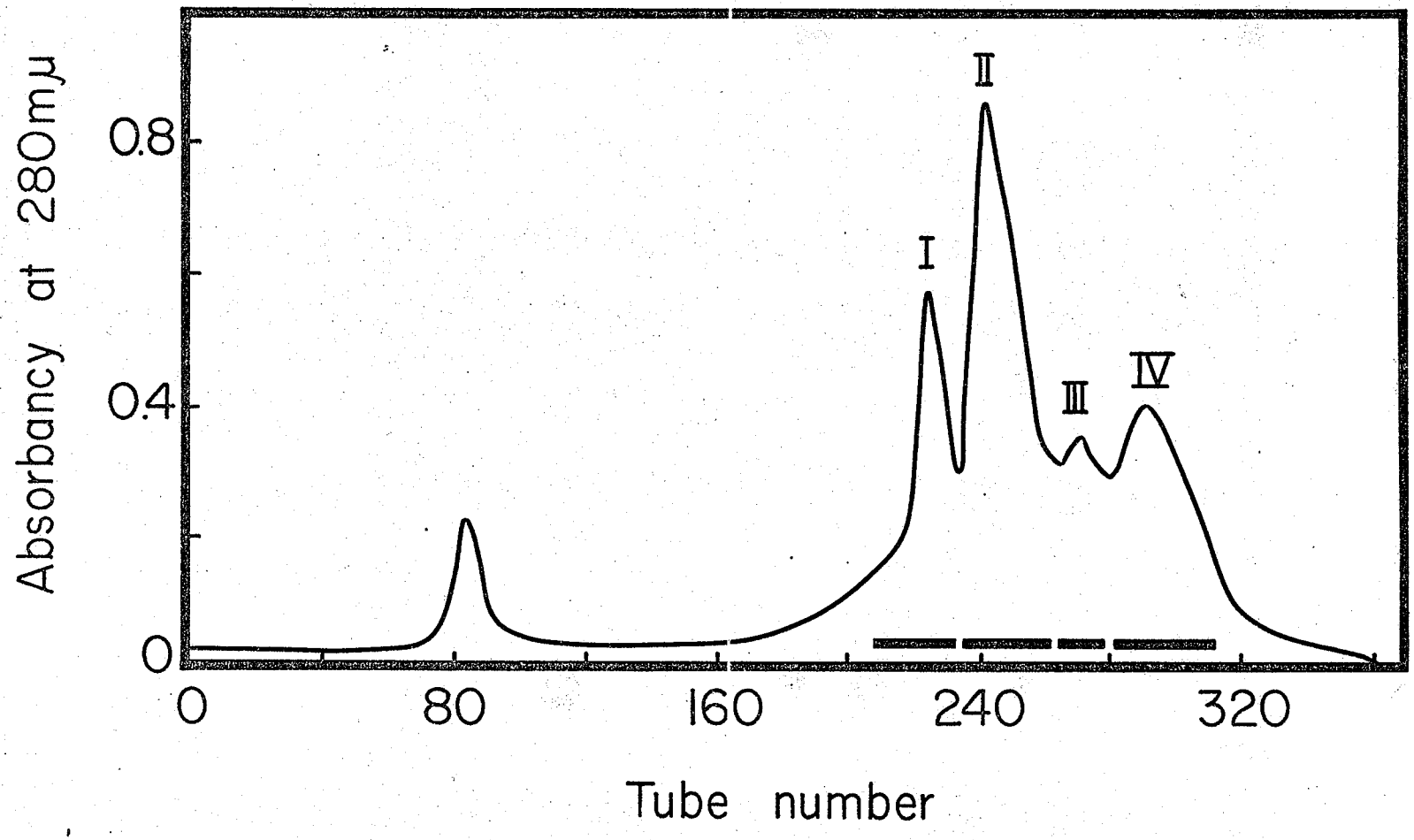
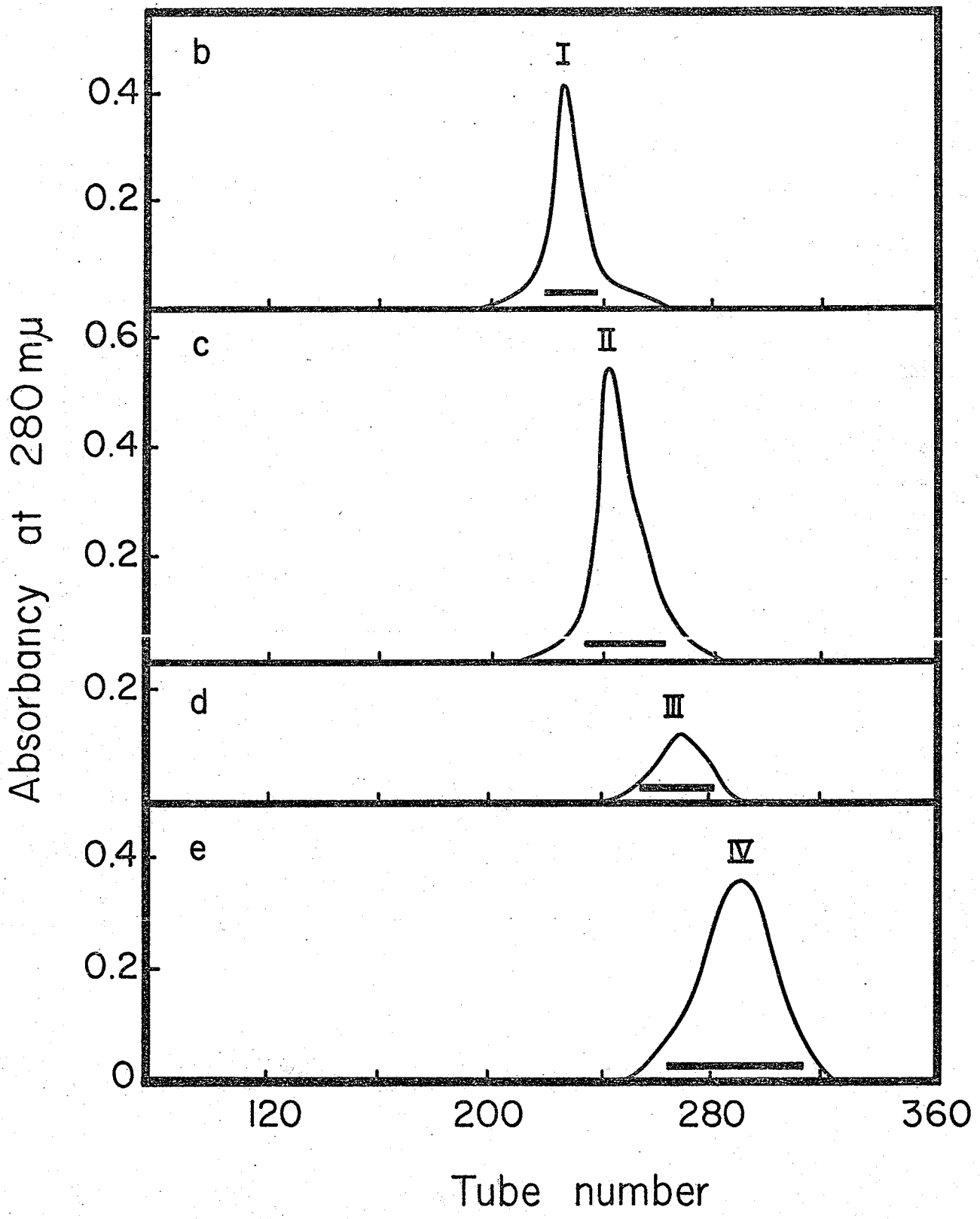

Absorbancy at 280mμ
0.8
0.4
0
0
80
160
240
320
Tube number
I
II
III
IV

Figure 2b, c, d, e Rechromatography of LBI fractions I, II, III and IV on DEAE-cellulose.

The experimental conditions are the same as those described in Figure 2a.


b
I
0.4
0.2
c
II
0.6
0.4
0.2
d
III
0.2
e
IV
0.4
0.2
0
120
200
280
360
Absorbancy at 280 mμ
Tube number

Figures 3 and 4 Standard curves for trypsin and chymotrypsin

The amount of turkey ovomucoid added to 30-40 µg of enzyme is plotted against the transmittance change/unit time (the units of the rate of transmittance change were arbitrarily chosen and are not absolute. The rate in transmittance change is a measure of enzyme activity). The substrate for trypsin was TAMe and the substrate for chymotrypsin was BTEE.

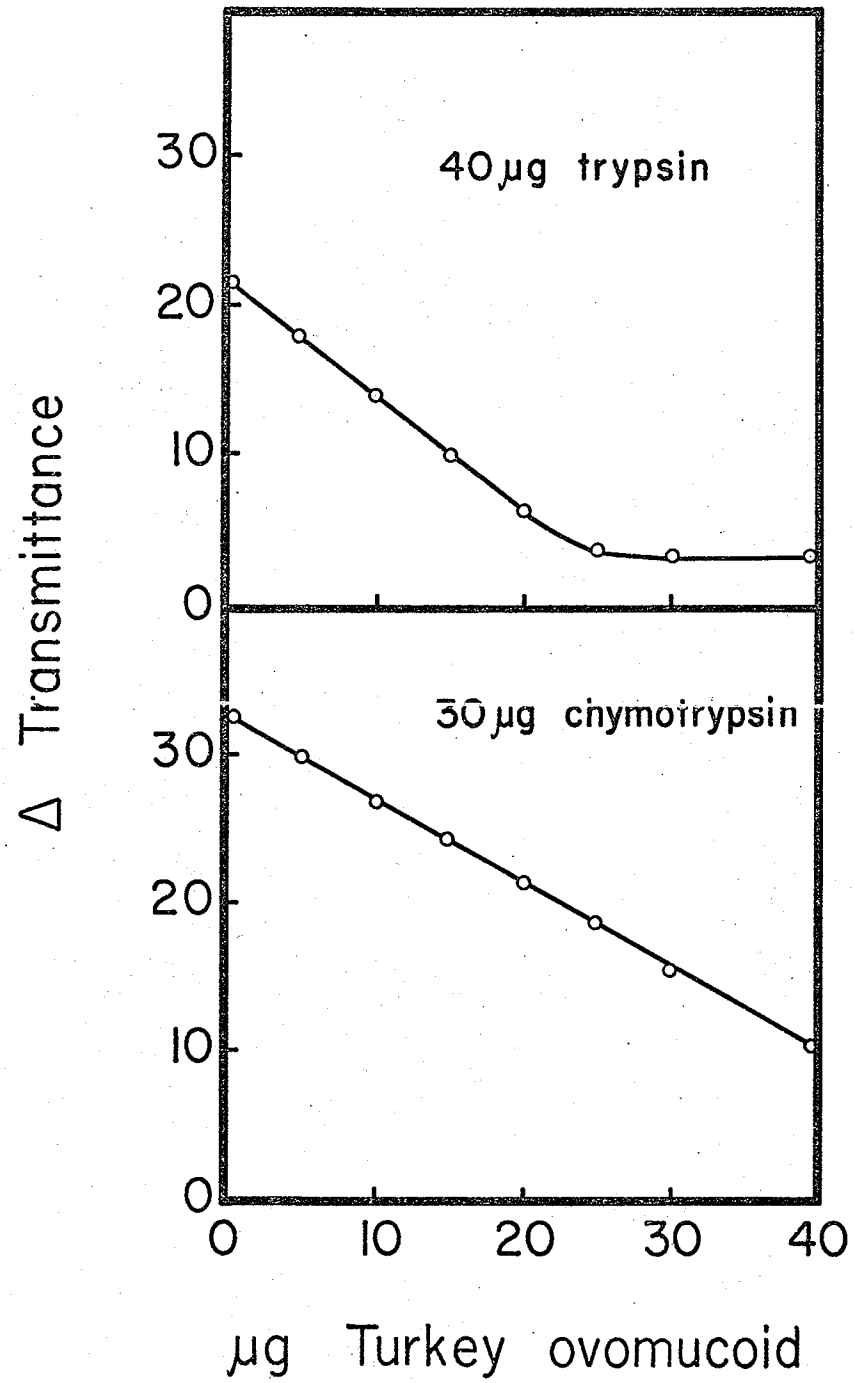

40 µg trypsin
30 µg chymotrypsin
Δ Transmittance
30
20
10
0
0
10
20
30
40
µg Turkey ovomucoid

Figure 5 Patterns of polyacrylamide gel electrophoresis of LBI fractions I, II, III and IV.

Electrophoresis was performed in 0.03M glycine buffer, pH 8.3, at a constant current of 2.5 ma per tube with the positive electrode in the lower chamber. The gels were stained in amido black.

1, Sample of commercial LBI

2, Fraction I

3, Fraction II

4, Fraction III

5, Fraction IV

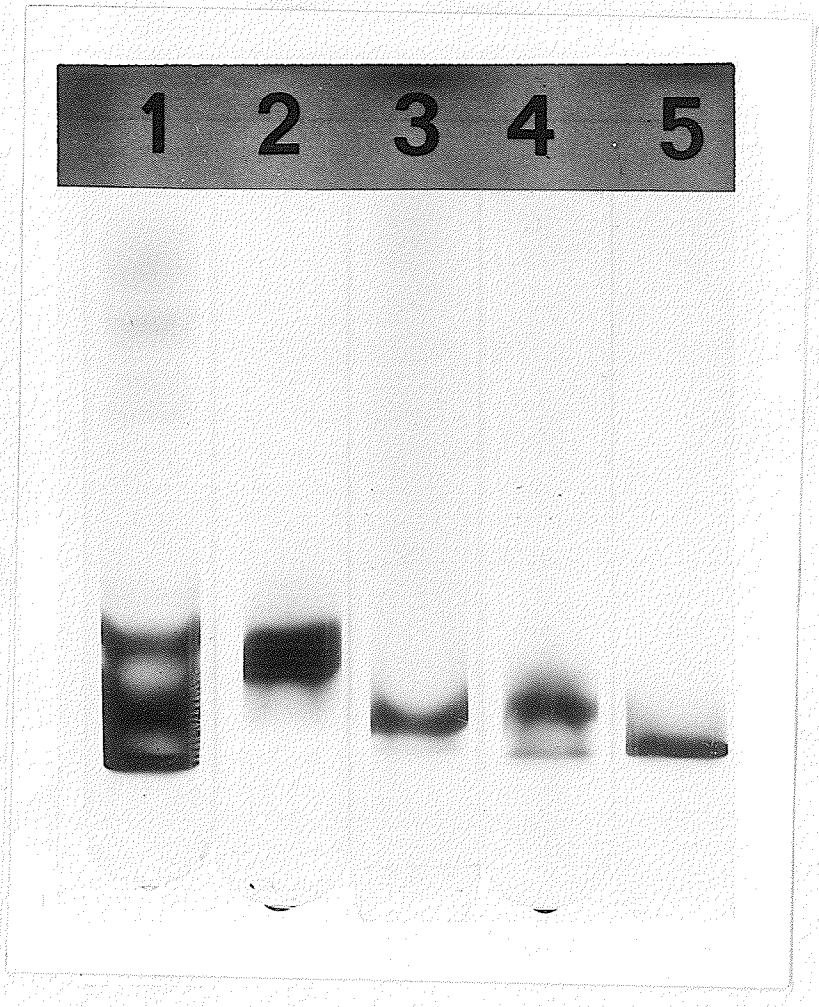

1 2 3 4 5

Figure 6 Patterns of polyacrylamide gel electrophoresis of LBI IV and RCAM-LBI IV in SDS and mercaptoethanol.

Electrophoresis was performed in phosphate buffer pH 7.0 at a constant current of 8 ma per tube with the positive electrode in the lower chamber. The gels were stained in Coomasie brilliant blue. From left to right -

LBI IV in mercaptoethanol

LBI IV

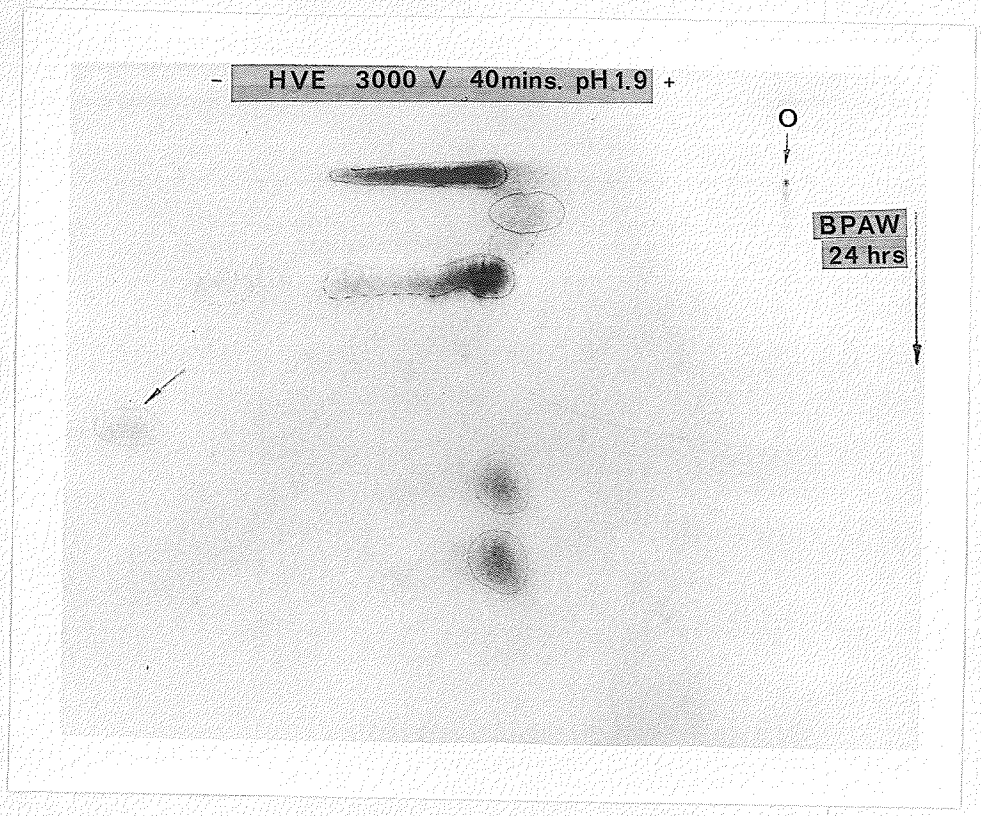
RCAM-LBI IV in mercaptoethanol

RCAM-LBI IV

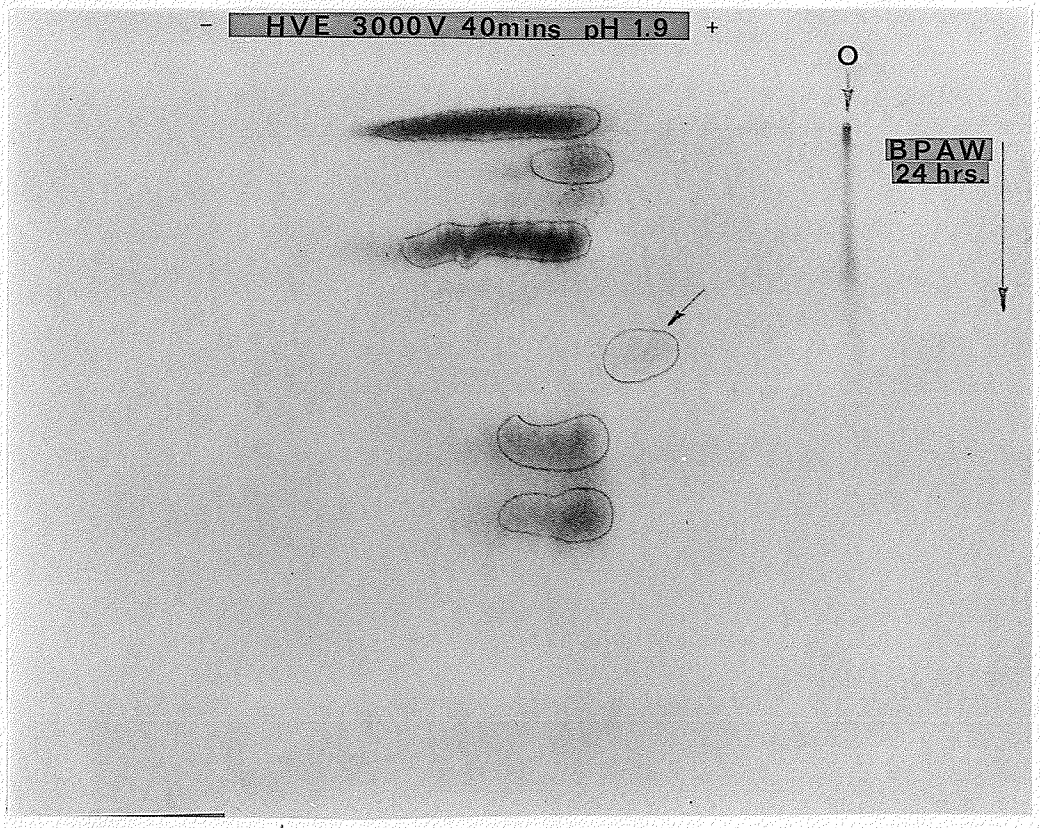
Details are described in the text.

Figure 7 Peptide maps of tryptic digests of LBI I and IV.

0 is the site of application of sample. The peptide maps of the two digests are very similar except for the peptides indicated by the arrows.



Fraction I



Fraction IV

Figure 8 Digestion curve of RCAM-LBI IV (180 mg) with trypsin.

Details are given in the text.

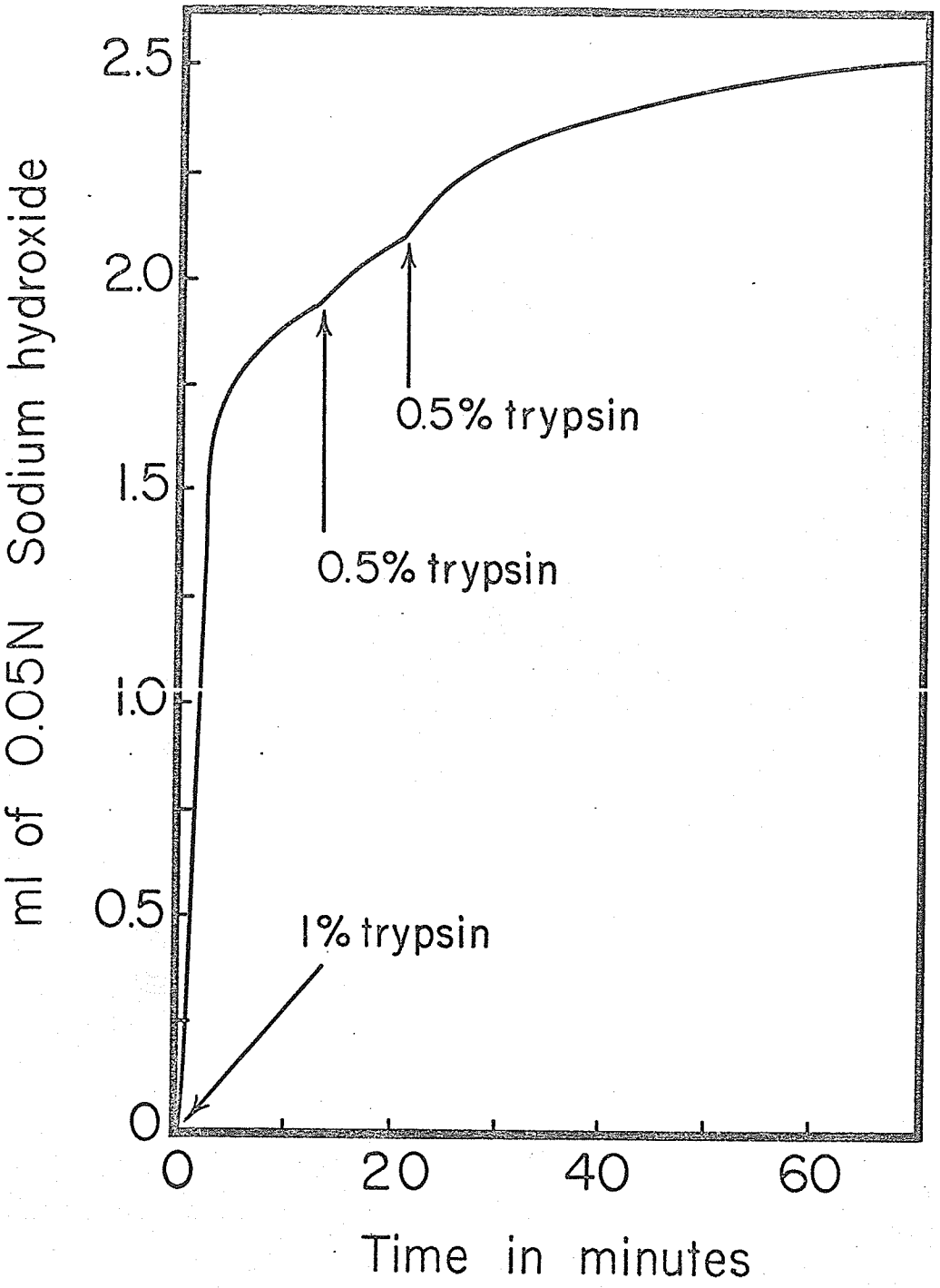

2.5
2.0
1.5
1.0
0.5
0
ml of 0.05N Sodium hydroxide
0.5% trypsin
0.5% trypsin
1% trypsin
0
20
40
60
Time in minutes

Figure 9 Elution pattern of peptides from a tryptic digest of RCM-LBI IV on Dowex 50-X2.

The digest (150 mg) was applied to a column (1.9 x 100 cm) of Dowex 50-X2 with pyridine-acetic acid buffers. The flow rate was 40 ml/hr and fractions of 4.5 ml were collected. Samples of 200 µl each from alternate tubes were analyzed after alkaline hydrolysis by the ninhydrin method. The details are described in the text.

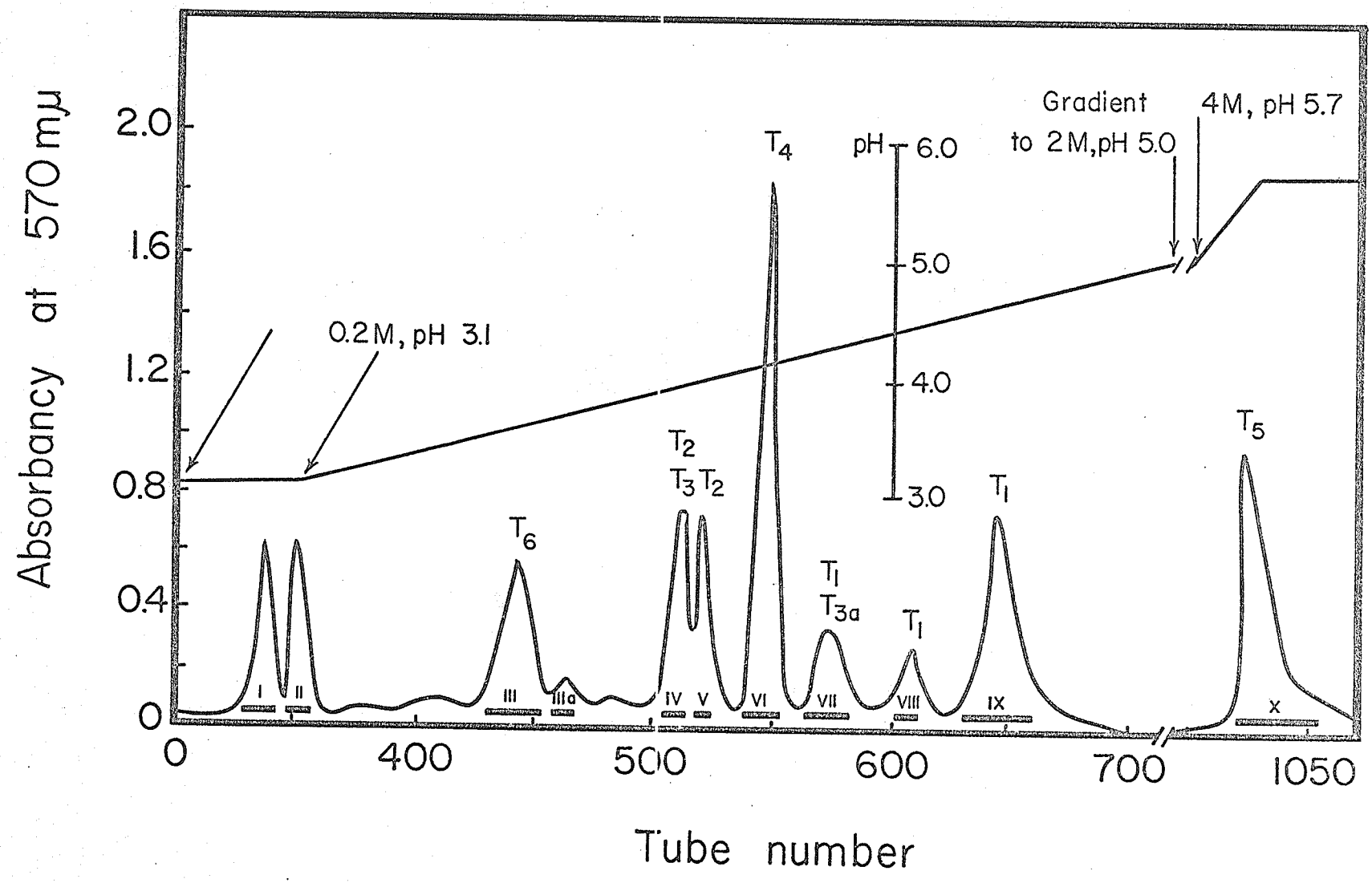
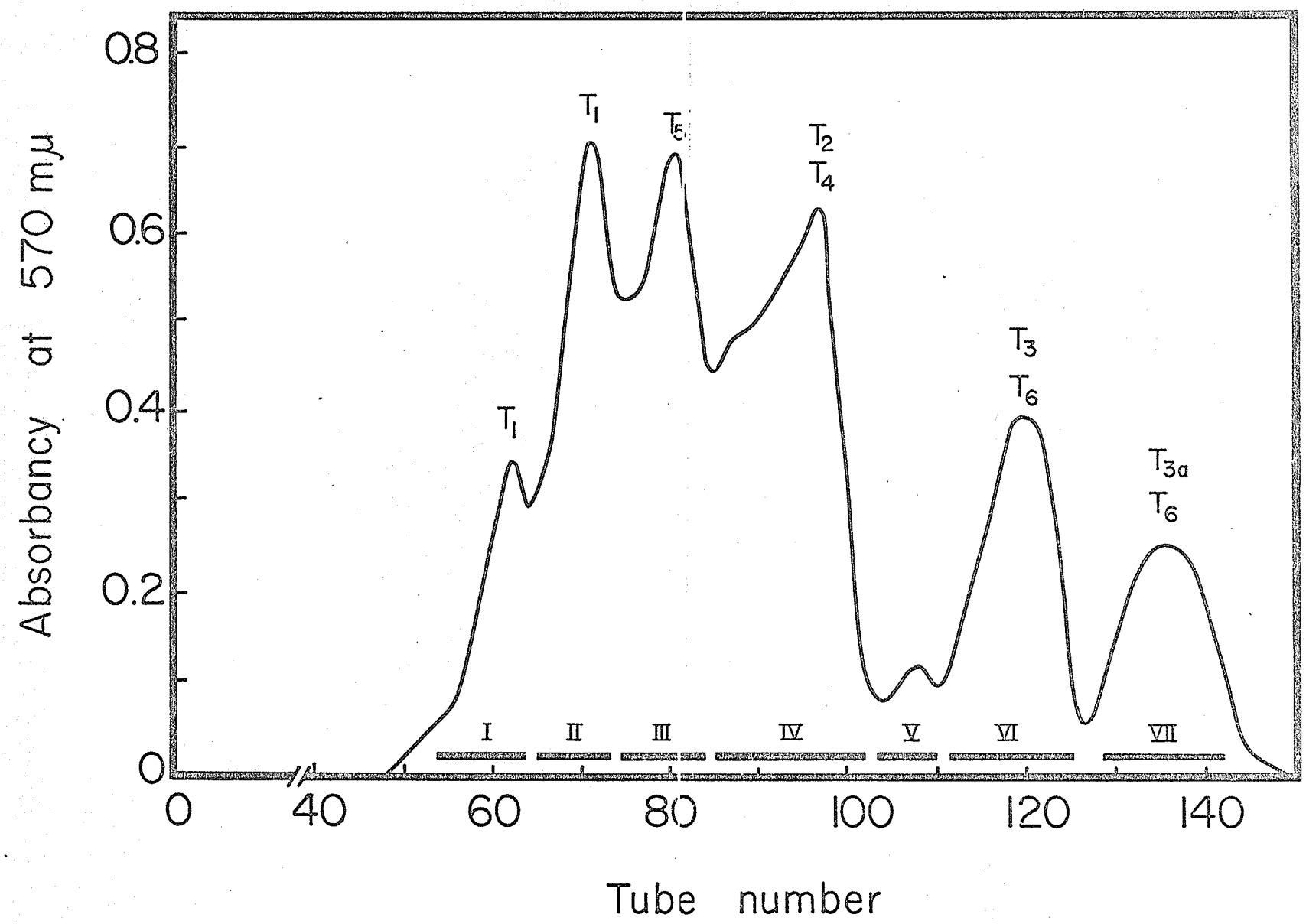

Absorbancy at 570 mμ
2.0
1.6
1.2
0.8
0.4
0
0
400
500
600
700
1050
Tube number
0.2 M, pH 3.1
Gradient
to 2 M, pH 5.0
4 M, pH 5.7
pH
6.0
5.0
4.0
3.0
T4
T6
T2
T3 T2
T1
T3a
T1
T1
T5
I
II
III
IIIa
IV
V
VI
VII
VIII
IX
X

Figure 10 Elution pattern of peptides from a tryptic digest of RCAM-LBI IV on Bio-Gel P-6.

The digest (180 mg) was applied to a column (2.5 x 90 cm) of Bio-Gel P-6. The eluant was 10% acetic acid, the flow rate was 21 ml/hr and fractions of 3.7 ml were collected. Samples of 50 µl each from alternate tubes were analyzed after alkaline hydrolysis by the ninhydrin method. The details are described in the text.


Absorbancy at 570 mμ
0.8
0.6
0.4
0.2
0
0
40
60
80
100
120
140
Tube number
T1
T5
T2
T4
T1
T3
T6
T3a
T6
I
II
III
IV
V
VI
VII

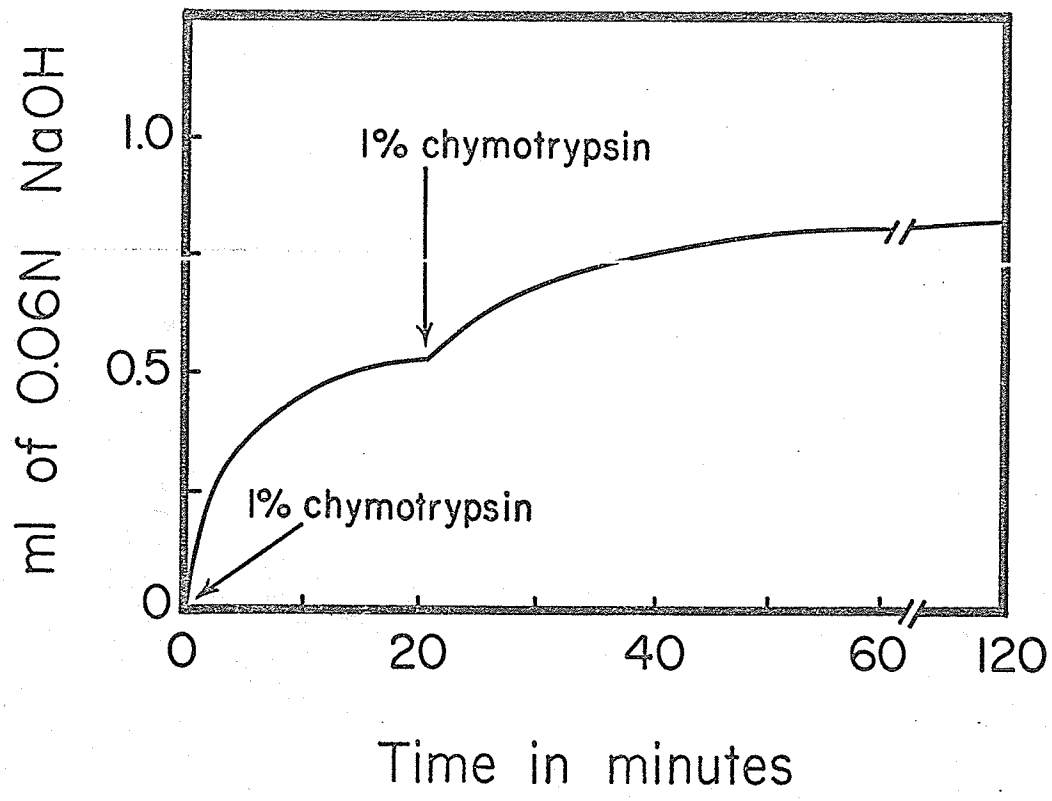

ml of 0.06N NaOH
1.0
0.5
0
1% chymotrypsin
1% chymotrypsin
0
20
40
60
120
Time in minutes

Figure 12 Elution pattern of peptides from a chymotryptic digest of RCM-LBI IV on Dowex 50-X2.

The digest (180 mg) was fractionated on a column (1.9 x 100 cm) of Dowex 50-X2 with pyridine-acetic acid buffers. The flow rate was 40 ml/hr and fractions of 4.5 ml were collected. Samples of 100 µl each from alternate tubes were analyzed after alkaline hydrolysis by the ninhydrin method. The details are described in the text.

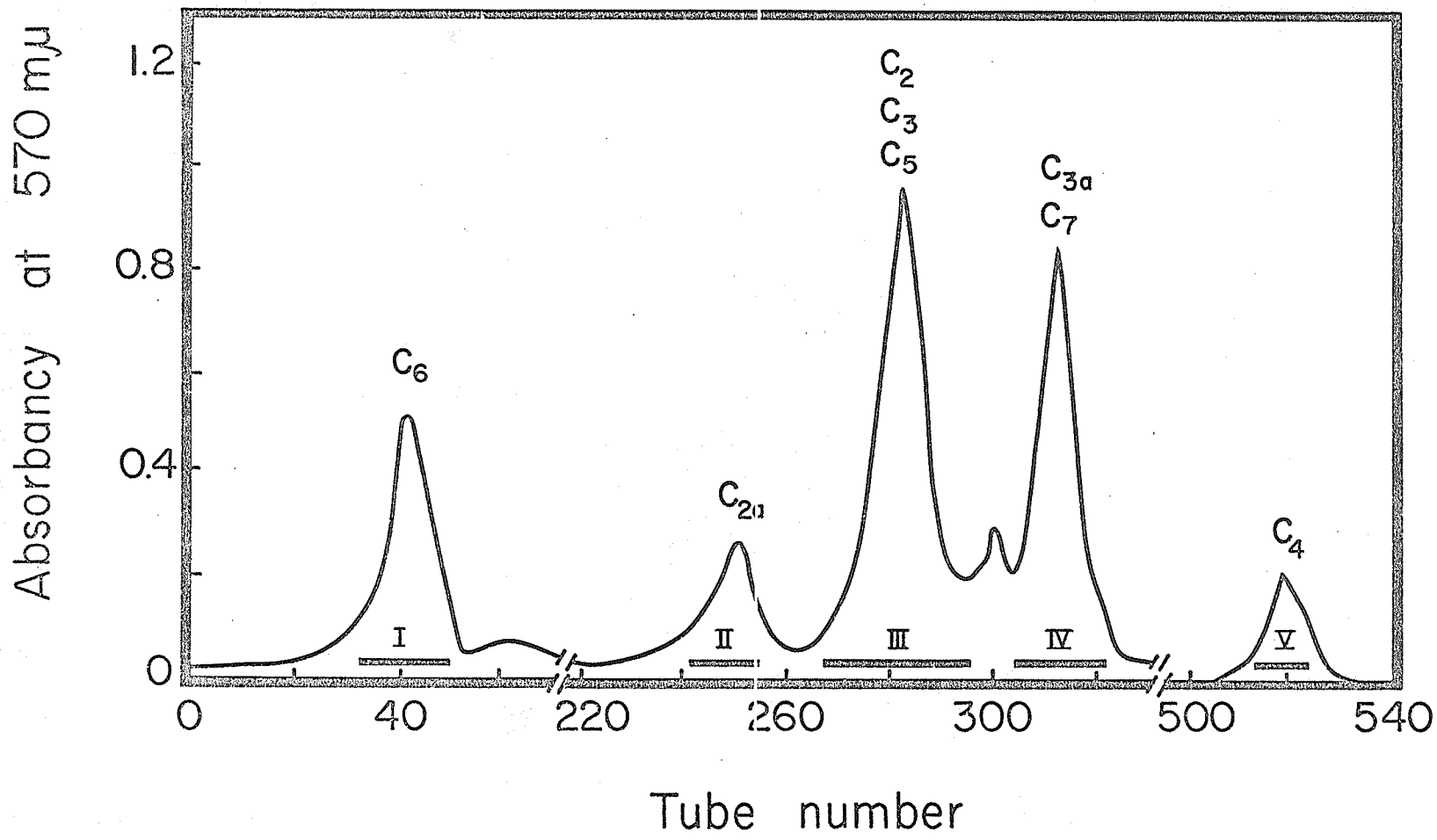

Absorbancy at 570 mμ
1.2
0.8
0.4
0
C6
C2a
C2
C3
C5
C3a
C7
C4
I
II
III
IV
V
0
40
220
260
300
500
540
Tube number

Figure 13 Gel filtration on Bio-Gel P-4 of Peak III from Figure 12.

The column (2.5 x 90 cm) was equilibrated and developed with 0.1M NH_4HCO_3, pH 7.8 at a flow rate of 37 ml/hr. Fractions of 4.5 ml were collected. Samples of 200 µl each from alternate tubes were taken to dryness before alkaline hydrolysis and were analyzed by the ninhydrin method.

Figure 14 Gel filtration on Bio-Gel P-2 of Peak III-2 from Figure 13.

The column (2.5 x 90 cm) was equilibrated and developed with 10% acetic acid at a flow rate of 37 ml/hr. Fractions of 3.5 ml were collected. Samples of 200 µl each from alternate tubes were analyzed after alkaline hydrolysis by the ninhydrin method.

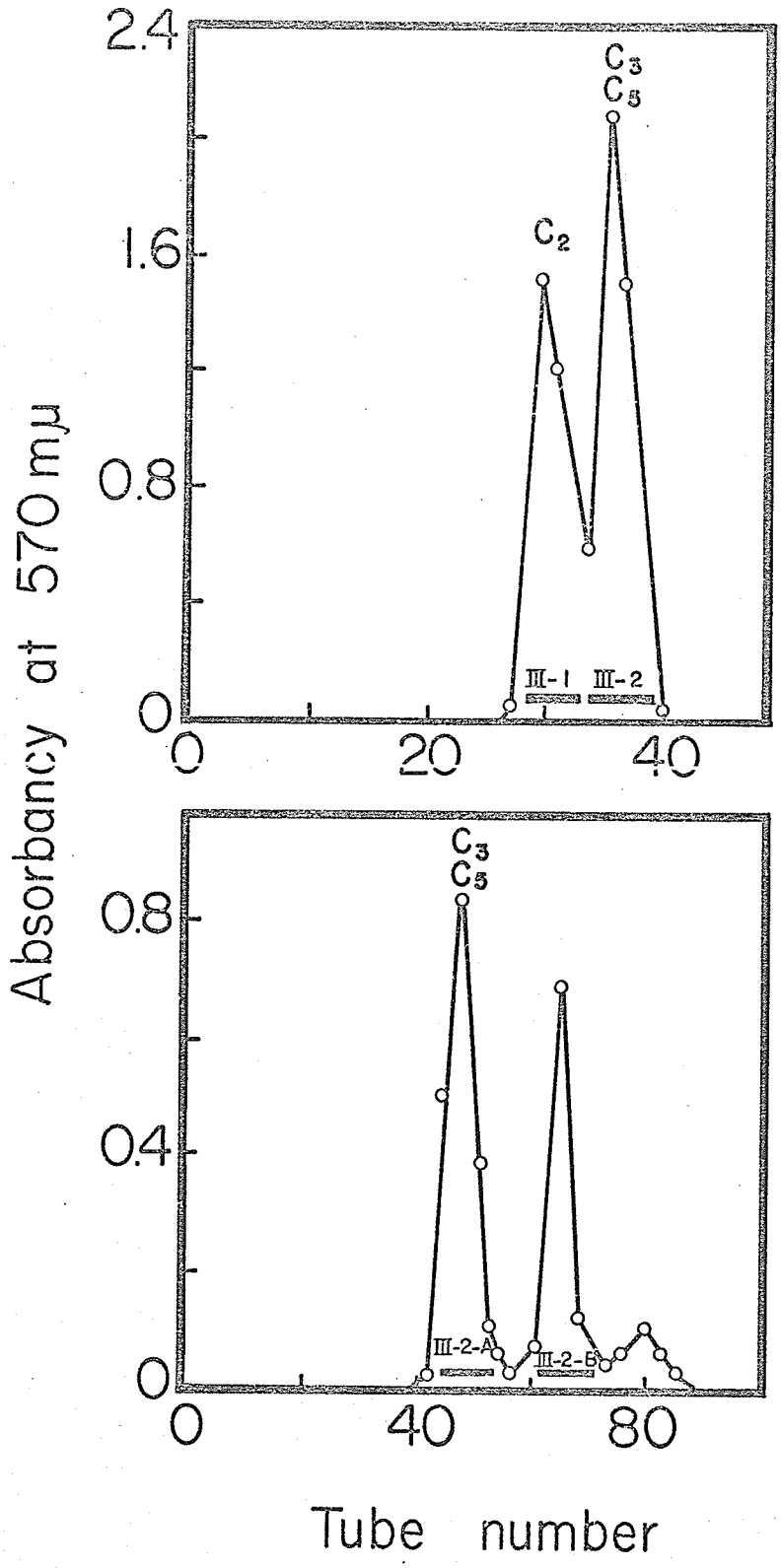

2.4
1.6
0.8
0
C_3
C_5
C_2
III-1
III-2
0
20
40
0.8
0.4
0
C_3
C_5
III-2-A
III-2-B
0
40
80
Absorbancy at 570 mμ
Tube number

Figure 15 Digestion curve of GRCM-LBI IV (130 mg) with trypsin.

Details are given in the text.

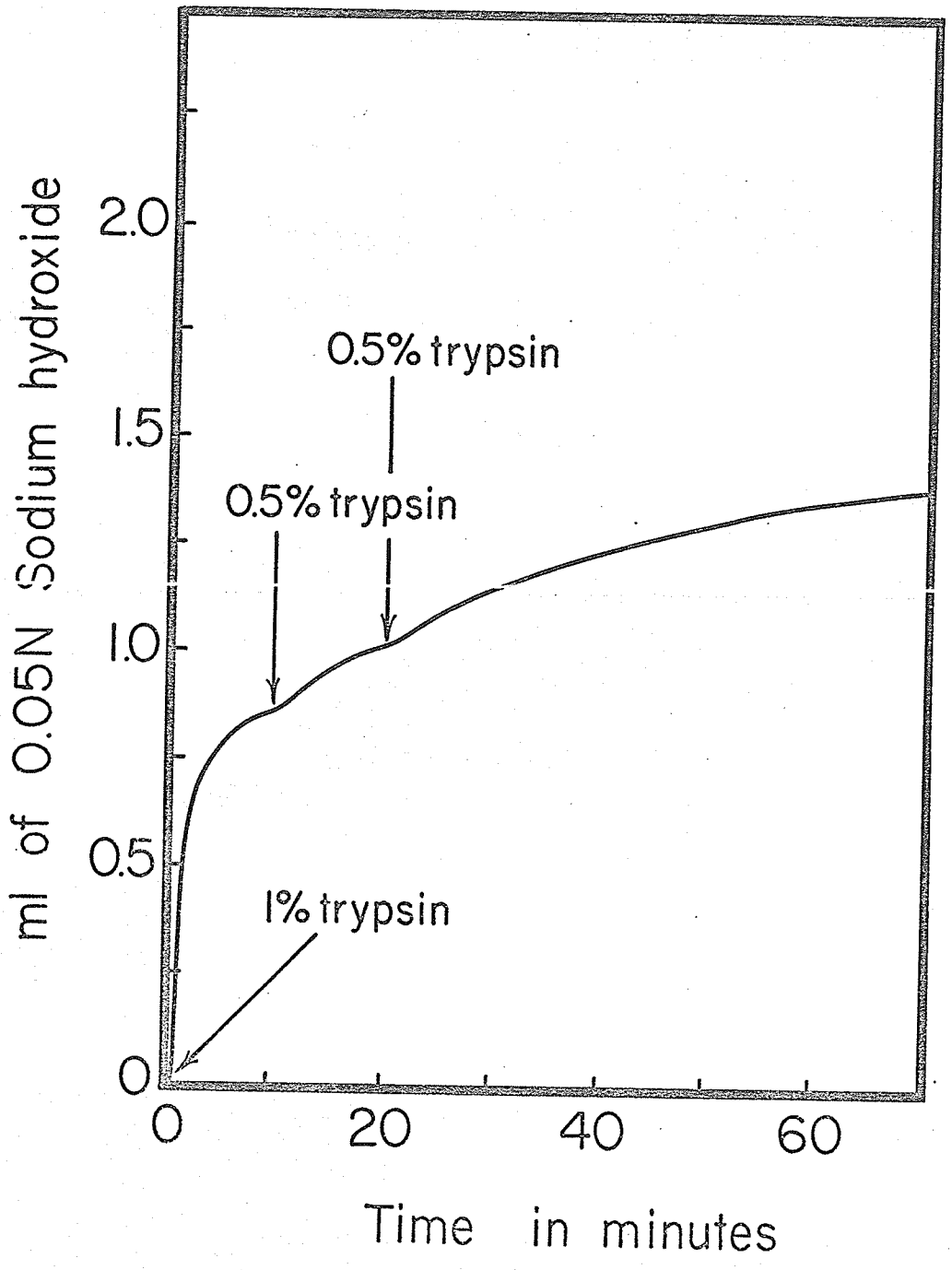

ml of 0.05N Sodium hydroxide
2.0
1.5
1.0
0.5
0
0.5% trypsin
0.5% trypsin
1% trypsin
0
20
40
60
Time in minutes

Figure 16 Elution pattern of peptides from a tryptic digest of GRCM-LBI IV.

The digest (130 mg) was applied to a column (2.5 x 80 cm) of Bio-Gel P-6, equilibrated in 10% acetic acid which was also the eluant. The flow rate was 34 ml per hour and fractions of 3.5 ml were collected. Samples of 100 µl each from alternate tubes were analyzed after alkaline hydrolysis by the ninhydrin method.

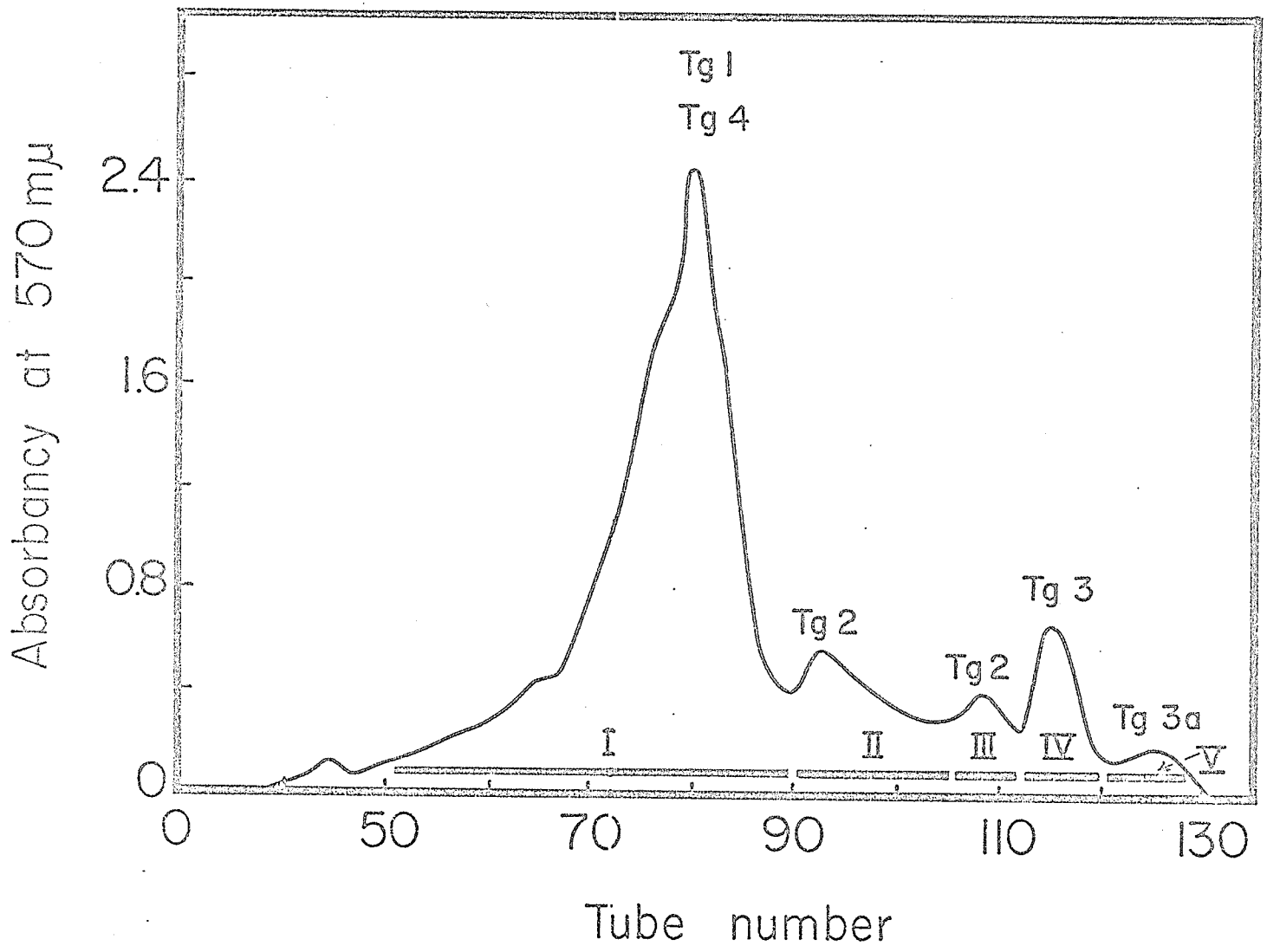

Tg 1
Tg 4
Tg 2
Tg 2
Tg 3
Tg 3a
I
II
III
IV
V
Absorbancy at 570 mμ
2.4
1.6
0.8
0
0
50
70
90
110
130
Tube number

Figure 17 Gel filtration on Sephadex G-25 of Peak I from Figure 16.

The column (2.5 x 180 cm) was equilibrated and developed with 10% acetic acid at a flow rate of 6.8 ml/hr. Fractions of 2.1 ml were collected. Samples of 50 µl each from alternate tubes were analyzed after alkaline hydrolysis by the ninhydrin method.

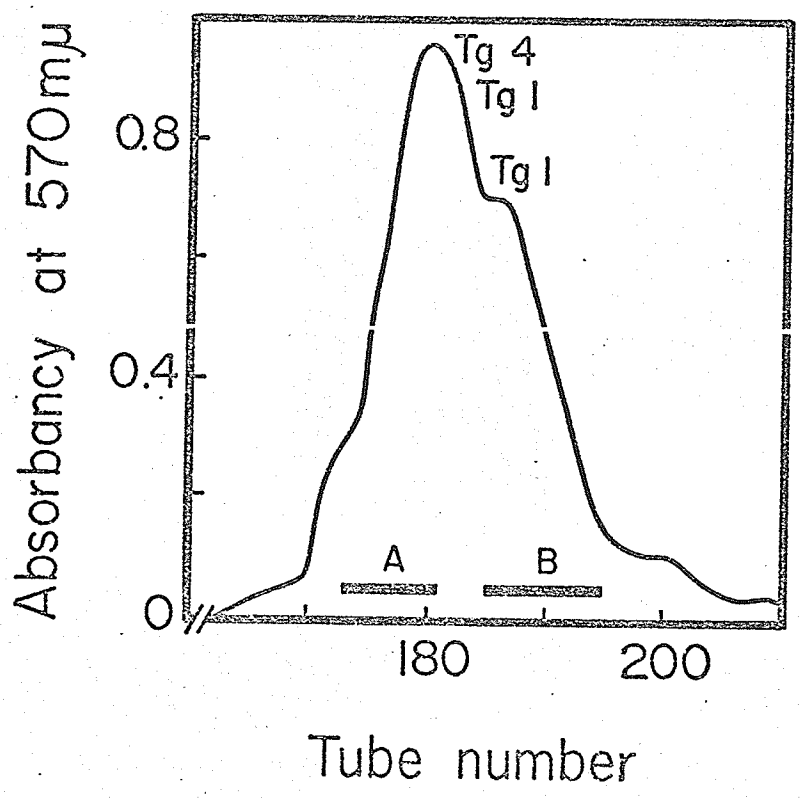
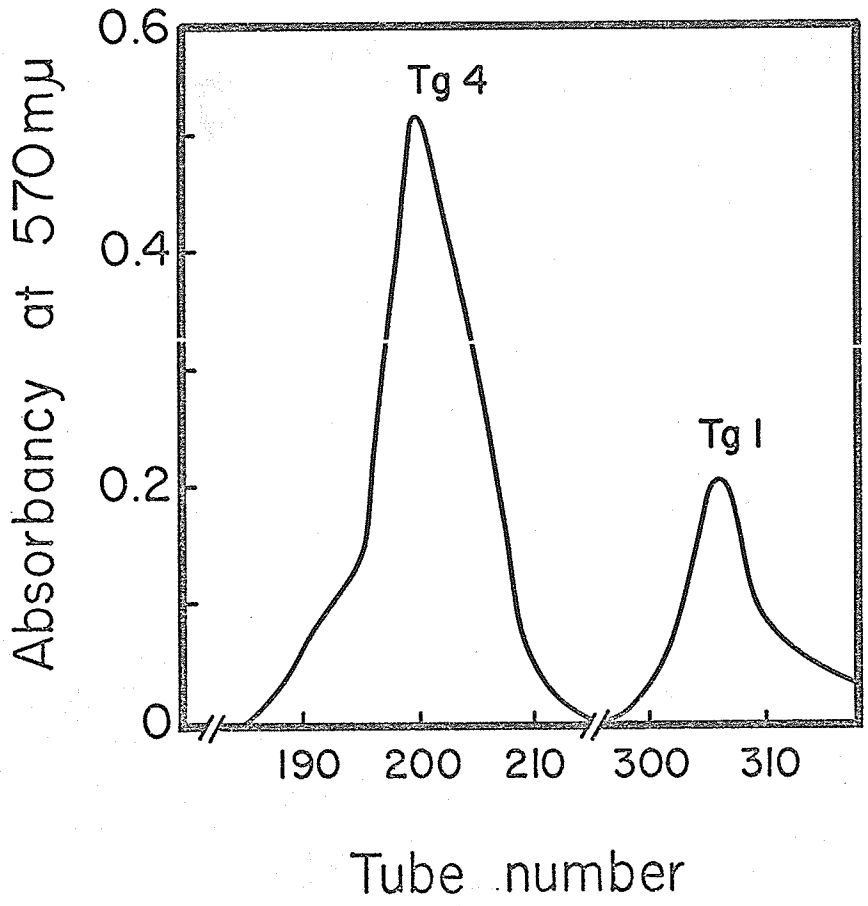

Absorbancy at 570mμ
0.8
0.4
0
Tg 4
Tg I
Tg I
A
B
180
200
Tube number

Figure 18 Elution pattern of Peak I (see Figure 17) on Dowex 50-X2.

Peak I (from Figure 17) was fractionated on a column (1.4 x 60 cm) of Dowex 50-X2 with pyridine-acetic acid buffers. The flow rate was 30 ml/hr and fractions of 3.4 ml were collected. Samples of 50 µl each from alternate tubes were analyzed after alkaline hydrolysis by the ninhydrin method. The details are described in the text.


0.6
0.4
0.2
0
Absorbancy at 570mμ
Tg 4
Tg 1
190
200
210
300
310
Tube number

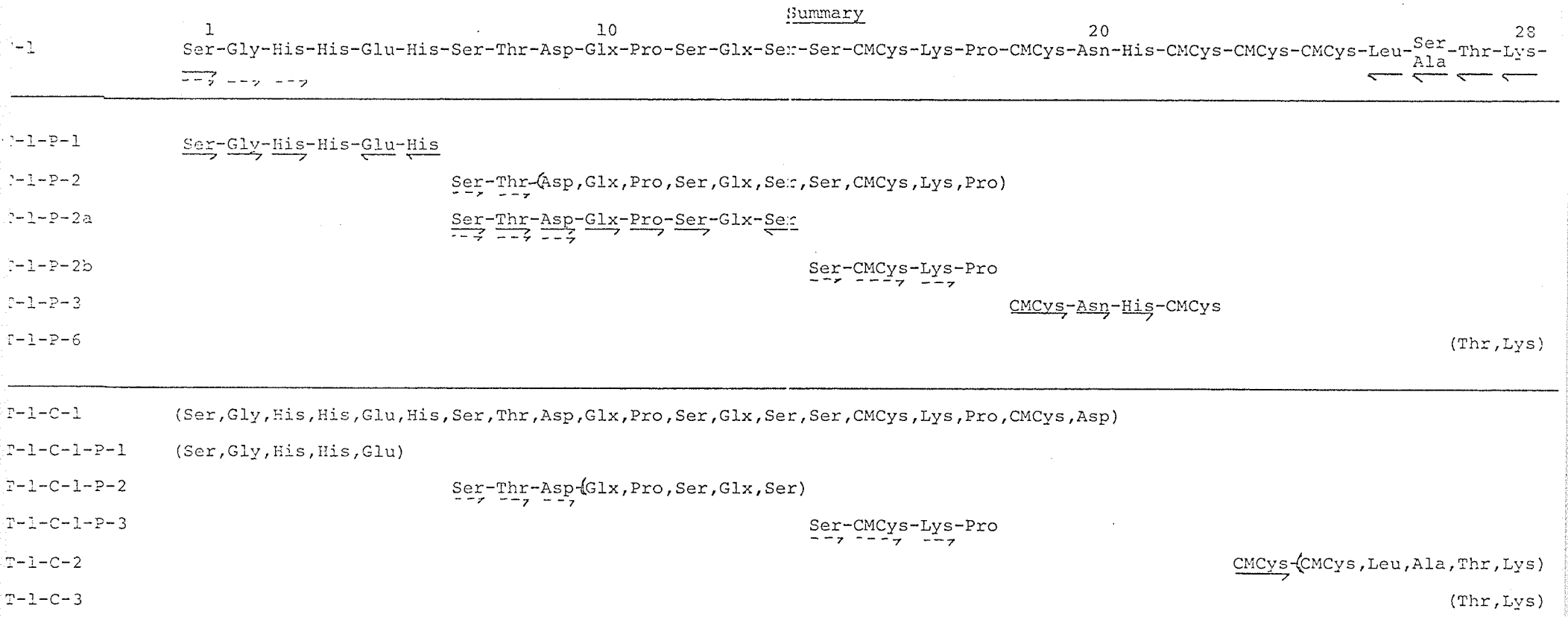


Figure 19 Amino acid sequence of Peptide T-1

Figure 20 Amino acid sequence of LBI IV

The sequence was established from the overlapping peptides of the tryptic digest (T), chymotryptic digest (C), and tryptic digest of GRCM-LBI IV. Vertical arrows indicate the points of cleavage.

FIGURE 20. AMINO ACID SEQUENCE OF LBI - IV

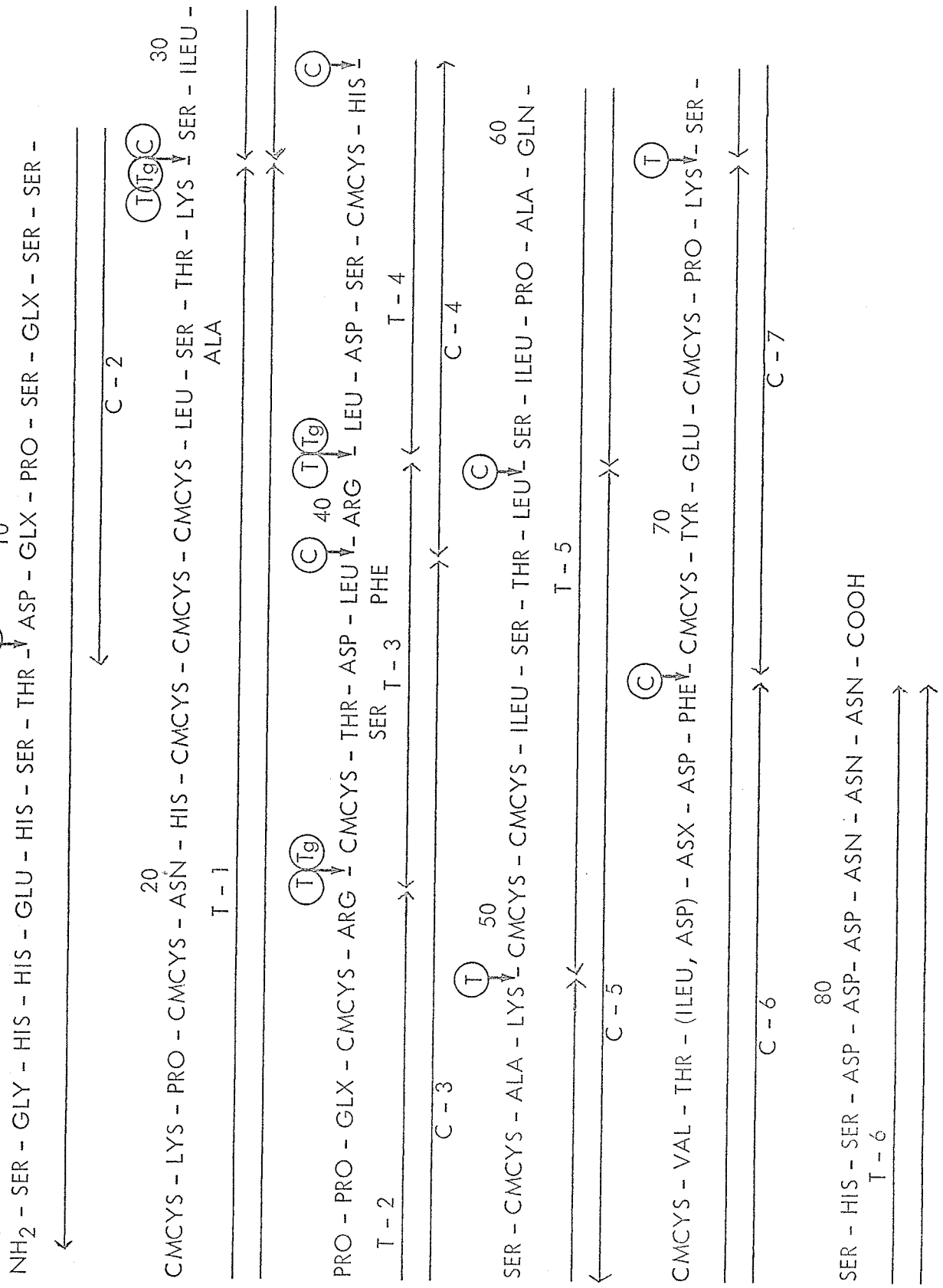


TABLE I

Amino acid compositions of Inhibitors of Plant Origin

	Soybean Inhibitors			Lima bean Inhibitor[d]			
	Kunitz Inhibitor[a]	Bowman-Birk Inhibitor[b]	1.9 S Inhibitor[c]	I	II	III	IV
Lysine	11	5	10	4	4	4	4
Histidine	2	1	2	5	3	6	6
Arginine	9	2	4	2	2	2	2
Aspartic Acid	29	12	23	12	14	13	13
Threonine	8	2	4	4	3	5	5
Serine	13	9	18	12	12	15	13
Glutamic Acid	21	7	14	6	5	7	7
Proline	10	6	12	6	6	7	7
Glycine	18	0	0	1	0	1	1
Alanine	9	4	8	3	3	4	3
Half Cystine	4	14	26	12	14	16	14
Valine	12	1	2	1	1	1	1
Methionine	3	1	2	0	0	0	0
Isoleucine	14	2	4	4	4	5	4
Leucine	16	2	4	3	3	3	3
Tyrosine	4	2	4	1	1	2	1
Phenyl-alanine	9	2	4	1	1	2	2
Tryptophan	2	0	0.6	0	0	0	0
Total Residues	194	78	141	77	76	93	86
Molecular Weight	21,500	7,975	16,400	8408	8291	9892	9423

[a] Wu and Scheraga (1962a)

[b] Frattali (1969)

[c] Yamamoto and Ikenaka (1967)

[d] Jones et al. (1963)

TABLE II

Amino Acid Compositions of Inhibitors of Animal Origin

Amino Acid	BPTI[a]	Kazal Inhibitor[b]	Porcine Inhibitor I[c]	Porcine Inhibitor II[c]
Lysine	4	3	4	4
Histidine	0	0	0	0
Arginine	6	3	2	2
Aspartic Acid	5	7	4	4
Threonine	3	4	6	5
Serine	1	2	6	5
Glutamic Acid	3	7	7	6
Proline	4	4	5	4
Glycine	6	5	4	4
Alanine	6	1	1	1
Half Cystine	6	6	6	6
Valine	1	4	4	4
Methionine	1	1	0	0
Isoleucine	2	3	3	3
Leucine	2	4	2	2
Tyrosine	4	2	2	2
Phenylalanine	4	0	0	0
Total	58	56	56	52
Molecular Weight	6513	6155	6024	5609

[a] Kassell et al. (1963)

[b] Greene et al. (1966)

[c] Greene et al. (1968)

TABLE III

Reactive Sites of Inhibitors

Inhibitor	Reactive Site	Reference
Soybean (Kunitz)	Arg_{64}-Ileu	Ozawa and Laskowski (1966)
BPTI (Bovine)	Lys_{15}-Ala	Kassell and Laskowski (1965)
Colostrum (Bovine)	Lys_{18}-Ala	Čechová and Muszyńská (1970)
Kazal Inhibitor(Bovine)	Arg_{18}-Ileu	Rigbi and Greene (1968)
Porcine Inhibitor I	Lys_{18}-Ileu	Tschesche et al. (1969)
Chicken Ovomucoid	Arg-Ala	Ozawa and Laskowski (1966)
Corn	Arg-Ileu	Hochstrasser and Werle (1969)
Rye	Arg-Ala	Hochstrasser and Werle (1969)
Wheat	Arg-Ala	Hochstrasser and Werle (1969)
Peanut	Arg-Ala	Hochstrasser et al. (1969c)
Lima Bean	Lys-X	Stevens (1969)

TABLE IV

Inhibitory Activity of Inhibitors from Lima Beans Against Trypsin and α-chymotrypsin with Ester Substrates

Inhibitor Fraction	Enzyme Inhibited	
	Trypsin	Chymotrypsin
	mg/mg inhibitor	
I	2.35	1.38
II	2.52	1.45
III	2.09	1.04
IV	2.50	1.50

TABLE V

Amino acid compositions of trypsin inhibitors from lima beans. The results are expressed as molar ratios of the amino acids and were obtained by a duplicate analyses of 22 hour hydrolysates.

Amino Acid	Fraction I 22h (Ave)	Integer	Fraction II 22h (Ave)	Integer	Fraction III 22h (Ave)	Integer	Fraction IV 22h (Ave)	Integer
Lysine	4.87	5	4.30	4	4.00	4	3.90	4
Histidine	5.60	6	5.85	6	5.79	6	5.65	6
Arginine	1.67	2	1.85	2	1.86	2	1.86	2
CM Cysteine[b]	14.40	14	14.00	14	16.00	16	14.20	14
Aspartic Acid	14.30	14	15.60	16	12.50	13	13.3	13
Threonine[a]	5.00	5	5.75	6	5.56	6	5.32	5
Serine[a]	12.20	12	13.00	13	11.25	12	13.75	14
Glutamic Acid	6.00	6	6.10	6	5.66	6	6.89	6
Proline	5.55	6	5.92	6	5.55	6	6.02	6
Glycine	1.00	1	1.00	1	1.00	1	1.00	1
Alanine	3.45	3	3.06	3	2.17	2	2.30	3
Valine	1.00	1	1.06	1	0.79	1	0.54	1
Methionine	0	0	0	0	0	0	0	0
Isoleucine	3.87	4	4.31	4	3.52	4	3.72	4
Leucine	2.95	3	3.49	3	3.61	4	3.38	3
Tyrosine	0.80	1	0.90	1	0.83	1	1.05	1
Phenyl-alanine	1.02	1	1.10	1	1.17	1	1.36	1
Total Residues		84		87		85		84

[a] The values for serine and threonine were obtained by extrapolation to zero time.

[b] Values of carboxymethylcysteine were determined on samples of protein that had been carboxymethylated.

TABLE VI

Amino acid compositions of tryptic peptides from RCM-LBI IV isolated from Dowex 50-X2.

The results are expressed as molar ratios of the amino acids and were obtained by analysis of 22 hour hydrolysates. No corrections have been made for destruction of serine, threonine and tyrosine during hydrolysis. The numbers in parentheses are numbers of residues obtained by sequence analyses.

Amino Acid	T-1	T-2	T-3	T-3a	T-4	T-6	RCM-LBI IV
Lysine	2.00(2)				1.00(1)		3.90
Histidine	4.00(4)				1.00(1)	1.00(1)	5.65
Arginine		0.63(1)	1.00(1)	0.73(1)			1.86
CM Cysteine	4.86(5)	1.25(1)	1.23(1)	1.00(1)	2.52(2)		14.20
Aspartic Acid	2.42(2)		1.01(1)	1.07(1)	1.08(1)	5.75(6)	11.90
Threonine	2.20(2)		0.92(1)				4.60
Serine	4.70(5)	1.00(1)		1.00(1)	2.00(2)	2.90(3)	10.70
Glutamic Acid	3.48(3)	1.10(1)					6.38
Proline	1.94(2)	2.45(2)					5.62
Glycine	1.00(1)						1.00
Alanine	0.29(1)				1.00(1)		2.12
Valine							0.92
Isoleucine		1.00(1)					3.62
Leucine	0.74(1)		0.94(1)		1.00(1)		3.05
Tyrosine							1.28
Phenyl-alanine				1.05(1)			1.30
Amount µmole	6	6.6	5.4	1.2	8.64	9	
Yield	46%	51%	41%	9.2%	66.5%	69%	
Total Residues	28	7	5	5	9	10	

TABLE VII

Amino acid compositions of tryptic peptides from RCAM-LBI IV isolated from Bio-Gel P-6. The results are expressed as molar ratios of the amino acids and were obtained by analyses of 22 hour hydrolysates. No corrections have been made for destruction of serine, threonine, tyrosine during hydrolysis. Destruction of carboxymethylcysteine during hydrolysis resulted in their identification as half cystines or cysteic acid in some samples*. The numbers in parentheses are numbers of residues obtained by sequence analyses. These peptides account for the entire sequence of the protein.

Amino Acid	T-1	T-2	T-3	T-3a	T-4	T-5	T-6
Lysine	2.04(2)				1.00(1)	1.47(1)	
Histidine	3.76(4)				0.75(1)		1.00(1)
Arginine		0.93(1)	0.50(1)	0.88(1)			
CM Cysteine	4.86(5)	1.00(1)*	1.00(1)	1.27(1)	1.64(2)*	5.10(5)	
Aspartic Acid	2.20(2)		0.88(1)	1.00(1)	1.12(1)	3.30(3)	6.00(6)
Threonine	2.04(2)		1.10(1)			2.06(2)	
Serine	4.70(5)	1.28(1)		1.00(1)	2.20(2)	3.43(2)	3.00(3)
Glutamic Acid	3.20(3)	1.14(1)				2.80(2)	
Proline	1.94(2)	1.85(2)				2.30(2)	
Glycine	1.00(1)						
Alanine	0.40(1)				1.09(1)	1.00(1)	
Valine						0.54(1)	
Isoleucine		0.86(1)				1.65(3)	
Leucine	0.71(1)		1.07(1)		1.12(1)	1.00(1)	
Tyrosine						0.69(1)	
Phenyl-alanine				0.55(1)		0.79(1)	
Amount µmole	10	5	3	3.1	8.04	1	13.5
Yield	54%	27.8%	16.7%	17.2%	44%	5%	75%
Total Residues	28	7	5	5	9	25	10

TABLE VIII

Amino acid compositions of chymotryptic peptides RCM-LBI IV isolated from Dowex 50-X2. The results are expressed as molar ratios of the amino acids and were obtained by analyses of 22 hr hydrolysates. No corrections have been made for destruction of serine, threonine and tyrosine during hydrolysis. The numbers in parentheses are the nearest integer values. These peptides account for the entire sequence of the protein with the exception of residues 1-8 which were recovered in very small amounts in Peak III-2-B (See Figure 14) so that no further work was done on them.

Amino Acid	C-2	C-2a	C-3	C-3a	C-4	C-5	C-6	C-7
Lysine	1.82(2)	1.00(1)				1.14(1)		1.00(1)
Histidine	1.01(1)	0.93(1)			1.00(1)			1.19(1)
Arginine			0.85(1)	1.00(1)	0.84(1)			
CM Cysteine	4.73(5)	4.00(4)	1.86(2)	2.00(2)	0.90(1)	2.96(3)	1.35(1)	2.07(2)
Aspartic Acid	2.76(2)	1.83(2)	0.98(1)	1.06(1)	1.07(1)		3.24(3)	5.15(5)
Threonine	1.25(1)		1.00(1)			1.13(1)	0.75(1)	
Serine	3.32(3)	2.92(3)	1.06(1)	2.00(2)	1.05(1)	2.14(2)	1.09(1)	2.77(3)
Glutamic Acid	2.16(2)	1.96(2)	1.04(1)	1.07(1)			1.15(1)	0.91(1)
Proline	2.20(2)	1.67(2)	2.00(2)	2.30(2)			1.02(1)	0.85(1)
Alanine	0.41(1)					1.10(1)	1.12(1)	
Valine							0.83(1)	
Isoleucine			0.99(1)	1.00(1)		0.67(1)	1.77(2)	
Leucine	1.00(1)	0.81(1)	1.00(1)		1.00(1)	1.00(1)		
Tyrosine								0.13(1)
Phenyl-alanine				1.00(1)			0.94(1)	
Amount μmole	2.4	1.07	0.83	0.1	2.6	1.6	2.4	3.5
Yield	13%	6%	4.6%	0.5%	14.5%	8.9%	13.4%	19.4%
Total Residues	20	16	11	11	6	10	13	15

TABLE IX

Amino Acid compositions of tryptic peptides from GRCM-LBI IV isolated from Bio-Gel P-6.

The results are expressed as molar ratios of the amino acids and were obtained by analyses of 22 hrs. No corrections have been made for destruction of serine, threonine and tyrosine during hydrolysis. The numbers in parentheses are the nearest integer values.

Amino Acid	Tg-1	Tg-2	Tg-3	Tg-3a	Tg-4	GRCM-LBI IV
Lysine						0
Histidine	4.00(4)				2.00(2)	5.70
Arginine		1.00(1)	1.00(1)	1.04(1)		2.27
Homoarginine	1.95(2)				2.00(2)	4.00
Aspartic Acid	3.54(3)		1.10(1)	1.10(1)	9.50(10)	12.70
Threonine	2.00(2)		0.93(1)		2.20(2)	5.60
Serine	5.16(5)	1.00(1)		1.06(1)	6.50(7)	13.00
Glutamic Acid	3.82(4)	1.07(1)			2.10(2)	7.30
Proline	2.70(3)	2.00(2)			2.00(2)	7.30
Glycine	1.00(1)					1.00
Alanine	0.40(1)				2.00(2)	2.70
Valine					1.00(1)	1.00
Isoleucine		0.83(1)			3.10(3)	4.10
Leucine	0.90(1)		1.00(1)		2.00(2)	3.90
Tyrosine					1.10(1)	1.30
Phenyl-alanine				0.94(1)	1.18(1)	1.79
CM Cysteine	5.30(5)	1.00(1)	1.00(1)	0.69(1)	6.95(7)	14.01
Total Residues	31	7	5	5	44	87

TABLE X

Amino Acid Sequence of Peptide T-1 (residues 1 through 28)

Composition: Lys,2.04(2); His,3.70(4); CMCys,4.86(5); Asp,2.20(2); Thr,2.04(2); Ser,4.70(5); Glu,3.20(3); Pro,1.94(2); Gly,1.00(1); Ala,0.40(1); Leu,0.71(1)

Sequence	Ser-Gly-His-His-Glu-His-Ser-Thr-Asp-Glx-Pro-Ser-Glx-Ser-Ser-CMCys-Lys-Pro-CMCys-Asn-His-CMCys-CMCys-CMCys-Leu-Ala-Thr-Lys-
Co B + Co A	4 hrs: Lys,1; Thr,0.95; Ala,0.49; Leu,0.73
LAP	Ser,1; Gly,0.29; His,0.75
Edman degradation	
Step 1	Residue: Ser,3.40(4); Gly,1.30(1); His,2.10[a]; Lys,1.00[a]; CMCys,2.45; Asp,2.05(2); Thr,2.15(2); Glu,2.69(3); Pro,2.02(2); Ala,0.49(1) Leu,0.54(1)
Step 2	Residue: Ser,3.64(4); Gly,0.41; His,2.03[a]; Lys,1.69(2)[a]; CMCys,4.45(5); Asp,2.24(2); Thr,1.91(2); Glu,3.08(3); Pro,2.24(2); Ala,0.47(1) Leu,0.75(1)
Pronase Digest of T-1	

([a] values are low)

P-1	His,2.80(3); Ser,0.76(1); Gly,1.00(1); Glu,1.00(1)
Sequence	Ser-Gly-His-His-Glu-His-
Co A	His,1; Glu,0.48
Edman degradation	
Step 1	Residue: Ser,0.40; Gly,1.00(1); His,3.00(3); Glu,1.50(1)
Step 2	Residue: Ser,0.30; Gly,0.40; His,2.50(3); Glu,1.30(1)
Step 3	Residue: Ser,0.44; Gly,0.34; His,1.50(2); Glu,1.30(1)
Step 4	Residue: Ser,0.44; Gly,0.54; His,1.33(1); Glu,1.00(1)
P-2	Ser,3.60(4); Thr,1.10(1); Asp,1.10(1); Glu,2.08(2); Pro,1.92(2); CMCys,1.00(1); Lys,1.00(1)
Sequence	Ser-Thr-Asp-Glx-Pro-Ser-Glx-Ser-Ser-CMCys-Lys-Pro T-1-P-2a / T-1-P-2b
APM	Ser,1.00; Thr,0.50 (Duplicate)
Co A + Co B	No digestion
Pronase digest of P-2	
P-2a	Ser,2.50(3); Thr,0.96(1); Asp,1.00(1); Glu,1.89(2); Pro,1.08(1)

TABLE X (continued)

Sequence	Ser-Thr-Asp-Glx-Pro-Ser-Glx-Ser-
Co A	Ser (+)
LAP	Ser,1.00; Thr,0.63; Asp,0.36
Edman degradation	
Step 1	Residue: Ser,2.00(2); Thr,0.90(1); Asp,1.05(1); Glu,2.00(2); Pro,1.00(1)
Step 2	Residue: Ser,2.07(2); Thr,0.40; Asp,0.99(1); Glu,1.97(2); Pro,1.02(1)
Step 3	Residue: Ser,2.06(2); Thr,0.42; Asp,0.62; Glu,2.02(2); Pro,1.00(1)
Step 4	Residue: Ser,1.70(2); Thr,0.24; Asp,0.38; Glu,1.34(1); Pro,1.00(1)
Step 5	Residue: Ser,1.72(2); Thr,0.26; Asp,0.36; Glu,1.25(1); Pro,0.62
Step 6	Residue: Ser,1.19(1); Thr,0.21; Asp,0.32; Glu,1.00(1); Pro,0.37
P-2b	Lys,1.01(1); Cysteic,0.80(1); Ser,0.99(1); Pro,0.95(1)
Sequence	Ser-CMCys-Lys-Pro-
LAP	Ser,1.00; CMCys,0.54; Lys,0.35
P-3	His,1.10(1); CMCys,2.02(2); Asp,1.10(1)
Sequence	CMCys-Asn-His-CMCys
APM	Asn (+), CMCys (+)

Edman degradation	
Step 1	Residue: CMCys,1.00(1); Asp,1.05(1); His,1.00(1)
Step 2	Residue: CMCys,0.84(1); Asp,0.49; His,1.00(1)
Step 3	Residue: CMCys,0.83(1); Asp,0.46; His,0.50
P-6	Thr,1.00(1); Lys,1.00(1)
Chymotrypsin digestion of T-1	
C-1	Lys,1.00(1); His,3.40(4); CMCys,2.50(3); Asp,2.18(2); Thr,1.39(1); Ser,5.00(5); Glu,3.04(3); Pro,2.12(2); Gly,1.00(1)
Pronase digest of C-1	
P-1	His,1.45(2); Ser,0.82(1); Glu,0.79(1); Gly,1.00(1)
P-2	Asp,1.00(1); Thr,1.00(1); Ser,2.61(3); Glu,1.91(2); Pro,1.00(1)
LAP	Ser,1; Thr,0.85; Asp,0.15
P-3	Lys,1.01(1); CMCys,0.80(1); Ser,0.99(1); Pro,0.95(1)
LAP	Ser,1; CMCys,0.54; Lys,0.35
C-2	Lys,1.00(1); CMCys,2.00(2); Thr,1.00(1); Ala,0.58; Leu,0.71(1)
Edman degradation	
Step 1	Residue: Lys,1.05(1); CMCys,0.88(1); Thr,1.50(1); Ala,0.62(1); Leu,0.74(1)

Summary

	1	10	20	28
T-1	Ser-Gly-His-His-Glu-His-Ser-Thr-Asp-Glx-Pro-Ser-Glx-Ser-Ser-CMCys-Lys-Pro-CMCys-Asn-His-CMCys-CMCys-CMCys-Leu-Ser/Ala-Thr-Lys-			
T-1-P-1	Ser-Gly-His-His-Glu-His			
T-1-P-2		Ser-Thr-(Asp,Glx,Pro,Ser,Glx,Ser,Ser,CMCys,Lys,Pro)		
T-1-P-2a		Ser-Thr-Asp-Glx-Pro-Ser-Glx-Ser		
T-1-P-2b			Ser-CMCys-Lys-Pro	
T-1-P-3			CMCys-Asn-His-CMCys	
T-1-P-6				(Thr,Lys)
T-1-C-1	(Ser,Gly,His,His,Glu,His,Ser,Thr,Asp,Glx,Pro,Ser,Glx,Ser,Ser,CMCys,Lys,Pro,CMCys,Asp)			
T-1-C-1-P-1	(Ser,Gly,His,His,Glu)			
T-1-C-1-P-2		Ser-Thr-Asp-(Glx,Pro,Ser,Glx,Ser)		
T-1-C-1-P-3			Ser-CMCys-Lys-Pro	
T-1-C-2				CMCys-(CMCys,Leu,Ala,Thr,Lys)
T-1-C-3				(Thr,Lys)

Figure 19 Amino acid sequence of Peptide T-1

TABLE XI

Amino acid sequence of Peptide T-2 (residues 29 through 35)

Composition: Ser,1.00(1); Ileu,1.00(1); Pro,2.45(2); CMCys,1.25(1); Glu,1.10(1); Arg,0.63(1)

Sequence	Ser-Ileu-Pro-Pro-Glx-CMCys-Arg
Co B	4 hrs: Arg,1.00
Co B + Co A	18 hrs: Arg, 1.00
Edman degradation	
Step 1	PTH-Ser
	Residue: Ser,0.13; Ileu,1.00(1); Pro,2.00(2); CMCys,0.83(1); Glu,1.11(1); Arg,0.82(1)
Step 2	PTH-Ileu
	Residue: Ser,0.18; Ileu,0.23; Pro,2.16(2); CMCys,1.00(1); Glu,1.33(1); Arg,0.98(1)
Step 3	PTH-Pro
	Residue: Ser,0.08; Ileu,0.17; Pro,1.00(1); CMCys,0.67(1); Glu,1.3(1); Arg,1.00(1)

Step 4	PTH-Pro
	Residue: Ser,0.09; Ileu,0.20; Pro,0.44; CMCys,0.67(1); Glu,1.00(1); Arg,0.60(1)
Partial Acid Hydrolysis of T-2	
PA-1	Cysteic, 0.44(1); Glu,1.00(1)
PA-2	Glu,1.00(1)
PA-3	Ser,0.68(1); Ileu,1.00(1); Pro,2.16(2); Glu,1.10(1)

Summary

T-2	Ser-Ileu-Pro-Pro-Glx-CMCys-Arg
T-2-PA-1	(Glu,CMCys)
T-2-PA-2	Glu
T-2-PA-3	(Ser,Ileu,Pro,Pro,Glu)

TABLE XII

Amino acid sequence of Peptide T-3 (Residues 36 through 40)

Composition: CMCys,1.23(1); Thr,0.92(1); Asp,1.01(1); Leu,0.94(1); Arg,1.00(1)

Sequence	CMCys-Thr-Asp-Leu-Arg
Co B + Co A	4 hrs: Arg (+1)
Co B + Co A	18 hrs: Arg, 1.00; Leu,0.25
Edman degradation	
Step 1	Residue: CMCys,0; Thr,0.89(1); Asp,1.00(1); Leu,0.79(1)*
Step 2	PTH-Thr
	Residue: CMCys,0; Thr,0.24; Asp,1.00(1); Leu,1.09(1)*
Step 3	PTH-Asp
	Residue: CMCys,0; Thr,0.21; Asp,0.37; Leu,1.00(1)*
APM	Asp (+1), CMCys (+1), Thr (+), Leu (+), Arg (+)

* The hydrolysates were analyzed on the long column only

TABLE XIII

Amino acid sequence of Peptide T-3a(residues 36 through 40)

Composition: CMCys,1.00(1); Ser,1.00(1); Asp,1.07(1); Phe,1.05(1); Arg,0.73(1)

Sequence	CMCys-Ser-Asp-Phe-Arg
Co B	4 hrs: Arg,1.00; Phe,0.73
Edman degradation	
Step 1	PTH-CMCys
	Residue: CMCys,0; Ser,0.82(1); Asp,0.93(1); Phe,1.00(1)*
Step 2	PTH-Ser
	Residue: CMCys,0; Ser,0.28; Asp,1.00(1); Phe,0.88(1)*
Step 3	Residue: CMCys,0; Ser,0.28; Asp,0.54; Phe,0.88(1)*
APM	Asp (+), CMCys (+), Ser (+), Phe (+), Arg (+)

* The hydrolysates were analyzed on the long column only

TABLE XIV

Amino acid sequence of Peptide T-4 (residues 41 through 49)

Composition: Leu,1.00(1); Asp,1.08(1); Ser,2.00(2); CMCys,2.52(2); His,1.00(1); Ala,1.00(1); Lys,1.00(1)

Sequence	Leu-Asp-Ser-CMCys-His-Ser-CMCys-Ala-Lys
Co B	4 hrs: Lys,1.00; Ala,0.81; CMCys,0.48
Co B + Co A	16 hrs: Lys,1.00; Ala,1.00; CMCys,1.00; Ser,0.41
Edman degradation	
Step 1	PTH-Leu
	Residue: Leu,0.00; Asp,1.00(1); Ser,1.70(2); CMCys,1.90(2); His,0.41(1); Ala,1.06(1); Lys,0.40(1)
Step 2	PTH-Asp
	Residue: Leu,0; Asp,0.41; Ser,2.00(2); CMCys,1.82(2); His,0.84(1); Ala,1.00(1); Lys,0.28(1)
Step 3	Residue: Leu,0; Asp,0.24; Ser,1.00(1); CMCys--[a]; His,0.30(1); Ala,1.00(1); Lys,0.29(1)
Step 4	Residue: Leu,0; Asp,0.20; Ser,1.11(1); CMCys,1.22(1); His,0.27(1); Ala,1.00(1); Lys,0.50(1)

APM	Leu,1.32(1); Asp,1.20(1); Ser,1.99(2); CMCys,1.55(2); His,0.62(1); Ala,1.00(1); Lys,0.62(1)
Papain digest of T-4	
Pn-1	Asp,1.00(1); Ser,1.00(1); Cysteic,1.12(1)
Pn-5	Ser,1.00(1); Cysteine and Cysteic,1.20(1); Ala,1.17(1)
Pn-6	Ser,1.16(1); CMCys,1.00(1); Ala,1.00(1)
Pn-7	Leu,1.00(1); Asp,1.16(1); Ser,1.12(1)
Pn-8	Leu,1.00(1); Asp,1.23(1)

Summary

T-4	Leu-Asp-Ser-CMCys-His-Ser-CMCys-Ala-Lys
T-4-Pn-1	(Asp,Ser,CMCys)
T-4-Pn-5	(Ser,CMCys,Ala)
T-4-Pn-6	(Ser,CMCys,Ala)
T-4-Pn-7	(Leu,Asp,Ser)
T-4-Pn-8	(Leu,Asp)

[a] Carboxymethylcysteine might have been oxidized during acid hydrolysis to cysteic acid. Unfortunately cyteic acid was not evaluated in the analysis.

TABLE XV

Amino acid sequence of Peptide T-6 (residues 75 through 84)

Composition: Ser,2.90(3); His,1.00(1); Asp,5.75(6)

Sequence	Ser-Ser-His-Ser-Asp-Asp-Asp-Asn-Asn-Asn
Co A	Asn (+)
Edman degradation	
Step 1	Residue: Ser,2.46(2); His,0.42(1); Asp,6.00(6)
Step 2	Residue: Ser,1.85(1); His,0.91(1); Asp,6.00(6)
Step 3	Residue: Ser,1.73(1); His,0.63; Asp,6.00(6)
Aminopeptidase M digestion of T-6	
	2 hrs: Ser,1.18; His, 0
	10 hrs: Ser,2.28; His,0.35
	24 hrs: Ser,2.36; His,0.43
	30 hrs: Ser,2.46; His,0.71
	48 hrs: Ser,3.00; His,1.00+ Residue
Edman degradation of T-6-A-1	
Step 1	PTH-Asp
Step 2	PTH-Asp
Step 3	PTH-Asp
Step 4	PTH-Asn
Step 5	PTH-Asn
	Residue: Asn

Summary

T-6 Ser-Ser-His-Ser-Asp-Asp-Asp-Asn-Asn-Asn

T-6-A-1 Asp Asp Asp Asn Asn Asn

TABLE XVI

Subtractive Edman Degradation of Peptide T-6-A-1

	Electrophoretic Mobility at pH 4.7*	Ninhydrin Colour	Conclusion
Aspartic Acid	1	Turquoise	
Asparagine	-0.07	Orange	
Peptide T-6-A-1	1.71	Blue	Asp-Asp-Asp-Asn-Asn-Asn
Step 1	1.08	Blue	Asp-Asp-Asn-Asn-Asn
Step 2	0.86	Blue	Asp-Asn-Asn-Asn
Step 3	-0.07	Grey-Blue	Asn-Asn-Asn
Step 4	-0.08	Orange	Asn-Asn

* Electrophoretic mobilities are expressed relative to aspartic acid (=1.0)

TABLE XVII

Amino acid sequence of Peptide C-2 (residues 9 through 28)

Composition: Lys,1.86(2); His,1.01(1); CMCys,4.73(5); Asp,2.76(2); Thr,1.25(1); Ser,3.32(3); Glu,2.16(2); Pro,2.20(2); Ala,0.41(1); Leu,1.00(1)

Sequence	Asp-(Glx, Pro, Ser, Glx, Ser, Ser, CMCys, Lys, Pro, CMCys, Asn, His, CMCys)-CMCys-CMCys-Leu-$\frac{\text{Ser}}{\text{Ala}}$-Thr-Lys
Co B + Co A	3 hrs: Lys,1.00; Thr,0.55; Ser,0.27; Ala,0.069; Leu,0.29
Edman degradation	
Step 1	PTH-Asp
	Residue: CMCys,2.50*; Asp,1.42(1); Thr,1.00(1); Ser,2.20(2); Glu,1.50(2); Ala,0.36; Leu,1.00(1); Pro,1.55(2); Lys,0.31
Partial Acid Hydrolysis of C-2	
PA-1	Asp,1.08(1); Glu,1.06(1); Pro,0.98(1)
PA-2	Asp,1.00(1); Glu,1.86(2); Pro,0.93(1); Ser,1.20(1)

Edman degradation	
Step 1	Residue: Asp,0.13; Glu,2.00(2); Pro,1.00(1); Ser,1.20(1)
PA-3	Ser,0.98(1); Cysteic,1.20(1); Lys,1.00(1)
PA-4	Ser,2.14(2); Cysteic,1.00(1); Lys,0.60(1)
PA-5	Lys,1.85(2); His,0.94(1); Cysteic,3.20(3)*; Asp,1.30(1); Thr,1.06(1); Ser,2.30(2); Glu,0.62(1); Pro,1.20(1); Ala,0.33; Leu,0.66(1)
PA-6	Cysteic,2.00(2); Ser,1.00(1); Leu,1.00(1)
PA-7	Lys,1.00(1); Thr,0.98(1); Ser,0.60(1)
PA-8	Ser
PA-9	Asp

Summary

C-2 Asp-(Glx,Pro)-Ser-Glx-(Ser,Ser,Cys,Lys) (Pro,Cys,Asn,His,Cys)-Cys-Cys-Leu-Ser/Ala-Thr-Lys

C-2-PA-1 Asp-(Glu,Pro)

C-2-PA-2 Asp-(Glu,Pro)-Ser-Glu

C-2-PA-3 (Ser,Cys,Lys)

C-2-PA-4 (Ser,Ser,Cys,Lys)

C-2-PA-5 (Glu, Ser,Ser,Cys,Lys,Pro,Cys,Asp,His,Cys) Cys-Cys-Leu-Ser/Ala-Thr-Lys

C-2-PA-6 (Cys,Cys,Leu,Ser)

C-2-PA-7 Ser-Thr-Lys

C-2-PA-8 Ser

* Value for CMCys is low

TABLE XVIII

Amino acid sequence of Peptide C-3 (residues 29 through 39)

Composition: Ser,1.06(1); Ileu,0.99(1); Pro,2.00(2); CMCys,1.86(2); Glu,1.04(1); Arg,0.40(1); Thr,1.00(1); Asp,0.98(1); Leu,1.00(1)

Sequence	Ser-Ileu-Pro-Pro-(Glx, CMCys, Arg)-CMCys-Thr-Asp-Leu
Co A + Co B	16 hrs: Leu,1.00; Asp,0.92; Thr,0.62; CMCys,0.53
Edman degradation	
Step 1	PTH-Ser
	Residue: Ser,0.40; Ileu,1.00(1); Pro,2.10(2); Glu,1.10(1); CMCys,1.19(2)[a]; Arg,0.85(1); Thr,1.10(1); Asp,1.05(1); Leu,1.00(1)
Step 2	PTH-Ileu
	Residue: Ser,0.42; Ileu,0.48; Pro,1.90(2); Glu,0.97(1); CMCys--[a] Arg,0.83(1); Thr,1.03(1); Asp,1.05(1); Leu,1.00(1)

Step 3	PTH-Pro Residue: Ser,0.31; Ileu,0.27; Pro,1.50(1); Glu,1.20(1); CMCys,0.87(1)[a]; Arg,0.82(1); Thr,1.10(1); Asp,1.09(1); Leu,1.00(1)
Step 4	PTH-Pro Residue: Ser,0.56; Ileu,0; Pro,0; Glu,1.20(1); CMCys,0.65(1)[a]; Arg,0.89(1); Thr,1.00(1); Asp,1.10(1); Leu,1.00(1)

[a] Carboxymethylcysteine values are low in the Edman degradation. Oxidation of this residue to cysteic acid might have occurred during acid hydrolysis. Unfortunately cysteic acid was not evaluated in these analyses.

TABLE XIX

Amino acid sequence of Peptide C-4 (residues 40 through 45)

Composition: Arg,0.84(1); Leu,1.00(1); Asp,1.07(1); Ser,1.05(1); CMCys,0.90(1); His,1.00(1)

Sequence	Arg-Leu-Asp-Ser-CMCys-His
Co A + Co B	16 hrs: His,1.00; CMCys,0.80; Ser,0.65
Edman degradation	
Step 1	PTH-Arg
	Residue: Arg,0; Leu,1.00(1); Asp,1.04(1); Ser,0.95(1); CMCys,0.60(1); His,0.50(1)
Step 2	PTH-Leu
	Residue: Arg,0; Leu,0.10; Asp,1.10(1); Ser,1.00(1); CMCys,0.55(1); His,0.55(1)
Step 3	PTH-Asp
	Residue: Arg,0; Leu,0; Asp,0.33; Ser,1.00(1); CMCys,0.77(1), His,0.27(1)

TABLE XX

Amino acid sequence of Peptide C-5 (residues 46 through 55)

Composition: Ser,2.14(2); CMCys,2.96(3); Ala,1.10(1); Lys,1.14(1); Ileu,0.67(1); Thr,1.13(1); Leu,1.00(1).

Sequence	Ser-CMCys-Ala-Lys-CMCys-CMCys-Ileu-Ser-Thr-Leu ← C-5-T-1 → ←C-5-T-2 →
Co A + Co B	16 hrs: Leu,1.00; Thr,0.30; Ser,0.24
Tryptic Digest	
T-1	Ser,1.00(1); CMCys,0.50(1); Ala,1.00(1); Lys,1.00(1)
Edman degradation	
Step 1	Residue: Ser,0.28; CMCys,0.50(1); Ala,1.00(1); Lys,1.00(1)
Step 2	Residue: Ser, 0; CMCys,0; Ala,0.80(1); Lys,1.00(1)
T-2	CMCys,2.00(2); Ileu,1.00(1); Ser,1.00(1); Thr,1.00(1); Leu,1.00(1)
Edman degradation	

Step 1	Residue: CMCys,1.00(l); Ileu,1.00(l); Ser,0.10; Thr,1.08(l); Leu,1.44(l)
Step 2	Residue: CMCys, 0; Ileu,1.00(l); Ser,0.10; Thr,1.29(l); Leu,1.21(l)
Step 3	Residue: CMCys, 0; Ileu,0.20(l); Ser,0.25; Thr,1.00(l); Leu,1.00(l)
Step 4	Residue: CMCys, 0; Ileu,0.30; Ser,0.45; Thr,1.00(l); Leu,1.00(l)

Summary

C-5 Ser-CMCys-Ala-Lys-CMCys-CMCys-Ileu-Ser-Thr-Leu

C-5-T-1 Ser-CMCys-Ala-Lys

C-5-T-2 CMCys-CMCys-Ileu-(Ser, Thr, Leu)

TABLE XXI

Amino acid sequence of Peptide C-6 (residues 55 through 68)

Composition: Ser,1.09(1); Ileu,1.77(2); Pro,1.02(1); Ala,1.12(1); Glu,1.15(1); CMCys,1.35(1); Val,0.83(1); Thr,0.75(1); Asp,3.24(3); Phe,0.94(1)

Sequence	Ser-Ileu-Pro-Ala-Gln-CMCys-Val-Thr-(Ileu, Asp)-Asx-Asp-Phe
Co A	16 hrs: Phe, 0.12
LAP	4 hrs: Ser, 0.17
APM	16 hrs: Ser, 0.32
Edman degradation	
Step 1	Residue: Ser,0.20; Ileu,2.00(2); Pro,1.00(1); Ala,1.00(1); Glu,1.00(1); Cysteine,0.50(1); Val,0.8(1); Thr,1.00(1); Asp,3.20(3); Phe,1.00(1)
Step 2	Residue: Ser,0.31; Ileu,1.17(1); Pro,0.84(1); Ala,1.00(1); Glu,1.13(1); CMCys,0.40(1); Val,1.00(1); Thr,1.07(1); Asp,3.10(3); Phe,1.13(1)
Step 3	Residue: Ser,0.50; Ileu,1.00(1); Pro,0; Ala,1.00(1); Glu,1.00(1); CMCys,0.50(1); Val,0.85(1); Thr,1.00(1); Asp,3.14(3); Phe,1.07(1)
Pronase digest of C-6	
P-1	Ser,1.00(1); Ileu,0.70(1); Pro,0.50(1); Ala,0.80(1); Glu,1.00(1); CMCys,0.70(1)

P-2	Ser,1.00(1); Ileu,1.00(1); Pro,0.95(1); Ala,1.02(1); Glu,1.05(1)
P-3	Glu,1.33(1); Cysteic,1.00(1)
APM	Gln (+)
P-4	Cysteic,0.95(1); Val,1.00(1)
P-5	Val,1.00
P-6(acidic)	Asp,1.14(1); Ileu,1.00(1)
P-7(acidic)	Asp,3.00(3); Ileu,1.00(1)
P-8(acidic)	Asp,1.00(1); Phe,0.95(1)
Chymotrypsin digest of C(6+7)	
C-6a	Ser,1.00(1); Ileu,1.00(1); Pro,0.83(1); Ala,0.76(1); Glu,0.90(1); CMCys,1.40(1); Val,0.40(1); Thr,0.56(1)
C-6b	CMCys,0.93(1); Val,0.63(1); Thr,0.97(1); Ileu,1.04(1); Asp,3.32(3); Phe,0.91(1)
Edman degradation	
Step 1	Residue: Sample was lost
Step 2	Residue: CMCys,0; Val,0; Thr,0.82(1); Ileu,1.00(1); Asp,3.08(3); Phe,1.00(1)
Step 3	Residue: CMCys,0; Val,0; Thr,0.53; Ileu,1.00(1); Asp,3.00(3); Phe,1.00(1)

TABLE XXI (continued)

Summary

C-6	Ser-Ileu-Pro-Ala-Gln-CMCys-Val-Thr-(Ileu, Asp)-Asx-Asp-Phe
C-6-P-1	(Ser,Ileu,Pro,Ala,Glu,CMCys)
C-6-P-2	(Ser,Ileu,Pro,Ala,Glu)
C-6-P-3	(Gln,CMCys)
C-6-P-4	(CMCys,Val)
C-6-P-5	Val
C-6-P-6	(Ileu,Asp)
C-6-P-7	(Ileu,Asp,Asp,Asp)
C-6-P-8	Asp,Phe
C-6a	(Ser,Ileu,Pro,Ala,Glu,CMCys,Val,Thr)
C-6b	CMCys[a]-Val-Thr-(Ileu, Asp, Asp)-Asp-Phe

[a] Carboxymethylcysteine is placed in position 60 on consideration of the overlapping Peptides C-6-P-3, C-6-P-4

TABLE XXII

Amino acid sequence of Peptide C-7 (residues 69 through 84)

Composition: CMCys,2.07(2); Tyr,0.13(1); Glu,0.91(1); Pro,0.85(1); Lys,1.00(1); His,1.19(1); Asp,5.15(5); Ser,2.77(3)

Sequence	CMCys-Tyr-Glu-CMCys-Pro-Lys-(Ser, Ser, His, Ser, Asp, Asp, Asp, Asn, Asn)
Co B + Co A	3 hrs or 16 hrs: Asn (+)
LAP	CMCys,1; Tyr,0.53; Glu,0.33
Edman degradation	
Step 1	Residue: CMCys,0.85(1); Tyr,0.17(1); Glu,1.00(1); Pro,1.21(1); Lys,0.35(1); His,0.69(1); Asp,5.00(5); Ser,2.88(3)
Step 2	Residue: CMCys,0.60(1); Tyr,0; Glu,1.00(1); Pro,1.20(1); Lys,0.25(1); His,0.25(1); Asp,4.75(5); Ser,2.50(3)
Step 3	Residue: CMCys,0.40(1); Tyr,0; Glu,0.60; Pro,1.00(1); Lys,0.73(1); His,0.22(1); Asp,5.30(5); Ser,2.56(3)

Tryptic Digest	
T-1	CMCys,1.70(2); Tyr,0.20(1); Glu,1.04(1); Pro,0.96(1); Lys,1.00(1)
Co B + Co A	16 hrs: Lys,1.00(1)
LAP	CMCys,1.00; Tyr,0.50; Glu,0.25
Edman degradation	
Step 1	Residue: CMCys,1.00(1); Tyr,0.10(1), Glu,1.32(1); Pro,0.75(1); Lys,0.96(1)
T-2	His,1.00(1); Asp,5.20(5); Ser,2.86(3)

Summary

C-7 CMCys-Tyr-Glu-CMCys-Pro-Lys-[Ser,(3); Asp,(5); His,(1)]

C-7-T-1 CMCys-Tyr-Glu-CMCys-Pro-Lys

C-7-T-2 [Ser,(3); Asp,(5); His,(1)]

APPENDIX

STRATEGY OF AMINO ACID SEQUENCE DETERMINATION

A. General

The work of Sanger and his associates in 1951 on the structure of insulin was an important contribution to protein chemistry. It was the first instance in which the primary structure of a protein was completely established. Since then, great progress has been made in the development and application of methods for the determination of the amino acid sequence of a protein. The rationale and methodology of amino acid sequence determination has been extensively reviewed in a number of books and review articles (Schroeder, 1968; Leggett Bailey, 1967; Hirs, 1967).

As a preliminary to the analysis of the amino acid sequence of a protein, it is essential to establish its purity, molecular weight and amino acid composition. Various methods of analysis have been described, but the automatic amino acid analyzer introduced in 1956-1958 by the Rockefeller group (Spackman et al., 1958) has been the key instrument for a quantitative analyses of amino acids obtained from a hydrolysate of a protein.

To elucidate the amino acid sequence of the protein under study, the following steps are necessary: partial hydrolysis of the protein into peptides; isolation in pure form of the resulting peptides; determination of the sequences of the smaller peptides and the deduction of the entire sequence from

sequences of peptides of overlapping structure. These four points will be discussed below.

B. Preparation of Peptides

i) Denaturation

Most proteolytic enzymes show only minimal action toward native proteins, so it is usual to denature the protein before enzymatic digestion. The denaturing agent used will depend on the protein under study. Strong solutions of urea (8M) and guanidine hydrochloride (6M) have been successfully used in a number of cases.

Cleavage of disulfide bonds

The disulfide bonds of a protein contribute to its conformation, and these bonds have to be disrupted before enzymic digestion of the protein. They may be cleaved by either reduction or oxidation.

Reduction is commonly achieved with mercaptoethanol in 8M urea or 5M guanidine hydrochloride (Hirs, 1967). The resulting sulfhydryl groups may be blocked by alkylation with iodoacetic acid or iodoacetamide (Hirs, 1967), or modification with ethyleneimine (Cole, 1967).

Performic acid is the reagent commonly used to oxidize disulfide bonds. Cystine and cysteine are converted to cysteic acid and methionine is converted to methionine sulfone (Hirs, 1967).

ii) Partial hydrolysis of proteins

a) Enzymatic hydrolysis

Once a protein is denatured it is ready for enzymatic cleavage. The object is to hydrolyze certain bonds as completely and specifically as possible. Cleavage has to be specific since random cleavage will reduce the final yield of any peptide. Several proteolytic enzymes of high purity are commercially available.

Trypsin

This is the most specific proteolytic enzyme in common use. Its action is limited to peptide bonds associated with the carboxyl groups of lysine and arginine. Even then, because the substrate is complex, the various susceptible bonds may be cleaved at different rates and a Lys-Pro or Arg-Pro bond is rarely cleaved by the enzyme. Experimental conditions for trypsin digestion have been described by Smyth (1967). The pattern of tryptic action can be changed by chemical modification of the protein. Additional trypsin sensitive bonds can be introduced by aminoethylation of sulfhydryl groups (Cole, 1967). On the other hand, modification of lysine or arginine residues will reduce the number of trypsin sensitive bonds. The ε-NH_2 group of lysine may be blocked by trifluoroacetylation (Goldberger and Anfinsen, 1962); amidination (Hunter and Ludwig, 1962) or cyanoethylation (Riehm and Scheraga, 1966). In these, tryptic action is limited to arginine residues. Conversely, modification of arginine residues with cyclohexanedione (Toi et al., 1965) results in tryptic cleavage at lysine residues only.

Chymotrypsin

This enzyme splits bonds associated with the carboxyl groups of tryptophan, tyrosine, phenylalanine and leucine. Cleavage may also occur at bonds associated with methionine, aspargine, glutamine, histidine, threonine and lysine (Schroeder, 1968).

Other enzymes

Because of their wide specificity, other proteolytic enzymes such as papain, subtilisin, thermolysin and pronase are rarely used for initial digest on proteins to be sequenced. Digestion with such an enzyme would result in a very complex mixture of peptides which would be very difficult to separate. Furthermore, the yields of individual peptides would be very low.

Proteases of wide specificity are however, commonly used for the further hydrolysis of the larger peptides obtained by other methods of cleavage.

b) Chemical cleavage

The most widely used method for chemical cleavage is the reaction with cyanogen bromide (Gross, 1967). The carboxyl end of methionine residues is cleaved by the reagent; bond splitting is accompanied by conversion of methionine to homoserine.

C. Separation and Purification of Peptides

A variety of separation methods have been used, including ion exchange chromatography, gel filtration, paper

chromatography and high voltage paper electrophoresis. Separations on Dowex-50 (a strong cation exchanger) and Dowex-1 (a strong anion exchanger), using volatile buffers as the eluants, have been extremely useful (Schroeder, 1967).

In gel filtration, the separation largely takes place on the basis of size and a wide selection (Sephadex series, Bio-Gel series) of materials is available for this procedure. Since the separation can be carried out in such solvents as 50% aqueous acetic acid, this technique can be very useful for the separation of peptides which are insoluble under the conditions used in ion exchange chromatography.

Peptides may also be separated by paper chromatography, paper electrophoresis or a combination of both. These techniques have been used very successfully in a number of cases and the methodology has been extensively reviewed (Bennett, 1967; Leggett Bailey, 1967). Purification of peptides on paper is limited by the amount of sample that can be applied to the paper and also by the difficulty to recover certain peptides in good yield from paper. Losses of up to 50%-75% for each chromatogram are not uncommon.

It can be seen that many methods are available for the separation of peptides. Most often, a single method does not allow complete separation of all the peptides of a digestion mixture and one has to resort to a combination of several methods. The major problem still remains with the purification of large, relatively insoluble peptides.

D. End Group and Peptide Sequence Determination

i) Carboxyterminal

a) Enzymatic hydrolysis

Commercially available carboxypeptidases are very useful for the determination of carboxyterminal residues. Carboxypeptidases act on proteins or peptides and sequentially release the amino acids from the carboxyterminal end of the molecule. Carboxypeptidase A does not release lysyl, arginyl or prolyl residues and carboxypeptidase B will remove lysyl and arginyl residues. By taking advantage of the specificity of these enzymes and by doing time course studies of the amino acids released by them, information concerning the sequence in the carboxyterminal region of the peptide can be gained.

b) Chemical

The hydrazinolytic method of Akabori et al. (1956) may be employed for the determination of the carboxy-terminal amino acid of proteins and peptides. Hydrazine reacts with proteins and peptides to produce hydrazides of all amino acids except the COOH-terminal, which can be isolated and identified.

ii) Aminoterminal

a) Chemical

There are various chemical methods for identifying the aminoterminal residue of proteins or peptides. The principle behind these methods is to attach to the amino-terminal amino acid some group that will identify it in the

presence of all other amino acids after complete hydrolysis of the protein. Three methods commonly employed in this fashion are: the DNP-method of Sanger (1945) using the reagent 1-fluoro-2,4-dinitrobenzene, the cyanate method of Stark (1963), and the dansyl chloride procedure of Gray (1967). The last named method is the most sensitive, since it requires only ~1 mμmole of material. In these three methods, only the amino-terminal amino acid is identified and all the other residues in the protein are recovered as free amino acids in the hydrolysate.

Sequential degradation by the Edman method

The only chemical technique for the sequential removal and identification of amino acids starting from the aminoterminal is the Edman degradation (1950). It consists of coupling the aminoterminal α-amino group with phenylisothiocyanate to give the phenylthiocarbamyl derivative at pH 9.0-9.5. Anhydrous acid can then cleave the peptide bond between the first and second residue resulting in the anilino-thiazolinone derivative of the aminoterminal residue and a new peptide containing one less amino acid. The anilino-thiazolinone derivative can be extracted and recycled to give the phenylthiohydantoin derivative and the residual peptide is ready for the next step of the degradation.

The sequential degradation of a peptide can be monitored by direct chromatographic identification of the PTH derivative removed after each cycle. However an unequivocal identification of certain PTH derivatives is sometimes difficult.

Alternatively one can take a sample of the residual peptide after each cycle of the Edman degradation and determine its amino acid composition after acid hydrolyses. This method, known as the substractive Edman degradation, is more dependable, but it is more costly in material and it does not allow differentiation between Asp and Asn or Glu and Gln.

The sequential degradation and dansylation (Gray, 1967) technique is analagous to the substractive degradation but it employs a direct identification of the end group of the residual peptide by the dansyl chloride technique. The advantage of this method is the high sensitivity of the dansyl chloride reaction but the limiting factor is the capacity of the electrophoresis equipment used in identifying the DNS-amino acid.

b) Enzymatic

Leucine aminopeptidase (LAP) can be of great help in completing the sequence of a peptide. It sequentially releases amino acids starting from the NH_2-terminal end of a peptide. A time course study of the hydrolysis can give some information about the aminoterminal sequence. The enzyme does not hydrolyze bonds involving proline residues and enzymic action will stop at a peptide bond aminoterminal to the imide bond (Light, 1967).

E. Overlaps

To determine the sequence of a protein it is necessary to fragment the molecule by some method, examine

the peptides, fragment the molecule by another method, examine these peptides and then attempt to put the puzzle together on the basis of partial identities in the various pieces.

F. Determination of Disulfide Bridges

In the previous paragraphs, the cystinyl residues have been considered in their two halves and have been placed in the sequence as cysteic acid, carboxymethylcysteine or some other derivative. In final analysis, it is necessary to know the positions of the disulfide bridges.

The number of disulfide bridges in a given protein can be determined from the difference between the total number of cysteine residues (determined as cysteic acid after oxidation or as S-carboxymethylcysteine after reduction and alkylation) and the total number of free sulfhydryl groups present in the protein.

Free sulfhydryl groups in proteins are estimated by use of sulfhydryl reagents such as p-chloromercuribenzoate, (PCMB), N-ethylmaleimide (NEM) or Ellman's reagent (DTNB).

Disulfide bridges

In order to determine the location of the disulfide bridges, the protein has to be fragmented under conditions such that the disulfide bridges remain intact, with no disulfide interchange occurring. The most effective procedure is the initial digestion of the protein with pepsin at pH 2.0 and subsequent cleavage with trypsin and chymotrypsin

at pH 6.5. The low pH prevents disulfide interchange and adequately denatures the protein for the attack of pepsin after which the molecule becomes susceptible to the other two enzymes.

For the isolation of cystine containing peptides the so called "diagonal" procedure of Brown and Hartley (1963) is commonly employed. The digest is submitted to electrophoresis on paper at pH 6.5 and the dried electropherogram is exposed to performic acid vapour and dried again. A second electrophoresis is made at a 90° angle to the first under the same conditions and the peptides are detected with ninhydrin. All the peptides will lie on the diagonal except the cystine-containing peptides each of which has been converted to two cysteic-acid containing peptides by performic acid (these will have a different charge and hence a different electrophoretic behaviour). The spots that lie in pairs off the diagonal are cysteic acid peptides. A preparative procedure for the isolation of these cysteic acid peptides will be the next step and an examination of the amino acid composition of the cysteic acid peptide pairs will enable one to identify the position of the disulfide bridges in the protein.

LITERATURE CITED IN APPENDIX

Akabori, S., Ohno, K., Ikenaka, T., Okada, Y., Hanafusa, H., Haruna, I., Tsugita, A., Sugae, K., and Matsushima, T. (1956), Bull. Chem. Soc. Japan 29,507.

Bennett, J. C. (1967), Methods Enzymol. 11,330.

Brown, J. R., and Hartley, B. S. (1963), Biochem. J. 89,59P.

Cole, R. D. (1967), Methods Enzymol. 11,315.

Edman, P. (1950), Acta Chem. Scand. 4,283.

Goldberger, R. F., and Anfinsen, C. B. (1962), Biochemistry 1,401.

Gray, W. R. (1967), Methods Enzymol. 11,139.

Gross, E., and Witkop, B. (1962), J. Biol. Chem. 237,1856.

Hirs, C. H. W., Ed. (1967), Methods in Enzymology, Vol.11, New York, N. Y., Academic Press.

Hirs, C. H. W. (1967), Methods Enzymol. 11,197.

Hirs, C. H. W. (1967), Methods Enzymol. 11,199.

Hunter, M. J., and Ludwig, M. L. (1962), J. Am. Chem. Soc. 84,3491.

LeggettBailey, J. (1967), Techniques in Protein Chemistry, 2nd. Ed., Amsterdam, Elsevier Pub. Co.

Light, A. (1967), Methods Enzymol. 11,426.

Riehm, J. P., and Scheraga, H. A. (1966), Biochemistry 5,93.

Sanger, F. (1945), Biochem. J. 39,507.

Schroeder, W. (1967), Methods Enzymol. 11,351.

Schroeder, W. A. (1968), The Primary Structure of Proteins, New York, N. Y., Harper & Row, Publishers.

Smyth, D. G. (1967), Methods Enzymol. 11,214.

Spackman, D. H., Stein, W. H., and Moore, S. (1958), Anal. Chem. 30,1190.

Stark, G. R., and Smyth, D. G. (1963), J. Biol. Chem. 238, 214.

Toi, K., Bynum, E., Norris, E., and Itano, H. A. (1965), J. Biol. Chem. 240, PC 3455.