Studies on the structure and function of plant protein inhibitors of trypsin and α-amylase

Campos, Francisco de Assis Paiva

STUDIES ON THE STRUCTURE AND FUNCTION OF PLANT PROTEIN INHIBITORS OF TRYPSIN AND α-AMYLASE

by

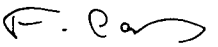
Francisco de Assis de Paiva Campos

(M.Sc. Fortaleza, Brazil)

A thesis submitted for the degree of Doctor of Philosophy
in the University of Durham

Department of Botany, December 1983

This thesis is entirely the result of my work. It has not been accepted for any other degree, and is not being submitted for any other degree.



F.A.P.CAMPOS

To Dr. José Xavier Filho

ABSTRACT

Studies of the primary structure of protein inhibitors of proteolytic enzymes and α-amylase were undertaken. The complete amino acid sequence of an α-amylase inhibitor and a bifunctional trypsin/α-amylase inhibitor from seeds of ragi (Eleusine coracana Gaertn.) and of the proteinase inhibitor PI-I from the Tracy cultivar of soybean (Glycine max L. Merr) were determined.

The bifunctional trypsin/α-amylase inhibitor from ragi seeds was shown to be a single polypeptide of 122 amino acids with a molecular weight of 13400. The two reactive (trypsin inhibitory) sites were also determined. Sequence comparisons revealed that this inhibitor seems to be divergently related to other trypsin and α-amylase inhibitors and also to the reserve protein from castor bean. In addition it is proposed that a new inhibitor family should be added to the existing ones (Laskowski and Kato, 1980) to accomodate this bifunctional inhibitor and its related proteins. The secondary structure of this inhibitor was also predicted.

The α-amylase inhibitor from ragi seeds was shown to be a single polypeptide of 95 amino acids with a molecular weight of 9300. The existence of two homologous regions in the amino acid sequence of this inhibitor seemed to indicate that the inhibitor molecule has arisen by a process of gene duplication. Sequence comparisons revealed that this inhibitor has no homology to any other α-amylase inhibitors, proteolytic enzyme inhibitors or any other plant protein of known primary structure. In addition the secondary structure of the α-amylase inhibitor from ragi seeds was also predicted.

The amino acid sequence of the proteinase inhibitor PI-I from the Tracy cultivar of soybean (Glycine max L. Merr) was shown to be identical to another proteinase inhibitor PI-II) from the same cultivar, which have been previously sequenced by Kashlan (1980).

ACKNOWLEDGMENTS

I express my gratitude to Dr. M. Richardson for the supervision of my work and for his kind advice and criticism. I am also grateful to him and his wife Mrs. Jean Richardson for their kind interest in my personal well being in Durham.

I also extend my gratitude to Mr. J. Gilroy and Dr. J. Yarwood for their generous help and interest in my work. Working with them was always a pleasurable experience.

Thanks are due to my colleague Mr. Hans Liesenberg for writing the computer programs utilised in this work. I am also grateful to the Duty Advisers from the Computer Unit of Durham University for their help.

My thanks to my brother Dr. F.I.P. Campos for taking the trouble of looking after my personal interests in Fortaleza.

During this work I have been financially supported by CAPES (Coordenação de Aperfeiçoamento do Pessoal do Ensino Superior) from the Brazilian Ministry of Education and by Universidade Federal do Ceará. I am very grateful to both.

I am grateful to my colleagues from the Departamento de Bioquímica e Biologia Molecular from the Universidade Federal do Ceará, Brazil, specially to Drs. I. L. Ainouz, J. Xavier Filho, M.G.S. Lima and R.A. Moreira for their continuous encouragament.

Some technical assistance from Mr. D. Bown, Mr. P. Miller, Mr. P. Preston, Mr. A. Reid and Mr. P. Sidney is gratefully acknowledged.

Thanks also to Dr. A. Gatehouse for her kind attention to me.

I wish to thank Prof. D. Boulter for the provision of research facilities in the Botany Department.

Finally, I would like to express my gratitude to my parents, brothers and sisters for their everlasting support and interest in my education. My gratitude to them has no limits.

ABREVIATIONS

The abreviations used in this thesis are as recommended in "Instructions to Authors", Biochem J. (1975) 131, 1-20.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

1. INTRODUCTION

The central role played by enzymes in almost every biological phenomenon has led to studies on the ways in which their biological activities are controlled. One of the ultimate goals of these studies is the hope that a clear understanding of enzymatic control may eventually lead to the manipulation of many biological processes as a means of controlling disease, enhancing agricultural productivity and to the controlled production for industrial or pharmaceutical purposes of biological substances.

In the last two decades much evidence has been accumulated indicating that protein inhibitors of enzymes are widely utilised in nature as a means of controlling unwanted enzymatic activity. A list of plant protein inhibitors of enzymes is presented in Table I. While information on the occurrence, distribution and biological function of protein inhibitors of enzymes others than proteinases and α-amylases is scanty, the amount of information on protein inhibitors of proteinases and α-amylases is considerable, although the level of understanding of the chemistry of these inhibitors far exceeds the level of understanding of their biological role. This introduction will be limited to a discussion of the protein inhibitors of proteinases and α-amylases, with particular emphasis on those of plant origin.

1.1. Protein Inhibitors of Proteinases

The protein inhibitors of proteinases can be grouped into two completely dissimilar classes: the macroglobulins and the inhibitors with class-specific reactive sites.

TABLE I - Range of enzymes known to be inhibited by proteins from plants and microorganisms

Enzyme Class	Enzyme	Source	References
Oxidoreductases	Alcohol dehydrogenase (E.C. 1.1.1.2)	Maize root	Lai and Scandalios, 1982
	NADH-Nitrate reductase (E.C. 1.6.6.3)	Soybean	Jolly and Tolbert, 1978
	Catalase (E.C. 1.11.1.6)	Maize	Sorensen and Scandalios, 1980
	Peroxidase (E.C. 1.11.17)	Avocado	Vanvelyveld and Bester, 1978
	Cinnamic-4-hydroxylase (1.14.13.11)	Gherkins	Billett et al., 1978
Transferases	Ornithine aminotransferase (E.C. 2.6.1.13)	Legumes	Dulloo, 1980
	Phospho-inositol kinase (E.C. 2.7.1.67)	Mungbean	Majunder and Biswas, 1973
Hydroxylases	Lipase (E.C. 3.1.1.3)	Green pepper	Kim et al., 1977
	Nuclease (E.C. 3.1.4.9)	Yeast	Ozumi et al., 1976
	DNase (E.C. 3.1.4.5)	Tobacco	Szopa and Wagner, 1980
	α-Amylase (E.C. 3.2.1.1)	Various sources	See Table II
	Cellulase (E.C. 3.2.1.4)	Slime mould	Jones and Gutta, 1981
	Acid invertase (E.C. 3.2.1.26	Potato	Anderson et al., 1980
	Endopolygalacturonase (E.C. 3.2.1.39)	Red kidney bean	Fisher et al.
	Proteolytic enzymes (e.g. trypsin, chymotrypsin, elastase)	Various sources	See Table II
	Urease (E.C. 3.5.1.5)	Melon	Malhotra and Rani, 1978
	ATPase (E.C. 3.6.1.3)	Yeast	Brooks and Senior, 1971
Lyases	Isocitrate lyase (E.C. 4.1.3.1)	Banana	Surendranatham and Nair, 1978
	Pectin lyase (E.C. 4.2.2.2)	Cucumber	Bock et al., 1975
	Endopectin lyase (E.C. 4.2.2.10)	Cucumber	Bock et al., 1975
	L-Phenylalanine ammonia lyase (E.C. 4.3.1.5)	Barley	Podstolski, 1981
	Guanylate cyclase (E.C. 4.6.1.1)	Bitter melon	Vesely et al., 1977

The macroglobulins are high molecular weight proteins, apparently present in the plasma of all mammals and in some invertebrates (Travis and Salvesen, 1983) but at this time have not been reported in any plants. These proteins seem to have an important role in the processing of plasma proteinases in the blood. They can form complexes with almost any proteinase, but in the complex the proteinase remains catalytically active against small synthetic substrates and this activity can be inhibited by small protein inhibitors of proteinases such as the pancreatic trypsin inhibitor (Kunitz) but not by larger ones such as the soybean trypsin inhibitor (Kunitz) (Laskowski and Kato, 1980; Travis and Salvesen, 1983). A "trap" mechanism was proposed for these inhibitors by Barrett and Starkey (1973). According to this mechanism, the proteinase hydrolyzes one or more particular susceptible peptide bonds in the inhibitor and triggers a conformational change in the inhibitor which traps the enzyme molecule. According to this mechanism the critical segment in the inhibitor must be rather long since it must contain peptide bonds matching the specificity requirements of the various proteinases. In this chemical sense the macroglobulins are so markedly different from all other known inhibitors that some scientists question whether they should be put in this class at all (Laskowski and Kato, 1980; Travis and Salvesen, 1983).

On the other hand, the inhibitors with class-specific reactive sites comprise a very broad class of proteins, which are of widespread occurrence in plants, animals and microorganisms (Pusztai, 1967; Liener and Kakade, 1969, 1980; Laskowski and Sealock, 1971; Ryan, 1973, 1979, 1981; Royer, 1975; Richardson, 1977, 1980, 1981; Laskowski and Kato, 1980; Whitaker, 1981). As in the case of the macroglobulins, these inhibitors form complexes with proteinases, but the enzymatic activity

of the proteinase in the complex is completely abolished and, at least in the well characterized systems, a particular inhibitory reactive site can inhibit only proteinases belonging to a single one of the six mechanistic classes proposed by Ryan and Walker-Simmons (1981).

1.1.1. Inhibitors with Class-Specific Reactive Sites

Substances inhibiting proteolytic enzymes have long been known. A first indication of the occurrence of such inhibitors is found in the literature as early as 1878 (Fredericq, 1878). The first indication of a proteinase inhibitor in plants was given by Read and Haas (1938). This was soon followed by the demonstration of the inhibition of trypsin by extracts of Glycine max by Ham and Sandstedt (1944) and the first isolation of a plant protease inhibitor from Glycine max by Kunitz (1945, 1946). A year later the first systematic study of plant protease inhibitors was made by Borchers and Ackerson (1947). These early studies were mainly restricted to the inhibition of trypsin and, to a lesser degree of chymotrypsin, at the time the most studied proteolytic enzymes. This fact, incidentally, has led to difficulties in interpreting the studies on the physiological role of these proteins in plants, as only a few serine endoproteinases have been found in plants (Ryan, 1981; Ryan and Walker-Simmons, 1981).

Inhibitors with class-specific reactive sites are often found as components of the cytoplasm, secretions and intercellular fluids of many organs and tissues in animals (Laskowski and Kato, 1980) and in plants are usually found in seeds, tubers, leaves and roots (Table II). The inhibitors of plant origin will be the only ones to be considered here.

Table II - Plant protein inhibitors of proteolytic enzymes

Mechanistic Set	Enzyme	Inhibitor Source	Part of Plant	References
Serine endopeptidases	Trypsin	Hordeum vulgare (barley)	Embryo	Mikola and Kirsi, 1972
		Hordeum vulgare	Endosperm	Mikola and Kirsi, 1972
		Solanum melongena (eggplant)	Exocarp	Kanamori et al., 1975
		Allium sativum (garlic)	Bulb	Sumathi and Pattabiraman, 1975
		Amorphophallus companulatus	Tuber	Sumathi and Pattabiraman, 1975
		Anacardium occidentale (cashew nut)	Nut	Xavier-Filho and Ainouz, 1977
		Bambusaria arundinaria (bamboo)	Sprouts	Suguira, 1975
		Lycopersicum esculentum (tomato)	Leaves	Nelson et al., 1981
		Eleusine coracana (ragi)	Seeds	Shivaraj and Pattabiraman, 1981
		Scopolia japonica	Callus	Sakato et al., 1975
		Vigna unguiculata (cowpea)	Pods	Carasco and Xavier-Filho, 1981
	Chymotrypsin	Hordeum vulgare (Barley)	Shoots	Kirsi and Mikola, 1977
		Ipomoea batata	Tuber	Balls and Ryan, 1963
	Trypsin/chymotrypsin	Phaseolus vulgaris (kidney bean)	Seeds	Puzstai, 1968
		Glycine max (soybean)	Seeds	Odani and Ikenaka, 1972
		Eleusine coracana (ragi)	Seeds	Shivaraj and Pattabiraman, 1982
	Subtilisin	Hordeum vulgare (barley)	Seeds	Yoshikawa et al., 1976
		Vigna unguiculata (cowpea)	Seeds	Vartak et al., 1980
		Hevea brasiliensis	Latex	Archer, 1983
	Enteropeptidase	Phascolus vulgaris (red kidney bean)	Seeds	Jacob et al., 1983
Sulfhydryl endopeptidase	Papain	Vigna unguiculata (cowpea)	Seeds	Vartak et al., 1980
	Bromelain	Ananus sativum (pineapple)	Stem	Reddy et al., 1975
Acid proteinase	Cathepsin D	Solanum tuberosum (potato)	Tubers	Keilova and Tomasak, 1976
Metalo exopeptidase	Carboxypeptidase A	Solanum tuberosum (potato)	Leaves	Graham and Ryan, 1981
	Carboxypeptidase A	Lycopersicum esculentum (tomato)	Fruit	Hass and Ryan, 1980[a]

Based on mechanistic considerations, six sets of proteolytic enzymes can be identified. Hartley (1960) originally defined four such sets: the serine, thiol, metal and acid proteinases. This classification was modified by Walsh (1975) who subdivided the metaloenzymes into an endopeptidase and an exopeptidase set, and later Ryan and Walker-Simmons (1981) proposed the subdivision of serine proteinases into two sets: the serine endopeptidases and the serine exopeptidases, giving a total of six mechanistic sets. Apart from inhibitors of metaloendoproteinases and serine exopeptidases, inhibitors of all of the remaining mechanistic sets have been found in plants (Table II). A detailed knowledge of the mechanism of inhibition is available only for the serine endoproteinase and metaloexopeptidase inhibitors (Laskowski and Kato, 1980; Rees and Lipscomb, 1982). For the remaining inhibitors there is no definitive model for their mechanism of action.

The overall mechanism of the serine endoproteinase/inhibitor interaction, including only those intermediates whose existence has been definitely shown, can be written (Laskowski and Kato, 1980) as:

$$E + I \rightleftharpoons L \rightleftharpoons C \rightleftharpoons X \rightleftharpoons L^* \rightleftharpoons E + I^* \quad \text{(Eq. 1)},$$

where E is the enzyme, I and I*, virgin and modified inhibitors, respectively, L and L*, noncovalently (rapidly dissociable) complexes of E with I and I* respectively, X is the relatively long-lived intermediate in the E + I* reaction and C is the stable enzyme-inhibitor complex. Detailed discussion on this mechanism may be found in a recent review by Laskowski and Kato (1980).

The only metaloexopeptidase inhibitor whose mechanism of action is known is the carboxypeptidase A inhibitor from potatoes

(PCI), isolated by Ryan et al (1974) and sequenced by Hass et al (1975). The refined crystal structure of the carboxypeptidase A/carboxypeptidase A inhibitor (CPA/PCI) complex at 2.5Å resolution was determined by Rees and Lipscomb (1980, 1982), who proposed the following mechanism of inhibition:

$$CPA + PCI \rightleftharpoons (CPA/PCI) \rightleftharpoons (CPA/PCI + Gly) \rightleftharpoons CPA + PCI^* + Gly \quad \text{(Eq. 2a)}$$

$$CPA + PCI^* \rightleftharpoons (CPA/PCI) \quad \text{(Eq. 2b)}$$

Mechanism 2a refers to the inhibition of carboxypeptidase A by the virgin and mechanism 2b refers of the inhibition of carboxypeptidase A by the modified inhibitor.

Like the serine endoproteinase inhibitors, the reactive site of the carboxypeptidase inhibitor is also hydrolyzed (Hass and Ryan, 1980b; Rees and Lipscomb, 1982). Unlike the carboxypeptidase A inhibitor, however, where the cleavage of Gly^{39} (the C-terminal amino acid) goes essentially to completion, the equilibrium constant between the modified and virgin serine endoproteinase inhibitor species is near unity (Laskowski and Kato, 1980). For the cleaved species to retain inhibitory activity, the two fragments must be held together by disulphide bridges or non-covalent interactions. This behaviour contrasts with the carboxypeptidase A case: both PCI and PCI* are equally effective inhibitors, even though Gly^{39} has been removed from the latter species. The differences in these two classes of proteinase inhibitors may reflect the fact that the loss of the peptides from a cleaved serine endoproteinase inhibitor prevents a significant portion of the binding interactions with the corresponding protease (Rees and Lipscomb,1982).

The protein inhibitors of proteinases others than serine endoproteinases still remain an understudied group, although their occurrence in plants is well documented. Inhibitors of sulfyhydryl proteinases have been isolated from pineapple stems (Reddy et al., 1975), potato tubers (Hoff et al., 1972), corn (Abe et al., 1980), Bauhinia seeds (Goldstein et al., 1973) and mung beans (Baumgartner and Chrispeels, 1976). Inhibitors of acid proteinases were isolated from potatoes (Keilova and Tomasek, 1976) and Bauhinia seeds (Goldstein et al., 1973). Metalocarboxypeptidase inhibitors were isolated from potato tubers (Ryan et al., 1974) and tomato fruits (Hass and Ryan, 1980).At the present there is no recorded example of a metaloendoproteinase inhibitor in plants.

The inhibitors of serine endoproteinases are the best known group of class-specific reactive site inhibitors. Their mechanism of action is fairly well understood and the primary structures of many of them have been determined (Richardson, 1977, 1980, 1981; Croft, 1980; Ramshaw, 1982). Despite being a class of proteins whose primary structures vary from one another, they all appear to obey the same mechanism of interaction/inhibition of their relevant proteases (Eq. 1). The elucidation of this mechanism of action was achieved by chemical modification studies (Sealock and Laskowski, 1969; Leary and Laskowski, 1973; Kowalski et al., 1974; Jering and Tschesche, 1976; Odani and Ikenaka, 1978), which helped to elucidate the role of the amino acids at the reactive site and their surroundings on the inhibition and also by both kinetic and X-ray crystalographic analysis (Ruhlman et al., 1973; Janin et al., 1974; Bolognasi et al., 1981; Ako et al., 1974; Gaier et al., 1981; Read et al., 1983).

An identification system for the amino acids at the reactive

site of the serine endoproteinase inhibitors (Scherchter and Berger, 1967) is given below:

reactive site

$$-P_3-P_2-P_1 \;\downarrow\; P_1'-P_2'-P_3'-$$

In the reaction of the inhibitor with the protease, the inhibitor behaves as a good substrate. A peptide bond on the surface of the inhibitor (the reactive site) where the P_1 residue matches the specificity requirement of the protease, specifically interacts with the active site of the protease. However, due to the fact that the values of K_{cat} amd K_M are much lower than the values for normal substrates, at low concentration of inhibitor and neutral pH the hydrolysis of the peptide bond is extremely slow and the system behaves as if it were a simple equilibrium between the enzyme on one hand and the complex on the other (Laskowski and Sealock, 1971; Laskowski and Kato, 1980; Read et al., 1983). It should be noted that the hydrolysis of the reactive site peptide bond does not go to completion. At neutral pH the equilibrium constant between virgin and modified inhibitor is near unity (Laskowski and Kato, 1980). The cleavage of the reactive site peptide bond creates a problem for the purification of inhibitors by affinity chromatography, as it may alter their electrophoretic mobilities, isoelectric points and, most importantly, creates difficulties during any subsequent sequencing of the inhibitor. However, this difficulty can be overcome by using anhydrotrypsin or anhydrochymotrypsin as the column ligand. In these derivatives of trypsin and chymotrypsin the serine at the reactive site is converted to dehydroalanine by reactions which selectively remove water from the serine R group (Ako et al., 1974). Despite being catalytically inactive,

anhydrotrypsin and anhydrochymotrypsin can still form stable complexes with their inhibitors (Ako et al., 1974; Yung and Trowbridge, 1980).

It was initially thought that the only reason why the modified inhibitor (with its reactive site peptide bond hydrolyzed) did not dissociate to give two peptide chains was due to the fact that the reactive site was localised in a disulphide loop. However, it seems that the positioning of the reactive site in a peptide loop is not an essential requirement for the non-dissociation of the inhibitors, as inhibitors such as the yeast proteinase B inhibitor 2 (Maier et al., 1979), the leech (Hirudo medicinalis) trypsin inhibitor (Seemuller et al., 1980) and a lysine-rich protease inhibitor from barley (Svendsen and Jonassen, 1980) are devoid of intramolecular disulphide bridges and apparently obey the standard mechanism of inhibition, which suggests that strong non-covalent bonding can substitute for a disulphide bridge. Furthermore, it has been recently demonstrated that carboxymethylation of the two cysteines forming the only disulphide bridge in the potato inhibitor I had no effect on the inhibitory activity of this inhibitor (Plunkett and Ryan, 1980). Incidentally, the existence of these cysteine-free protease inhibitors, which also have an unusual thermal stability, also indicates that the resistance of most proteinase inhibitors to heat denaturation (Richardson, 1977; Laskowski and Kato, 1980; Ryan, 1981) is not dependent only on the tight packing of the molecules due to disulphide bonding.

The number of proteinase inhibitors which have been isolated and partially characterized is enormous. This creates difficulties in making comparisons of the specificities, thermodynamic and kinetic parameters of the enzyme-inhibitor interactions. In order to facilitate

these comparisons, Laskowski and Kato (1980) proposed to assign each inhibitor to a family. The assignment of an inhibitor to a family was based primarily on amino acid sequence homology. As the proposed classification was concentrated on the serine endoproteinase inhibitors, a knowledge of the reactive site, details of the topological relationships between their disulphide bridges and evidence that the standard mechanism was obeyed, would be necessary before definitely assigning an inhibitor to a certain family. The inhibitor families, as proposed by Laskowski and Kato (1980) are presented in Table III.

The assignment of an inhibitor to a family (or otherwise to assign an inhibitor to a new family) is not always a straightforward task. Difficulties arise when a particular inhibitor has weak homology with inhibitors belonging to different families. The decision whether the homology is due to divergent or convergent evolution is a difficult one. It is thought that the inhibitors within a family are divergently related, while the families are apparently convergently related, but highly distant divergent relationships cannot be discarded and this is in fact suspected for the Streptomyces subtilisin inhibitor family and the pancreatic secretory trypsin inhibitor (Kazal) family (Laskowski and Kato, 1980). Thus, the placement of an inhibitor within a family, or in other words, to decide to which others inhibitors a particular inhibitor related in terms of evolution, may depend not only on a knowledge of its primary structure. It is now suspected that three-dimensional structures of proteins change more slowly than amino acid sequences during evolution (Phillips et al., 1983). In this context it is interesting to note that detailed crystalographic studies of members of three inhibitor families revealed that their three-dimensional structures are strikingly different (Laskowski and

TABLE III - Families of protein inhibitors that inhibit serine endoproteinases and obey the standard mechanism, as proposed by Laskowski and Kato (1980).

I. Bovine pancreatic trypsin inhibitor (Kunitz) family
II. Pancreatic secretory trypsin inhibitor (Kazal) family
III. Soybean trypsin inhibitor (Kunitz) family
IV. *Streptomyces* subtilisin inhibitor family
V. Soybean proteinase inhibitor (Bowman-Birk) family
VI. Potato inhibitor I family
VII. Potato inhibitor II family
VIII. *Ascaris* trypsin inhibitor fmaily
IX. Other families

Kato, 1980; McPhallen et al., 1983).

The number of inhibitor families originally proposed by Laskowski and Kato (1980) was recently added to by the creation of a new family to accomodate two small trypsin inhibitors from squash seeds (Wilusz et al., 1983). These are the smallest proteinase inhibitors so far known, being composed of 29 amino acids (molecular weight 3300) and seem to obey the standard mechanism. One may conclude that further families will have to be added as more sequence information becomes available (see Results and Discussion section).

Just as the number of inhibitor families is likely to increase in the future, so will the members of each of the existing families; the number of members of the potato inhibitor I family (Richardson, 1974; Richardson and Cossins, 1974) was recently expanded by the inclusion of two cysteine-free protease inhibitors from barley (Svendsen et al.,1982) and a cysteine-free inhibitor from leech (Hirudo medicinalis) (Seemuller et al., 1980). The inclusion of the leech inhibitor in this family is particularly interesting as, besides being a cysteine-free inhibitor, it is also the first inhibitor of animal origin to be shown to have homology with a plant protease inhibitor (Figure 1). The inhibitors of the potato inhibitor I family also show a small degree of similarity with a cysteine-free protease A inhibitor 2 from yeast (Maier et al., 1979). This similarity,however, is not strong enough to suggest divergent evolution.

Similarly, until recently the soybean trypsin inhibitor (Kunitz) family was composed of a single member. As sequence data became available for the inhibitors from seeds of Peltophorum africanum (Joubert, 1981), silktree (Albizzia julibrissin) (Odani

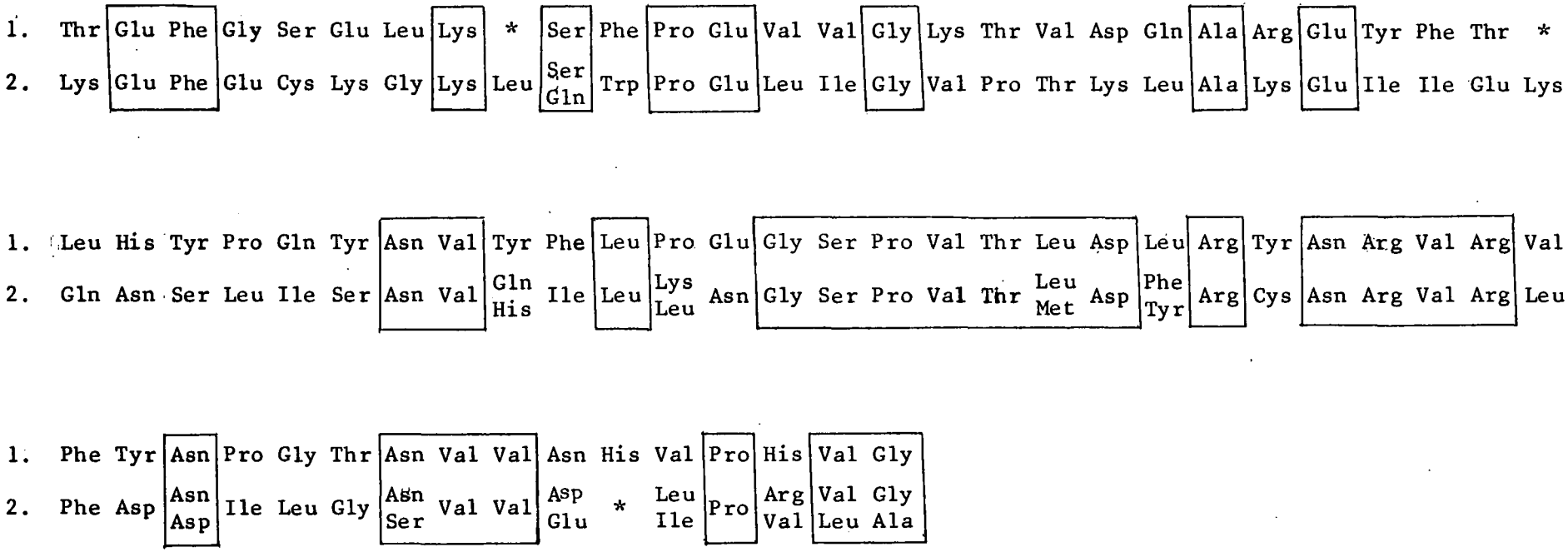


FIGURE 1 - Comparison of the amino acid sequences of eglin (an elastase-cathepsin G inhibitor) from leech (Hirudo medicinalis) (1) (Seemuller et al., 1980) and the polymorphous chymotrypsin inhibitor I from potatoes (2) (Richardson and Cossins, 1974). Identical residues are enclosed in boxes; (*) indicates gap.

et al., 1980), acacia (Acacia elata) (Kortt and Jernyn, 1981), Erythrina humana (Joubert, 1982), Acacia siberana (Joubert, 1983) and winged bean (Psophocarpus tetragonolobus) (Kortt et al., 1983), all of these inhibitors were seen to belong to this family. Apart from the winged bean inhibitor all of the others inhibit trypsin and chymotrypsin, while the soybean inhibitor inhibits only trypsin. The inhibitors from the silktree and acacia are both composed of subunits linked by disulphide bridges (Odani et al., 1980).

Biological Role of Proteinase Inhibitors

Despite the advances in the knowledge of the chemistry of the class-specific reactive site inhibitors, their possible biological role is still largely unknown, and at best, presumptive. Although their gross physiological function may be the prevention of unwanted proteolysis, a detailed physiological function is still not known. Possibly the problem lies in the fact that for most inhibitors so far studied, the true target enzyme or the enzyme that is most strongly inhibited is not known. As Laskowski and Kato (1980) put it " they become proteins in search of a function, rather than proteins isolated to account for a previously discovered biological function."

The possible role of proteinase inhibitors in heightening the fitness of plants against the attack of phytofagous animals is illustrated by the work of Professor Ryan's group (Green and Ryan,1972; McFarland and Ryan, 1974; Walker-Simmons and Ryan, 1977; Graham and Ryan, 1981; Cleveland and Black, 1982; Plunkett et al., 1982) who showed that the wounding of tomato leaves or potato leaves leads to a great increase in the levels of a polysacharide hormone, proteinase

inhibitor inducing factor (PIIF), which in turn leads to a huge increase in the levels of serine and metalocarboxypeptidase inhibitors. It was also established that PIIF activity is generally distributed throughout the plant kingdom (McFarland and Ryan, 1974). Another piece of evidence for the role of proteinase inhibitors as protective agents, comes from the work of Gatehouse et al.(1979,1980) and Gatehouse and Boulter (1983) who showed that the increased resistance of the seeds of the cultivar TVU 2027 of Vigna unguiculata is due, at least in part, to elevated levels of trypsin inhibitors. This is the clearest example to date that the increased resistance of one seed variety over another toward insect attack may be due to elevated trypsin inhibitor content. Similar observations have been made before by Weiel and Hapner (1976) who also suggested that varietal differences in the levels of proteinase inhibitors in barley plants, may possibly be related to their differential susceptibility to grasshopper (Camula pelucida) damage. Peng and Black (1976) also showed that the proteinase inhibitor activity of blight resistant varieties of tomato plants increased following infection by Phytophthora infestans .

The release into the external medium of proteinase inhibitors during the germination of seeds seems to be a general phenomenon, at least among legume seeds (Wilson, 1980). David et al.(1977) and Wilson et al. (1982) hypothesized that the release of proteinase inhibitors during germination may play an important role in the establishment of a "normal" microbial complement in the rhizosphere of legumes and probably other plant families. Such an effect may be both due to the inhibition of pathogenic organisms and to the stimulation of beneficial symbionts.

The indications that protein inhibitors of proteinases together with lectins, tannins, lignin, cellulose and unassimilable starch, may play a role in plant defence are strong (Janzen, 1980, 1981a, 1981b). However, it seems that the relevance of each one of these compounds for a given plant is not absolute, but instead may vary according to the ability of a particular predator to bring about chemical or physical detoxification and there are many indications that protein inhibitors of proteinases are toxic for some organisms, but completely harmless for others. The rodent Liemys salvini can live with a sole diet of seeds of Enterolobium cyclocarpum, a mimosaceous legume that contains substantial amounts of proteinase inhibitors, but no lectins. This rodent is also able to live with a diet containing 25% of soybean trypsin inhibitor, but will die if its diet contains only black bean (Phaseolus vulgaris) seeds which have proteinase inhibitors and a high content of lectins, but are without alkaloids or uncommon amino acids (Janzen, 1981). Janzen et al (1977) also found that the southern cowpea weavils (Calosobruchus spp) could ingest trypsin inhibitors in concentrations of up to 5% of the diet without adverse effect. However, proteinase inhibitors from potatoes are lethal at this same concentration in the diet. Similarly, purified maize inhibitors at 2-5% in the diet have no effect on young corn borer larvae (Steffens et al., 1978), but larval growth is impaired by soybean trypsin inhibitor.

The inhibitory activities of most known proteinase inhibitors are primarily directed against the digestive proteinases found in animals and microorganisms and only a few are known that inhibit plant proteinases (Richardson, 1977; Ryan, 1981). The proteinase inhibitors

that inhibit endogenous proteinases may have a regulatory or protective role(s) in controlling these proteinases before or during germination. Such inhibitors have been isolated from wheat (Preston and Kruger, 1976); barley (Mikola and Enari, 1970; Kirsi and Mikola, 1971); mung bean (Baumgartner and Chrispeels, 1976); corn (Reed and Penner, 1978); scots pine (Pinus sylvestris) (Salmia and Mikola, 1980; Salmia, 1981); buckwheat (Fagopyrum esculentum) (Belozerskii, 1982) and cashew nuts (Annacardium occidentalis) (Xavier-Filho and Ainouz, 1977).

The inhibitors of endogenous proteinases which were more thoroughly studied, are those from barley (Kirsi and Mikola, 1971), mung bean (Baumgartner and Chrispeels, 1976; Chrispeels and Baumgartner, 1978) and scots pine (Salmia and Mikola, 1980; Salmia, 1981). Precise knowledge of the subcellular compartmemtation of inhibitors and proteases is available only in the mung bean case, where the inhibitor is known to be localized in the cytosol, whereas the major mung bean protease is localized in the protein bodies. The function of the inhibitor seems to be to protect the cytoplasm from the very high and unspecific proteinase activity present in protein bodies. Although data on the intracellular localization of the proteases and inhibitors from scots pine and barley is not available, indirect evidence suggest a similar compartmentation and function as in mung bean (Mikola, 1983). On the other hand, the inhibitors from cowpea (Royer et al., 1974), rice (Horigushi and Kitagashi, 1971) and lettuce (Shain and Maier, 1968) are thought to regulate endogenous proteinases before germination.

The large quantities of inhibitors present in seeds and in some cases their transitory presence in different organs in some plants may suggest a possible role as storage proteins (Ryan, 1973; Richardson, 1977, 1980, 1981). Carasco and Xavier-Filho (1981) found evidence for

a paralelism between the synthesis of reserve proteins and the synthesis of proteinase inhibitors in cowpea. The high cystine content of many inhibitors, makes them attractive as sulphur reserves for the developing seedling (Pusztai, 1972). In some storage tissues inhibitors of serine endoproteinases decrease during germination, being apparently utilized by the seed for growth (Xavier-Filho, 1973, 1974; Xavier-Filho and Negreiros, 1979; Baumgartner and Chrispeels, 1976; Freed and Ryan,1978). However as yet it has not been shown that this decrease corresponds to any specific biochemical or physiological event.

The findings discussed above, apparently indicate that, as a class of proteins, inhibitors of proteinases do not have a sole biological role and that, probably due to their peculiar properties they may, for example, double as a sulphur source for the germinating seed and as a protective agent against the attack of predators. In this connection, Adams and Rinne (1981) put forward the idea that the protein inhibitors of proteinases, together with lectins, β-amylases, urease and some proteinases, could be classified as "dispensable proteins" as opposed to the "undispensable proteins" such as glycolytic enzymes, ribosomal proteins and cytochromes. As some plants survive quite well without obvious examples of these "dispensable proteins" and also as their physiological function sometimes does not appear to be related to their intrinsic biochemical properties, these proteins may fulfill a non-specific role as part of the cellular protein complement and therefore could presumably be replaced by other proteins in other species.

1.2. Protein Inhibitors of α-Amylase

Protein inhibitors of α-amylase have been known to occur in plants since Chrzasczs and Janick (1933,1934) described their presence in wheat seeds. Since then many more have been isolated from various plants sources (Table IV). As with the protein inhibitors of proteinases, they seem to be widespread throughout the plant kingdom. However, in contrast to more than thirty years of extensive research on protein inhibitors of proteinases, until recently protein inhibitors of α-amylase have received little attention. As with proteinase inhibitors, the main motivation for the study of α-amylase inhibitors is their possibly deleterious effect on human and animal nutrition, as well as their potential use in clinical disorders such as obesity and diabetes mellitus.

Plant protein inhibitors of α-amylase have been reviewed Marshall (1975), Buonocore et al. (1977), Kassel (1978), Richardson (1980, 1981) and Whitaker (1981). A recent review by Warchalewski (1983) deals specifically with α-amylase inhibitors from cereals.

1.2.1. Occurrence and Distribution

The protein inhibitors of α-amylase appear to have been most extensively studied in the Leguminosae, Gramineae and Solanaceae, probably because of the large number of species in these families which form important food sources. The detection of α-amylase inhibitory activity is sometimes difficult because of the long time required for the formation of the enzyme-inhibitor complex (Marshall, 1975;Buonocore et al., 1977; Warchalewski, 1983) and also due to the high endogenous

Table IV - Protein inhibitors of α-amylase in plants

Inhibitor Source	Part of Plant	References
Avena sativa (oats)	Seeds	Elliot and Leopold, 1953
Cajanus cajan	Seeds	Shivaraj and Pattabiraman, 1976
Canavalia ensiformis (jack bean)	Seeds	Jaffé et al., 1973
Cicer aretinium (chick pea)	Seeds	Jaffé et al., 1973
Colocasia antiquorum	Tubers	Sharma and Pattabiraman, 1980
Colocasia esculenta (taro)	Tubers	Rao et al., 1967
Dioscorea alata (yam)	Tubers	Sharma and Pattabiraman, 1982
Dolichos biflorus (horse gram)	Seeds	Jaffé et al., 1973
Dolichos lablab (field bean)	Seeds	Jaffé et al., 1973
Eleusine coracana (ragi)	Seeds	Shivaraj and Pattabiraman, 1980
Hordeum vulgare (barley)	Seeds	Mundy et al., 1983
Mangifera indica (mango)	Fruit	Mattoo and Modi, 1970
Penisetum typhoideum (pearl millet)	Seeds	Chandrasekher et al., 1981
Pisum sativum (garden pea)	Seeds	Jaffe et al., 1973
Phaseolus vulgaris (red kidney bean)	Seeds	Powers and Whitaker, 1977
P. vulgaris (white kidney bean)	Seeds	Marshall and Lauda, 1975
P. lunatus (lima bean)	Seeds	Jaffe et al., 1973
P. coccineus (french bean)	Seeds	Jaffe et al., 1973
P. aborigenes (wild bean)	Seeds	Jaffe et al., 1973
Lens culinaris (lentils)	Seeds	Jaffe et al., 1973
Secale cereale (rye)	Seeds	Granum, 1978
Setaria italica (setaria)	Seeds	Chandrasekher et al., 1981
Sorghum bicolor (Sorghum)	Seeds	Chandrasekher et al., 1981
Triticum aestivum (wheat)	Seeds	Buonocore et al., 1977
Zea maiz (maize)	Seeds	Blanco-Labra and Iturbe-Chiñas, 1980

amylase activity in some plants (Warchalewski, 1983). These facts seems to be at least in part responsible for the initial failure to detect α-amylase inhibitors in some plants, such as soybean (Jaffe et al., 1973), maize, barley and oats (Kneen and Sandstedt, 1946), horsegram and mothbean (Subbulakshmi et al.,1976), and blackgram (Reddi and Salunkhe, 1980).

All of the protein inhibitors of α-amylase so far studied were isolated from seeds, fruit and tubers (Marshall, 1975;Buonocore et al., 1977; Kassel, 1978; Warchalewski, 1983) but information on the distribution within these tissues is scanty, although at least in cereals they are thought to be concentrated in the endosperm, associated with starch granules (Buonocore et al., 1977; Blanco-Labra and Iturbe-Chinas, 1980; Mundy et al., 1983). As for their subcellular localization no information is available.

1.2.2. Specificity and Mechanism of Action

The in vitro specificity of most protein inhibitors of α-amylase seems to be directed against α-amylases other than the endogenous ones present in the plants from which the inhibitors were extracted. A uniform specificity pattern among inhibitors from legumes and cereals does not exist, although inhibition of human pancreatic and salivary α-amylase and α-amylases from insects is commonly found (Table V). However it is not always clear whether this is due to the absolute specificity of the inhibitor or to analytical shortcomings, such as the non-fulfilment of optimal conditions of pH, ionic strength and temperature required by each α-amylase in order for inhibition to be detected. Also the very slow rate of complex formation between the

TABLE V - Specificities of α-amylase inhibitors from plants

Inhibitor Source	α-Amylase Tested	References
Phascolus vulgaris (red kidney bean)	Porcine pancreas (+), human pancreas (+), human saliva (+), Tenebrio molitor (+), Bacilus subtilis (-), Aspergillus oryzae (-), barley (-), red kidney bean (-)	Powers and Whitaker, 1977
Zea maiz (maize)	Maize (+), Bacilus subtilis (+), Tribolium castaneum (+), Sitophilus zeamais (+), Rhizpertha dominica (+), rye (-) wheat (-), triticale (-), sorghum (-), human saliva (-), hog pancreas (-), Aspergillus oryzae (-), Bacilus subtilis (-)	Blanco-Labra and Iturbe-Chinas, 1981
Phaseolus vulgaris (red kidney bean)	Tenebrio molitor (+), Tribolium castaneum (+), Tribolium confusum (+), Anagasta khuniela (+), Sitophilus granarius (larvae) (+), Sitophilus granarius (adult) (-)	Powers and Culberston, 1982; Jaffe et al., 1973
Phaseolus vulgaris (white kidney bean)	Hog pancreas (+), human saliva (+), human pancreas (+), Bacilus amyloliquifaciens (-), Bacilus subtilis (-), rye (-), Bacilus licheniformis (-), Aspergillus oryzae (-) barley malt (-)	Marshall and Lauda, 1975
Hordeum vulgare (barley)	Barley (+)	Mundy et al., 1983
Hordeum vulgare (barley)	Malted barley α-amylase II (+), germinated wheat α-amylase II (+), malted barley α-amylase I (-), germinated wheat α-amylase I (-)	Weselake et al., 1983
Triticum aestivum (wheat-0.19 inhibitor)	Human saliva (+), Tenebrio molitor (+), rice weavel (+), yeallow mealworm (+), confused flour beetle (+), american coakroach (+), german coakroach (+), sawtoothed grain beetle (+), wheat (-), Aspergillus oryzae (-), Bacilus subtilis (-) guinea pig (-), turkey (-)	Silano et al., 1973, Silano et al., 1975; Granum and Whitaker, 1977
Triticum aestivum (wheat-0.28 inhibitor)	Rice weavil (+), yellow mealworm (+), confused flor beetle (+), german coakroach (+), american coakroach (+), sawtoothed grain beetle (+), human saliva (-), guinea pig (-), chicken pancreas (-)	Silano et al., 1975
Tricitum aestivum (wheat-0.55 inhibitor)	Human saliva (+), hog pancreas (-), Bacilus subtilis (-), Aspergillus oryzae (-)	Granum and Whitaker, 1977
Setaria italica	Human saliva (+), human pancreas (-), bovine pancreas (-)	Chandrasekher et al., 1981
Sorghum bicolor (sorghum)	Human saliva (+), human pancreas (+), porcine pancreas (+), bovine pancreas (-)	Chandresekher et al., 1981
Dioscorea alata	Human saliva (+), human pancreas (+), porcine pancreas (+), Bacilus subtilis (-), Aspergillus oryzae (-)	Sharma and Pattabiraman, 1982
Colocasia antiquorum	Human saliva (+), human pancreas (+), hog pancreas (+), Bacillus subtilis (-), Aspergillus oryzae (-)	Sharma and Pattabiraman, 1980
Eleusine coracana (ragi)	Human pancreas (+), human saliva (+), hog pancreas (+), Bacilus subtilis (-)	Shivaraj and Pattabiraman, 1980

(+)- indicates positive inhibition

(-)- Indicates negative inhibition

α-amylase inhibitor and the α-amylase (from 5 minutes to 2 hours) is a factor which, if overlooked, may produce misleading results (Marshall, 1975; Buonocore et al., 1977). Thus it may not be accidental that reports of the inhibition of plant α-amylases by endogenous inhibitors have only recently started to appear (Warchalewski, 1977; MacGregor et al., 1983; Mundy et al., 1983; Blanco-Labra and Iturbe-Chinas, 1981).

The reaction between the different α-amylases and their inhibitors involves the formation of an enzyme-inhibitor complex, where the catalytic activity of the enzyme towards substrates such as starch and amylose, is completely abolished. The formation of a stable enzyme-inhibitor complex has been demonstrated in several cases (Shivaraj and Pattabiraman, 1981; Powers and Whitaker, 1977; Marshall and Lauda, 1975; Sharma and Pattabiraman, 1982). However, detailed information about the mechanism of complex formation is scarce. The possibility exists that the formation of the complex may be triggered by the recognition by the α-amylase of a carbohydrate moiety in the inhibitor (Buonocore et al., 1981), but strong evidence for this has not yet been presented.

Unfortunately considerable uncertainty exists in the literature concerning the kinetics of inhibition and also over the assay conditions used to measure it. Furthermore, the determination of the value of kinetic parameters such as the inhibition constant (K_i), have been based on models constructed for the analysis of interactions between proteinases and their inhibitors (Green and Work, 1953; Bieth, 1974, 1980), which in all cases are of a competitive nature (Laskowski and Kato, 1980), while competitive inhibition as yet has not been demonstrated for any inhibitor/α-amylase system. On the contrary, most inhibitors of α-amylase seem to inhibit α-amylase in a non-competitive

way, although uncompetitive inhibition has been claimed for some inhibitors (Buonocore et al., 1977). However the fact that inhibitors from cereals (Mundy et al., 1983), tubers (Chandrasekher et al., 1981) and legumes (Marshall, 1975) do bind to amylase in the absence of substrate and furthermore the demonstration of an enzyme-inhibitor-substrate ternary complex in cereals (Mundy et al., 1983), and legumes (Whitaker, 1981), indicates that a non-competitive mode of inhibition may be prevalent among protein inhibitors of α-amylase. Also adding to this the fact that the rate of complex formation is very slow (Marshall, 1975; Buonocore et al., 1977; Warchalewski, 1983) one may speculate that the active site of the α-amylase is not affected by the inhibitor, but the binding of the inhibitor to the α-amylase triggers a slow conformational change in the α-amylase active site so that it does not form the enzyme-substrate complex at its normal rate and, once formed, the enzyme-substrate complex does not decompose at normal rates to yield products.

As some α-amylase inhibitors have different specificities towards human salivary and pancreatic α-amylase (Table V), they are being used for differential isoenzyme assays in human serum and urine (O'Donnel et al., 1977).

1.2.2. Structure

The well characterized α-amylase inhibitors from plants are proteins commonly found to contain carbohydrate moieties (Marshall, 1975; Buonocore et al., 1977; Richardson, 1980, 1981). Their molecular weights range from 12000 to 60000. The inhibitors with molecular weight above 12000 generally give rise to subunits of 12000 molecular

weight under dissociating conditions (Powers and Whitaker,1977; Marshall, 1975; Buonocore et al., 1977; Granum, 1978; Granum and Whitaker, 1977).

The carbohydrate content of protein inhibitors of α-amylase varies from 0.9% in the wheat inhibitors (Silano et al., 1975; Buoncore et al., 1981) to 64% in the yam α-amylase inhibitor (Sharma and Pattabiraman, 1982). The high carbohydrate content of some inhibitors is thought to be responsible for the difficulties in ascertaining the molecular weight by molecular exclusion chromatography as their elution from the column tends to be delayed by interactions between their carbohydrate moiety and the column gel (Shivaraj and Pattabiraman, 1981).

The endosperm of wheat seeds is known to contain a number of closely related protein inhibitors of α-amylase. Three of these inhibitors are well characterized: the 0.53, 0.28 and 0.19 inhibitors, named after their electrophorectic mobilities in polyacrylamide gels. The 0.53 and 0.19 inhibitors have molecular weights around 24000, but under dissociating conditions are separated into two subunits of aproximately 12000 molecular weight. The two subunits which form the 0.53 inhibitor are identical (Maeda et al., 1983) and the two subunits which form the 0.19 inhibitor also seem to be identical (Petrucci et al., 1978). The 0.28 inhibitor has a molecular weight of 13400. The sequences of the 24 N-terminal residues in the 0.19 (Redman, 1976) and 0.28 (Petrucci et al., 1978) inhibitors show that nine out of the 24 N-terminal residues occupy identical positions. Further evidence for the structural homology between the wheat α-amylase inhibitors was provided by the complete amino acid sequences of the 0.28 inhibitor (Richardson and Kashlan, 1981; Kashlan and Richardson, 1981) and the

0.53 inhibitor (Maeda et al., 1983) which showed that these two inhibitors are highly homologous (Figure 2). The sequence of the 0.53 inhibitor is notable in having an odd number (9) of cysteine residues and in having phenylalanine and histidine, residues which were thought to be absent in the wheat inhibitors (Richardson and Kashlan, 1981; Kashlan and Richardson, 1981; Buonocore et al., 1977). Recently Richardson (personnal communication), has shown that the homology of the 0.19 inhibitor and the 0.53 and 0.28 inhibitors is not restricted to the N-terminal residues, but extends throughout the whole sequence. Also recently, Aschauer et al. (1981) have reported the sequence of an α-amylase inhibitor from Streptomyces tendae (strain 4158). This inhibitor of molecular weight 7900 has no structural homology with the wheat 0.28 and 0.53 inhibitors.

The amino acid composition of the red kidney bean α-amylase inhibitor (Powers and Whitaker, 1977) is quite different from that of the α-amylase inhibitors from cereals, especially with respect to aspartic acid, threonine, serine, proline, methionine, isoleucine, tryptophan and arginine. Also notable is the very low content of cysteine residues and the absence of proline.

1.2.4. Biological Role

It was believed for some time that the α-amylase enzymes from higher plants were not inhibited by the α-amylase inhibitors found in the same tissues and therefore the inhibitors would not be involved in the control of starch metabolism in plants. However reports on the isolation of α-amylase inhibitors from wheat (Warchalewski, 1977), maize (Blanco-Labra and Iturbe-Chinas, 1981) and barley (Mundy et al.,

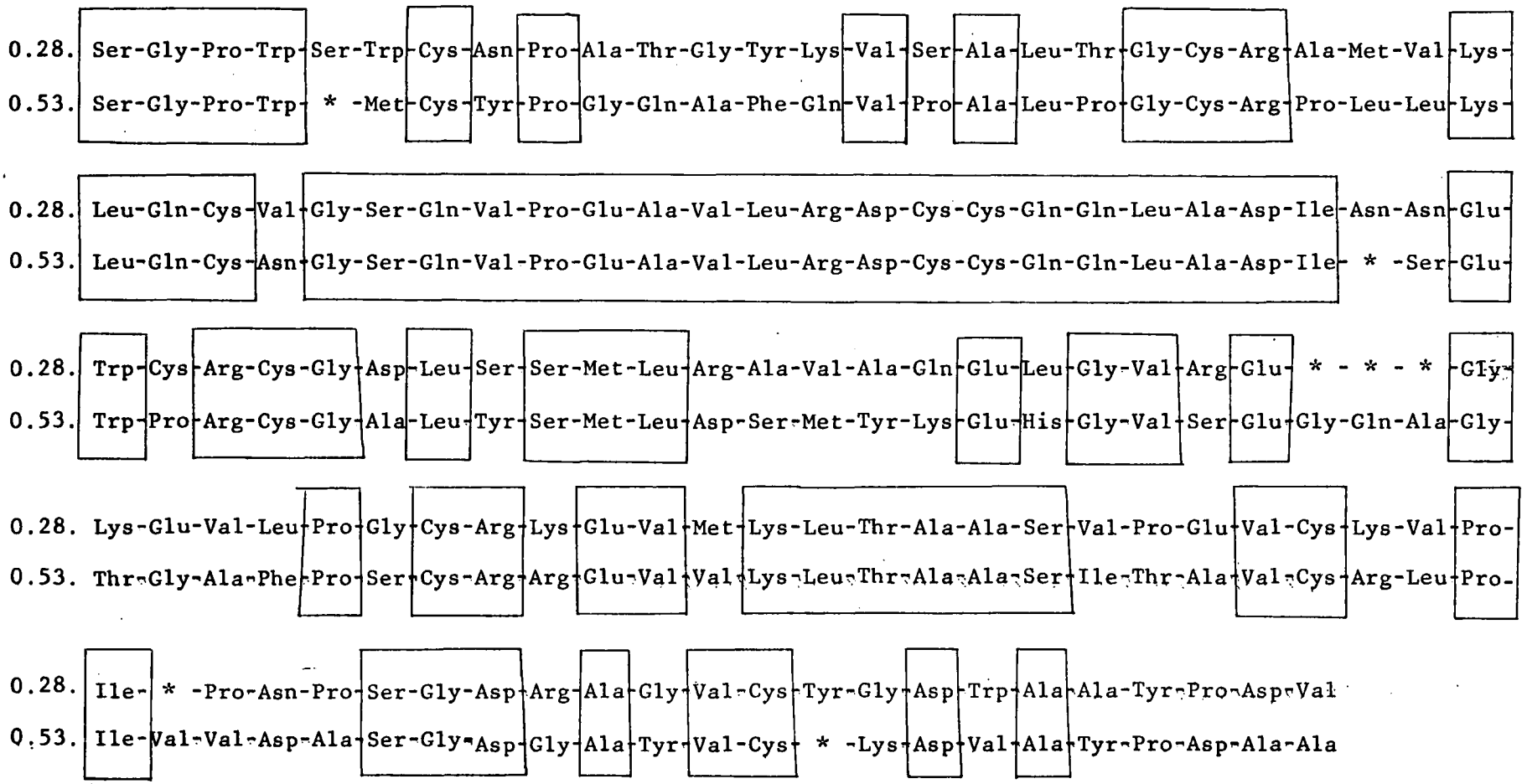


Figure 2 - Comparison of the amino acid sequences of 0.28 (Kashlan and Richardson, 1981) and 0.53 (Maeda et al 1983) α-amylase inhibitors from wheat. Homologous positions are enclosed in brackets. (*) indicates gap.

1983; Weselake et al., 1983) which inhibit α-amylases from these plants, show that an active in vivo role for these substances cannot be ruled out. The maize inhibitor is especially interesting as it seems to behighly specific for an α-amylase produced during germination and is also active against α-amylases from insects which are known to attack stored maize grain. The maize inhibitor is inactive towards α-amylases from wheat, barley, rye, Triticale, sorghum and also those from Bacillus subtilis and Aspergillus oryzae (Blanco-Labra and Iturbe-Chinas, 1981). These authors speculate that the inhibitor may help to prevent the biochemical events of starch breakdown which have been associated with the pre-harvest sprouting in cereals. The inhibitor isolated from barley (Mundy et al., 1983) apparently has a passive function during germination, but these authors speculate that this inhibitor plays a regulatory role during development of the seed by inhibiting endogenous α-amylase activity during starch granule synthesis and also may help to delay degradation caused by premature sprouting. Although the α-amylase inhibitor isolated from barley by Weselake et al. (1983) has not yet been characterized sufficiently to allow a comparison with the α-amylase/subtilisin inhibitor isolated from barley by Mundy et al. (1983) and other α-amylase inhibitors, it seems to inhibit specifically malted barley α-amylase II, and germinated wheat α-amylase II, having no inhibitory activity towards malted barley α-amylase I and germinated wheat α-amylase I. The effect of this inhibitor on animal α-amylase is not known. As the barley (Mundy et al., 1983) and maize (Balnco-Labra and Iturbe-Chinas, 1981) α-amylase inhibitors are also active against insect α-amylases, they may double as regulatory proteins and protective agents against seed predation. Although interesting, these results have to be viewed with caution as the possibility exists that

there are factors in plants which prevent the *in vivo* inhibitor/α-amylase association.

While in general inhibitors of α-amylase constitute a significant portion of the protein content of many cereals, legumes and tubers (Marshall, 1975; Buonocore *et al.*, 1977; Whitaker, 1981), their mobilisation during plant growth has not been the subject of much study. Preliminary studies on the inhibitors from *Colocasia antiquorum* tubers by Sharma and Pattabiraman (1980) have shown that the total α-amylase inhibitor activity decreases during plant growth. In wheat the inhibitor content also decreases after germination begins (Sandstedt and Beckford, 1946; Pace *et al.*, 1978).

With few exceptions the protein inhibitors of α-amylase are active against α-amylases from insects and microorganisms (Table V), and thus to assign to them a role of protecting seeds and tubers against the attack of insects and microorganisms is tempting. If they can be proved to be effective in this way, they would constitute an excellent index for the selection of varieties of plants resistant to pests. Unfortunately most of the studies on the inhibition of insect and microorganisms α-amylases by protein inhibitors of α-amylase from plants are restricted to the *in vitro* action of these inhibitors (Buonocore *et al.*, 1980; Powers and Culberston, 1982; Orland *et al.*, 1983) so that accurate assessement of the *in vivo* effects of the inhibitors cannot be made. However, Applebaum (1964) has shown that if wheat α-amylase inhibitors are incorporated into the semi-synthetic diet of *Tenebrio molitor* larvae, the larval growth is impaired. These results are supported by Yetter *et al.* (1979) who showed a correlation between the *in vivo* resistance of the seeds to the insect and the extent of *in vitro* inhibition of the insect larval α-amylase by the

inhibitor extracted from the same seeds.

The uncertainty in relation to the in vivo effect of α-amylase inhibitors on insects and microorganisms, is paralleled by the uncertainty concerning their role in human and animal nutrition. As many inhibitors are inactivated in vitro by pepsin (Saunders and Lang, 1973; Petrucci et al., 1978; Chandrasekher et al., 1981), it is thought that they would affect salivary α-amylase in vivo but not pancreatic α-amylase. However, the intake of purified α-amylase inhibitors from wheat by rats, humans, dogs and cats, results in a reduction of the rate of appearence of glucose in the blood (Marshall, 1975). Also, Marshall (1975) reported that rats which were fed with a diet containing wheat inhibitors which had been inactivated by autoclaving grew at a rate 15-20% faster than did rats fed on an untreated diet. Lang et al. (1974) showed that when rats were fed on a casein/starch diet in the presence and absence of purified α-amylase inhibitor reduction of growth rate and increased faecal starch levels were caused by the presence of inhibitors, and the magnitude of these changes was dependent on the inhibitor dosage. On the other hand, Shainkin and Birk (1970) and Savaiano et al. (1977) showed that rats have a normal growth after consuming massive quantities of red kidney bean α-amylase inhibitor which is without any detectable trypsin inhibitor or hemaglutinin activity.

The pharmacological and clinical significance of the α-amylase inhibitors have been under study for some time (Puls et al., 1980; Berchtold and Kieselbach, 1980; Frerichs et al., 1973; Silano, 1978; Puls and Keup, 1973). However the effectiveness of α-amylase inhibitors in the control of clinical disorders associated with carbohydrate metabolism is still controversial. This controversy is highlighted by

a recent report by Carlson et al. (1983) who claim that an α-amylase inhibitor preparation from kidney bean with potent inhibitory activity in vitro, which was being sold in the United States as an aid in controlling obesity, has no effect on starch degradation in humans.

The main purpose of the present study was to determine the complete primary structures of the bifunctional trypsin/α-amylase and the α-amylase inhibitors present in the seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) (Shivaraj and Pattabiraman, 1980, 1981). The trypsin/α-amylase inhibitor, at the time, was the first protein to be shown to have inhibitory activity against two completely unrelated classes of hydrolytic enzymes. It was thought that the determination of its primary structure would give important clues to unravelling the possible evolutionary relationship between inhibitors of α-amylase and inhibitors of proteolytic enzymes. Furthermore it was hoped that a comparison of the primary structure of the trypsin/α-amylase inhibitor with the primary structure of the α-amylase inhibitor would give some indications of the structural features of the α-amylase inhibitor which are responsible for the interaction with the α-amylases. Finally, a minor part of this work is dedicated to the sequencing of the protease inhibitor P-I from the Tracy cultivar of soybean (Glycine max).

2. MATERIALS AND METHODS

2.1. Materials

2.1.1. Biological Materials

Seeds of ragi (Indian finger millet, Eleusine coracana, Gaertn.) were obtained as a generous gift from the Tropical Products Institute, Culham, Oxford.

Ragi is a cereal crop of substantial economical importance (Hill, 1937; Johnston, 1958). Both upland and irrigated forms are grown mainly in Asia and Africa, where they sometimes constitute the staple diet of many people. In India it is one of the major crops, particularly during the rainy season. Ragi flour is used for soup, bread ,puddings and cakes, and also a fermented beverage is made from the grain (Hill, 1937; Johnston, 1958; Shivaraj and Pattabiraman, 1976).

2.1.2. Chemicals and Reagents

The following products were obtained from Sigma Chemical Company Ltd., Poole, Dorset: Trypsin (E.C. 3.4.21.4) Type XI, from porcine pancreas; α-amylase (E.C. 3.2.1.1.) Type I-A from hog pancreas; Cytochrome C, Type III, from horse heart; 3,5,-dinitrosalicylic acid; p-nitrophenol-p'-guanidino-benzoate; soluble starch; α-N-benzoyl-DL-arginine-p-nitroanilide-HCl (BAPNA).

The following products were obtained from British Drug House (BDH) Ltd., Poole, Dorset: Acrylamide; NN'-methylenebisacrylamide; coomassie brilliant blue and iodoacetic acid.

The following products were obtained from Worthington Biochemical Corporation, New Jersey, U.S.A.: Trypsin-TPCK treated (E.C. 3.4.21.4); α-chymotrypsin (E.C. 3.4.21.1) and Staphylococcus aureus (Strain V-8) protease (E.C. 3.4.21.19).

Thermolysin (E.C. 3.4.24.4) was obtained from Daiwa Kasei K.K., Osaka, Japan.

Proline-Specific endoproteinase (E.C. 3.4.21.26) was obtained from Seikagaku Kogyo Company Ltd., Tokyo, Japan.

Pyridine (Sequanal grade) and trifluoracetic acid (Sequanal grade) were obtained from Rathburn Chemicals Ltd., Walkerburn, Peebleshire.

Guanidine hydrochloride (Sequanal grade) was obtained from Pierce, Rockford, Illinois, U.S.A.

Ninhydrin (indantrione hydrate) and cyanogen bromide (CNBr) were obtained from Koch-Light Laboratories Ltd., Colnbrook, Bucks.

Arginyl-arginine was obtained from Cyclochemical Corporation, Los Angeles, U.S.A.

Phenylisothiocyanate (PITC) was obtained from Rathburn Chemicals Ltd., Walkerburn, Peebleshire.

4-NN-Dimethylaminobenzene-4-isothiocyanate (DABITC) was kindly supplied by Mr. John Gilroy and prepared as described by Chang et al. (1976) and Chang (1977), and also obtained from Fluka AG Chemische Fabrik, Buchs, Switzerland.

Ampholine carrier ampholytes (pH 2-11) were obtained from Serva, Heidelberg, West Germany.

Polyamide sheets were obtained from Cheng Chin Company Ltd.,

Taipei, Taiwan.

Sephadex G-10 and Sephadex G-50 were obtained from Pharmacia Fine Chemicals AB, Upsala, Sweden.

Biogel P-2, Biogel P-4 and Biogel P-6 were obtained from Bio-Rad Laboratories Ltd., London.

CM-52 cellulose was obtained from Whatman Biochemicals Ltd., Maidstone, Kent.

Buffer components and reagents, except as noted, were obtained from British Drug Houses Ltd. and were of analytical grade when necessary. Tris "Trizma base" was obtained from Sigma Chemical Company Ltd., Poole, Dorset.

2.2. Methods

2.2.1. Purification of the Trypsin/ α-Amylase and α-Amylase Inhibitors from Ragi (Indian Finger Millet, Eleusine coracana Gartn.)

The purification of both inhibitors was carried out essentially as described by Shivaraj and Pattabiraman (1980, 1981). One Kg of ragi seeds were ground in a mill using a 0.1mm sieve. The flour was homogeneized with 2 L of 0.15M NaCl and after stirring for 30 minutes, the slurry was centrifuged at 2000 x g for 30 min in a MSE 4L centrifuge. The residue was reextracted with 1 L of 0.15M NaCl, stirred for 30 minutes and the slurry centrifuged. Solid ammonium sulphate was slowly added with stirring to the combined supernatants to bring the saturation to 55% (w/v). After standing overnight at 4° C, the solution was centrifuged at 5000g for 30 minutes. The precipitate obtained was

dissolved in 250 ml of 2mM acetate buffer pH 5.0 containing 75mM NaCl and the solution dialysed against 20 volumes of the same buffer. The dialysed solution was centrifuged at 5000g for 30 minutes. The clear supernatant was designated as the ammonium sulphate fraction (ASF).

The ASF was subjected to chromatography on a column (1.4 x 20 cm) of CM-cellulose equilibrated with 2mM acetate buffer pH 5.0 containing 75mM NaCl. The column was eluted successsively with the equilibrating buffer, 20 mM acetate buffer 0.15M NaCl, and 20mM acetate buffer, 0.30M NaCl pH 5.0. Fractions of 10 ml were collected at a flow rate of 30 ml/h. The elution profiles were followed by measurement of the absorbance at 280nm and by assaying for the trypsin and α-amylase inhibitory activity of each fraction. The fractions eluted with the second elution buffer (20mM acetate buffer, 0.15M NaCl, pH 5.0) contained the trypsin/α-amylase inhibitory activity (Fraction 1) and the fractions eluted with the third buffer (20mM acetate buffer, 0.30M NaCl, pH 5.0) contained the α-amylase inhibitory activity (Fraction 2).

After dialysis overnight against fifty volumes of 2mM acetate buffer, 75mM NaCl, pH 5.0, Fraction 1 was concentrated by chromatography on a column of CM-cellulose (1.0 x 8.0 cm) equilibrated with the dialysis buffer. The initial column effluent was discarded and then the column was washed with 150 ml of the equilibrating buffer. The column bound proteins were eluted with 0.1M phosphate buffer, 0.15M NaCl, pH 6.9 and 5 ml fractions were collected at a flow rate of 30 ml/hour. The concentrated inhibitor solution obtained was subjected to chromatography on a column (2.6 x 90cm) of Sephadex G-50. Elution was performed using the same buffer at a flow rate of 20 ml/h and 5 ml fractions were collected. The fractions with trypsin/α-amylase inhibitory activity

were pooled and designated as Sephadex G-50 Fraction 1.

Similarly after dialysis against fifty volumes of 0.1M phosphate buffer, 0.15M NaCl, pH 6.9, Fraction 2 was subjected to chromatography on a Sephadex G-50 column (2.6 x 90cm), using the dialysis buffer as eluent. The fractions with α-amylase inhibitory activity were pooled and designated as Sephadex G-50 Fraction 2.

Both fractions were recromatographed on a column (1.0 x 6.0 cm) of CM-cellulose. The fractions were dialysed against 2mM acetate buffer, 75mM NaCl, pH 5.0 and applied to the column previously equilibrated with the dialysis buffer. The effluents were discarded and the columns eluted with the dialysis buffer. The effluents were discarded and the columns eluted with a linear gradient system formed between 100 ml of the equilibrium buffer and 100 ml of 2mM acetate buffer, 0.3M NaCl, pH 5.0. Fractions of 4 ml were collected at a flow rate of 20ml/h. When the Sephadex G-50 Fraction 1 was subjected to this process the fractions shown to contain trypsin and α-amylase inhibitory activity were pooled and designated as the trypsin/α-amylase inhibitor. Similarly the fractions with α-amylase inhibitory activity obtained from processing the Sephadex G-50 Fraction 2 were pooled and designated as the α-amylase inhibitor.

The homogeneity of both inhibitors was checked by subjecting them to SDS-polyacrylamide gel electrophoresis, isoelectric focusing and N-Terminal analysis by the dansyl-chloride method and by the DABITC/PITC double coupling method.

2.2.2. Purification of the Protease Inhibitor PI-I from the Tracy Cultivar of Soybean (Glycine max L. Merr)

The protease inhibitor PI-I from the Tracy cultivar of soybean, purified as described by David et al. (1977) was obtained as a gift from Dr. D. Foard (Department of Botany and Plant Pathology, Purdue University, West Lafayette, Indiana 47907, U.S.A.)

2.2.3. Enzyme and Inhibitor Assays

2.2.3.1. α-Amylase Inhibitor Assay

α-Amylase inhibitor activity was determined by the method of Bernfeld (1955).

assay mixture - 0.5 ml α-amylase (1.25 μg/ml in 0.05M acetate buffer pH 7.0)
0.5 ml starch solution (1% w/v in H_2O)
1.0 ml inhibitor solution (in 0.05M acetate buffer, pH 7.0)

2.0 ml

After incubation of the enzyme and inhibitor at 30°C for 25 minutes, the assay was started by the addition of the substrate. After 10 minutes the assay was terminated by the addition of 2.0 ml of dinitrosalicylate solution (prepared by dissolving 1g of dinitrosalicylic acid in 20 ml NaOH and 50 ml H_2O; 30 g of potassium tartrate (Rochelle Salt) was added and the solution volume made up to 100 ml). The mixture was then heated for 10 minutes in a boiling water bath and 10 ml of water was added to

the solution and after cooling, the absorption at 530 nm was measured. In every assay controls were established in order to determine the maximun and minimum digestion of starch under the experimental conditions, by incubating the enzyme solution only with starch and by incubating the inhibitor solution with the starch solution.

2.2.3.2. Trypsin Inhibitor Assay

Trypsin inhibitory activity was determined by measuring the inhibition of the hydrolysis of BAPNA by trypsin as proposed by Erlanger et al. (1961).

assay mixture - 0.5 ml trypsin (0.3 mg/ml in 2.5×10^{-3}M HCl)
0.5 ml inhibitor solution (in 0.05M Tris-HCl pH 8.2)
2.0 ml BAPNA (1.25×10^{-3}M in 0.05M Tris -HCl pH 8.2)
3.0 ml

After incubating the enzyme and inhibitor for 5 minutes at 37° C, the substrate was added. After 10 minutes the reaction was stopped by the addition of 0.5ml 30% acetic acid. The absorption of the solution was measured at 410nm.

2.2.3.3. Trypsin Active Site Titration

The concentration of active trypsin in the different solutions used for assay of trypsin activity was determined by active site titration of trypsin by p-nitrophenol-p'-guanidino-benzoate (NPGB), by the method of Chase and Shaw (1970). In these experiments a double beam recording spectrophotometer was used in order to allow the nonenzymatic breakdown of NPGB to be automatically subtracted by adding an equal

amount of the compound to the reference cuvette.

Titration procedure - Sodium veronal buffer 0.1M, pH 8.3 was pipetted into the reference cuvette and the enzyme sample and buffer were pipetted into the sample cuvette to a volume of 1.0ml. The cuvettes were placed in the spectrophotometer and the instrument (set at 410nm) zeroed. The cuvettes were then removed from the instrument and 5 µl of NPGB solution was quickly added to the sample cuvette, the solutions mixed and the cuvette replaced, and the instrument turned on. The concentration of active enzyme was then calculated from the increase in optical density thus determined and the molar coefficient of p-nitrophenol at this pH and wavelength.

2.2.4. Isoelectric Focusing

Isoelectric focusing experiments were carried out as described by Wrigley (1976). Samples for isoelectric focusing were prepared by dissolving the lyophylized material in distilled water at a concentration of 2mg/ml. Cytochrome C (pI 10.65) was used as a marker. Prior to loading, the gels were pre-run at aproximately 260V for 10 minutes to stabilise the pH gradient. The inhibitor samples were loaded centrally and the cytochrome C marker was loaded at both poles. The gel was run at aproximately 15W until the two marker samples had merged together and become focused; this took approximately 2 hours. Phosphoric acid 0.1M pH 1.5 formed the anode, whilst ethylenediamine formed the cathode. All attempts to stain the inhibitors after isoelectric focusing by using coomassie brilliant blue failed. Nevertheless prolonged immersion of the gels in 17.5% trichloroacetic acid, made the precipitation bands in the gels visible.

2.2.5. SDS-Polyacrylamide Gel Electrophoresis

The SDS-polyacrylamide gel electrophoresis experiments were carried out as described by Laemli (1970), and were used to assess the homogeneity of inhibitor preparations and also to determine the apparent molecular weight of proteins. 17% acrylamide gels were used.

Separating gel formulation:

22.5 ml Tris-HCl pH 8.8

34.5 ml acrylamide stock (30% acrylamide - o.135% bis-acrylamide

1.5 ml ammonium sulphate (15mg/ml)

20 µl TEMED

Stacking gel formulation:

2.5 ml Tris-HCl pH 6.8

2.0 ml acrylamide stock (30% acrylamide - 0.344% bis-acrylamide)

14.8 ml distiled water

0.5 ml ammonium sulphate (15 mg/ml)

20 µl TEMED

Lyophylized samples for electrophoresis were dissolved in 0.2M Tris-HCl containing 2% SDS and 10 % sucrose. Prior to electrophoresis bromophenol blue was added as a marker. Electrophoresis was initially carried out at 15 mA/gel and subsequently as 25mA/gel for 6-7 hours. After electrophoresis the total length of the gel and the distance of migration of the bromophenol blue marker were measured. The gel was then stained overnight with 0.25% coomassie brilliant blue in methanol/acetic acid and destained using a 5% methanol, 7% acetic acid mixture in water.

2.2.6. Protein Sequence Determination

2.2.6.1. Reduction and S-Carboxymethylation

The reduction and S-carboxymethylation of the inhibitors was carried out using a modification of the method proposed by Crestfield et al. (1963). The protein (10-30mg) was dissolved in 2 ml of 6M guanidine hydrochloride, 1M Tris-HCl, pH 8.5. 10-30 µl of 2-mercaptoethanol was added and the mixture left standing, under a constant stream of nitrogen for 3 hours at room temperature (18-20° C). Then 100-300 ul of iodoacetic acid (0.268g/ml in 1M NaOH) was added and the solution kept in the dark for 15 minutes under a nitrogen stream. The reduced and S-carboxymethylated protein was recovered by gel filtration (in the dark) of the solution on a column (30 x 1.5cm) of Sephadex G-10, equilibrated in 50mM pyridine acetate buffer, pH 5.4. The eluted protein was subsequently freeze-dried.

2.2.6.2. Chemical Cleavage with Cyanogen Bromide (CNBr)

CNBr cleavage of the inhibitors was carried out by the method of Steers et al. (1965). The reduced and carboxymethylated protein was dissolved in 1.0 ml of 70% (v/v) formic acid and an 20 fold molar excess of CNBr was added and the solution kept in the dark for 24 hours at room temperature and under a nitrogen stream. After this period the solution was diluted 10-fold with distilled water and the reaction terminated by lyophilisation.

2.2.6.3. Enzymatic Hydrolysis

2.2.6.3.1. Trypsin (E.C. 3.4.21.4)

The protein or peptide was dissolved in a minimal volume of 0.2M N-ethylmorpholine buffer, pH 8.5 and trypsin (TPCK- treated) was added to give a 2% (w/w) enzyme/substrate ratio. After incubation for 2 hours at 37° C, the reaction was stopped by lyophilisation.

Specificity - Trypsin catalyses the hydrolysis of peptide bonds involving the carboxyl group of lysine and arginine, except where the following residue is proline. Trypsin has low exopeptidase activity, especially aminopeptidase, and adjacent lysine or arginine residues usually provide alternative points of cleavage, with low yields of free lysine or arginine (Allen, 1981). Even when pure trypsin is used some chymotryptic activity can usually be detected (Inagami and Sturtvant, 1960).

2.2.6.3.2. Chymotrypsin (E.C. 3.4.21.1)

The conditions of digestion were the same as those used with trypsin.

Specificity - Chymotrypsin cleaves peptides and proteins at the carboxyl side of tryptophan, phenylalanine, tyrosine, leucine and methionine, with occasional cleavage at other sites (e.g. histidine, asparagine). If the following residue is proline cleavage does not occur (Allen, 1981).

2.2.6.3.3. Thermolysin (E.C. 3.4.24.4)

The conditions of digestion were the same as those used with trypsin, except that the buffer contained 5mM $CaCl_2$.

Specificity - Thermolysin catalyses the hydrolysis of peptide bonds involving the amino group of the hydrophobic amino acids leucine, isoleucine, methionine, phenylalanine, tryptophan, and valine. Other sites, at the amino side of alanine, tyrosine and threonine may also be cleaved (Matsubara, 1965).

2.2.6.3.4. Staphylococcus aureus (Strain V-8) Protease

The protein or peptide was dissolved in a minimal volume of 0.1M Na-phosphate buffer pH 7.8 and incubated for 16 hours at 37° C. The reaction was stopped by lyophilisation (Drapeau, 1977).

Specificity - If the digestion is performed in 50-100mM Na-Phosphate buffer pH 7.8 the V-8 protease specifically cleaves on the carboxyl side of glutamyl and aspartyl peptide bonds. The rate of digestion is low if glutamic acid or aspartic acid are within three residues from either the N- or C-terminus of a peptide containing lysine or arginine and Glu-Pro and Asp-Pro peptide bonds are not cleaved (Austen and Smith, 1976; Croft, 1980). If the buffer is changed to either NH_4HCO_3 (pH 7.8) or ammonium acetate (pH 4.0), cleavage only occurs at glutamyl residues.

2.2.6.3.5. Proline-Specific Endoproteinase (E.C. 3.4.21.26)

The protein or peptide was dissolved in a minimal volume of 0.1M Na-phosphate buffer, pH 7.0 and the enzyme was added to give a 2%

(w/w) enzyme/substrate ratio. After incubation for 1-2 hours at 37° C thereaction was stopped by lyophilisation (Yoshimoto et al., 1980).

Specificity - The proline-specific endoproteinase specifically cleaves Pro-X and occasionally Ala-X peptide bonds (Yoshimoto et al., 1980).

2.2.6.4. Purification of Protein and Peptide Mixtures

The purification of protein and peptide mixtures produced by enzymatic or chemical digestion was performed either by molecular exclusion chromatography, or by high voltage paper electrophoresis or by reverse-phase high performance liquid chromatography (RP-HPLC) or by a combination of these methods, as indicated.

2.6.4.1. Molecular Exclusion Chromatography

The hydrolysates were dissolved in the column eluent, as indicated, and applied to columns (0.8 x 190 cm) of Sephadex G-50, Biogel P-4 or Biogel P-6. The elution was performed at the flow rates indicated and 1.4ml fractions were collected. The detection of peptides in the fractions was done either by measuring the optical density of the fractions at 280 and/or 230 nm, or by removing small samples (10-30 µl) for N-terminal and semi-quantitative amino acid analysis using the dansyl-chloride method (see below).

2.2.6.4.2. High Voltage Paper Electrophoresis

The separation of peptide mixtures by high voltage paper electrophoresis was done as described by Thompson et al. (1970), at pH

6.5 (pyridine-acetic acid-water, 25:10:225 by volume) or at pH 1.9 (acetic acid-formic acid-water, 4:1:45 by volume) on a water-cooled flat-plate apparatus (107 x 15cm; The Locarte Co., London). Separation was achieved on Whatman paper (107 x 15 cm) using a voltage of 9 KV to give a current of 50-50 mA at 7 p.s.i. pressure for 120-150 minutes for the electrophoresis at pH 6.9 and by using a voltage of 9 KV to give a current of 30-50 mA for 60-120 minutes at 7 p.s.i. pressure for the electrophoresis at pH 1.9 .

The peptides were located using the cadmium ninhydrin reagent (cadmium acetate 100 mg, 10ml water, 5 ml glacial acetic acid, 100 ml acetone). A 1% (w/v) solution of ninhydrin in this solution was prepared, just before it was required (Heilman et al., 1957). 10% strips of the electrophoresis paper (guide-strips) were dipped in the cadmium ninhydrin reagent and dried at 105° C for 2-5 minutes. Some electrophoresis guide strips were also passed through freshly prepared Ehrlich reagent (2% w/v) 4-dimethylaminobenzaldehyde in 20% HCl in acetone (v/v). The pink ninhydrin spots became colorless and a purple colour appeared indicating a positive reaction with tryptophan.

At pH 1.9 peptide mobilities were measured from 1-dimethylamino-naphtalene-5-sulfonic acid and expressed relative to the distance of the dansyl-arginine standard. At pH 6.5 they were measured from a true neutral point determined as 4/11 of the distance from the standard dansyl-arginylarginine to the 1-dimethylaminonaphtalene-5-sulfonic acid.

Peptides were eluted with 20% (v/v) aqueous pyridine into screw-cap tubes (1 x 5cm). The samples were freeze-dried in vacuo over solid NaOH.

2.2.6.4.3. Reverse-Phase High Performance Liquid Chromatography (RP-HPLC)

RP-HPLC was performed in a Varian model 5000 HPLC fitted with a C-18 μ-bondapak column (0.5 x 25cm) (from HPLC Technology Ltd., Macclesfield, Cheshire) and using a linear gradient of 0 to 70% acetonitrile in 0.1% (v/v) trifluoracetic acid (Mahoney and Richardson, 1980). Mixtures of peptides were injected in 100 μl of 0.1% trifluoracetic acid (TFA) or in the case of mixtures which were insoluble in this solvent in 100 μl of 6M guanidine hydrochloride in 0.1% TFA. The presence of peptides in the eluting solvent was detected by measuring the absorbance at 214 nm. The peaks of absorbance at 214 nm were collected manually and lyophilised.

2.2.6.5. Amino Acid Analysis

2.2.6.5.1. Semi-Quantitative Amino Acid Composition of Peptides

Purified samples (30 - 50 nM) of peptides were dried in vacuo over solid NaOH in Durham tubes, 50 ul of 6M HCl was added and the tubes sealed with an oxygen flame, and hydrolysed for 18 hours at 105° C. The acid was removed by drying in vacuo over solid NaOH and the free amino acids were labelled by the dansyl-chloride method of Gray and Hartley (1963) and the dansyl-amino acids identified as described in Section 2.2.6.6.

2.2.6.5.2. Quantitative Amino Acid Composition of Proteins and Peptides

Samples of protein (0.2 - 0.5mg) or peptides (0.05-0.25

μmole) were hydrolysed with 5.7 M HCl at 105° C in evacuated pyrex tubes for 20-72 h. After hydrolysis the samples were dried in vacuo over solid NaOH and stored at -20° C until analysis.

The amino acid composition of the samples was determined using either a Locarte amino acid analyser, or, alternatively, in a Varian 5000 High Performance Liquid Chromatograph fitted with a Micropak Hydrolysate Amino Acid (AA) column (4mm x 15cm) in the sodium form, a post-column orthophthalaldehyde reactor system, and a Fluorichrom detector. The buffer system used was the Phix PicoBuffer System II (Pierce Chemical Co., Rockford, Saint Louis, U.S.A.), and the post-column reaction system was as described in the operators manual (Varian Associates, 1981).

2.2.6.6. N-Terminal Analysis of Protein and Peptides

The N-terminal analysis of peptides and proteins was performed using the dansyl-chloride method (Gray and Hartley, 1963; Gray, 1972). A sample of the peptide or protein (≃ 10nmol) was dissolved in 50% aqueous pyridine and transfered to a durham tube (30mm x 6mm, A. Gallenkamp Ltd., London) and dried in vacuo over solid NaOH. 10 μl of 0.2M $NaHCO_3$ was added and the solution dried again in vacuo over solid NaOH. 10 μl of a solution containing water and dansyl-chloride (5mg/ml acetone) in equal proportion was added, the tube sealed with parafilm and incubated for 50 minutes at 45° C. The dansylated protein or peptide was dried in vacuo over solid NaOH. 50 ul of 6M HCl was then added and the tube sealed using an oxygen flame and incubated for 18 hours at 105° C. After incubation the tube was broken open and the acid removed in vacuo over solid NaOH. The residue

was dissolved in 10 µl of 20% (v/v) aqueous pyridine and the dansyl amino acid was identified by thin layer chromatography (TLC), in the solvent systems described by Ramshaw et al (1970). The 10 µl solution was spotted on both sides of a polyamide sheet (15 x 15cm), 80% of the sample on the experimental side and 20% on the standard side, and dried under a hot-air draught. 1 µl of a chromatography marker solution containing 0.1mg/ml in 95% ethanol of the following amino acids: dansyl-proline, dansyl-isoleucine, dansyl-phenylalanine, dansyl-glutamic acid, dansyl-glycine, dansyl-serine and dansyl-arginine was applied to the standard side of the sheet. Frames of the type described by Smith (1958) were used to support up to five polyamide sheets for simultaneous separation. The solvents used to separate the amino acid derivatives were:

Solvent A - 1.5% (v/v) formic acid

Solvent B - Toluene-acetic acid, 9:1 (v/v)

Solvent C - Butylacetate-methanol-acetic acid, 30:20:1 (v/v/v)

The dansyl derivatives were identified by running the sheets in Solvent A for 45 minutes and after drying under a hot air draught, were run at right angles in relation to the first running with Solvent B for 45 minutes. After being dried the sheets were examined in a U.V. light (350nm) and the results recorded (Figure 3). The sheets were then run in Solvent C in the same direction as Solvent B, for 35 minutes, dried and examined under U.V. light (350 nm) and the results recorded (Figure 4). The co-chromatography of unknown samples with known standards was used to help in the identification of close pairs of derivatives.

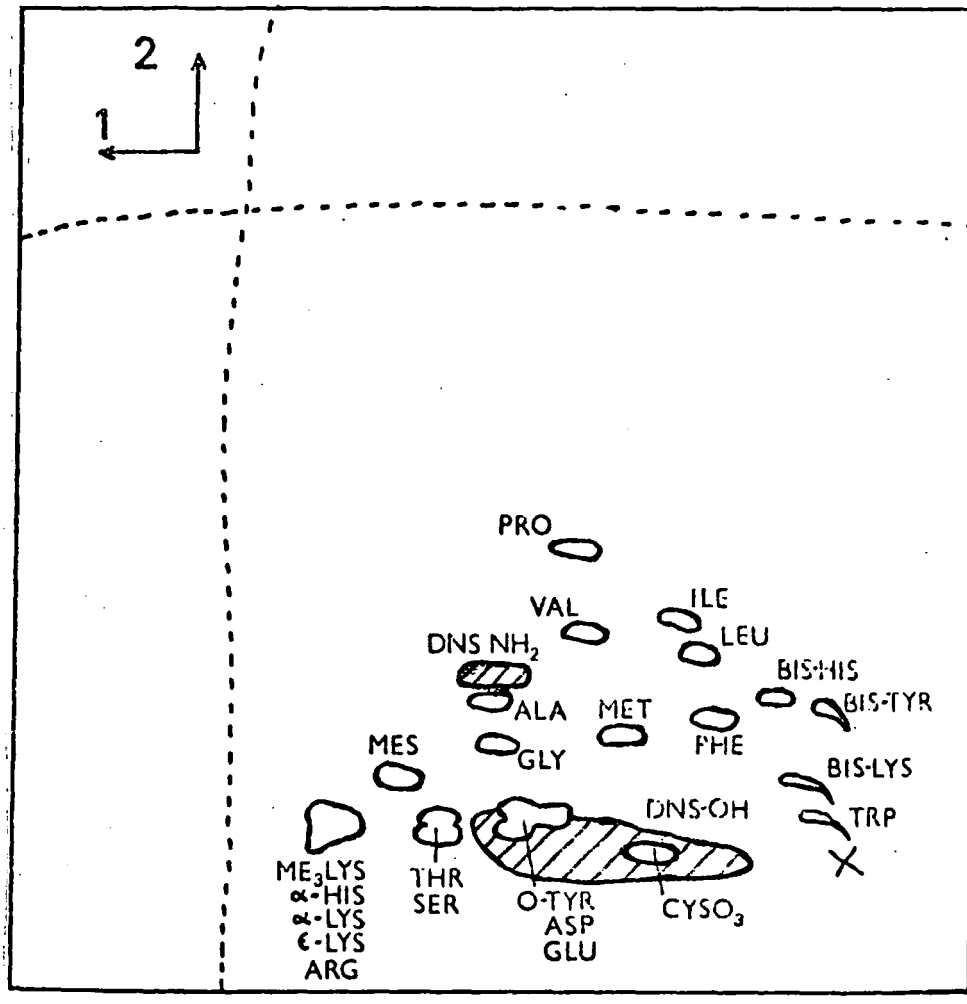


Figure 3 - Two dimensional separation of dansyl-amino acids on a polyamide sheet (Ramshaw et al., 1970). The solvents used for chromatographic separation are given in the text. Development was by solvent A in the first dimension and solvent B in the second dimension.

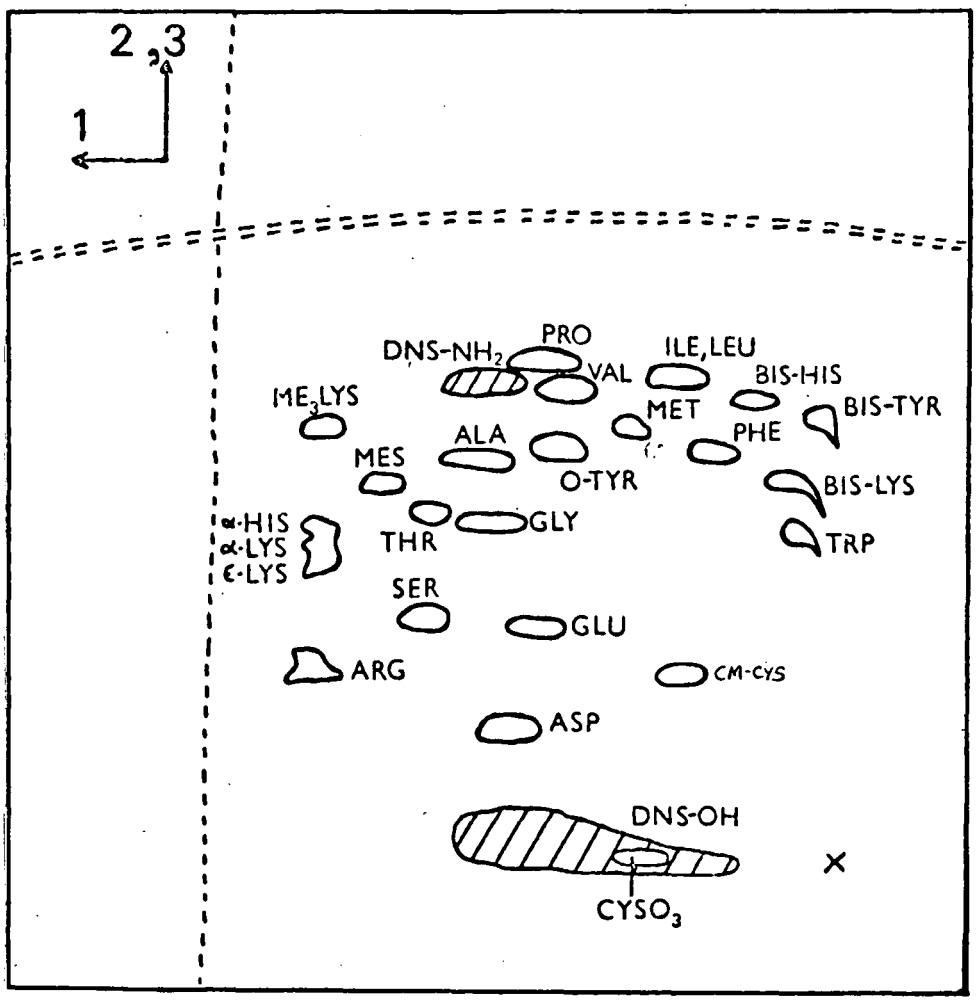


Figure 4 - Two dimensional separation of dansyl-amino acids on a polyamide sheet (Ramshaw et al., 1970). The solvent used for chromatographic separation are given in the text. Development was by solvent A in the first dimension and solvent B followed by solvent C in the second dimension.

2.2.6.7. C-Terminal Analysis Using Carboxypeptidase A

10 ul of carboxypeptidase A suspension (10mg/100 µl) diisopropyl-phosphofluoridate treated was washed twice with 2 ml of distilled water and 100 ul of 0.2M $NaHCO_3$ was added to the suspension which was then placed in an ice bath and 0.1M NaOH was added to just dissolve the enzyme. After the enzyme was dissolved, 0.1N HCl was added to neutralise the solution and the volume was made up to 1.5 ml with 0.2M N-ethylmorpholine-HCl buffer pH 8.5 .

Samples of the protein or peptide to be digested were dried in a durham tube in vacuo over solid NaOH. 20 µl of the carboxypeptidase A solution was added and the tube sealed with parafilm. The incubation was carried out for varying times (30 min. to 4 hours at 37º C) and the liberated amino acids were determined as their dansyl derivatives (see Section 2.2.6.6.).

2.2.6.8. Manual Sequencing Methods

2.2.6.8.1. "Dansyl-Edman" Method

This method is a variation of the phenylisothiocyanate (PITC) degradation procedure introduced and developed by Edman (1950, 1953, 1956; Ilse and Edman, 1963; Edman and Begg, 1967). In the "dansyl-Edman method the newly released N-Terminal residue at each cycle is identified as its dansyl derivative.

The peptide (0.1 - 0.5 nmol) was dissolved in 150 µl of 20% aqueous pyridine. Then 150 ul of a 5% solution of re-distilled phenylisothiocyanate (PITC) was added and the tube flushed with

oxygen-free nitrogen and capped. The samples were allowed to react at 45° C for 1 hour. Excess reagent and the volatile reaction by-products were then removed by drying *in vacuo* over solid NaOH. When the samples were completely dry, 200 µl of anhydrous TFA was added, the tube sealed with parafilm and incubated at 45° C for 30 min. The TFA was removed *in vacuo* over solid NaOH. The residue was dissolved in 200 ul of distilled water and extracted twice with 0.5ml of butyl acetate, and then dried *in vacuo* over solid NaOH. The residue was dissolved in 150 µl of 20% aqueous pyridine. 5-10% of this material was removed and a new degradation cycle was initiated with the remaining materi The material removed was utilised to identify the newly revealed N-terminal amino acid as described in Section 2.2.6.6.

2.2.6.8.2. *DABITC-PITC Double Coupling Method* (Chang *et al.*, 1978)

Samples of the peptide or protein (5 - 10 nmol) were placed in a glass tube (0.5 x 3.0cm) fitted with a ground glass stopper and dissolved in 80 µl of 50% (v/v) aqueous pyridine. 40 µl of the 4-NN-dimethylaminobenzene-4'-isothiocyanate (DABITC) solution (2.82 mg/ml pyridine) was added, the tube flushed with nitrogen, the contents mixed by agitation on a whirlmixer and incubated at 52° C for 50 minutes. 10 µl of PITC was then added, the mixture flushed again with nitrogen and incubated at 52° C for 30 minutes. The excess of reagents and by-products were removed by extracting the mixture three times with a 2:1 (v/v) solution of heptane/ethyl acetate. The organic phase was separated in each case by centrifugation, removed by pipette and discarded. The aqueous phase was dried down *in vacuo* over solid NaOH. The extraction of the 4-NN-dimethylaminobenzene-4'-thiazolinone-amino acid (DABTZ-amino acid) was done by adding to the

residue 50 µl of water and 200 µl of butyl acetate. After mixing on a whirlmixer for three seconds, the contents were centrifuged for 30 seconds at 1000g. The upper butyl acetate layer was removed with a pipette and the peptide in the aqueous phase was dried down in vacuo over solid NaOH and subjected to the next degradation cycle.

The butyl acetate extract was dried down in vacuo over solid NaOH and the residue was dissolved in 20 µl of water and 40 µl acetic acid saturated with HCl (or alternatively with 50 ul of 50% aqueous TFA). Conversion of the thiazolinones of the amino acids into the thiohydantoins was carried out in the 52°C heating block for 50 minutes, or by heating at 80° C for 10 minutes when the 50% aqueous TFA was used. The sample was dried in vacuo over solid NaOH and redissolved in a suitable volume (5 - 30µl) of 95% (v/v) ethanol.

The identification of the 4-NN-dimethylaminobenzene-4'-thiohydantoin-amino acid (DABTH-amino acid) was done by thin layer chromatography (TLC). The synthetic marker 4-NN-dimethylaminobenzene-4'-thiocarbamyl (DABTC)-diethylamine was used as a marker and spotted in the same position as the ethanol extract. The first dimensional separation was done using an acetic acid/water (2:1, v/v) solution and the second dimension separation with a toluene/n-hexane/acetic acid solution (2:1:1, v/v/v). After TLC the dried sheets were exposed to HCl vapours which led to the temporary formation of the characteristic colours of the DABTH-amino acids. The DABTH amino acids were identified by comparing the position of the coloured spots (blue, red or purple) with the blue marker (Figure 5).

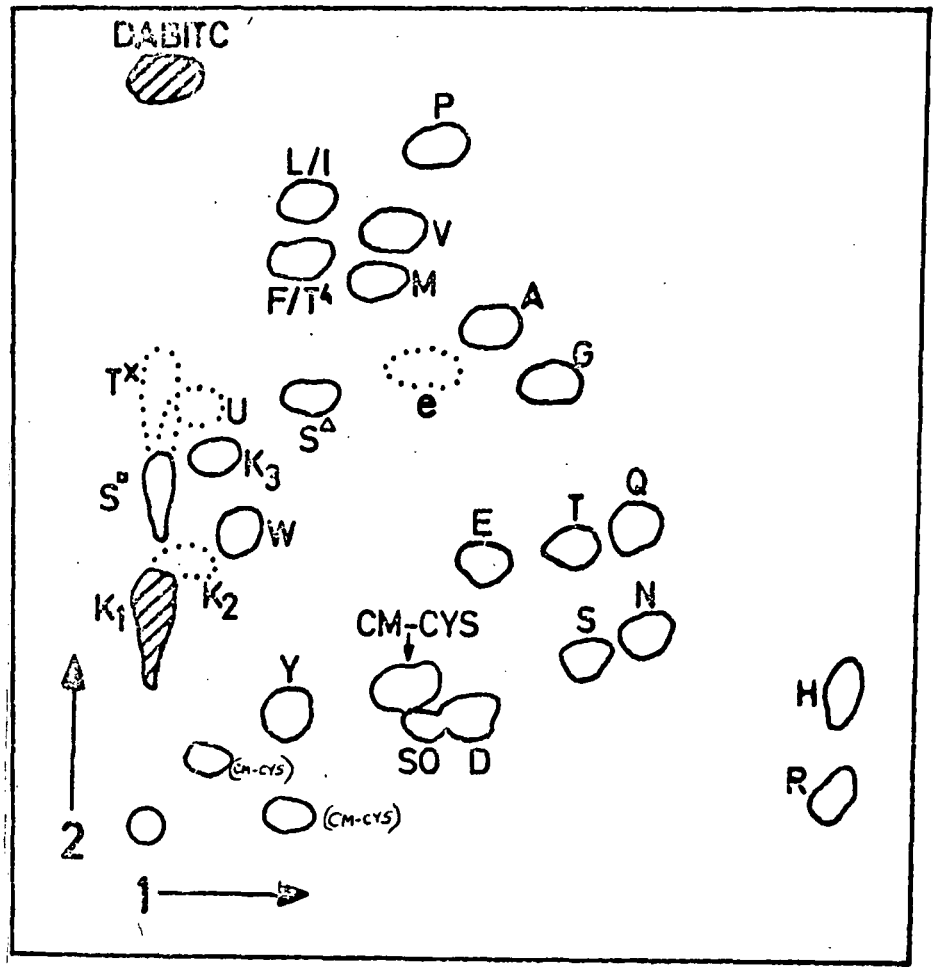


Figure 5 - Two-dimensional separation of DABTH-amino acids on a polyamide sheet (Chang et al., 1978). The solvents used for chromatographic separation are given in the text. The colours (after exposure to HCl vapours) are represented by solid areas (red), dotted areas (blue) and hatched areas (purple). (e) is the blue synthetic marker DABTC-diethylamine. (U) is a blue coloured thiourea formed by the coupling of PITC with hydrolyzed DABITC. The DABTH derivatives of the amino acids are denoted by the single letter code. T^{Δ}, DABTH-dehydrothreonine; S^{Δ}, S° and S^{σ}, products formed by -elimination of DABTH-serine; T^{X}, product formed after β-elimination of DABTH-threonine; K_1, α-DABTH-ϵ-DABTC-lysine; K_2- α-PTH-ϵ-DABTC-lysine; K_3, α-DABTH-ϵ-PTC-lysine. Other blue spots are sometimes visible, due to incomplete conversion to DABTH derivatives, particularly with the glycine derivative.

2.2.6.9. Nomenclature of Peptides

In the summarising data (Figures 35, 58, and 63) and in the text and tables, each peptide is characterized by its origin (i.e. from digestion with CNBr (CN), trypsin (T), chymotrypsin (Ch), Thermolysin (Th), Staphylococcus aureus (Strain V-8) protease (V), proline-specific endoproteinase (P), catalytic cleavage with trypsin (CC)) and by its elution order from gel filtration and/or reverse phase columns. For example, Th5-7 means a thermolysin peptide obtained by the separation of peak 5 (from gel filtration of the thermolysin digest of the relevant protein), by RP-HPLC; the number 7 indicates the order of elution of the peptide from the reverse phase column. Also, CN-Ch4 means a chymotryptic peptide obtained by the sub-digestion of a CNBr peptide with chymotrypsin; the number 4 indicates the order of elution of the peptide from the reverse-phase column.

2.2.6.10 - Identification of the Reactive (Trypsin Inhibitory) Site of the Trypsin/ α-Amylase Inhibitor (Ozawa and Laskowski, 1966; Richardson, 1974)

The native inhibitor was dissolved in 0.2M acetic acid containing 0.05M $CaCl_2$. Trypsin was added to give a 2% (w/w) solution. After adjusting to pH 2.5 with HCl the solution was incubated for 20 hours at 37° C. The reaction was stopped by lyophilisation. After reduction and S-carboxymethylation the fragments produced by catalytic cleavage were separated by RP-HPLC. The position of the reactive site in the primary structure of the inhibitor was deduced from the analysis of the N-terminal sequence of each fragment by the DABITC-PITC double coupling method (see Section 2.2.6.8.2.).

2.2.7. Sequence Comparisons

The alignment of amino acid sequences was performed using a computer program developed from Gotoh's algorithm (Gotoh, 1982) for aligning biological sequences (proteins and nucleic acids). Gotoh derived his algorithm from Waterman's algorithm (Waterman et al., 1976). In order to do so he made some simplifying assumptions, which were already used in practice by users applying Waterman's algorithm (Smith et al., 1981; Kanchisa, 1982). These simplifications allowed him to perform some algebraic manipulations of the equations used by Waterman et al. (1976) which eventually reduced the order of the algorithm's complexity.

The "best" alignment of two sequences is the one with the smallest distance measure (DM), according to a given set of parameters. The distance measure is defined in such a way that the greater its value the less the number of matching residues. The insertion of gaps and their size also increase the distance measure.

The functions are used in order to evaluate the distance measure:

a) - The function d(a,b) represents the contribution of two aligned residues to the distance measure. It can be interpreted as a penalty for a mismatch. So d(a,b) is equal to zero (no penalty) if a=b and is equal to $c > 0$ otherwise; and

b) - the function w(k) defines the penalty for a gap where k represents the size or length of the gap. The function is defined by Gotoh (1982) as $w(k) = u \times k + v$, where $u \geq 0$ and $v \geq 0$. This means that v is the initial cost or penalty for introducing a

gap and u is the increment to that penalty for each gap size unit.

The distance measure is now defined as the sum of the functions d applied to all aligned residues plus the sum of the function w applied to all gaps. Thus the value of the distance measure of a certain alignment depends on the values given to the three parameters c, u and v. The definitions of the functions d and w also imply that the greater the value of u in relation to the value of c the harder becomes the insertion of a gap. Additional difficulty might also be given by a greater value for v. Then defining these values the user must keep in mind the likelihood or not of the occurrence of gaps.

The alignment program operates in two major steps. During the first the "best" alignment is determined according to Gotoh's algorithm. At the second one the statistical significance of the alignment's similarity is determined. In order to determine the significance of the alignment, both biological sequences are shuffled 100 times (the residues of each sequence are randomly rearranged) and for each shuffle the smallest distance measure is determined for the same parameters c, u and v as used before. The mean (μ) and the standard deviation (σ) are derived from those 100 minimum distance measures of the randomly rearranged biological sequences can be described by a normal distribution with mean (μ) and standard deviation (σ). At a next step the distance measure (x) of the originally unshuffled sequences is derived in terms of standard deviations units ($z = (\mu - x)/\sigma$). An alignment is said to be significant if $z \leq 3.0$. This means that the probability of the alignment occurring by chance is less than 0.135×10^{-2} according to Table 35 in Dayhoff (1979).

2.2.8. Prediction of Secondary Structure

The prediction of the secondary structure of the inhibitors was performed by the computerized method of Robson (Garnier et al., 1978). This is the most widely utilised method for predictions of protein secondary structure from the amino acid sequence (Kabsch and Sander, 1983) and has the advantage over the also popular method of Chou and Fasman (1974, 1978, 1979), as it is presented as an algorithm and not , as in the Chou and Fasman predictive model, as a set of rules (some of them qualitative rather than quantitative). Thus, in the Robson's method unambiguous and objective predictions are guaranteed.

The choice of the directional method of Robson was also made on consideration of the fact that it also has a higher rate of predictive success than the method of Chou and Fasman. Kabsch and Sander (1983) have demonstrated that one may expect a success rate of about 50% with the Chou and Fasman method and 55 to 56% with the Robson's method. A similar conclusion was reached by Busetta and Hospital (1982).

The Robson method allows for a four state prediction, in which each residue is unambiguously assigned to one conformational state of α-helix (H), extended chain (E), reverse turn (T) or random coil (C). A complete account of the theory which justifies the formal correcteness of the method as well as an analysis of its accuracy, can be found in Robson (1974), Robson and Suzuki (1976) and Garnier et al. (1978). An outline of the method is given below (Garnier et al., 1978).

The method was built on the consideration that the prediction

of S_j (the conformational state of the j^{th} residue in the sequence), depends not only on the type of one or several residues, but on the whole sequence. Thus the most general statement concerning the information for the conformation of the j^{th} residue is thus

$$I\ (S_j; R_1, R_2, \ldots\ldots\ldots R_{last}),$$

which reads as the "information which the residue at the first, second and so on up to the last residue carry about the conformation of the j^{th} residue". As the model allows for the prediction of four conformational states H (α-helix), E (extended chain), T (reverse turn) and C (random coil), the predictive procedure leads to four values for information associated with each residue (R):

$$I\ (S_j : H; R_1, R_2, \ldots R_{last})$$

$$I\ (S_j : E; R_1, R_2, \ldots R_{last})$$

$$I\ (S_j : T; R_1, R_2, \ldots R_{last})$$

$$I\ (S_j : C; R_1, R_2, \ldots R_{last})\ .$$

Whichever of these values is highest defines the conformational state predicted. Each of these four information values may be estimated by inspection of the amino acid sequence. The actual procedure for carrying out this inspection and assigning the values depends on how $i(S_j; R_1, R_2, \ldots R_{last})$ is expanded. In their studies Garnier et al. (1978) have shown that the effect of one residue type on the conformation of residues up to eight residues distant plays a predominant role, while choosing shorter separation distances neglects significant information. The equation below was then proposed as a simple aproximation:

$$I\ (S_j; R_1, R2, \ldots R_{last}) \simeq \sum_{m=-8}^{m=+8} I\ (S_j; R_{j+m}),$$

(Equation 1)

in which terms in the expansion containing more than one R_j parameter have been neglected and interactions between residues separated by more than eight positions in the amino acid sequence are neglected.

The measure of $I(S_j;R_{j+m})$ represents the information which the type of residue at j+m carries about the conformational state of the j^{th} residue. Since there are 20 types of residue R and four conformational states, there are 20 x 4 differents parameters for each separation m. Since $-8 \leq m \leq 8$, including m=0, there are effectively 17 separations. If m is negative, the information concerns a residue on the N-terminal side of R_j. If m is positive, the information concerns a residue on the C-terminal side of R_j. If m is zero, the information is $I(S_j;R_j)$, which is the information the residue carries about its own conformation. There is thus a total of 20 x 4 x 17 parameters, which were calculated by the authors. The calculation of these parameters was based on the statistical (information theory) analysis of 25 proteins of known sequence and conformation.

The prediction of the conformational state of each j residue thus, involves the evaluation of equation 1 for each conformational state, and then choosing the conformation with the hightest information content.

It was also shown that the predictions may be greatly improved by introducing different decision constants for all four states. The values of each decision constant, which have to be subtracted from the information measure at each residue, were established by tracing the highest percentage of correctly predicted residues for the four conformations by varying the decision constant. These studies showed that the higher the content of a given

conformation, the more negative the value of the optimal value of the decision constant. The value of the optimal decision constant was calculated as the average of the optimal decision constant of the proteins having an α-helix content less than 20%, between 20% and 50% and over 50%. The optimization of the predictive model through the use of decision constants requires the prior measurement of secondary structure contents (e.g. through circular dichroism). When these data are not available the choice of the decision constants can still be made from a preliminary prediction with all decision constants equal to zero.

Finally, the optimized predictive model when applied to 26 proteins of known conformation, led to 66.1% on average of the residues observed as α-helix being correctly predicted as helix, 55.2% as extended chain, 42.1% as reverse turn but only 28.6% as coil.

3. RESULTS AND DISCUSSION

3.1. Purification of the Trypsin/ α-Amylase and α-Amylase Inhibitors from Ragi Seeds

The trypsin/ α-amylase and α-amylase inhibitors from ragi seeds were purified as described by Shivaraj and Pattabiraman (1980, 1981). Both inhibitors were obtained in a pure state, with yields of 25 and 17 mg per 500 g of ragi flour, for the trypsin/ α-amylase and α-amylase inhibitor, respectively . These yields were almost the same as the ones obtained by Shivaraj and Pattabiraman (1980,1981). The homogeneity of both inhibitors was ascertained by SDS-polyacrylamide gel electrophoresis, isoelectric focusing and by N-terminal analysis by the dansyl-chloride and by the DABITC-PITC double coupling method, which in every case confirmed their homogeneity.

From the SDS-polyacrylamide gel electrophoresis (SDS-PAGE) experiments the molecular weights of the inhibitors were estimated to be 13400 for the trypsin/ α-amylase inhibitor and 10200 for the α-amylase inhibitor$_A$ (FIGURE 5a) These values are in good agreement with the values calculated when the amino acid sequences of both inhibitors were determined. Previously Shivaraj and Pattabiraman (1981) had estimated the molecular weight of the trypsin/ α-amylase inhibitor as 14300 and the α-amylase inhibitor as 16500.

3.2. The Complete Amino Acid Sequence of the Trypsin/ α-Amylase Inhibitor from Seeds of Ragi (Indian Finger Millet, Eleusine coracana Gaertn.)

3.2.1. Amino Acid Analysis

The results of the amino acid analysis of the trypsin/α-amylase inhibitor from ragi are shown in Table VI. The results obtained are in very good agreement with the composition of amino acids calculated from the sequence (Figure 35) with the only exceptions beign the values for Asp,Met, Tyr and Val. It is possible that some part of the high value obtained for Asp in the analysis was contributed by the Cys residues which were not completely resolved by the column system employed. The low values for Met and Tyr were thought to be due to partial destruction of these amino acids during hydrolysis. The slightly low value for Val could be due to the high resistance to acid hydrolysis of peptide bonds involving this amino acid.

It should be noted that the amino acid composition of the ragi inhibitor reported here shows strong similarity to the very recently published data on the composition of another trypsin/α-amylase inhibitor from the same tissue by Manjunath et al (1983) except for the values of Lys, Ser, Cys and Leu.

3.2.2. N-Terminal Analysis

The N-terminal analysis of the reduced and carboxymethylated protein was carried out by the DABITC-PITC double coupling method. The sequence obtained is given below:

Ser-Val-Gly-Thr-Ser-Cys-Ile-Pro-Gly-Ala-Ile-Pro-His-Asn-Pro-Leu-Asp

The identities of Ile^{7} and Ile^{12} were ascertained by comparison with peptides Th3-7 and Th3-8.

Table VI - The amino acid composition of the trypsin/α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.).

Amino Acid	Analysis	Sequence
Asp	6.59	5
Thr	6.15	6
Ser	6.59	6
Glu	12.23	13
Gly	12.02	12
Ala	11.02	11
Val	8.38	9
Met	1.62	3
Ile	5.00	5
Leu	11.70	12
Tyr	1.26	2
Phe	1.89	2
Lys	3.02	2.5
His	2.89	3
Arg	9.98	10
Trp	ND	1
Cys	ND	10
Pro	ND	9.5

ND= Not determined.

3.2.3. C-Terminal Analysis

All attempts to identify the C-terminal residues of the native and reduced and carboxymethylated trypsin/ α-amylase inhibitor by the use of carboxypeptidase A were unsuccessful. The reasons for this became apparent when the C-terminal peptides (Th5-6, Ch3-6, CN-Ch3 and CN-T7) were identified and their analysis revealed that the C-terminal sequence was Gly-Glu. Both of these amino acids are only released very slowly by carboxypeptidase A (Allen, 1981).

3.2.4. Cyanogen Bromide (CNBr) Digestion

20 mg of the reduced and carboxymethylated protein were subjected to CNBr cleavage , under the conditions described in Section 2.2.6.2. The fragments obtained were purified by RP-HPLC, separating aproximately 1 mg at each run (Figure 6). Although cross-contamination among some peptides was observed, all peaks collected could at least be partially sequenced. These cross-contaminations are attributed to the existence of microheterogeneity within some peptides thus altering their elution time and to the incomplete reduction and/ or carboxymethylation of the protein. The sequence of the peptides obtained, together with their elution time and position in the sequence are given in Table VII.

Peptide CN-1 (positions 1 to 10)

Ser-Val-Gly-Pro-Ser-Cys-Ile-Pro-Gly-HSE

The identity of Ile^{7} was ascertained by comparison with peptide Th3-7 (see below) and by semi-quantitative amino acid analysis by the dansyl-chloride method.

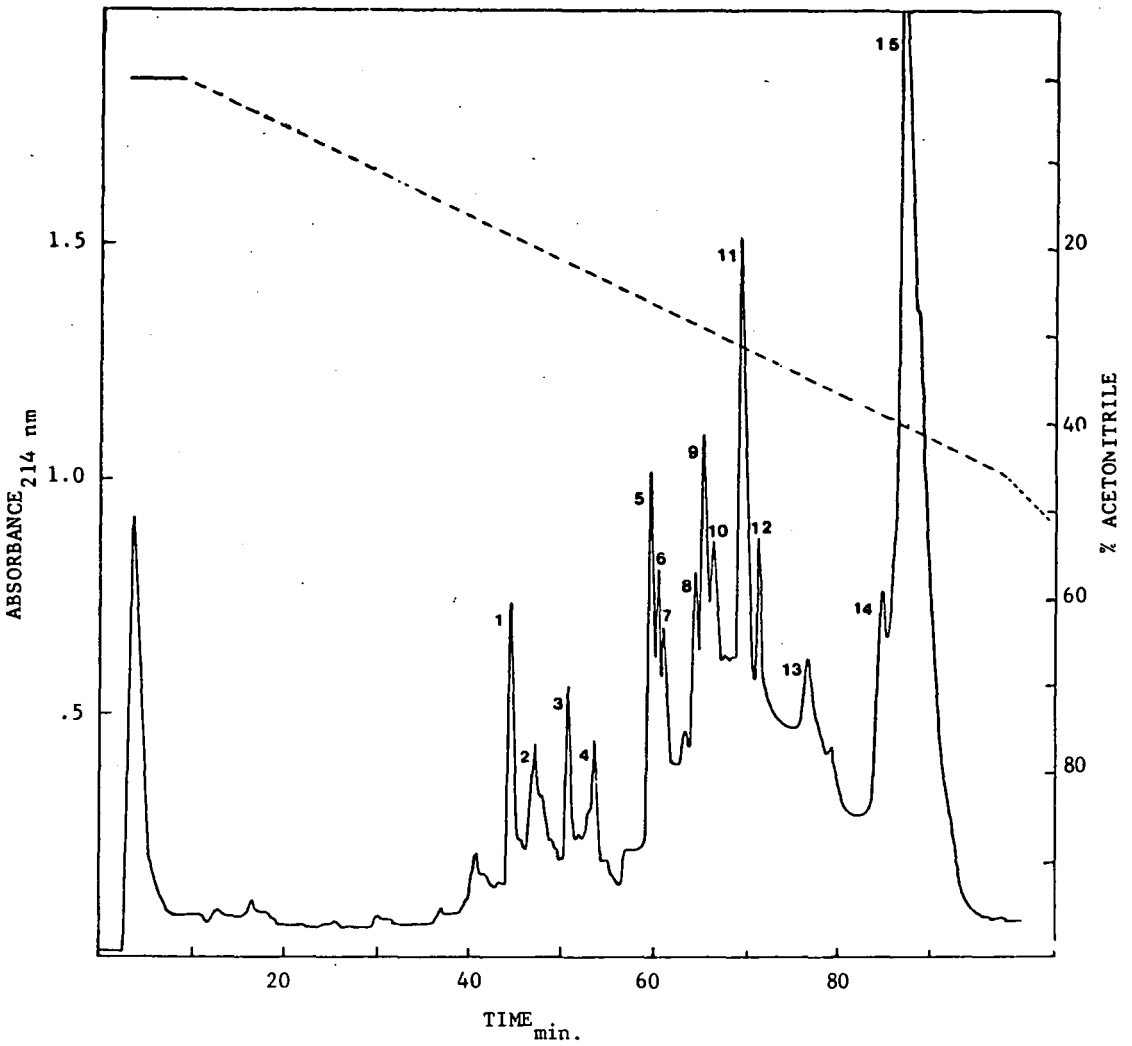


Figure 6 - RP-HPLC separation of peptides obtained by the treatment of the trypsin/ α-amylase inhibitor from ragi seeds (Indian finger finger millet, Eleusine coracana Gaertn.) with CNBr. Separation was performed on a u-Bondapak column (0.5 x 25cm) with a linear gradient of 0-70% acetonitrile (v/v) in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions corresponding to peptides which were fully or partially sequenced are numbered. Dotted line indicate the concentration of acetonitrile.

Table VII - Peptides obtained by the cleavage of the trypsin/α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) with CNBr.

Peptide no. and Position in Sequence	Elution Time	Sequence
CN-1 (1-10)	43	Ser-Val-Gly-Thr-Ser-Cys-Ile-Pro-Gly-HSE
CN-2 (23-40)	46	Tyr-Val-Ser-Thr-Arg-Thr-Cys-Gly-Val-Gly-Pro-Arg-Leu-Ala-Thr-Gln-Glu-HSE
CN-3 (11-22)	49	Ala-Ile-Pro-His-Asn-Pro-Leu-Asp-Ser-Cys-Arg-Trp
CN-4 (23-40)	53	Tyr-Val-Ala-Lys-Arg-Ala-Cys-Gly-Val-Gly-Pro-Arg-Leu-Ala-Thr-Gln-Glu-HSE
CN-5 (11-22)	58	Ala-Ile-Pro-His-Asn-Pro-Leu-Asp-Ser-Cys-Arg Trp
CN-6 (11-22)	60	Ala-Ile-Pro-His-Asn-Pro-Leu-Asp-Ser-Cys-Arg Trp
CN-7 (11-22)	61	Ala-Ile-Pro-His-Asn-Pro-Leu-Asp-Ser-Cys-Arg Trp
CN-8 (41-64)	64	Lys-Leu-Arg-Cys-Cys-Arg-Gln-Leu-Glu-Ala-Ile Pro-Ala-Tyr-Cys-Arg-Cys-Glu-Ala-Val-Arg-Ile Leu-HSE
CN-9 (41-64)	65	Lys-Leu-Arg-Cys-Cys-Arg-Gln-Leu-Glu-Ala-Ile Pro-Ala-Tyr-Cys-Arg-Cys-Glu-Ala-Val-Arg-Ile Leu-HSE
CN-10 (41-64)	67	Lys-Leu-Arg-Cys-Cys-Arg-Gln-Leu-Glu-Ala-Ile Pro-Ala-Tyr-Cys-Arg-Cys-Glu-Ala-Val-Arg-Ile Leu-HSE
CN-11 (41-64)	69	Lys-Leu-Arg-Cys-Cys-Arg-Gln-Leu-Glu-Ala-Ile Pro-Ala-Tyr-Cys-Arg-Cys-Glu-Ala-Val-Arg-Ile Leu-HSE

Table VII (cont.)

Peptide no. and Position in Sequence	Elution Time	Sequence
CN-12 (41-64)	71	Lys-Leu-Arg-Cys-Cys-Arg-Gln-Leu-Glu-Ala-Ile-Pro-Ala-Tyr-Cys-Arg-Cys-Glu-Ala-Val-Arg-Ile-Leu-HSE
CN-13 (65-122)	76	Asp-Gly-Val-Val-Thr-$\frac{\text{Ser}}{\text{Pro}}$-Ser-Gly-Gln-His-Glu-Gly-Arg-Leu-Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe-Ala-Pro-Lys-Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu
CN-14 (65-122)	84	Asp-Gly-Val-Val-Thr-$\frac{\text{Ser}}{\text{Pro}}$-Ser-Gly-Gln-His-Glu-Gly-Arg-Leu-Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe-Ala-Pro-Lys-Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu
CN-15 (65-122)	86	Asp-Gly-Val-Val-Thr-$\frac{\text{Ser}}{\text{Pro}}$-Ser-Gly-Gln-His-Glu-Gly-Arg-Leu-Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro Arg-Gln-Val-Gln-Arg-Ala-Phe-Ala-Pro-Lys-Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu

Peptides CN-3, CN-5, CN-6 and CN-7 (positions 11 to 22)

Ala-Ile-Pro-His-Asn-Pro-Leu-Asp-Ser-Cys-Arg-Trp

The identity of Ile^{12} was ascertained by subjecting the peptides to one cycle of degradation by the "Dansyl-Edman" method. The multiplicity of these peptides was thought to be due to the incomplete reduction and/or carboxymethylation of Cys^{19}.

Peptide CN-2 (positions 23-40)

Tyr-Val-Ser-Thr-Arg-Thr-Cys-Gly-Val-Gly-Pro-Arg-Leu-Ala-Thr-Gln-Glu-HSE

Peptide CN-4 (positions 23-40)

Tyr-Val-Ala-Lys-Arg-Ala-Cys-Gly-Val-Gly-Pro-Arg-Leu-Ala-Thr-Gln-Glu-HSE

Peptides CN-2 and CN-4 arose as a result of the anomalous cleavage of the Trp^{22}-Tyr peptide bond by the CNBr treatment. These peptides were separated in approximately equal amounts (Figure 6). They differ only at positions 25, 26 and 28. Confirmation of the significant level of heterogeneity in these positions was obtained from the analysis of other relevant peptides obtained from the digestion of the inhibitor with different proteolytic enzymes (see below).

Peptides CN-8, CN-9, CN-10, CN-11 and CN-12 (positions 41-64)

Lys-Ala-Arg-Cys-Cys-Arg-Gln-Leu-Glu-Ala-Ile-Pro-Ala-Tyr-Cys-Arg-Cys-Glu-Ala-Val-Arg-Ile-Leu-HSE

As for peptides CN-3, CN-5, CN-6 and CN-7, the above peptides provide further examples of different elution times for the same peptide. As no indication of microheterogeneity was detected, this phenomenon is probably due to the incomplete reduction and/or carboxymethylation of Cys^{44} and Cys^{45}, and also probably due to the presence of homoserine or homoserine lactone as the C-terminal residue.

Peptides CN-13, CN-14 and CN-15 (positions 65 to 122)

Asp-Gly-Val-Val-Thr-$\frac{Pro}{Ser}$-Ser-Gly-Gln-His-Glu-Gly-Arg-Leu-Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe-Ala-Pro-Lys-Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu

As peptides CN-13, CN-14 and CN-15 were easily isolated from the CNBr digestion of the trypsin/α-amylase inhibitor by RP-HPLC, their sub-digestion with proteolytic enzymes (see below) was used as a convenient means of establishing many important overlappings in the sequence and also to confirm the amino acid(s) residue(s) which occupy positions where heterogeneity was suspected. When refering to these peptides the term CNBr C-terminal fragment and the abbreviation CN is collectively applied to them.

The treatment of the inhibitor with CNBr led to cleavages at the three methionine residues in the molecule (Met^{10}, Met^{40}, and Met^{64}). Also cleavage occurred at the only tryptophan residue in the molecule. Although the cleavage of Trp-X peptide bonds by CNBr has been reported earlier (Blumenthal et al., 1975; Ozols and Gerard, 1977; Kashlan and Richardson, 1981), it is not known whether this cleavage was caused by the CNBr itself or was a result of prolonged

incubation in 70% formic acid. It is possible that the cleavage may have occurred due to the high concentration of CNBr used in the digestion of the inhibitor, as such conditions are known to favour cleavage at Trp-X peptide bonds (Han et al., 1983).

The sequencing of the CNBr fragments provided evidence for heterogeneity at positions 25, 26 and 28 and the variant forms were isolated separately. Apparently both forms are present in equal proportion. Heterogeneity was also detected at position 70, but in this case, the variant forms were not isolated separately; this may be due to the poor resolution of large peptides under the RP-HPLC conditions here utilized. The heterogeneities at positions 25, 26, 28 and 70 were later confirmed by the analysis of peptides arising from the cleavage of the inhibitor with different proteolytic enzymes.

3.2.5. Thermolysin Digestion

The thermolysin digestion of a 25 mg sample of the reduced and carboxymethylated protein was performed as described in Section 2.2.6.3.3. The fragments produced were separated by gel filtration on a Biogel P-4 column, equilibrated in 0.05M pyridine-acetate buffer, pH 5.4 (Figure 7). Further purification was achieved by subjecting each of the peaks from the Biogel P-4 column to RP-HPLC. The HPLC chromatograms of the separation of peaks 1, 3, 4 and 5 are shown in Figures 8, 9, 10, and 11, respectively The peptides present in peaks Th2, Th6 and Th7 were not resolved by RP-HPLC. The sequences, elution times and position in the sequence of each of the peptides purified are presented in Table VIII.

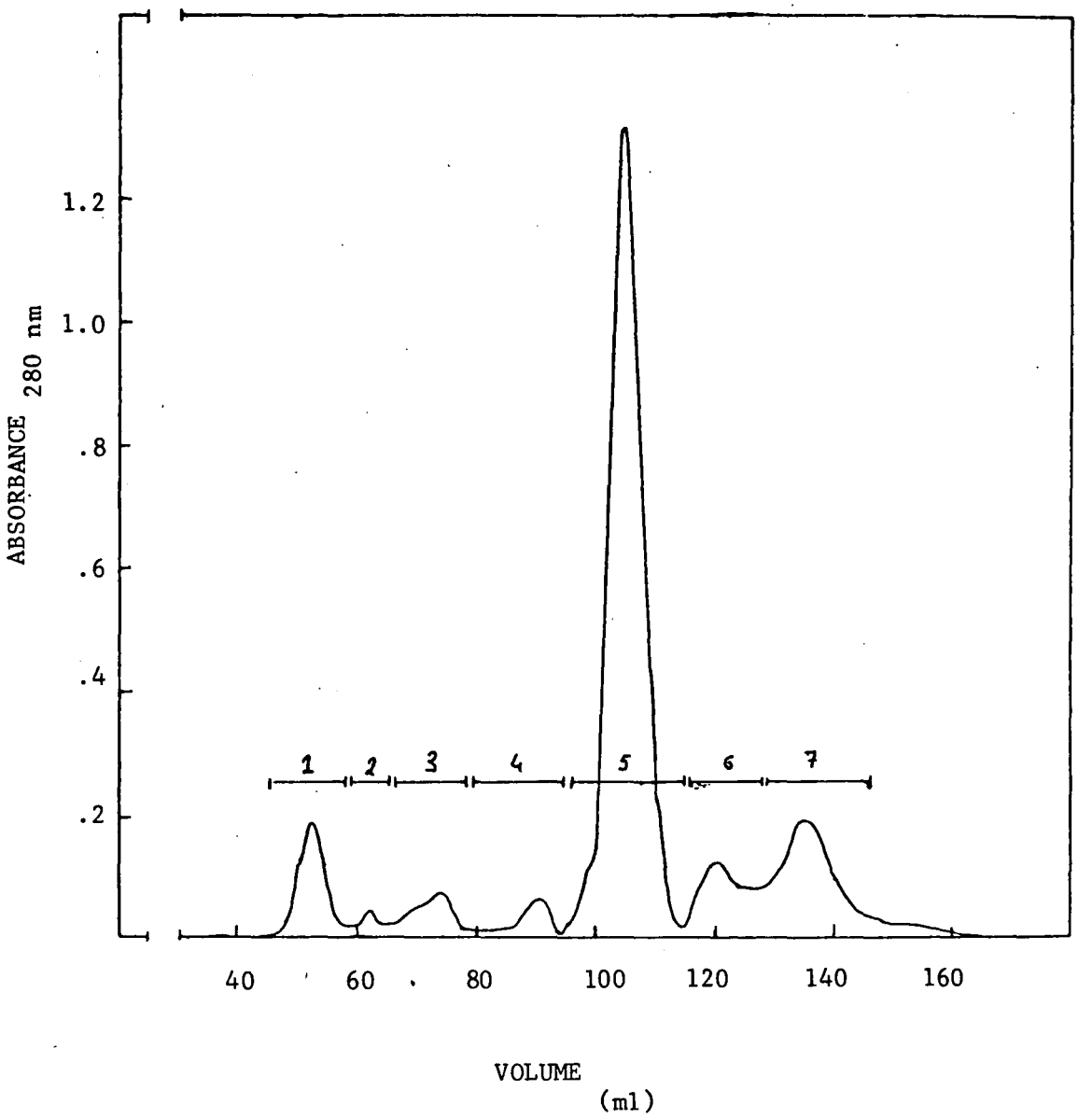


Figure 7 - Chromatography of the products of the digestion by thermolysin of the trypsin/ α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) on a Biogel P-4 column (0.8 x 190cm) in 0.05M pyridine acetate buffer, pH 5.4 (flow rate 6 ml/h., fraction size 1.4 ml).

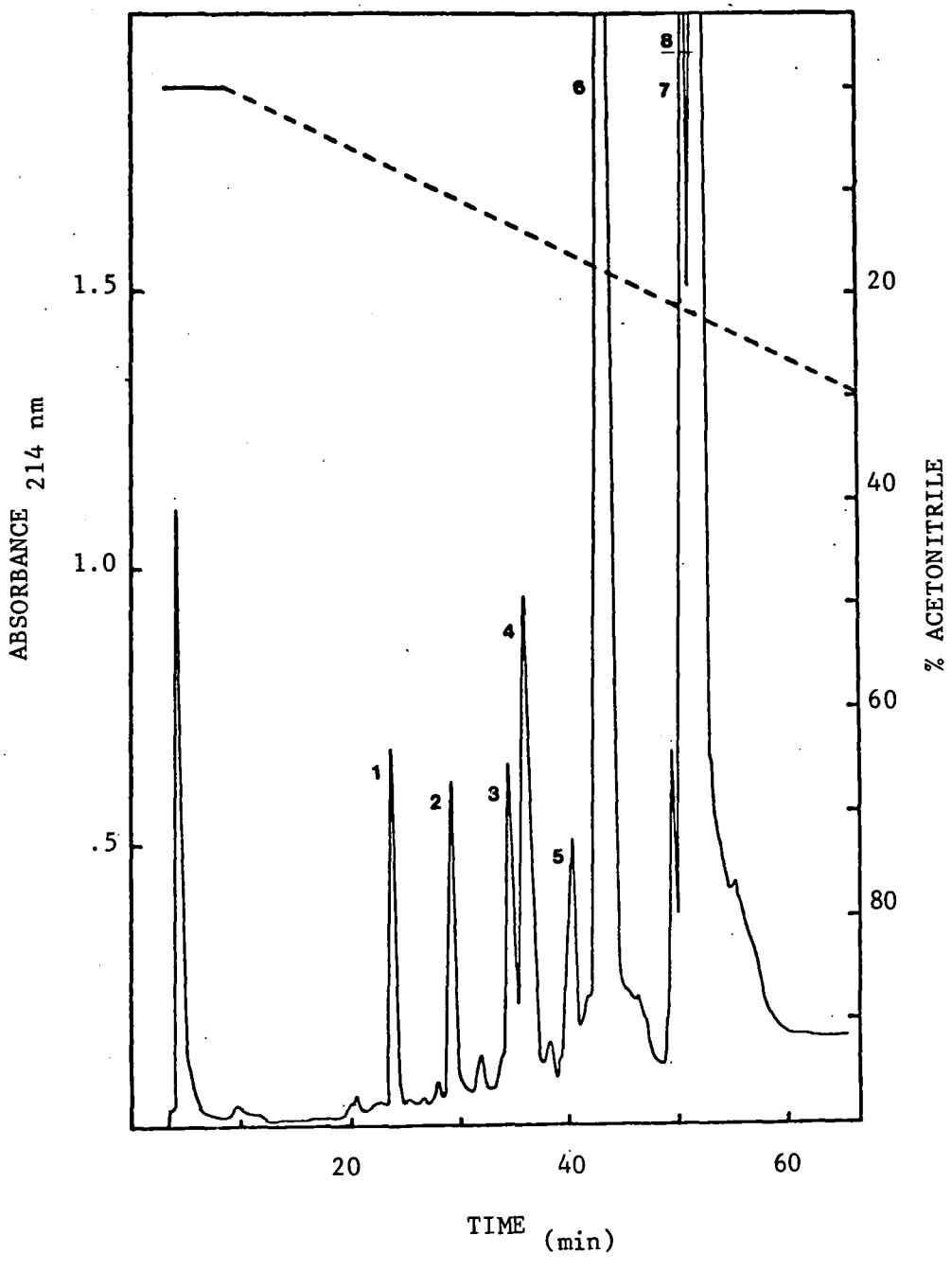


Figure 8 - RP-HPLC separation of fraction Th-1 from the thermolysin digestion of the trypsin/α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

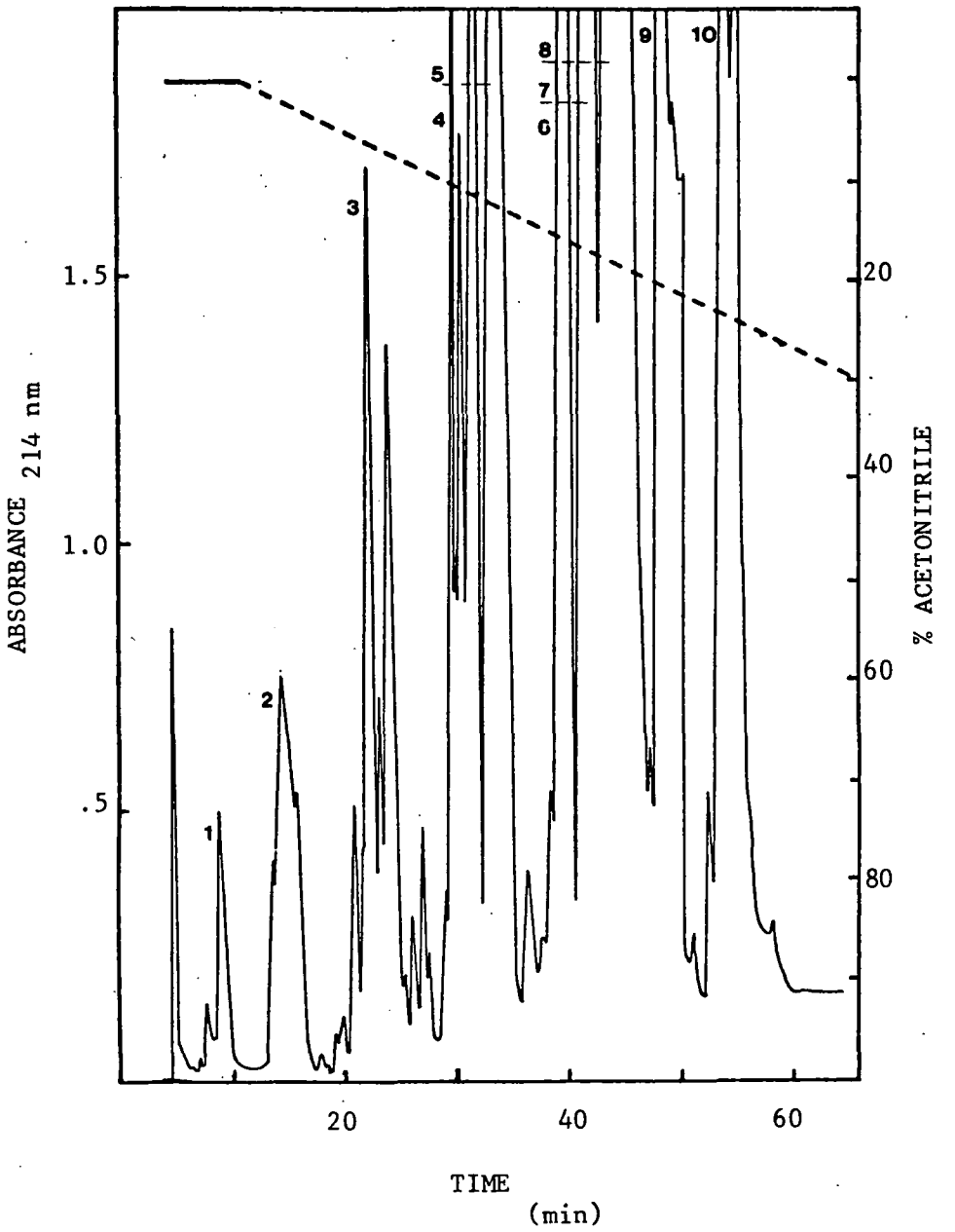


Figure 9 - RP-HPLC separation of fraction Th-3 from the thermolysin digestion of the trypsin/α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 u-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

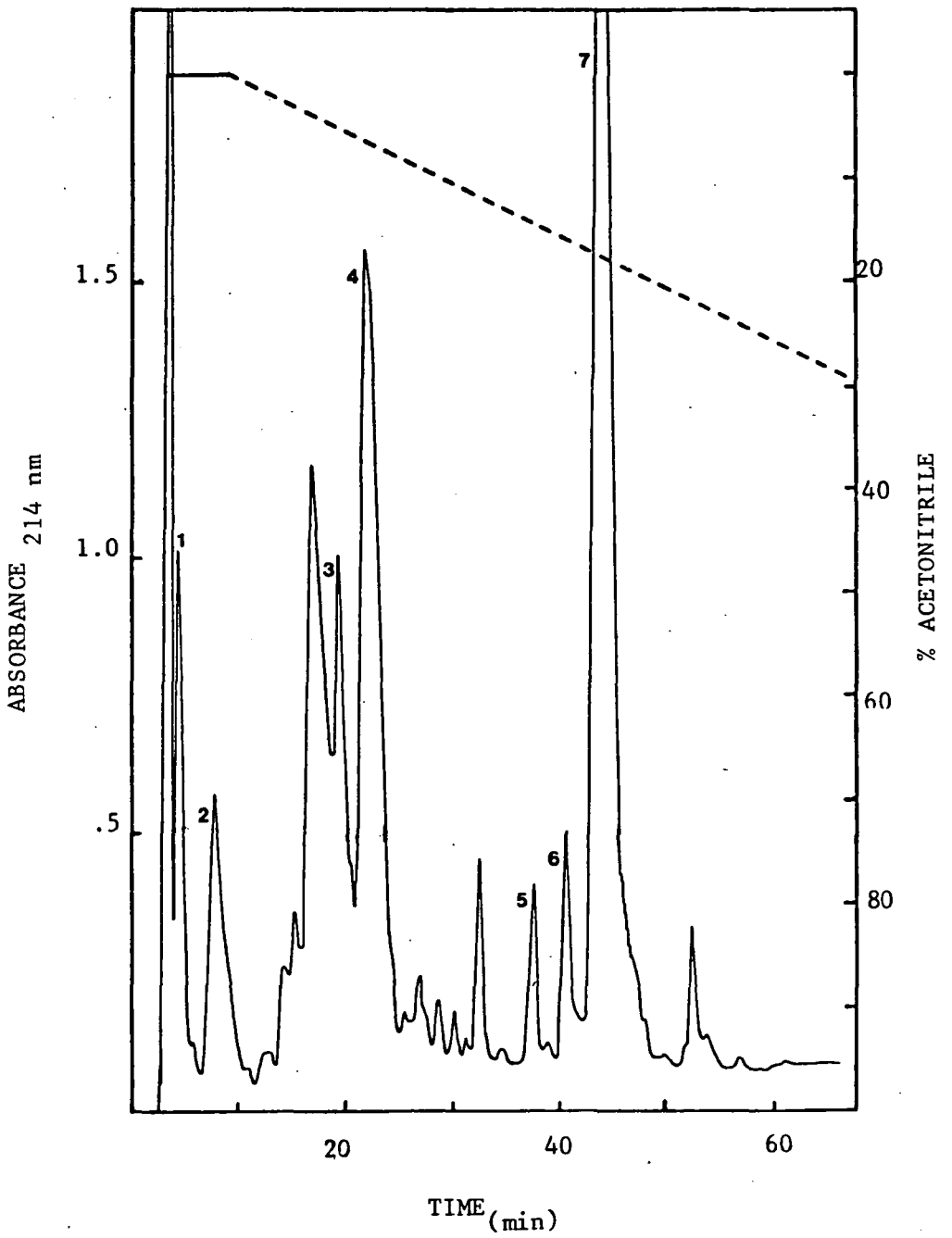


Figure 10 - RP-HPLC separation of fraction Th-4 from thermolysin digestion of the trypsin/ α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions corresponding to peptides which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

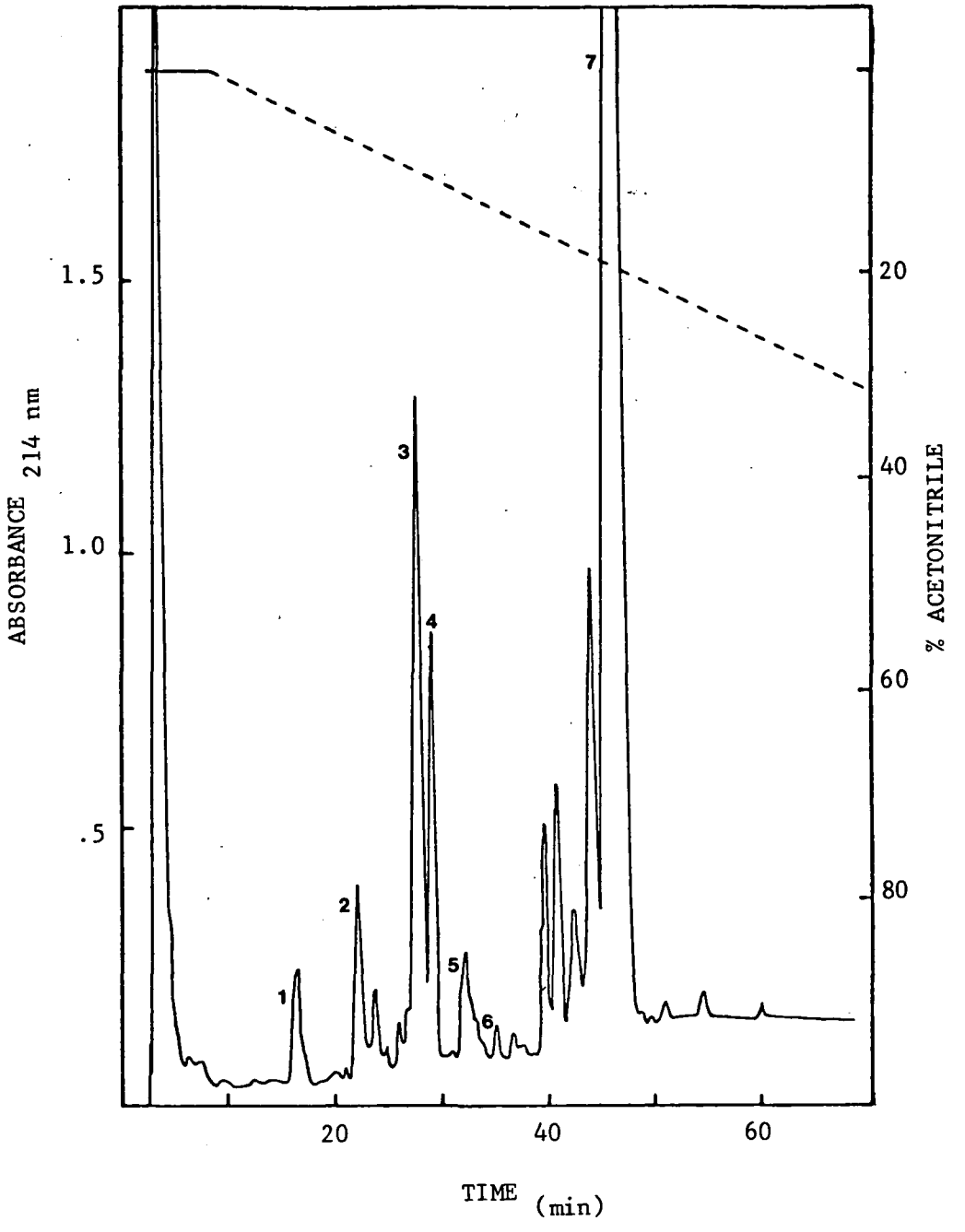


Figure 11 - RP-HPLC separation of fraction Th-5 from the thermolysin digestion of the trypsin/ α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 u-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% trifluoracetic acid at a flow rate of 0.1% ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table VIII - Peptides obtained by the cleavage of the trypsin/α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) with thermolysin.

Peptide no. and Position in Sequence		Elution Time	Sequence
Th1-1	(67-77)	23	Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu-Gly-Arg
Th1-2	(67-77)	28	Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu-Gly-Arg
Th1 3	(64-77)	34	Met-Asp-Gly-Val-Val-Thr-Pro-Ser-Gly-Gln-His-Glu-Gly-Arg
Th1-4	(67-78)	36	Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu-Gly-Arg-Leu
Th1-5	(63-69)	42	Leu-Met-Asp-Gly-Val-Val-Thr
Th1-6	(79-88)	43	Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln
Th1-7	(78-88)	50	Leu-Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln
Th1-8	(48-58)	54	Leu-Glu-Ala-Ile-Pro-Ala-Tyr-Cys-Arg-Cys-Glu
Th3-1	(24-27)	8	Val-Ala-Lys-Arg-Thr
Th3-2	(70-77)	14	Ser-Ser-Gly-Gln-His-Glu-Gly-Arg
Th3-3	(42-47)	22	Ala-Arg-Cys-Cys-Arg-Gln
Th3-4	(98-104)	30	Val-Thr-Glu-Val-Glu-Cys-Asn
Th3-5	(11-16)	33	Ala-Ile-Pro-His-Asn-Pro
Th3-6	(97-104)	40	Leu-Val-Thr-Glu-Val-Glu-Cys-Asn
Th3-7	(1-9)	41	Ser-Val-Gly-Thr-Ser-Cys-Ile-Pro-Gly
Th3-8	(10-16)	43	Met-Ala-Ile-Pro-His-Asn-Pro
Th3-9	(50-58)	48	Ala-Ile-Pro-Ala-Tyr-Cys-Arg-Cys-Glu
Th3-10	(48-58)	54	Leu-Glu-Ala-Ile-Pro-Ala-Tyr-Cys-Arg-Cys-Glu
Th4-1	(36-40)	5	Ala-Thr-Gln-Glu-Met-Lys

Table VIII - (Cont.)

Peptide no. and Position in Sequence		Elution Time	Sequence
Th4-2	(89-92)	8	Val-Gln-Arg-Ala
Th4-3	(35-39)	19	Leu-Ala-Thr-Gln-Glu
Th4-4	(93-96)	22	Phe-Ala-Pro-Lys
Th4-5	(62-66)	37	Ile-Leu-Met-Asp-Gly
Th4-6	(31-35)	41	Val-Gly-Pro-Arg
Th4-7	(108-114)	43	Ile-His-Gly-Gly-Pro-Phe-Cys
Th5-1	(105-107)	16	Leu-Ala-Thr
Th5-2	(60-61)	22	Val-Arg
Th5-3	(23-30)	28	Tyr-Val-Ser-Thr-Arg-Cys-Gly
Th5-4	(54-58)	29	Tyr-Cys-Arg-Cys-Glu
Th5-5	(60-63)	32	Val-Arg-Ile-Leu
Th5-6	(118-122)	35	Leu-Gly-Ala-Gly-Glu
Th5-7	(17-22)	45	Leu-Asp-Ser-Cys-Arg-Trp

Peptide Th3-7 (positions 1 to 9)

Ser-Val-Gly-Thr-Ser-Cys-Ile-Pro-Gly

The identity of Ile^{7} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide Th3-8 (positions 10-16)

Met-Ala-Ile-Pro-His-Asn-Pro

The identity of Ile^{12} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by subjecting the peptide to two cycles of degradation by the "Dansyl-Edman" method.

Peptide Th5-7 (positions 17-21)

Leu-Asp-Ser-Cys-Arg-Trp

The identity of Leu^{17} was ascertained by N-terminal analysis of the peptide by the dansyl-chloride method and through semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide Th3-5 (positions 11 to 16)

Ala-Ile-Pro-His-Asn-Pro

The identity of Ile^{12} was ascertained by subjecting the peptide to one cycle of degradation by the "Dansyl-Edman" method and by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide Th5-3 (positions 23-30)

Tyr-Val-Ser-Thr-Arg-Thr-Cys-Gly

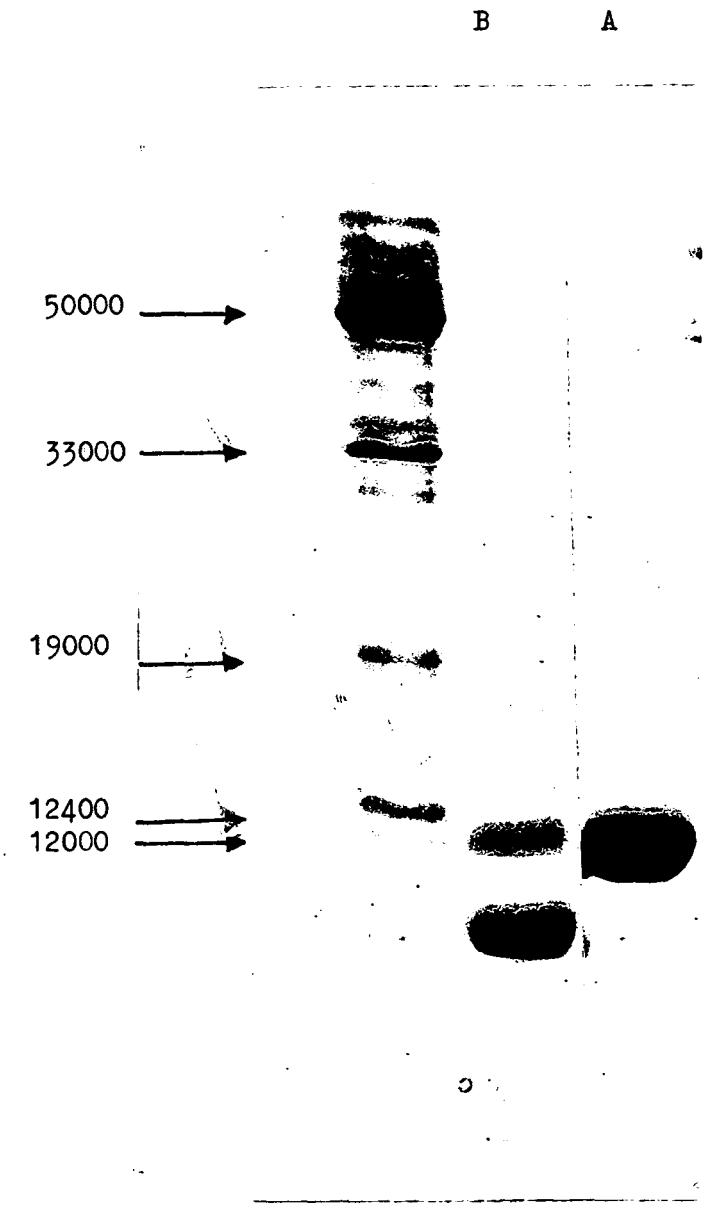


Figure 5a - SDS-polyacrylamide gel electrophoresis of the bifunctional trypsin/α-amylase inhibitor (A) and the α-amylase inhibitor (B) from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) in 17% gel (see text for details). The 12000, 19000, 33000 and 50000 vicilin subunits and cytochromec (12400) were used as molecular weight markers.

Peptide Th3-1 (positions 24-27)

Val-Ala-Lys-Arg-Thr

Peptides Th5-3 and Th3-1 confirm the heterogeneity at positions 25 and 26, as showed earlier by the CNBr peptides CN-2 and CN-4.

Peptide Th4-6 (positions 31-34

Val-Gly-Pro-Arg

Peptide Th4-3 (positions 35-39)

Leu-Ala-Thr-Gln-Glu

The identity of Leu^{35} was ascertained by N-terminal analysis by the dansyl-chloride method, and by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide Th4-1 (positions 36-41)

Ala-Thr-Gln-Glu-Met-Lys

Peptide Th3-3 (positions 42-47)

Ala-Arg-Cys-Cys-Arg-Gln

Peptides Th1-8 and Th3-10 (positions 48-58)

Leu-Glu-Ala-Ile-Pro-Ala-Tyr-Cys-Arg-Cys-Glu

The identities of Leu^{48} and Ile^{51} were ascertained by subjecting these peptides to three cycles of degradation by the "Dansyl-Edman" method, and by N-terminal analysis by the dansyl-chloride method.

Peptide Th3-9 (positions 50-58)

Ala-Ile-Pro-Ala-Tyr-Cys-Arg-Cys-Glu

The identity of Ile^{51} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by subjecting the peptide to one cycle of degradation by the "Dansyl-Edman" method.

Peptide Th5-4 (positions 54-58)

Tyr-Cys-Arg-Cys-Glu

Peptide Th5-2 (positions 60-61)

Val-Arg

Peptide Th5-5 (positions 60-63)

Val-Arg-Ile-Leu

The identities of Ile^{62} and Leu^{63} were ascertained by semi-quantitative peptide to three cycles of degradation by the "Dansyl-Edman" method and by comparison with peptides Th4-5, Th1-5 and Ch3-1.

Peptide Th4-5 (positions 62-66)

Ile-Leu-Met-Asp-Gly

The identities of Ile^{62} and Leu^{63} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by N-terminal analysis by the dansyl-chloride method.

Peptide Th1-5 (positions 63-69)

Leu-Met-Asp-Gly-Val-Val-Thr

The identity of Leu^{63} was ascertained by N-terminal analysis by the dansyl-chloride method and by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide Th1-3 (positions 64-77)

Met-Asp-Gly-Val-Val-Thr-Pro-Ser-Gly-Gln-His-Glu-Gly-Arg

Peptides Th1-1 and Th1-4 (positions 67-78)

Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu-Gly-Arg-Leu

Peptide Th1-1 is a shorter version of Th1-4, lacking the C-terminal Leu^{78}. The identity of Leu^{78} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by comparison with peptide Th1-6.

Peptide Th1-2 (positions 67-77)

Val-Val-Thr-Pro-Ser-Gly-Gln-His-Glu-Gly-Arg

Peptide Th1-2 differs from Th1-1 only in having a proline at residue 70 instead of serine, and thus provides evidence for heterogeneity at this position, which have been previously detected in the analysis of peptides CN-13, CN-14 and CN-15. As would be expected, these two peptides have different elution times (Figure 9). Further evidence for heterogeneity at position 70 will be given later.

Peptide Th3-2 (positions 70-77)

Ser-Ser-Gly-Gln-His-Glu-Gly-Arg

This peptide which arose from the cleavage of the peptide bond Thr^{69}-Ser, provides confirmation of the ability of thermolysin to specifically cleave this peptide bond, as established before by Emmers et al. (1966).

Peptide Th1-7 (positions 78-88)

Leu-Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln

The identities of Leu^{78}, Leu^{79} and Leu^{82} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by subjecting the peptide to four cycles of degradation by the

"Dansyl-Edman" method.

Peptide Th1-6 (positions 79 to 88)

Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln

The identity of Leu^{79} and Leu^{82} were ascertained by subjecting the peptide to three cycles of degradation by the "Dansyl-Edman" method and by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide Th4-2 (positions 89-92)

Val-Gln-Arg-Ala

Peptide Th4-4 (positions 93-96)

Phe-Ala-Pro-Lys

Peptide Th3-6 (positions Th3-6)

Leu-Val-Thr-Glu-Val-Glu-Cys-Asn

The identity of Leu^{97} was ascertained by N-terminal analysis by the dansyl-chloride method.

Peptide Th3-4 (positions 98-104)

Val-Thr-Glu-Val-Glu-Cys-Asn

Peptide Th5-1 (positions 105-107)

Leu-Ala-Thr

The identity of Leu^{105} was ascertained by N-terminal analysis by the dansyl-chloride method.

Peptide Th4-7 (positions 108-114)

Ile-His-Gly-Gly-Pro-Phe-Cys

The identity of Ile^{108} was ascertained by N-terminal analysis by the dansyl-chloride method.

Apart the cleavage of the peptide bond Thr^{69}-Ser, the other peptide bonds cleaved by the treatment of the ragi trypsin/α-amylase inhibitor with thermolysin were consistent with those previously reported (Matsubara, 1970; Allen, 1981). Cleavage at the amino side of serine residues has been previously reported (Emmens et. al., 1966) who found the peptide bond Thr^{99}-Ser in pike whale ribonuclease to be cleaved by thermolysin. However, as it will be showed later in the sequence of the ragi α-amylase inhibitor, thermolysin seems to be able to cleave other X-Ser and not only Thr-Ser peptide bonds.

3.2.6. Chymotrypsin Digestion

20 mg of the reduced and carboxymethylated inhibitor was used for the chymotrypsin digestion of the inhibitor. The fragments were separated on a Biogel P-4 column, equilibrated and eluted with 0.05M pyridine-acetate buffer, pH 5.4 . The fractions pooled (Figure 12) were subjected to RP-HPLC. The HPLC chromatograms of fractions Ch1,Ch2,Ch3,Ch4,Ch5,Ch6,Ch7 and Ch8 are shown in Figures 13,14,15,16,17,18, 19 and 20. A list of the peptides obtained, together with their elution times and positions in sequence is given in Table IX.

Peptide Ch2-8 (positions 1-14)

Ser-Val-Gly-Thr-Ser-Cys-Ile-Pro-Gly-Met-Ala-Ile-Pro-His

The identities of Ile^{7} and Ile^{12} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by comparison with peptides Th3-7 and Th3-8. The cleavage of the peptide bond Met^{10}-Ala in this peptide also gave rise to peptides Ch3-5 (Ser-Val-

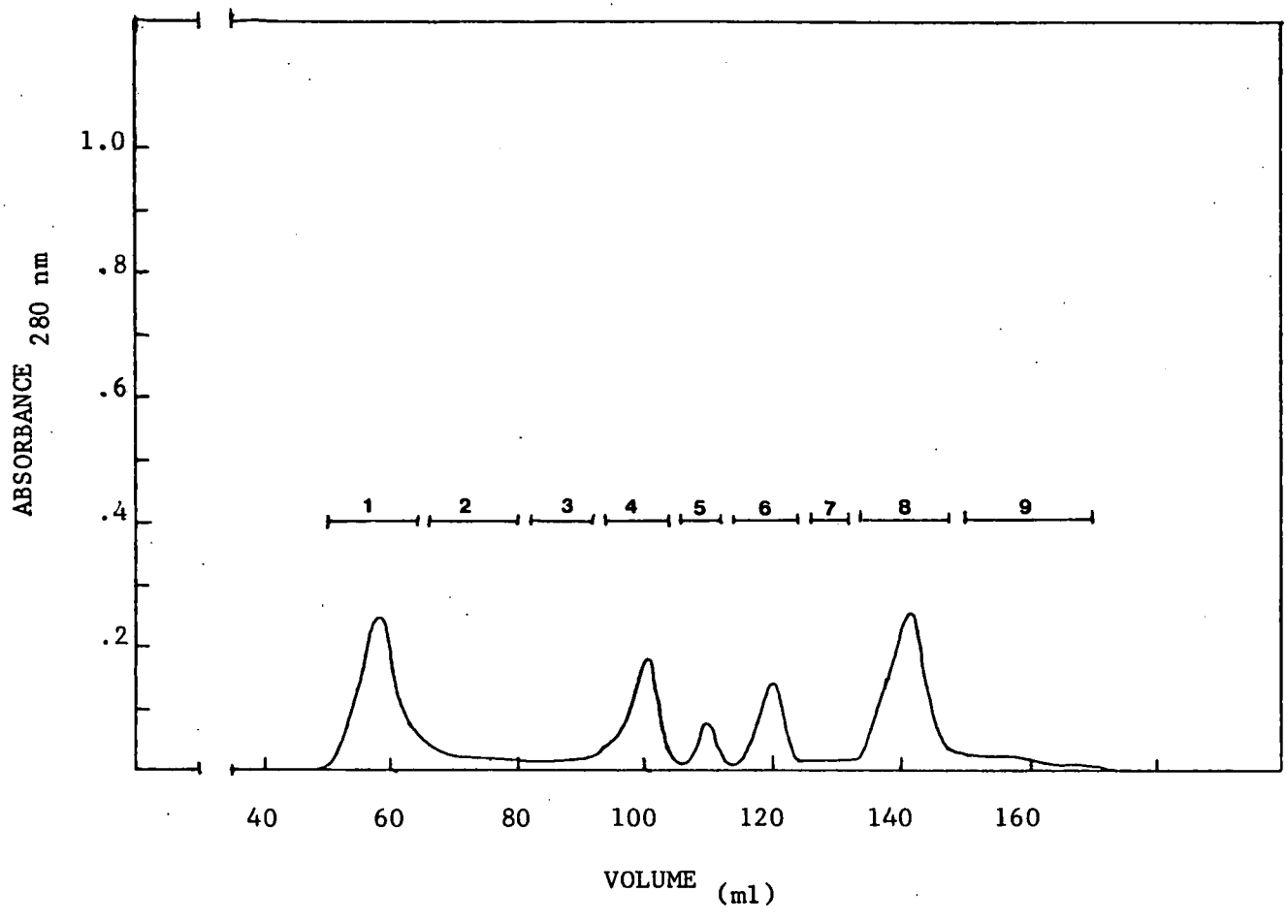


Figure 12 - Chromatography of the products of the digestion by chymotrypsin of the trypsin/α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) on a Biogel P-4 column (0.8 x 190c.m) in 0.05M pyridine acetate buffer, pH 5.4 (flow rate 6 ml/h., fraction size 1.4 ml).

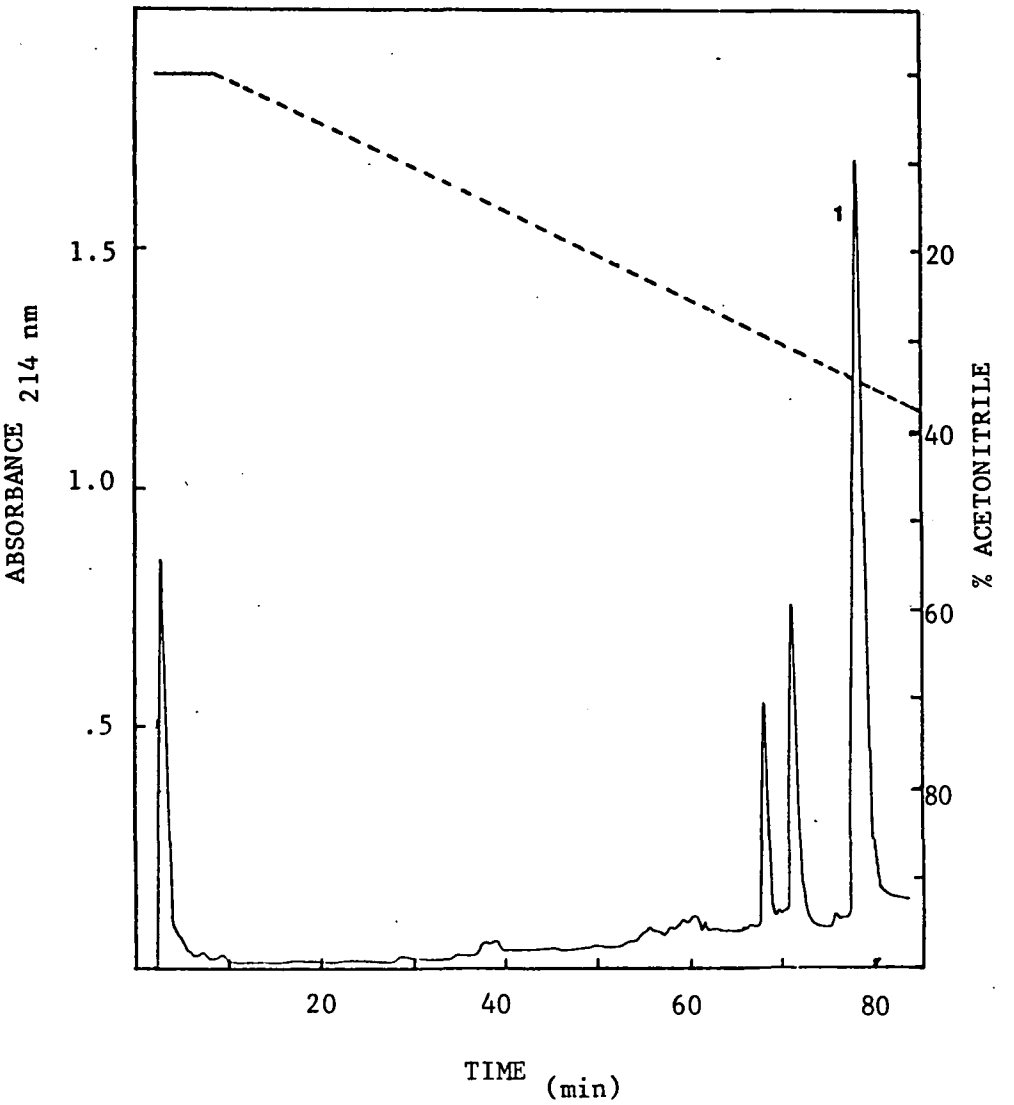


Figure 13 - RP-HPLC separation of fraction Ch-1 from the chymotrypsin digestion of the trypsin/ α-amylase inhibitor from the seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 u-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

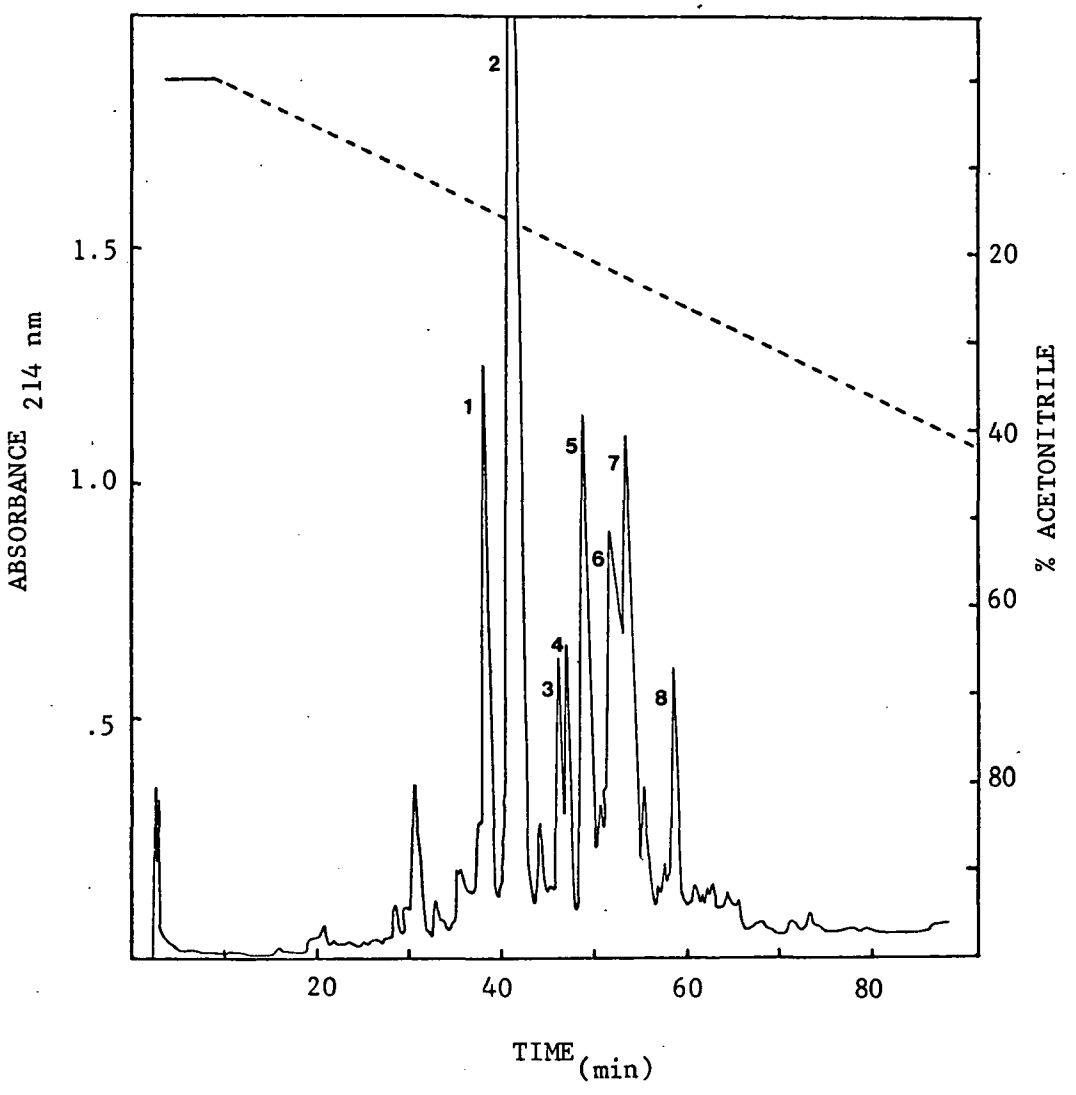


Figure 14 - RP-HPLC separation of fraction Ch-2 from the chymotrypsin digestion of the trypsin/ α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

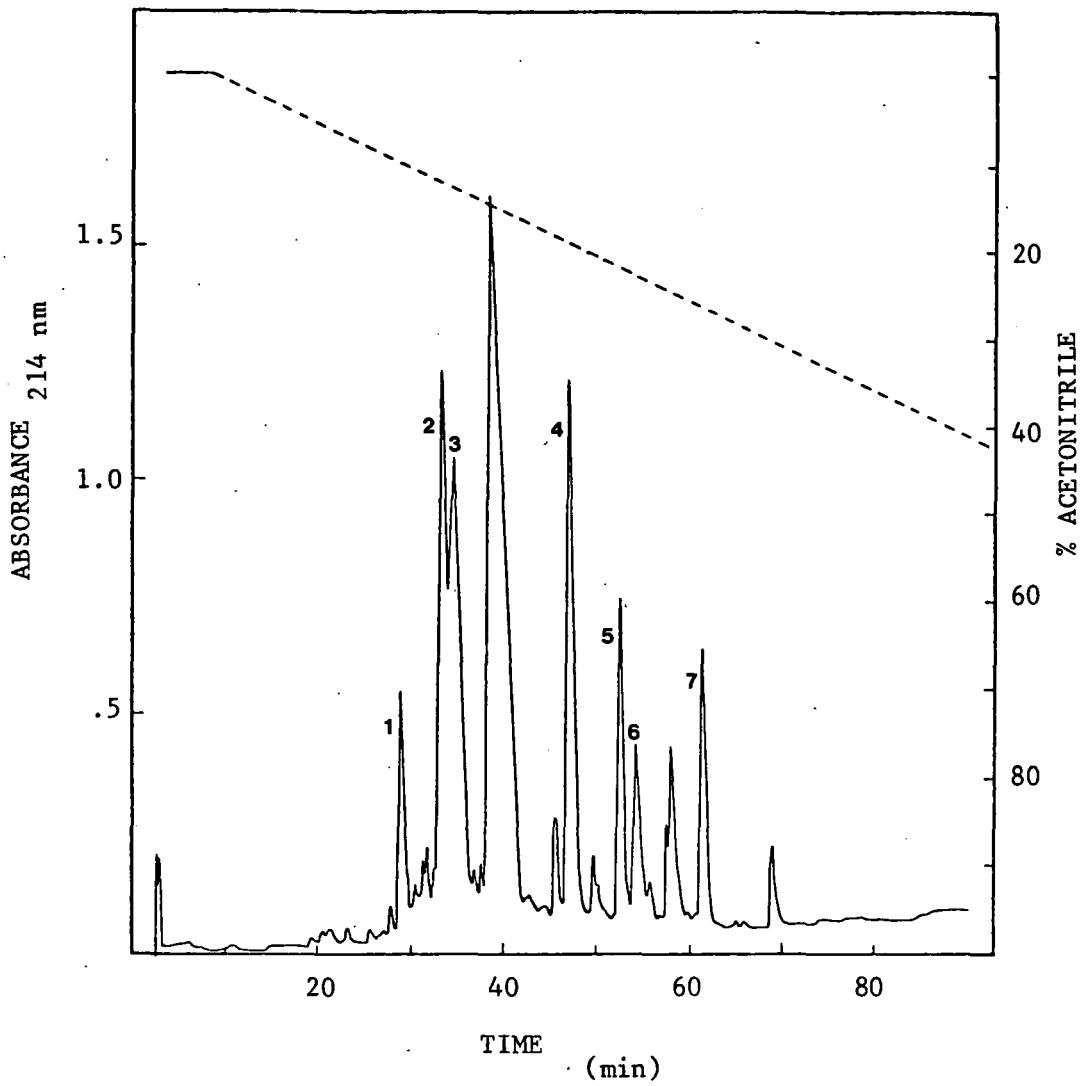


Figure 15 - RP-HPLC separation of fraction Ch-3 from the chymotrypsin digestion of the trypsin/ α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

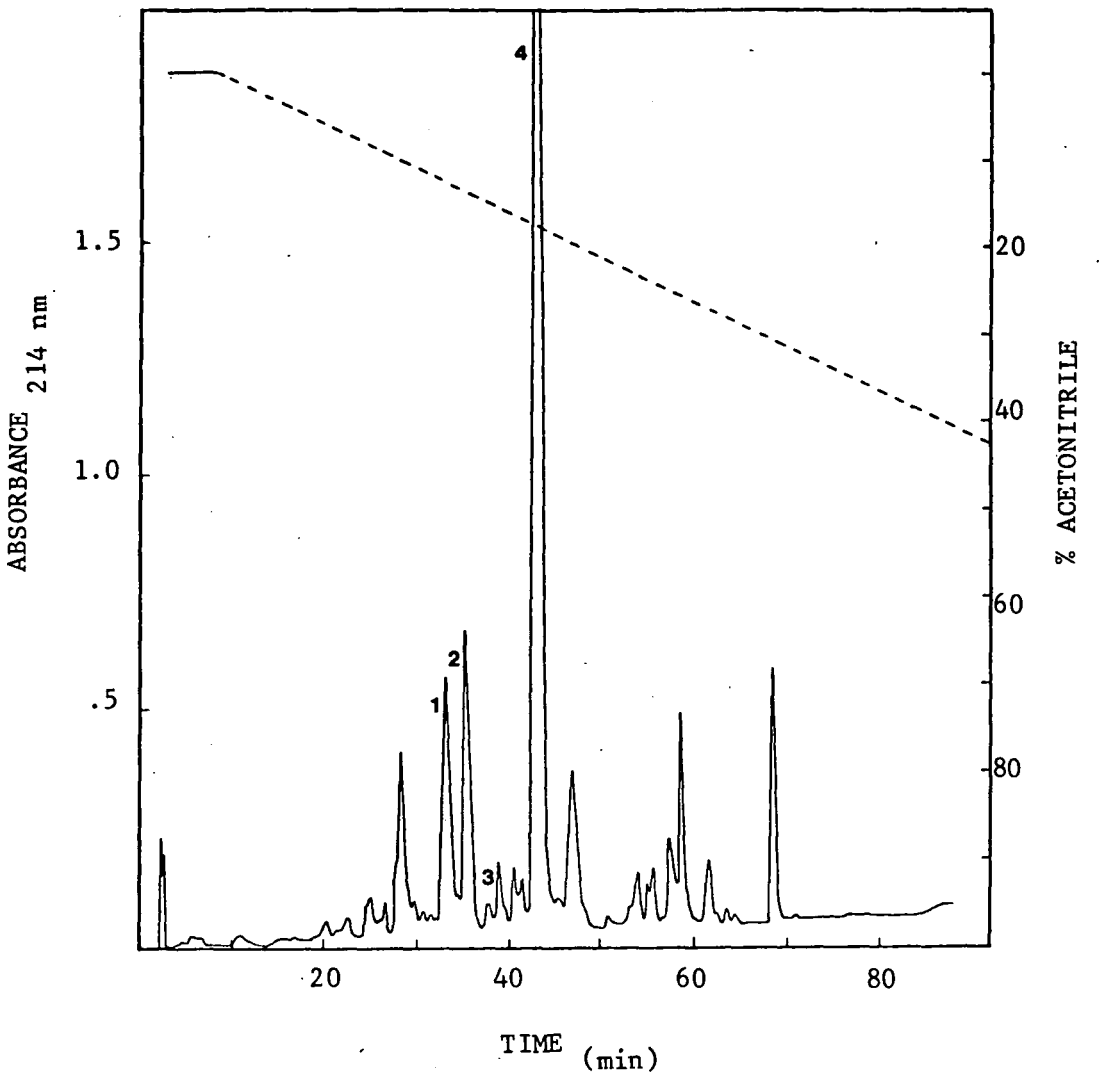


Figure 16 - RP-HPLC separation of fraction Ch-4 from the chymotrypsin digestion of the trypsin/ α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially seuqenced are numbered. Dotted line indicates concentration of acetonitrile.

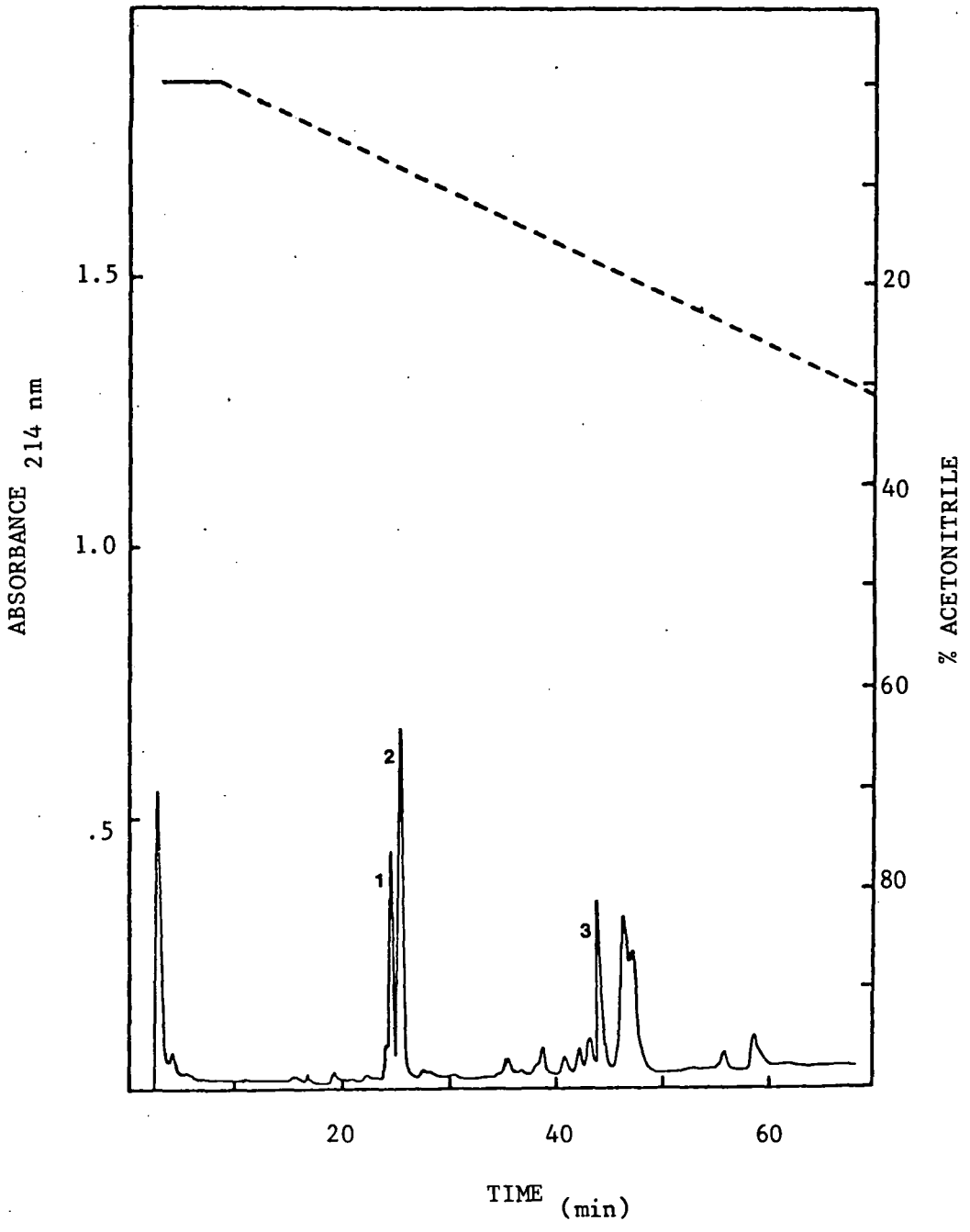


Figure 17 - RP-HPLC separation of fraction Ch-5 from the chymotrypsin digestion of the trypsin/ α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

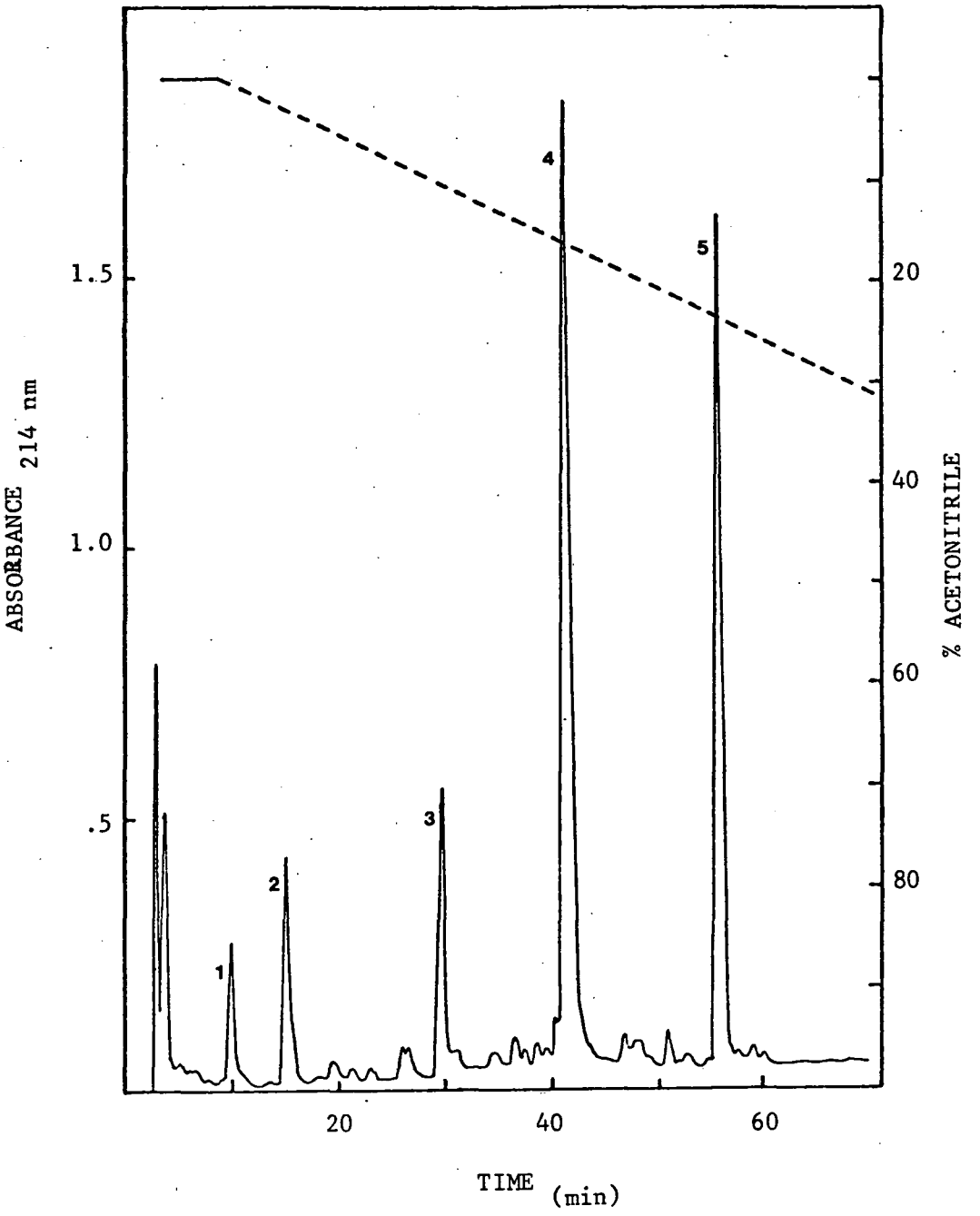


Figure 18 - RP-HPLC separation of fraction Ch-6 from the chymotrypsin digestion of the trypsin/α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially seuqenced are numbered. Dotted line indicates concentration of acetonitrile.

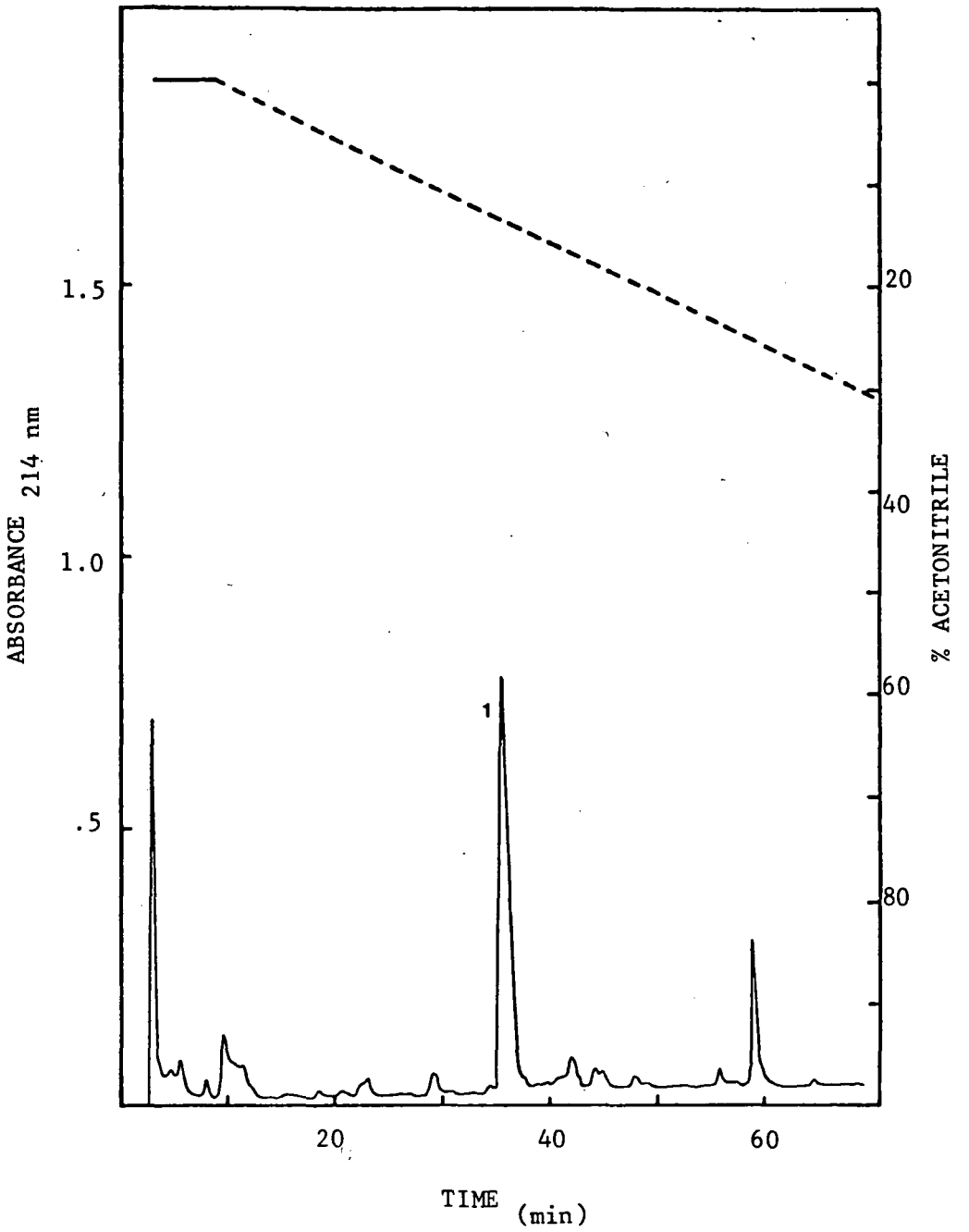


Figure 19 - RP-HPLC separation of fraction Ch-7 from the chymotrypsin digestion of the trypsin/α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

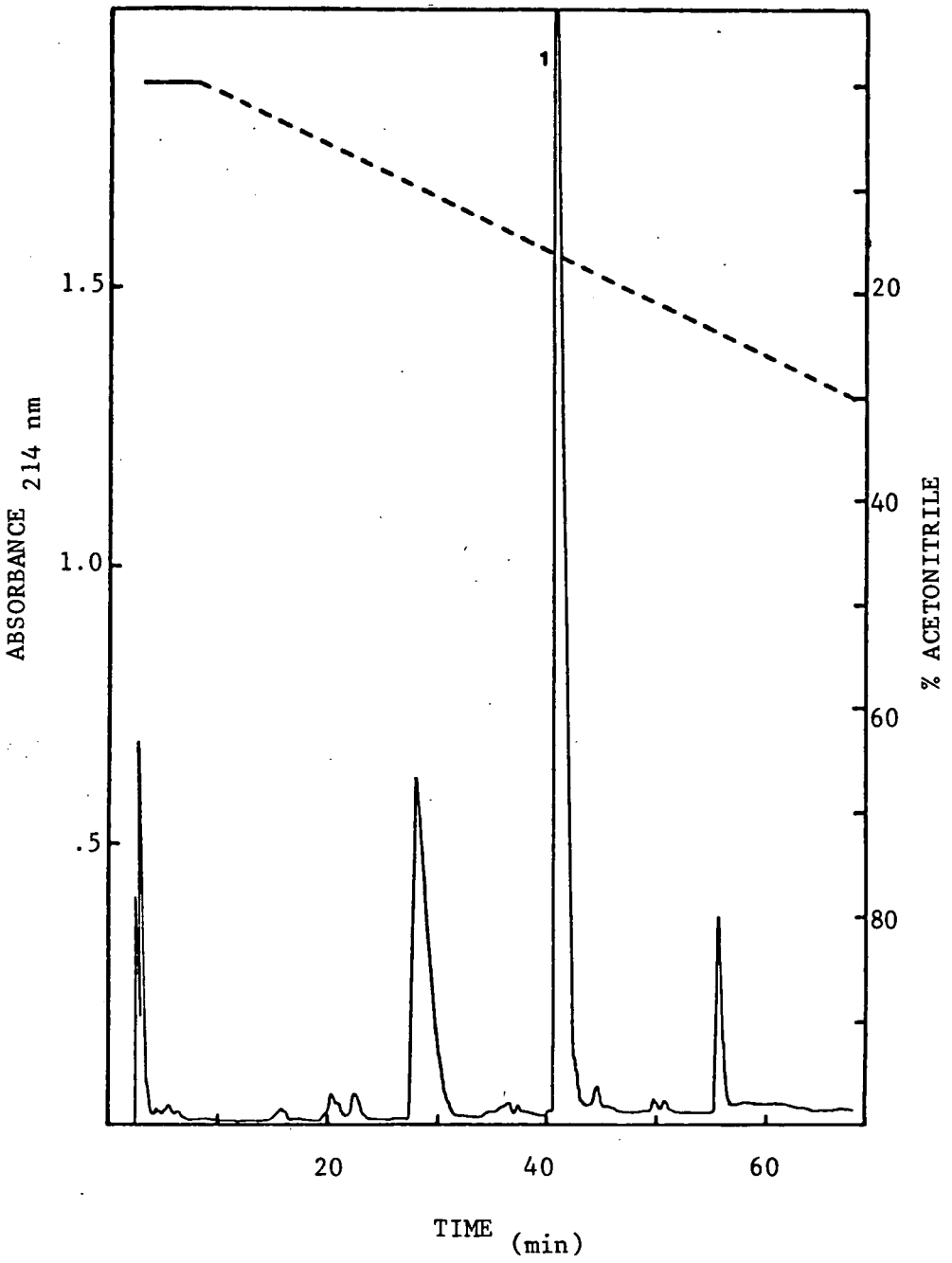


Figure 20 - RP-HPLC separation of fraction CH-8 from the chymotrypsin digestion of the trypsin/α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table IX - Peptides obtained by the digestion with chymotrypsin, of the trypsin/α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.)

Peptide No. and Position in Sequence		Elution Time	Sequence
Ch1-1	(94-122)	77	Ala-Pro-Lys-Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu
Ch2-1	(80-93)	38	Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe
Ch2-2	(80-93)	41	Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe
Ch2-3	(79-93)	47	Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe
Ch2-4	(29-40)	49	Cys-Gly-Val-Gly-Pro-Arg-Leu-Ala-Thr-Gln-Glu-Met
Ch2-5	(55-63)	51	Cys-Arg-Cys-Glu-Ala-Val-Arg-Ile-Leu
Ch2-6	(55-63)	53	Cys-Arg-Cys-Glu-Ala-Val-Arg-Ile-Leu
Ch2-7	(24-40)	55	Val-Ser-Thr-Arg-Thr-Cys-Gly-Val-Gly-Pro-Arg-Leu-Ala-Thr-Gln-Glu-Met
Ch2-8	(1-14)	58	Ser-Val-Gly-Thr-Ser-Cys-Ile-Pro-Gly-Met-Ala-Ile-Pro-His
Ch3-1	(70-78)	29	Ser-Ser-Gly-Gln-His-Glu-Gly-Arg-Leu
Ch3-2	(41-48)	33	Lys-Ala-Arg-Cys-Cys-Arg-Gln-Leu
Ch3-3	(24-35)	34	Val-Ala-Lys-Arg-Ala-Cys-Gly-Val-Gly-Pro-Arg-Leu

Table IX - (Cont.)

Peptide No. and Position in Sequence	Elution Time	Sequence
Ch3-4 (11-17)	47	Ala-Ile-Pro-His-Asn-Pro-Leu
Ch3-5 (1-10)	53	Ser-Val-Gly-Thr-Ser-Cys-Ile-Pro-Gly-Met
Ch3-6 (114-122)	54	Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu
Ch3-7 (11-23)	61	Ala-Ile-Pro-His-Asn-Pro-Leu-Asp-Ser-Cys-Arg-Trp-Tyr
Ch4-1 (35-41)	33	Leu-Ala-Thr-Gln-Glu-Met-Lys
Ch4-2 (42-48)	35	Ala-Arg-Cys-Cys-Arg-Gln-Leu
Ch4-3 (23-35)	39	Tyr-Val-Ser-Thr-Arg-Thr-Cys-Gly-Val-Gly-Pro-Arg-Leu
Ch4-4 (49-54)	42	Glu-Ala-Ile-Pro-Ala-Tyr
Ch5-1 (27-35)	26	Arg-Ala-Cys-Gly-Val-Gly-Pro-Arg-Leu
Ch5-2 (11-14)	28	Ala-Ile-Pro-His
Ch5-3 (64-78)	44	Met-Asp-Gly-Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu-Gly-Arg-Leu
Ch6-1 (24-28)	10	Val-Ser-Thr-Arg-Thr
Ch6-2 (114-117)	14	Cys-Leu-Ser-Leu
Ch6-3 (91-97)	29	Arg-Ala-Phe-Ala-Pro-Lys-Leu
Ch6-4 (18-23)	41	Asp-Ser-Cys-Arg-Trp-Tyr
Ch6-5 (15-22)	55	Asn-Pro-Leu-Asp-Ser-Cys-Arg-Trp
Ch7-1 (110-113)	35	Gly-Gly-Pro-Phe
Ch8-1 (18-23)	41	Asp-Ser-Cys-Arg-Trp-Tyr

Gly-Thr-Ser-Cys-Ile-Pro-Gly-Met) and Ch5-2 (Ala-Ile-Pro-His).

Peptide Ch3-7 (positions 11-23)

Ala-Ile-Pro-His-Asn-Pro-Leu-Asp-Ser-Cys-Arg-Trp-Tyr

The sequencing of this peptide by the DABITC/PITC double coupling method went until Trp^{22}. Tyr^{23} was placed by comparison with peptides Ch6-4 and Ch8-1 and also by semi-quantitative amino acid analysis by the dansyl-chloride method. The cleavage of the peptide bond Leu^{17}-Asp in Ch3-7 also gave rise to peptides Ch3-4 (Ala-Ile-Pro-His-Asn-Pro-Leu) and Ch6-4 and Ch8-1 (Asp-Ser-Cys-Arg-Trp-Tyr). The treatment of peptides Ch3-7, Ch6-4 and Ch8-1 with the Erlich reagent, confirmed the presence of tryptophan in it.

Peptide Ch6-5 (positions 15-22)

Asn-Pro-Leu-Asp-Ser-Cys-Arg-Trp

The identity of Leu^{17} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by subjecting the peptide to two cycles of degradation by the "Dansyl-Edman" method. The presence of tryptophan in this peptide was confirmed by the Erlich reaction.

Peptide Ch4-3 (positions 23-35)

Tyr-Val-Ser-Thr-Arg-Thr-Cys-Gly-Val-Gly-Pro-Arg-Leu

Peptide Ch2-7 (positions 24-40)

Val-Ser-Thr-Arg-Thr-Cys-Gly-Val-Gly-Pro-Arg-Leu-Ala-Thr-Gln-Glu-Met

The cleavege of the peptide bond Thr^{28}-Cys gave rise to peptides Ch6-1 (Val-Ser-Thr-Arg-Thr) and Ch2-4 (Cys-Gly-Val-Gly-Pro-Arg-Leu-Ala-Thr-Gln-Glu-Met). The identity of Leu^{35} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by comparison

with peptides Th4-3 and Ch4-1.

Peptide Ch3-3 (positions 24-35)

Val-Ala-Lys-Arg-Ala-Cys-Gly-Val-Gly-Pro-Arg-Leu

Peptide Ch5-1 (positions 27-35)

Arg-Ala-Cys-Gly-Val-Gly-Pro-Arg-Leu

Peptides Ch2-7, Ch3-3 and Ch5-1 once more provide confirmation for the heterogeneity at positions 25 (Ser/Ala), 26 (Lys/Thr) and 28 (Thr/Ala), which have been indicated before by peptides CN-13, CN-14, CN-15, Th1-1 and Th1-2.

Peptide Ch4-1 (positions 35-41)

Leu-Ala-Thr-Gln-Glu-Met-Lys

The identity of Leu^{35} was ascertained by N-terminal analysis by the dansyl-chloride method.

Peptide Ch3-2 (positions 41-48)

Lys-Ala-Arg-Cys-Cys-Arg-Gln-Leu

Peptide Ch4-2 (positions 42-48

Ala-Arg-Cys-Cys-Arg-Gln-Leu

Peptide Ch4-4 (positions 49-54)

Glu-Ala-Ile-Pro-Ala-Tyr

The identity of Ile^{51} was ascertained by semi-quantitative amino acid analysis and by comparison with peptides Th1-8 and Th3-10.

Peptides Ch2-5 and Ch2-6 (positions 55-63)

Cys-Arg-Cys-Glu-Ala-Val-Arg-Ile-Leu

The identity of Ile^{62} and Leu^{63} was ascertained by comparison with peptides

Th1-5 and Th4-5, and by the specificity of chymotrypsin.

Peptide Ch5-3 (positions 64-78)

Met-Asp-Gly-Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu-Gly-Arg-Leu

Peptide Ch3-1 (positions 70-78)

Ser-Ser-Gly-Gln-His-Glu-Gly-Arg-Leu

Peptide Ch2-3 (positions 79-93)

Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe

The identities of Leu^{79} and Leu^{82} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method, by N-terminal analysis by the dansyl-chloride method and by comparison with peptides Th1-6.

Peptide Ch2-1 and Ch2-2 (positions 80-93)

Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe

The identity of Leu^{82} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide Ch6-3 (positions 91-97)

Arg-Ala-Phe-Ala-Pro-Lys-Leu

Peptide Ch6-3 arose from an anomalous chymotryptic cleavage at the peptide bond Gln^{90}-Arg. The identity of Leu^{97} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide Ch1-1 (positions 94-122)

Ala-Pro-Lys-Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu

The identity of Leu^{97}, Leu^{105}, Ile^{108}, Leu^{114}, Leu^{116} and Leu^{117}, were

ascertained by comparison with peptides Th3-6, Th5-1, Th4-7 and Ch3-6.

Peptide Ch7-1 (positions 110-113)

Gly-Gly-Pro-Phe

Peptide Ch6-2 (positions 114-117)

Cys-Leu-Ser-Leu

The identities of Leu^{115} and Leu^{117} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide Ch3-6 (positions 114-122)

Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu

The identities of Leu^{115}, Leu^{117} and Leu^{118} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by subjecting the peptide to four cycles of degradation by the "Dansyl-Edman" method.

3.2.7. Sub-Digestion of Peptides CN-13, CN-14 and CN-15 with Chymotrypsin

During the analysis of the peptides obtained from the hydrolysis of the ragi trypsin/α-amylase inhibitor with chymotrypsin no confirmation of the previously observed heterogeneity at position 70 (Pro/Ser) was obtained. In an attempt to find this confirmation, the C-terminal CNBr fragment (CN-13, CN-14 and CN-15, Figure 6) was subdigested with chymotrypsin. The fragments were purified directly by RP-HPLC (Figure 21). As can be seen in Table X, no peptides were found which could confirm the findings of the analysis of the thermolysin fragments. Instead, evidence was found for a

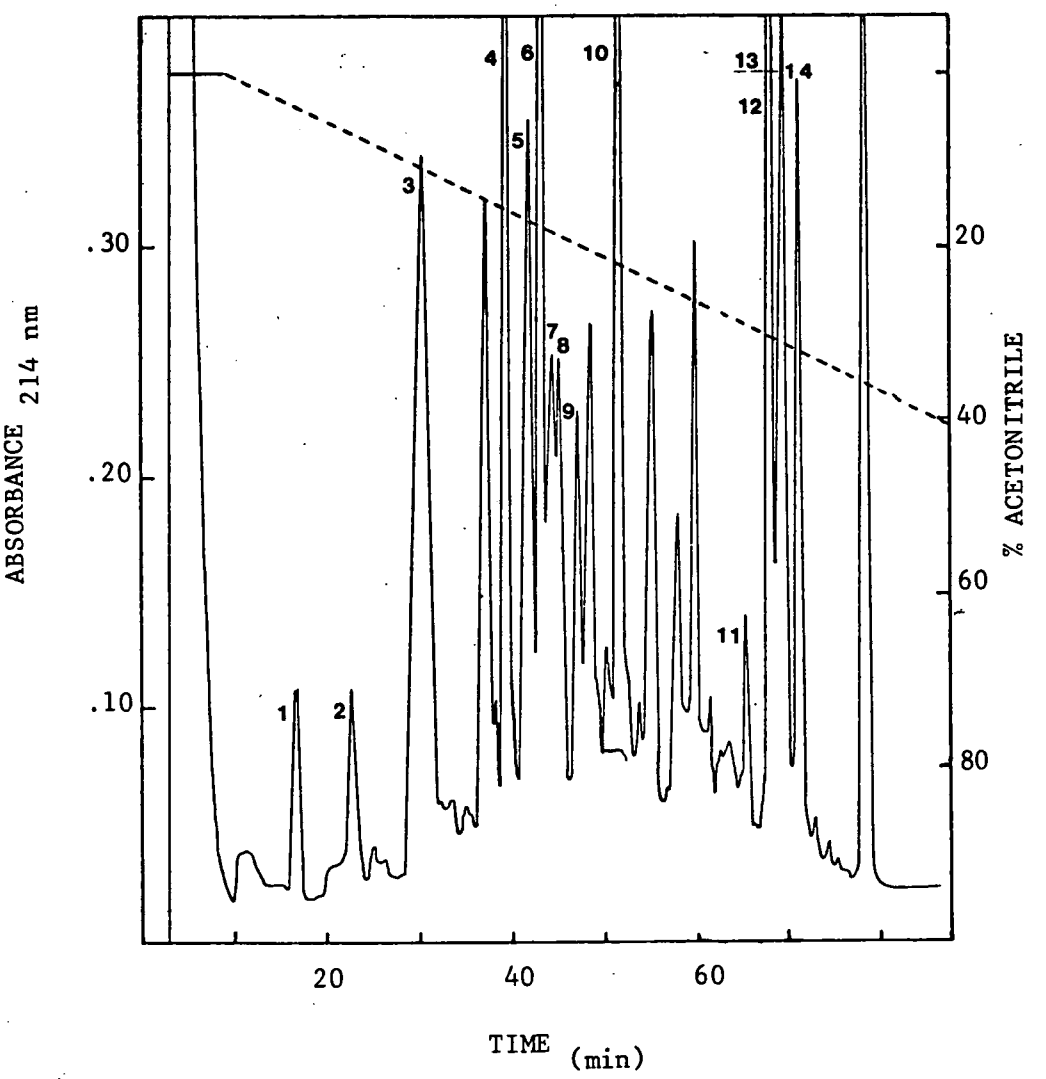


Figure 21 - RP-HPLC separation of peptides obtained by the digestion with chymotrypsin of the CNBr C-terminal peptide of the trypsin/α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.). The peptides were separated on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions corresponding to peptides which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table X - Peptides obtained by the digestion by chymotrypsin of the CNBr C-terminal fragment of the trypsin/ α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.)

Peptide No. and Position in Sequence		Elution Time	Sequence
CN-Ch1	(114-115)	16	Cys-Leu
CN-Ch2	(65-69)	22	Asp-Gly-Val-Val-Thr
CN-Ch3	(119-122)		Gly Ala-Gly-Glu
CN-Ch4	(80-93)	38	Gln-Asp-Leu-Ala-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe
CN-Ch5	(70-79)	41	Ser-Ser-Gly-Gln-His-Glu-Gly-Arg-Leu-Leu
CN-Ch6	(80-93)	42	Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe
CN-Ch7	(80-93)	43	Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe
CN-Ch8	(108-113)	44	Ile-His-Gly-Gly-Pro-Phe
CN-Ch9	(106-113)	47	Ala-Thr-Ile-His-Gly-Gly-Pro-Phe
CN-Ch10	(114-122)	51	Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu
CN-Ch11	(94-105)	65	Ala-Pro-Lys-Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu
CN-Ch12	(94-105)	68	Ala-Pro-Lys-Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu
CN-Ch13	(94-105)	69	Ala-Pro-Lys-Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu
CN-Ch14	(94-105)	71	Ala-Pro-Lys-Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu

microheterogeneity at position 83. Two peptides (CN-Ch4 and CN-Ch6) were sequenced and differ from one another only at position 83 (alanine in CN-Ch4 and proline in CN-Ch6).

The treatment of the inhibitor and its C-terminal CNBr fragment with chymotrypsin led to the expected cleavages at the carboxyl side of tyrosine, phenylalanine, methionine and leucine. Also observed were cleavages at threonine, glutamine, and histidine residues. Such cleavages at threonine, glutamine and histidine residues have often been observed previously (Croft, 1980). In addition some hydrolysis of the Arg-X and Lys-X peptide bonds was observed. (e.g. Arg^{34}-Leu and Lys^{41}-Ala), which was probably due to the contamination of the chymotrypsin preparation utilised by trypsin. These "anomalous" cleavages illustrate the difficulty in predicting accurately the sites of chymotrypsin cleavage, since it depends upon factors outside the immediate environment of the bond cleaved (Kaspar, 1970).

3.2.8. Staphylococcus aureus (Strain V-8) Protease Digestion

10 mg of the reduced and carboxymethylated ragi trypsin/ α-amylase inhibitor was digested with the V-8 protease, under the conditions described in Section 2.2.6.3.4. The fragments were separated on a Biogel P-4 column, equilibrated with 0.05M pyridine-acetate buffer, pH 5.4 . The peptides eluted from the column were monitored by measurement of theabsorbance of each tube collected at 280nm (Figure 22). Further purification was achieved by applying each peak from the Biogel column to RP-HPLC. The chromatograms of the RP-HPLC separation of peaks 2,3,4,5,6, and 7 are shown in

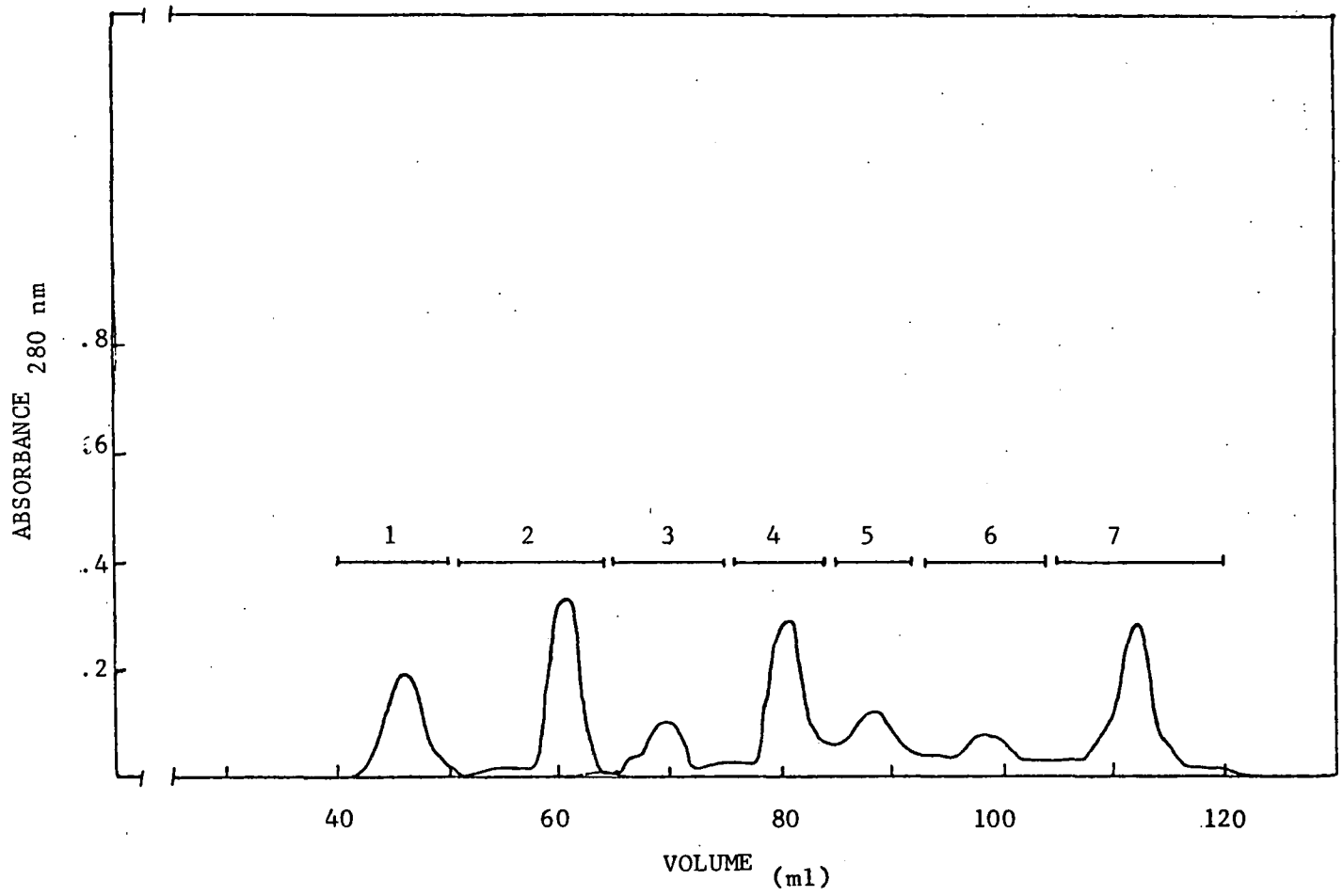


Figure 22 - Chromatography of the products of the digestion by Staphylococcus aureus (strain V-8) protease of the trypsin/ α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) on a Biogel P-4 column (0.8 x 190 cm) in 0.05M pyridine acetate buffer, pH 5.4 (flow rate 6 ml/h., fraction size 1.4 ml).

Figures 23, 24, 25, 26, 27 and 28. The sequences, elution time and positions in the sequence of the peptides obtained are presented in Table XI.

Peptide V2-1 (positions 1-34)

Ser-Val-Gly-Thr-Ser-Cys-Ile-Pro-Gly-Met-Ala-Ile-Pro-His-Asn-Pro-Leu-Asp-Ser-Cys-Arg-Trp-Tyr-Val-Ser-Thr-Arg-Thr-Cys-Gly-Val-Gly-Pro-Arg

The identity of Ile^{7} was ascertained by comparison with peptide Th3-7.

Peptide V5-1 (positions 35-39)

Leu-Ala-Thr-Gln-Glu

Peptide V5-1 arose from an anomalous cleavage at the peptide bond Arg^{34}-Leu. The identity of Leu^{35} was ascertained by N-terminal analysis by the dansyl-chloride method.

Peptide V5-4 (positions 40-46)

Met-Lys-Ala-Arg-Cys-Cys-Arg-Gln-Leu-Glu

Peptide V4-1 (positions 50-58)

Ala-Ile-Pro-Ala-Tyr-Cys-Arg-Cys-Glu

The identity of Ile^{51} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by comparison with peptides Th1-6 and Th3-10.

Peptide V7-1 (positions 59-65)

Ala-Val-Arg-Ile-Leu-Met-Asp

The identities of Ile^{62} and Leu^{63} were ascertained by comparison with peptides Th4-5 and Th1-5.

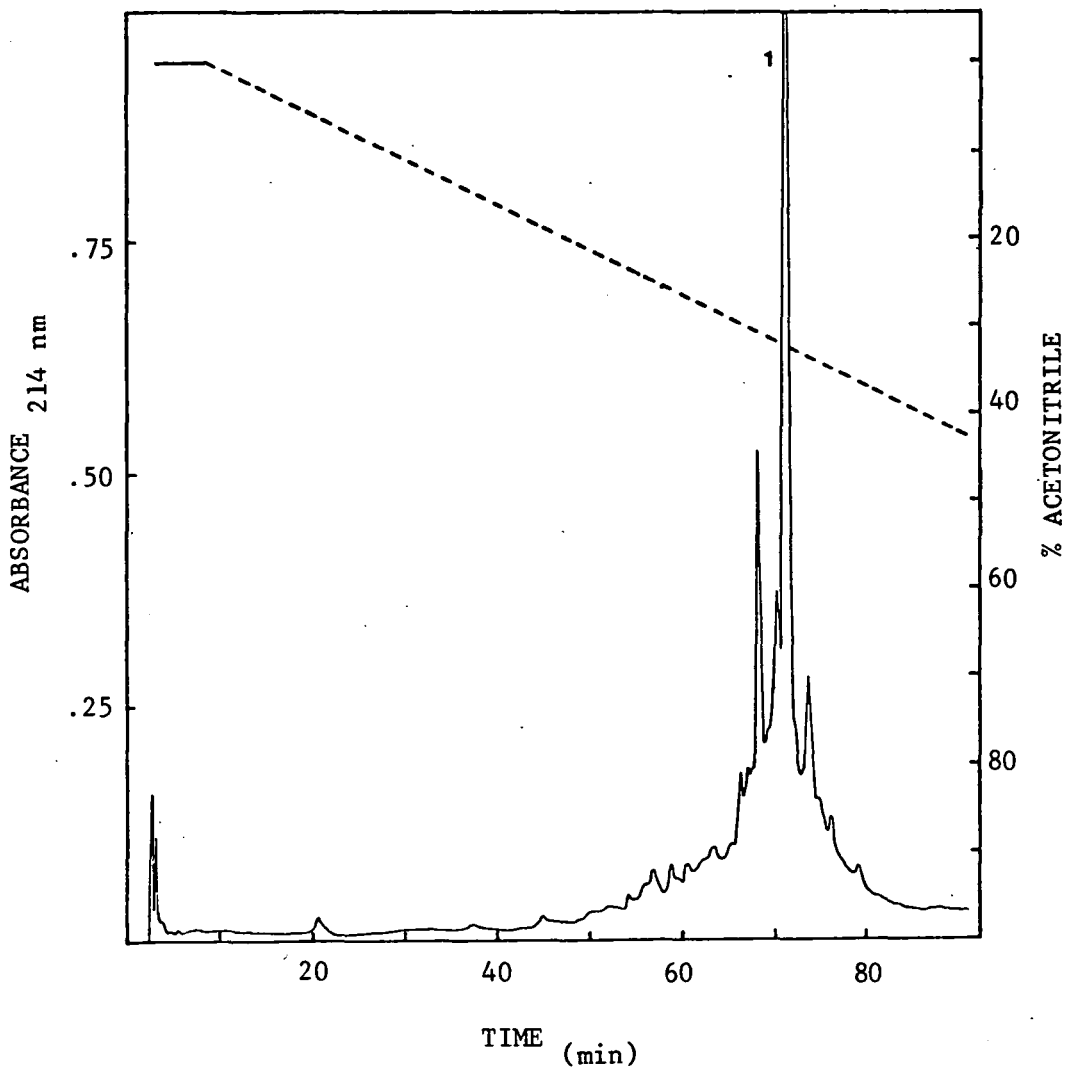


Figure 23 - RP-HPLC separation of fraction V-2 from the Staphylococcus aureus (strain V-8) protease digestion of the trypsin/α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

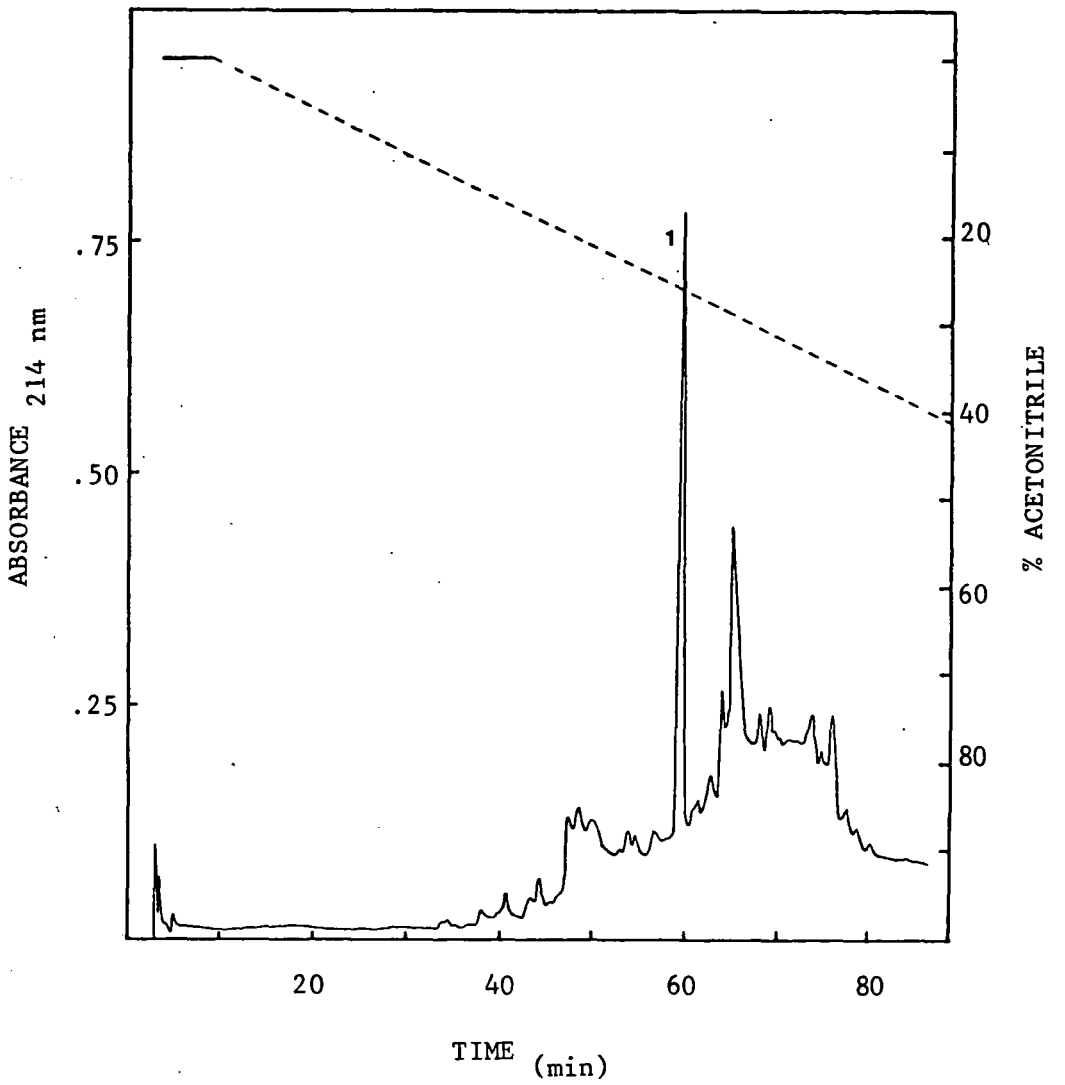


Figure 24 - RP-HPLC separation of fraction V-3 from the Staphylococcus aureus (strain V-8) protease digestion of the trypsin/α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 u-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

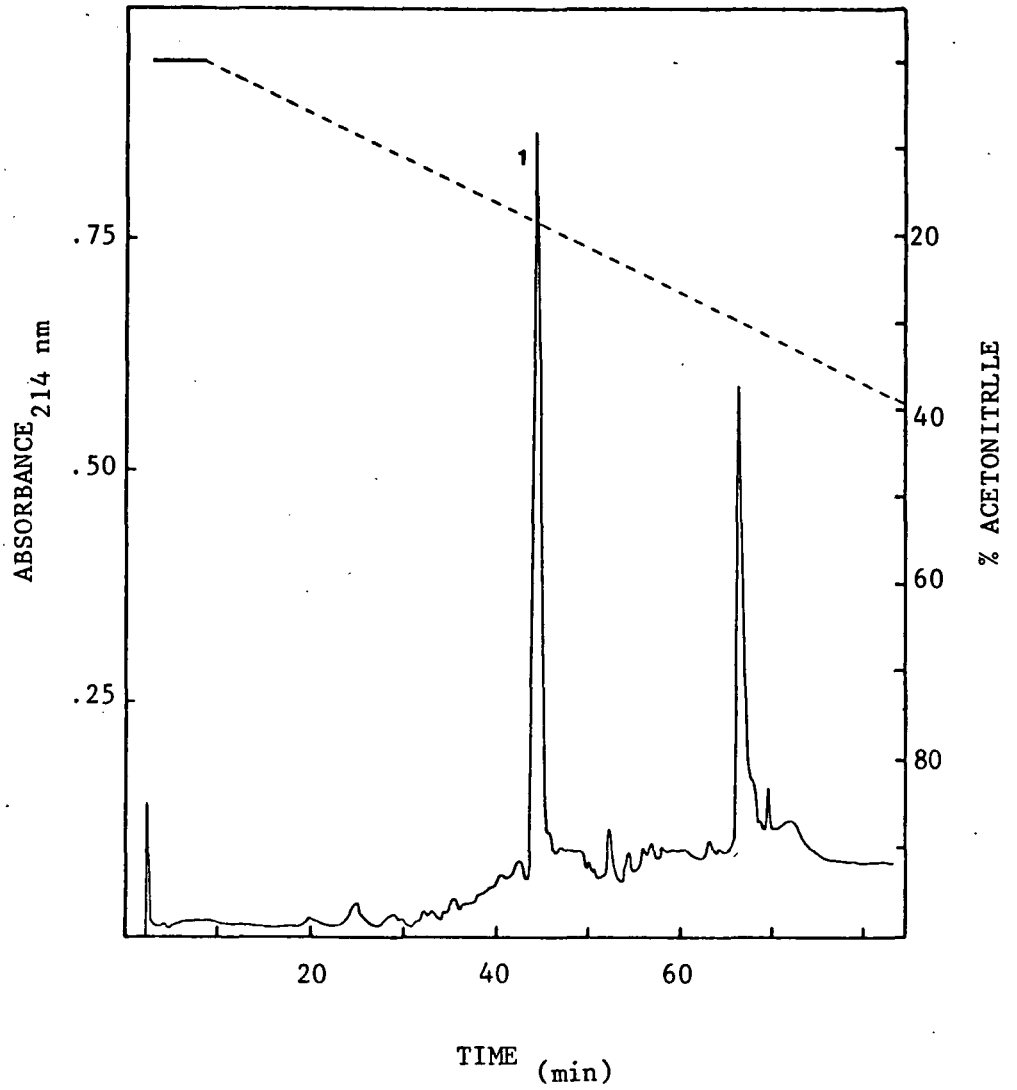


Figure 25 - RP-HPLC separation of fraction V-4 from the Staphylococcus aureus (strain V-8) protease digestion of the trypsin/ α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25 cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

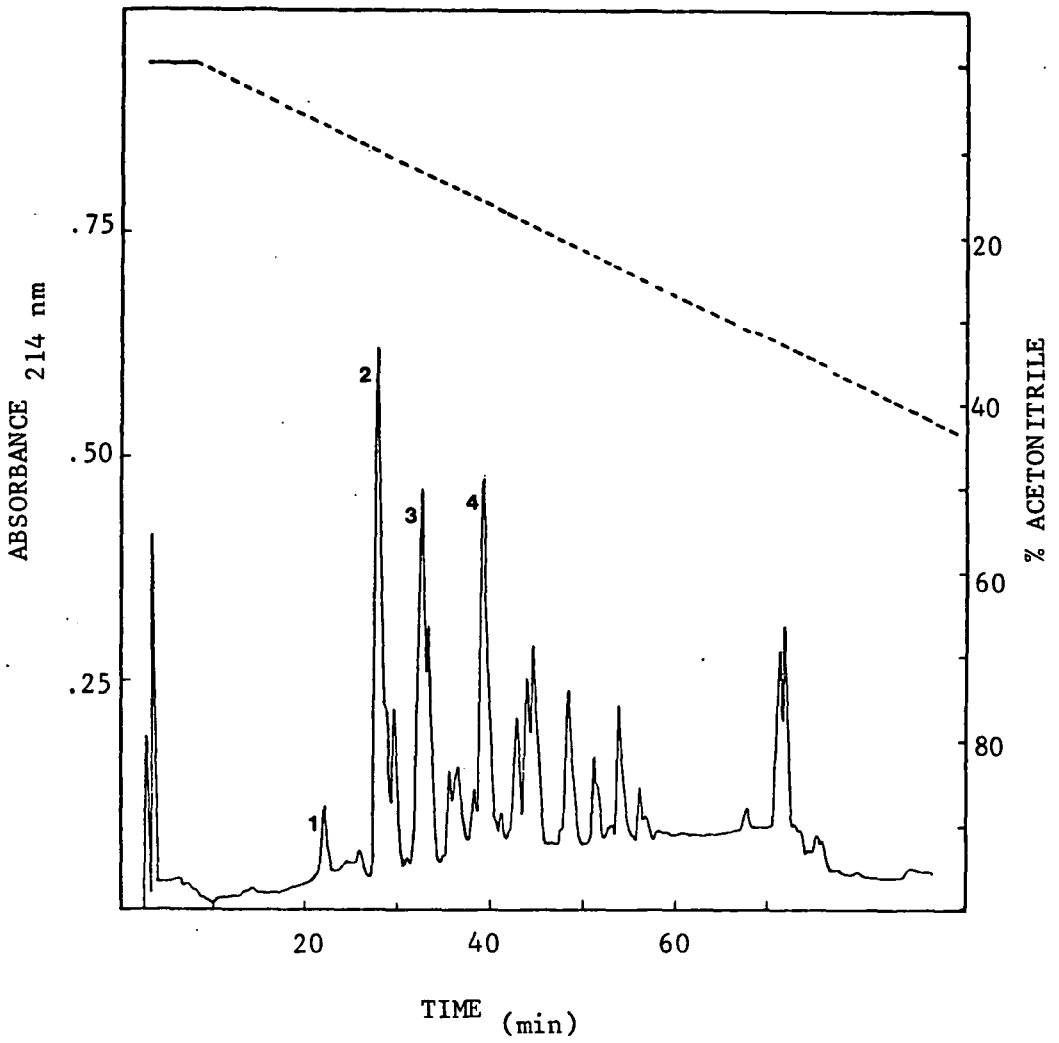


Figure 26 - RP-HPLC separation of fraction V-5 from the Staphylococcus aureus (strain V-8) protease digestion of the trypsin/ α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

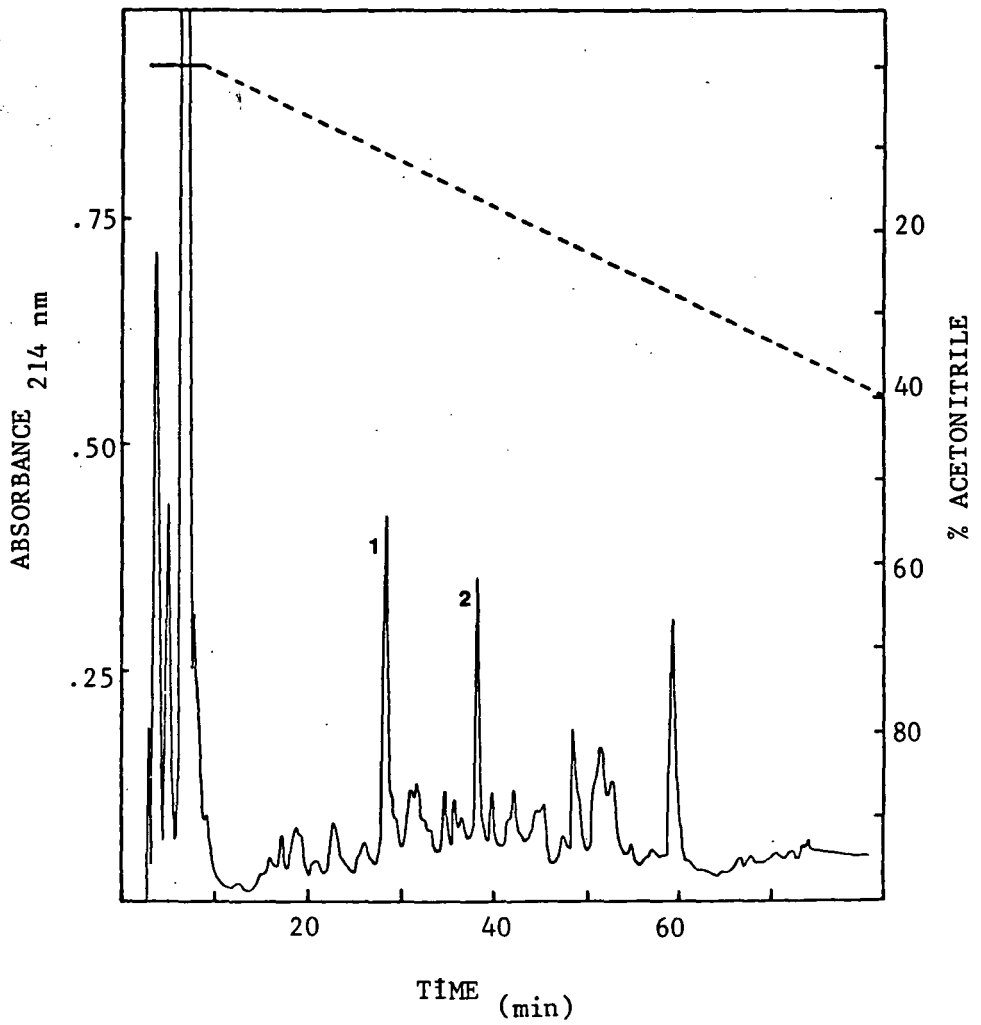


Figure 27 - RP-HPLC separation of fraction V-6 from the Staphylococcus aureus (strain V-8) protease digestion of the trypsin/ α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25 cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

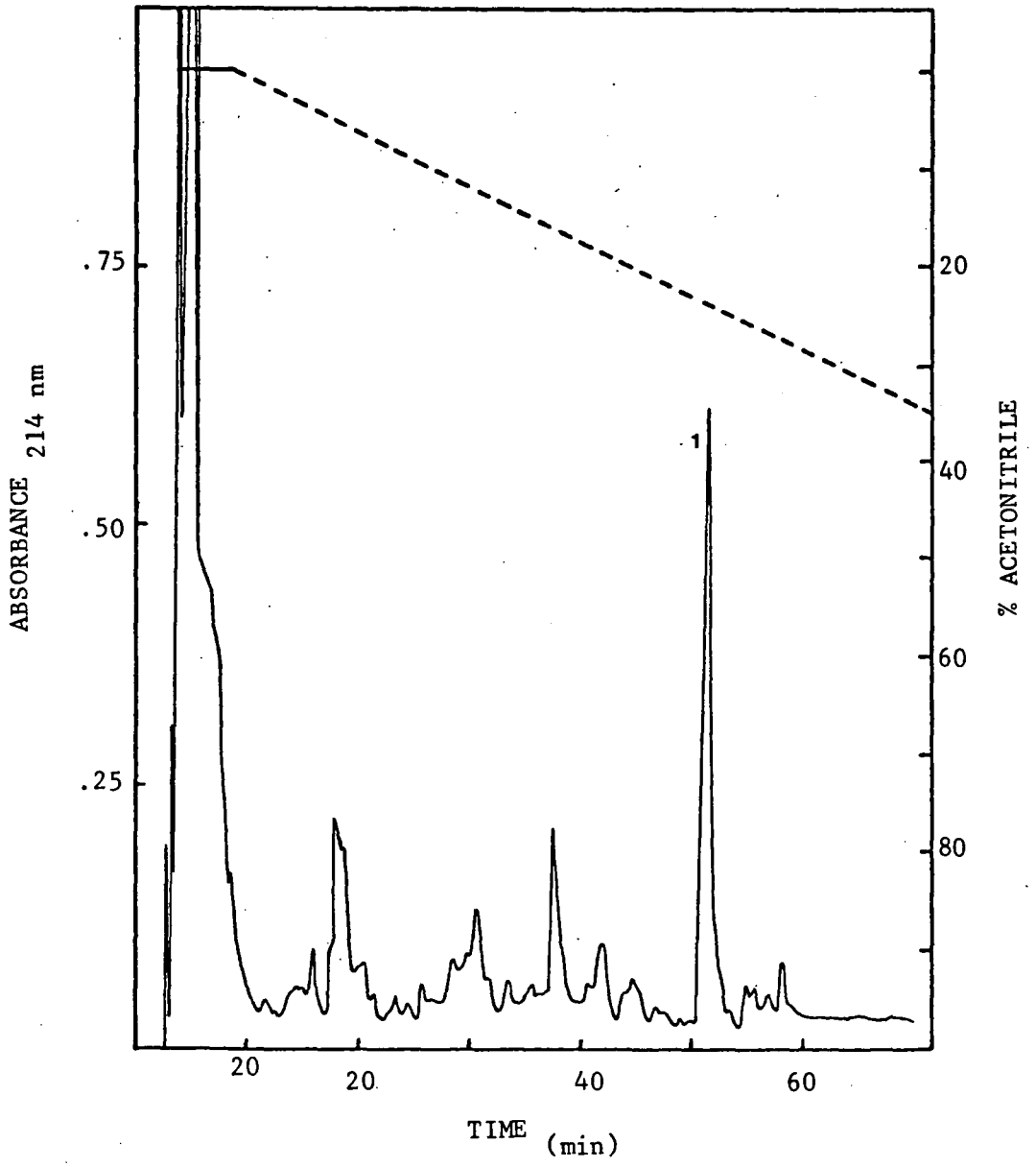


Figure 28 - RP-HPLC separation of fraction V-7 from the Staphylococcus aureus (strain V-8) protease digestion of the trypsin/α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.25 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table XI - Peptides obtained by the digestion of the trypsin/α-amylase inhibitor from the seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) with the Staphylococcus aureus (strain V-8) protease.

Peptide No. and Position in Sequence		Elution Time	Sequence
V2-1	(1-34)	71	Ser-Val-Gly-Thr-Ser-Cys-Ile-Pro-Gly-Met-Ala-Ile-Pro-His-Asn-Pro-His-Asn-Pro-Leu-Asp-Ser-Cys-Arg-Trp-Tyr-Val-Ser-Thr-Arg-Thr-Cys-Gly-Pro-Arg
V3-1	(82-100)	59	Leu-Ala-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe-Ala-Pro-Lys-Leu-Val-Thr-Glu
V4-1	(50-58)	44	Ala-Ile-Pro-Ala-Tyr-Cys-Arg-Cys-Glu
V5-1	(35-39)	22	Leu-Ala-Thr-Gln-Glu
V5-2	(66-75)	28	Gly-Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu
V5-3	(66-75)	32	Gly-Val-Val-Thr-Pro-Ser-Gly-Gln-His-Glu
V5-4	(40-49)	39	Met-Lys-Ala-Arg-Cys-Cys-Arg-Gln-Leu-Glu
V6-1	(66-75)	28	Gly-Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu
V6-2	(76-81)	38	Gly-Arg-Leu-Leu-Gln-Asp
V7-1	(59-65)	51	Ala-Val-Arg-Ile-Leu-Met-Asp

Peptides V5-2 and V6-1 (positions 66-75)

Gly-Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu

Peptide V5-3 (positions 66-75)

Gly-Val-Val-Thr-Pro-Ser-Gly-Gln-His-Glu

Peptides V5-2, V6-1 and V5-3 provide evidence for heterogeneity as position 70, as evidenced before with peptides CN-13, CN-14, CN-15, Th1-1 and Th1-2. As would be expected, peptides V5-2 and V5-3 have different elution times (Figure 26, Table XI).

Peptide V6-2 (positions 76-81)

Gly-Arg-Leu-Leu-Gln-Asp

The identities of Leu^{78} and Leu^{79} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide V3-1 (positions 82-100)

Leu-Ala-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg-Ala-Phe-Ala-Pro-Lys-Leu-Val-Thr-Glu

Peptide V3-1 provides additional evidence for heterogeneity at position 83. This position is occupied by proline in peptides Th1-6, Th1-7, Ch2-3 and by alanine in peptides CN-Ch4 and V3-1.

3.2.9. Sub-Digestion of Peptide CN-T11 with Staphylococcus aureus (Strain V-8) Protease

On examination of the results obtained from the digestion of the whole protein with the V-8 protease, it was apparent that some peptides expected to be derived from the C-terminal region of the

protein were not recovered. In order to obtain them, peptide CN-T11 obtained from the digestion of the C-terminal CNBr fragment (CN-13, CN-14 and CN-15 - see Figure 6 and Table VII) with trypsin (see Section 3.2.11.1 and Figure 31), was sub-digested with V-8 protease. The resulting fragments were separated through RP-HPLC (Figure 29). The sequences, elution times and positions in the sequence of the peptides obtained are shown in Table XII.

Peptide CN-T11-V2 (positions 97-100)

Leu-Val-Thr-Glu

The identity of Leu^{97} was ascertained by N-terminal analysis by the dansyl-chloride method.

Peptide CN-T11-V1 (positions 101-102)

Val-Glu

Peptide CN-T11-V4 (positions 103-122)

Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu

Peptides CN-T11-V3 and CN-T11-V5 (positions 101-122)

Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu

Apart from the cleavage of the bond Arg^{33}-Leu, the enzyme specificities observed were consistent with those previously reported (Allen, 1981). Cleavage at arginyl residues by the V-8 protease has been previously reported by Richardson et al. (1978) in the ox phospholipase A2.

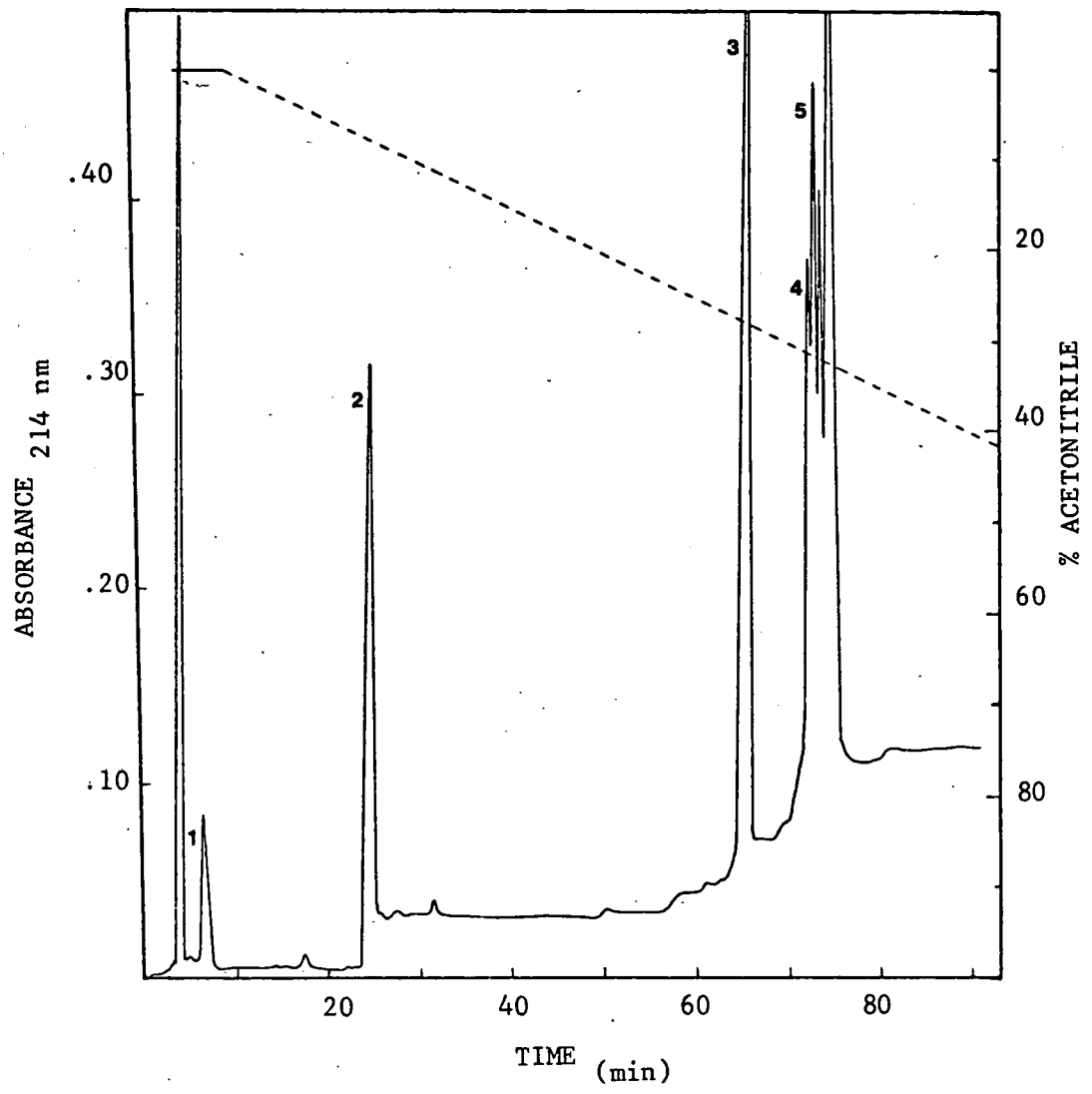


Figure 29 - RP-HPLC separation of peptides obtained by the digestion of the peptide CN-T11 (see Figure 31, Table XIV) with Staphylococcus aureus (strain V-8) protease. The peptides were separated on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table XII - Peptides obtained by the digestion of peptide CN-T11 with *Staphylococcus aureus* (strain V-8) protease

Peptide No. and Position in Sequence	Elution Time	Sequence
CN-T-V2 (97-100)	24	Leu-Val-Thr-Glu
CN-T-V3 (101-122)	65	Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu
CN-T-V1 (101-102)	6	Val-Glu
CN-T-V4 (103-122)	71	Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu
CN-T-V5 (101)122)	74	Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu

3.2.10. Proline-Specific Endoproteinase Digestion of the CNBr C-Terminal Fragment

1 mg of pooled fragments CN-13, CN-14 and CN-15 (see Figure 6, Table VII) was sub-digested with proline-specific endoproteinase, under the conditions described in Section 2.2.6.3.5. The peptides produced were separated by RP-HPLC (Figure 30). The sequences, elution times and positions in the sequence of the peptides obtained are presented in Table XIII.

Peptide CN-P5 (positions 65-76)

Asp-Gly-Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu-Gly

Peptide CN-P13 (positions 77-83)

Arg-Leu-Leu-Gln-Asp-Leu-Pro

The identities of Leu^{78},Leu^{79},Leu^{82} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by comparison with peptides Th1-6 and Th1-7. Peptides CN-P10 (Gln-Asp-Leu-Pro) and CN-P7 (Leu-Gln-Asp-Leu-Pro) were products of the cleavage of the peptide bonds Leu^{78}-Leu and Leu^{79}-Gln in peptide CN-P13.

Peptide CN-P1 (positions 84-86)

Gly-Cys-Pro

Peptide CN-P2 (positions 87-92)

Arg-Gln-Val-Gln-Arg-Ala

Peptide CN-P12 (positions 96-107)

Lys-Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr

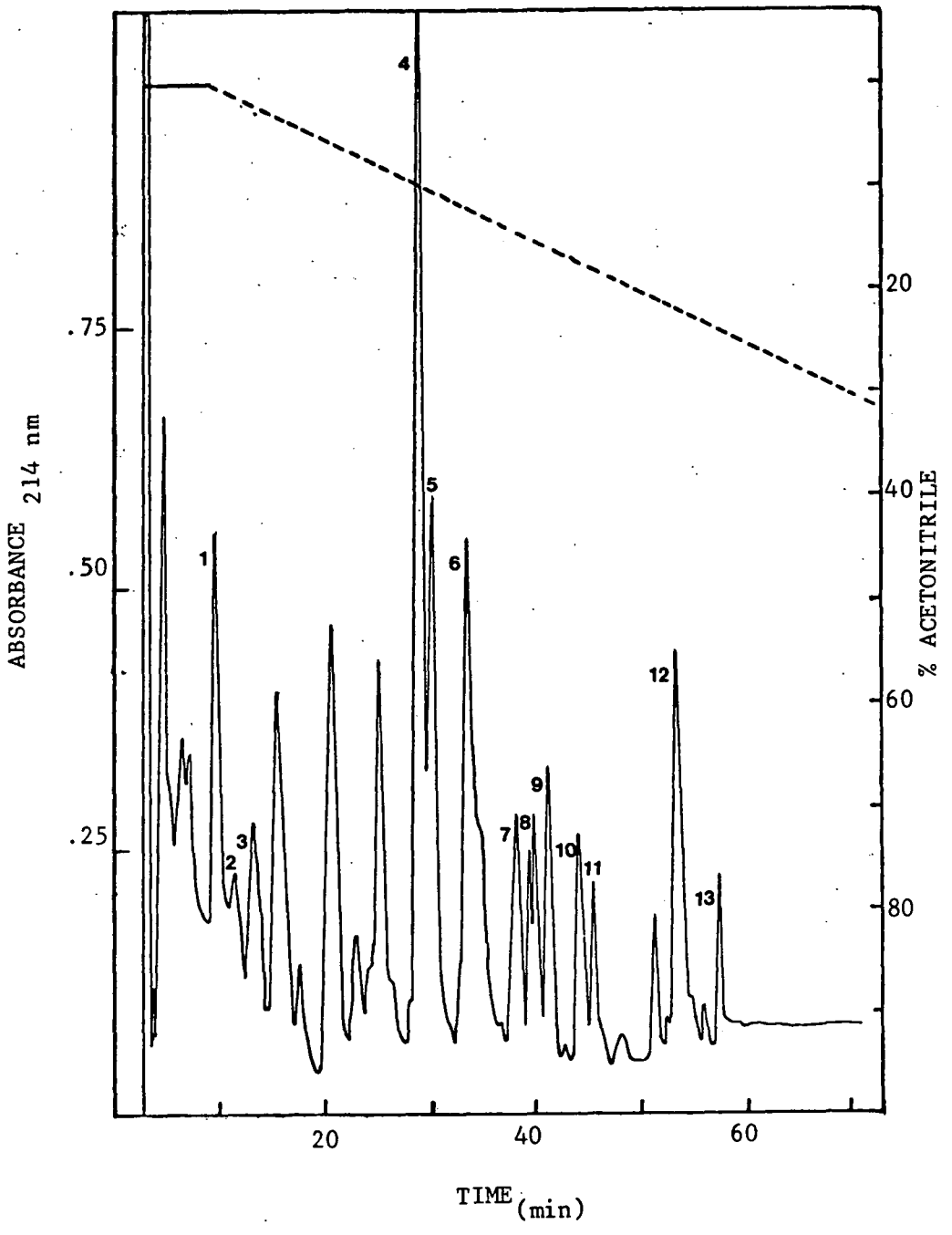


Figure 30 - RP-HPLC separation of peptides obtained by the digestion of the CNBr C-terminal peptide of the trypsin/α-amylase inhibitor from seeds of ragi (Indian finger millett, Eleusine coracana Gaertn.)with proline specific endoproteinase on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions corresponding to peptides which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table XIII- Peptides obtained by the cleavage with proline specific endoproteinase of the CNBr C-terminal fragment.

Peptide No. and Position in Sequence	Elution Time	Sequence
CN-P1 (84-86	11	Gly-Cys-Pro
CN-P2 (87-92)	12	Arg-Gln-Val-Gln-Arg-Ala
CN-P3 (98-102)	15	Ile-His-Gly-Gly-Pro
CN-P4 (96-99)	28	Lys-Leu-Val-Thr
CN-P5 (65-76)	30	Asp-Gly-Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu-Gly
CN-P6 (93-95)	32	Phe-Ala-Pro
CN-P7 (80-83)	34	Gln-Asp-Leu-Pro
CN-P8 (100-107)	39	Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr
CN-P9 (99-106)	41	Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr
CN-P10 (79-83)	43	Leu-Gln-Asp-Leu-Pro
CN-P11(113-116)	45	Phe-Cys-Leu-Ser
CN-P12 (96-107)	53	Lys-Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr
CN-P13 (77-83)	57	Arg-Leu-Leu-Gln-Asp-Leu-Pro

The identities of Leu^{97} and Leu^{105} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by comparison with peptides Th3-6 and Th4-11. Peptides CN-P4 (Lys-Leu-Val-Thr), CN-P8 (Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr) and CN-P9 (Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr) were products of the cleavage of the peptide bonds Thr^{99}-Glu and Val^{98}-Thr.

Peptide CN-P3 (positions 108-112)

Ile-His-Gly-Gly-Pro

The identity of Ile^{108} was ascertained by N-terminal analysis by the dansyl-chloride method, and by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide CN-P11 (positions 113-116)

Phe-Cys-Leu-Ser

The identity of Leu^{115} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method.

The specificity of the proline-specific endoproteinase was not as expected (Yoshimoto et al., 1980) in that in addition to cleavage of Pro-X and Ala-X peptide bonds, "anomalous" cleavages of Gly-X, Thr-X, Ile-X and Leu-X peptide bonds were also observed (Table XIII).

3.2.11. Trypsin Digestion

3.2.1.1. Trypsin Digestion of the C-Terminal CNBr Fragment

Fragments CN-13, CN-14 and CN-15 from the CNBr cleavage of the inhibitor (see Figure 6 and Table VII), were pooled and subjected

to digestion with trypsin under the conditions described in Section 2.2.6.3.1. The resulting fragments were separated by RP-HPLC (Figure 31). In Table XIV the sequences, elution times and positions in the sequence of the peptides obtained are presented.

Peptide CN-T5 (positions 47-56)

Gln-Leu-Glu-Ala-Ile-Pro-Ala-Tyr-Cys-Arg

The identity of Leu^{48} and Ile^{51} were ascertained by comparison with peptides Th1-8 and Th3-10 and by subjecting the peptide to four cycles of degradation by the "Dansyl-Edman" method.

Peptide CN-T1 (positions 57-61)

Cys-Glu-Ala-Val-Arg

Peptides CN-T5 and CN-T1 indicated that the CNBr C-terminal fragment used for this digest was contaminated with the fragments CN-8, CN-9, CN-10, CN-11 and CN-12 (which correspond to residues 41 to 64), or alternatively, indicates that the cleavage by CNBr of the peptide bond Met^{64}-Asp was not complete, and thus in the preparation used for the digestion, in addition to the CNBr C-terminal fragment (residues 65 to 122) there was an additional peptide corresponding to positions 41 to 122.

Peptide CN-T2 (positions 65 to 77)

Asp-Gly-Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu-Gly-Arg

Peptide CN-T4 (positions 78-87)

Leu-Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg

The identities of Leu^{78}, Leu^{79} and Leu^{82} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method.

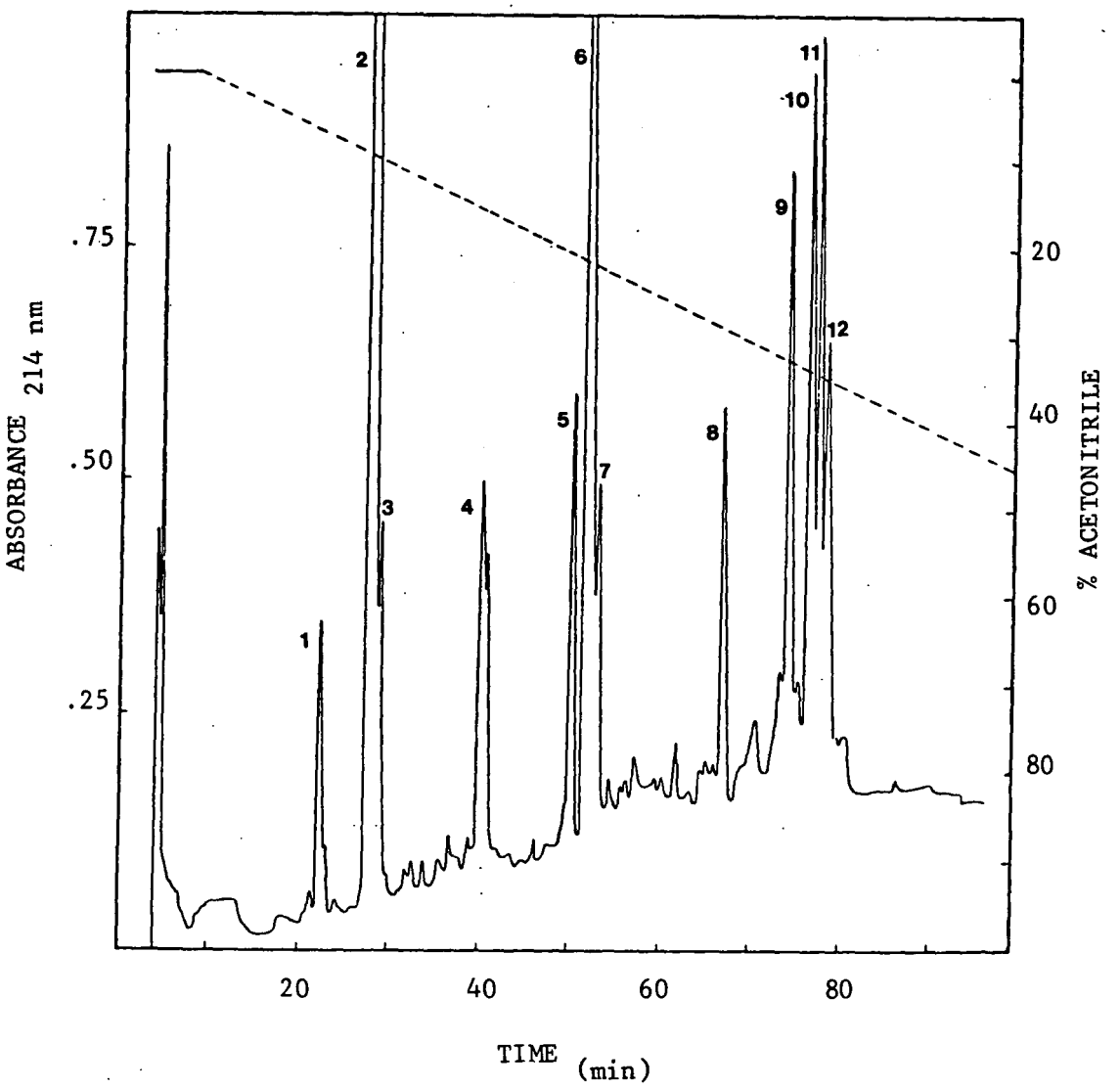


Figure 31 - RP-HPLC separation of the peptides obtained by the digestion of the CNBr C-terminal fragment of the trypsin/ α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) with trypsin. The peptides were separated on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate 0f 1.0 ml/min. Fractions corresponding to peptides which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table XIV - Peptides obtained by the digestion by trypsin of the CNBr C-terminal fragment of the trypsin/α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.)

Peptide No. and Position in Sequence		Elution Time	Sequence
CN-T1	(57-61)	22	Cys-Glu-Ala-Val-Arg
CN-T2	(75-77)	27	Asp-Gly-Val-Val-Thr-Ser-Ser-Gly-Gln-His-Glu-Gly-Arg
CN-T3	(92-96)	30	Ala-Phe-Ala-Pro-Lys
CN-T4	(78-87)	40	Leu-Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg
CN-T5	(47-56)	50	Gln-Leu-Glu-Ala-Ile-Pro-Ala-Tyr-Cys-Arg
CN-T6	(78-91)	52	Leu-Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln-Arg
CN-T7	(114-122)	54	Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu
CN-T8	(97-113)	67	Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe
CN-T9	(97-113)	73	Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe
CN-T10	(97-122)	76	Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu
CN-T11	(97-122)	78	Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu
CN-T12	(97-122)	81	Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr Ile-His-Gly-Gly-Pro-Phe-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu Gly-Ala-Gly-Glu

Peptide CN-T6 (positions 78-91)

Leu-Leu-Gln-Asp-Leu-Pro-Gly-Cys-Pro-Arg-Gln-Val-Gln-Arg

The identity of Leu^{78}, Leu^{79} and Leu^{82} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide CN-T3 (positions 92-96)

Ala-Phe-Ala-Pro-Lys

Peptides CN-T8 and CN-T9 (positions 97-114)

Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe

The identity of Leu^{97} was ascertained by N-terminal analysis by the dansyl-chloride method, and the identity of Leu^{106} and Ile^{109} were ascertained by comparison with peptides Th3-6, Th5-1 and Th4-7.

Peptides CN-T10, CN-T11 and CN-T12 (positions 97 to 122)

Leu-Val-Thr-Glu-Val-Glu-Cys-Asn-Leu-Ala-Thr-Ile-His-Gly-Gly-Pro-Phe-Cys-Leu-Ser-Leu-Leu-Gly-Ala-Gly-Glu

Peptide CN-T11 was sub-digested with Staphylococcus aureus (Strain V-8) protease. The results of this sub-digestion were described in Section 3.2.9.

Peptide CN-T7 (Positions 114-122)

The identities of Leu^{115}, Leu^{117} and Leu^{118} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide CN-T7 arose as the result of an anomalous tryptic cleavage at the peptide bond Phe^{113}-Cys. However, as the height of the peak indicates (Figure 31) this cleavage seems to have been only partial .

3.2.1.2. Trypsin Digestion of Catalytic Cleavage Fragment 1

In order to obtain a better overlapping between positions 21 to 34, the catalytic cleavage fragment 1 (CC1), which was shown to be the N-terminal fragment (see Table XVI) was subjected to hydrolysis with trypsin, under the conditions described in Section 2.2.6.3.1. The way in which this fragment was obtained is fully described in Section 3.2.12. The fragments obtained were directly separated by RP-HPLC (Figure 32). Table XV shows the sequences, elution times and positions in the sequence of the peptides obtained.

Peptides CC1-T3 and CC1-T4 (positions 1-21)

Ser-Val-Gly-Thr-Ser-Cys-Ile-Pro-Gly-Met-Ala-Ile-Pro-His-Asn-Pro-Leu-Asp-Ser-Cys-Arg

The identities of Ile^{7} and Ile^{12} were ascertained by comparison with peptides Th3-7 and Th3-8. The identity of Leu^{17} was ascertained by comparison with peptide Th5-8.

Peptide CC1-T2 (positions 22-27)

Trp-Tyr-Val-Ser-Thr-Arg

Peptide CC1-T1 (positions 28-34)

Thr-Cys-Gly-Val-Gly-Pro-Arg

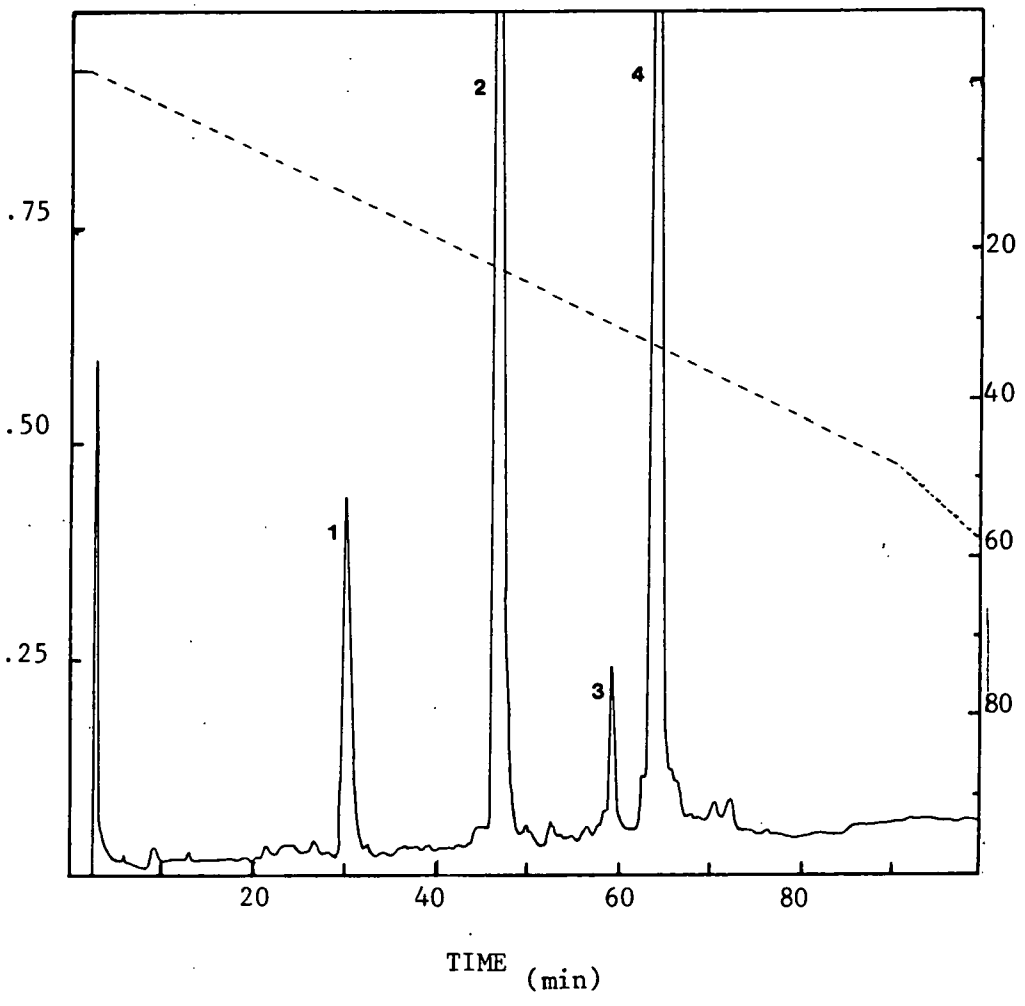


Figure 32 - RP-HPLC separation of peptides obtained by the digestion of catalytic cleavage fragment 1 (CC1) with trypsin.. The separation was performed on a C-18 u-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table XV - Peptides obtained by the cleavage with trypsin of the catalytic cleavage fragment 1 (CC1) of the trypsin/α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.)

Peptide No. and Position in Sequence		Elution Time	Sequence
CC1-T1	(28-34)	30	Thr-Cys-Gly-Val-Gly-Pro-Arg
CC1-T2	(22-27)	45	Trp-Tyr-Val-Ser-Thr-Arg
CC1-T3	(1-21)	59	Ser-Val-Gly-Thr-Ser-Cys-Ile-Pro-Gly-Met-Ala-Ile-Pro-His-Asn-Pro-Leu-Asp-Ser-Cys-Arg
CC1-T4	(1-21)	62	Ser-Val-Gly-Thr-Ser-Cys-Ile-Pro-Gly-Met-Ala-Ile-Pro-His-Asn-Pro-Leu-Asp-Ser-Cys-Arg

3.2.12. Identification of the Reactive (Trypsin Inhibitory) Sites of the Trypsin/α-Amylase Inhibitor from Ragi (Eleusine coracana Gaertn.)

The identification of the reactive (trypsin inhibitory) sites of the ragi trypsin/α-amylase inhibitor was performed by incubating 5 mg of the native inhibitor with a catalytic amount of trypsin (50 µg) (Ozawa and Laskowski, 1966; Richardson, 1974) for 20 hours at pH 2.5 The separation of the resulting fragments by RP-HPLC (Figure 33) was preceded by the reduction and carboxymethylation of the modified inhibitor. Each fragment collected was subjected to 8 sequencing cycles by the DABITC-PITC double coupling method (Table XVI) and the position in the sequence of each of them was unambiguously determined.

The three fragments obtained, which were identified as 1-34, 35-91 and 92-122, suggest the existence of two reactive sites in the inhibitor molecule, both having arginine at position P_1: Arg^{34}-Leu and Arg^{91}-Ala. The finding that the reactive site(s) involved an Arg-peptide bond is in agreement with earlier studies by Shivaraj and Pattabiraman (1981), which showed that the chemical modification of arginine residues by treating the inhibitor with cyclohexane-1,2-dione led to the almost complete loss of inhibitory activity. The Arg^{34}-Leu bond corresponds exactly to the position of the reactive site in the homologous barley trypsin inhibitor (Odani et al.,1983a) (Figure35) In addition both peptide bonds (Arg^{34}-Leu and Arg^{91}-Ala) appear to obey the proposed tentative rules for reactive sites (Laskowski and Kato, 1980).

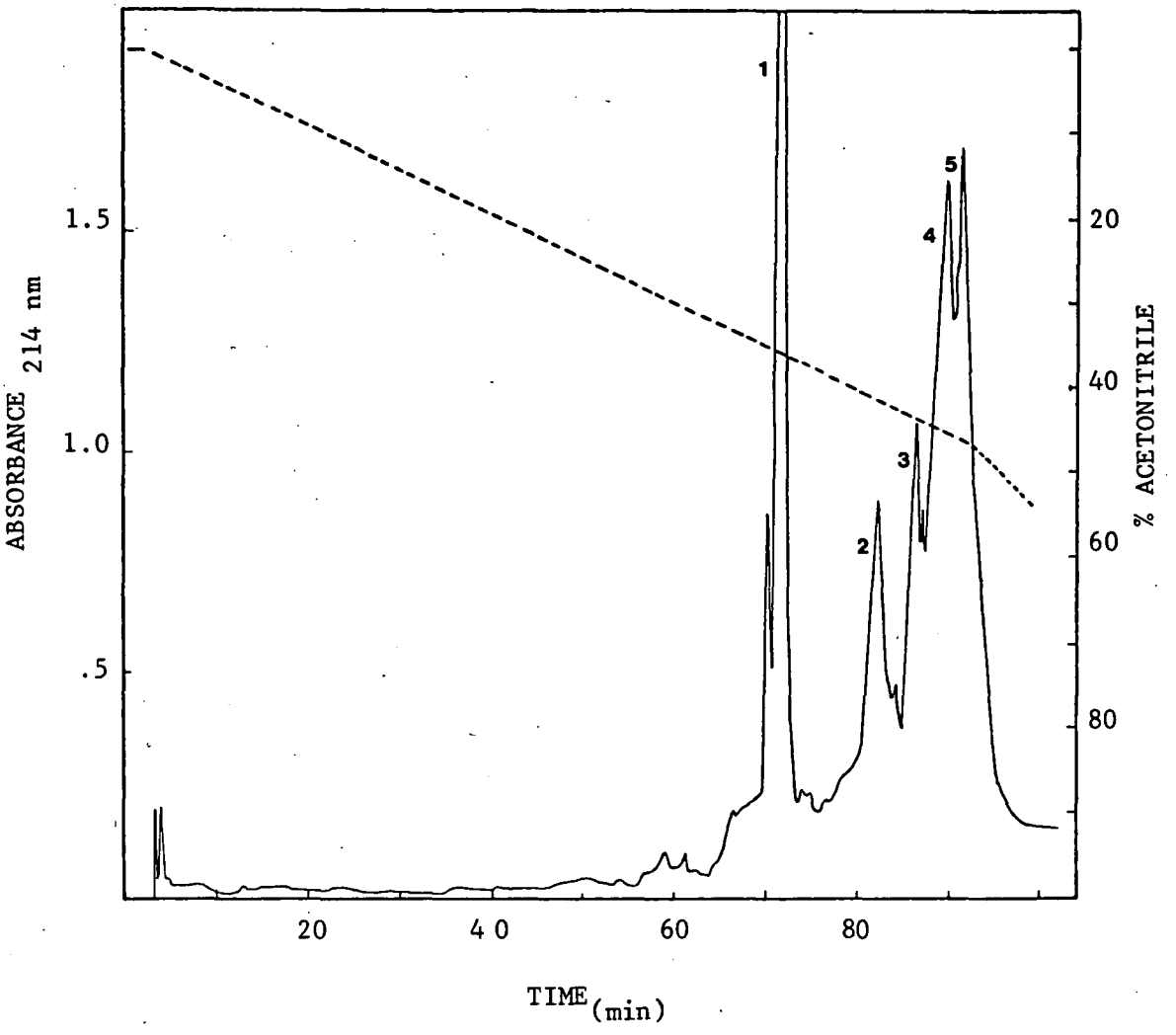


Figure 33 - RP-HPLC separation of peptides obtained by the cleavage of the native trypsin/ α-amylase inhibitor with trypsin (see text for details) on a C-18 u-Bondapak column (.05 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min.The N-terminal sequence of the peptides obtained is given in Table XVI. Dotted line indicates concentration of acetonitrile.

Table XVI - Peptides obtained by the catalytic cleavage of thr trypsin/ α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) with trypsin.

Peptide No. and Position in Sequence	Elution Time	Sequence
CC-1 (1-34)	72	Ser Val-Gly-Thr-Ser-Cys-Ile-Pro-...
CC-2 (92-122)	83	Ala-Phe-Ala-Pro-Lys-Leu-Val-Thr-...
CC-3 (92-122)	87	Ala-Phe-Ala-Pro-Lys-Leu-Val-Thr-...
CC-4 (35-91)	90	Leu-Ala-Thr-Gln-Glu-Met-Lys-Ala-...
CC-5 (35-91)	92	Leu-Ala-Thr-Gln-Glu-Met-Lys-Ala-...

Based on the elution of the enzyme-inhibitor complex from a Biogel P-200 column, Shivaraj and Pattabiraman (1981) suggested the formation of a 1:1 molar complex between the trypsin/α-amylase inhibitor and trypsin. Also, based on the same sort of experiment, they suggested that the inhibitor can bind simultaneously to trypsin and α-amylase, thus implying that the α-amylase inhibitory site is independent of the trypsin inhibitory site. Following the determination of the existence of two reactive (trypsin inhibitory) sites in the ragi trypsin/α-amylase inhibitor, attempts were made to determine through kinetic experiments, the molar ratio of the inhibitor/trypsin complex. As shown in Figure 34, the suggestion of Shivaraj and Pattabiraman (1981) of a 1:1 molar complex is confirmed.

The formation of a 1:1 molar complex between the trypsin/α-amylase inhibitor and trypsin is in apparent contradiction with the possible existence of two reactive (trypsin inhibitory) sites in the inhibitor molecule, as it would be expected that each inhibitory site would bind, simultaneously and independently, to one molecule of trypsin. The existence of more than one reactive site in the same inhibitor molecule is quite common among protein inhibitors of proteinases (Richardson, 1977, 1980, 1981; Laskowski and Kato, 1980; Ryan, 1981), and in all cases so far studied each reactive site can bind independently and simultaneously, to the enzyme they inhibit. Thus, the ragi trypsin/α-amylase inhibitor seems to be the first example of an inhibitor with two distinct but overlapping reactive sites, where possibly steric hindrance prevents association with two molecules of trypsin.

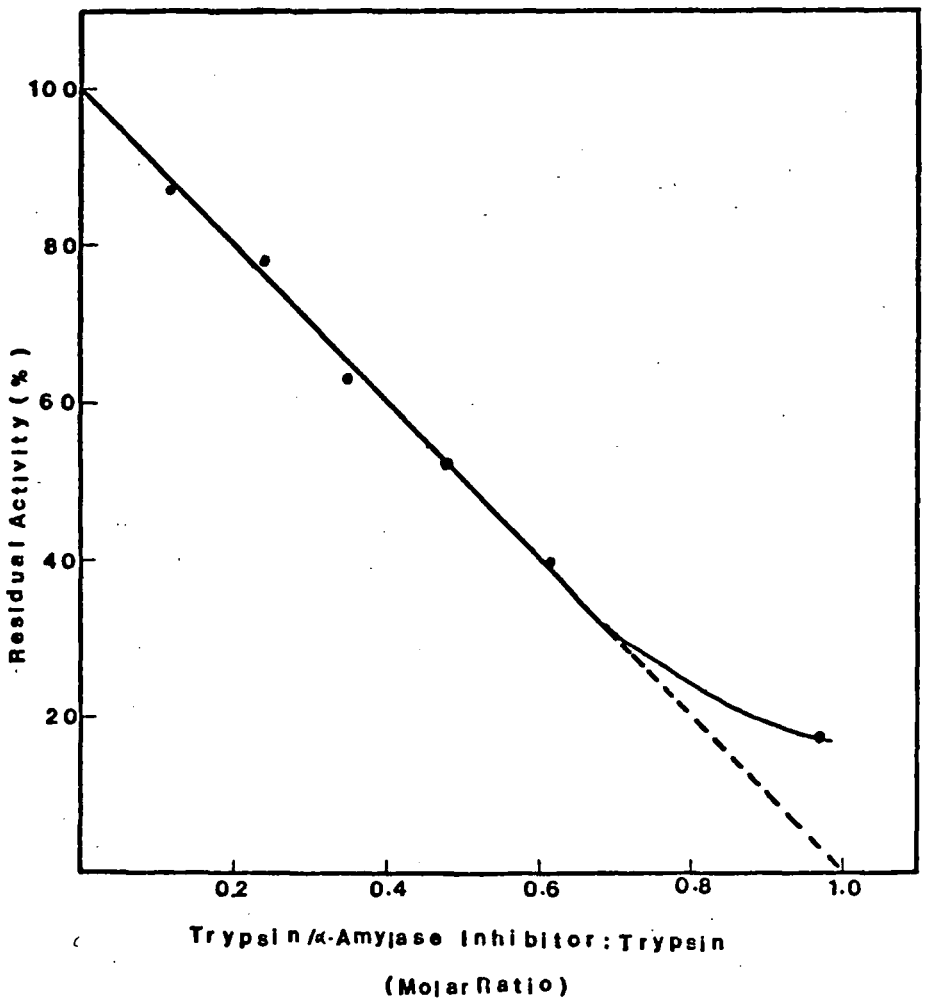


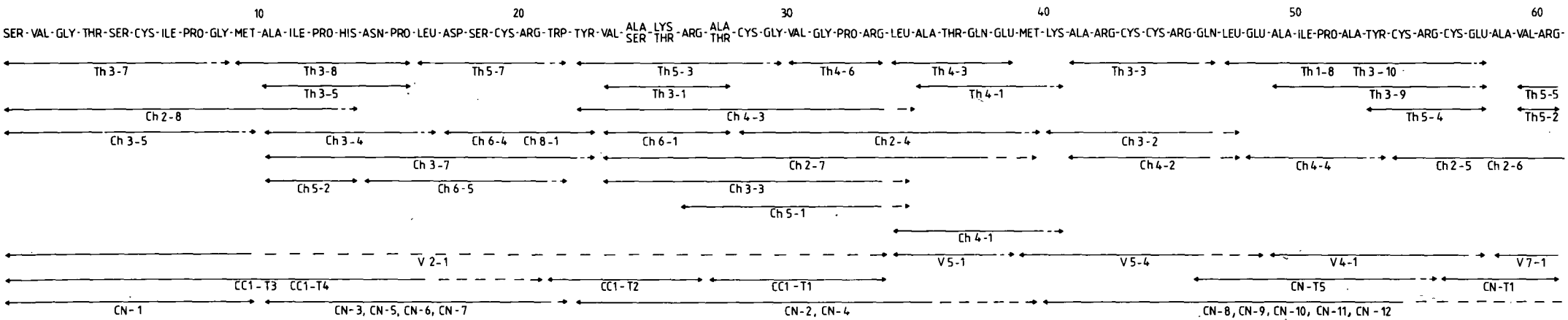
Figure 34 - Inhibition of trypsin by ragi trypsin/α-amylase inhibitor. A fixed amount of bovine trypsin was mixed with increasing amounts of the inhibitor, and the residual trypsin activity was measured using BAPNA as substrate (see text for details).

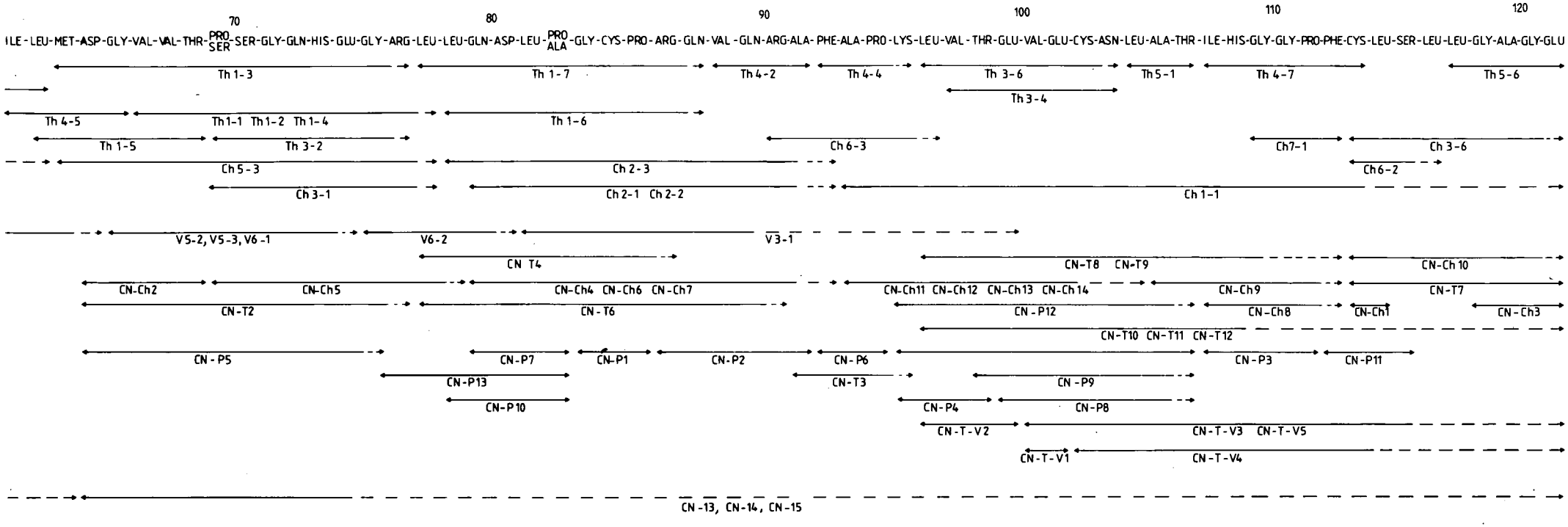
The complete amino acid sequence of the ragi trypsin/α-amylase inhibitor is shown in Figure 35. The protein consists of 122 amino acids which corresponds to a molecular weight of 13300, which is in good agreement with the molecular weight calculated from SDS-polyacrylamide gel electrophoresis. The sequence presented in Figure 35 is also in good agreement with the amino acid composition of the inhibitor (see Table VI).

The inhibitor was also shown to be heterogeneous at positions 25, 26, 28, 70 and 83. In most cases the variant forms were isolated and sequenced separately, thanks to the resolving power of the RP-HPLC system utilised, which is very sensitive to slight changes in the hydrophobicity character of the peptides being separated; the only exceptions were the large peptides with minor (single) changes (e.g. CN-13, CN-14 and CN-15) where such peptides were not resolved. Regarding the heterogeneities detected in the sequence of the ragi trypsin/α-amylase inhibitor, it is worth noting that the replacements Ala/Ser, Ala/Thr and Pro/Ala are the most frequently observed amino acid replacements in proteins (Doolitle, 1979).

All the amino acids in the sequence were unambiguously identified. As it is difficult to identify leucine and isoleucine by the DABITC-PITC double coupling method alone, these residues were also identified by semi-quantitative amino acid analysis, by N-terminal analysis or by subjecting the relevant peptide to a few degradative cycles of the "Dansyl-Edman" method, or to a combination of these three methods. The specificity of cleavage by chymotrypsin was also helpful in confirming the identity of certain leucine residues. Also the identification of aspartic acid/asparagine and glutamic acid/glutamine residues was unambiguous. Although during the DABITC-PITC double

Figure 35 - The complete amino acid sequence of the bifunctional trypsin/ α-amylase inhibitor. T = tryptic peptides, C = chymotryptic peptides, Th = thermolysin peptides, V = Staphylococcus aureus (strain V-8) peptides, CN = peptides resulting from digestion with CNBr, CC = catalytic cleavage peptides. (⟷) peptide sequenced by the DABITC-PITC double coupling method and/or the dansyl-Edman method; (⟷ — ⟶) regions of peptides which were not sequenced or gave unsatisfactory results.




70
80
90
100
110
120
ILE-LEU-MET-ASP-GLY-VAL-VAL-THR-PRO SER-SER-GLY-GLN-HIS-GLU-GLY-ARG-LEU-LEU-GLN-ASP-LEU-PRO ALA-GLY-CYS-PRO-ARG-GLN-VAL-GLN-ARG-ALA-PHE-ALA-PRO-LYS-LEU-VAL-THR-GLU-VAL-GLU-CYS-ASN-LEU-ALA-THR-ILE-HIS-GLY-GLY-PRO-PHE-CYS-LEU-SER-LEU-LEU-GLY-ALA-GLY-GLU
Th 1-3
Th 1-7
Th 4-2
Th 4-4
Th 3-6
Th 5-1
Th 4-7
Th 5-6
Th 3-4
Th 4-5
Th1-1 Th1-2 Th1-4
Th 1-6
Th 1-5
Th 3-2
Ch 6-3
Ch7-1
Ch 3-6
Ch 5-3
Ch 2-3
Ch6-2
Ch 3-1
Ch 2-1 Ch 2-2
Ch 1-1
V5-2, V5-3, V6-1
V6-2
V3-1
CN T4
CN-T8 CN-T9
CN-Ch 10
CN-Ch2
CN-Ch5
CN-Ch4 CN-Ch6 CN-Ch7
CN-Ch11 CN-Ch12 CN-Ch13 CN-Ch14
CN-Ch9
CN-T7
CN-T2
CN-T6
CN-P12
CN-Ch8
CN-Ch1
CN-Ch3
CN-T10 CN-T11 CN-T12
CN-P5
CN-P7
CN-P1
CN-P2
CN-P6
CN-P3
CN-P11
CN-P13
CN-T3
CN-P9
CN-P10
CN-P4
CN-P8
CN-T-V2
CN-T-V3 CN-T-V5
CN-T-V1
CN-T-V4
CN-13, CN-14, CN-15

coupling method it is known that deamidation of asparagine and glutamine residues can occur (Chang et al., 1978) this deamidation is only partial and the intensity of the spots corresponding to asparagine and glutamine residues on the TLC sheet are more conspicuous than the one of the deamidated forms. Another indirect confirmation ot the correct assignement of aspartic acid, asparagine, glutamic acid and glutamine residues, was provided by the fragments obtained from the hydrolysis of the inhibitor with the Staphylococcus aureus (strain V-8) protease: All of the peptide bonds involving aspartic acid and glutamic acid were cleaved, with the exception of Asp^{18}-Ser. The failure of the V-8 protease to cleave this peptide bond can be explained by the short time of incubation, as Glu-X peptide bonds are more readily cleaved than the Asp-X ones (Croft, 1981).

The many digestions and sub-digestion of the protein described above, although in some cases providing repetitive and redundant data, has assured that the fragments sequenced were correctly aligned. A further confirmation of the validity of the alignment shown in Figure 35 may be drawn from a comparison of the primary structures of the ragi trypsin/α-amylase inhibitor and the barley trypsin inhibitor (Odani et al., 1983a), a closely homologous protein (Figure 36).

3.2.13. Sequence Comparison

The search for proteins structurally related to the ragi trypsin/α-amylase inhibitor was conducted using the computer program described in Section 2.2.7. The value of the parameters d, u and v, as well as the distance measure and z (see Section 2.2.7. for

definitions) for each alignment are given in the legend of the relevant figures.

The alignment of the amino acid sequences of the ragi trypsin/α-amylase inhibitor and the barley trypsin inhibitor (Odani et al., 1983a), is shown in Figure 36. The homology between these two proteins is self evident. Identical residues are found at 66 out of the total of 122 (ragi) and 121 (barley) positions with all 10 cysteine residues being conserved. Also to be noted is the strong homology around one of the reactive sites of the ragi trypsin/α-amylase inhibitor (Arg^{34}-Leu) and the reactive site of the barley trypsin inhibitor. Such strong homology clearly indicates that these two proteins are related through a process of divergent evolution.

The ragi trypsin/α-amylase inhibitor also seems to have similarities with the maize trypsin inhibitor. The maize inhibitor was sequenced by Hochtrasser et al. (1970) and the sequence proposed has sixty five amino acids. Based on the fact that the total number of residues found in the tryptic digestion of the inhibitor represented only one third of the total amino acids determined from the molecular weight and percentage amino acid composition, they suggested that the native maize inhibitor was a polymer of a subunit which consisted of 65 amino acids residues. In Figure 37a the primary structure of the maize trypsin inhibitor is shown and in Figure 37b peptides derived from the treatment of the maize inhibitor with trypsin and chymotrypsin are aligned with segments of the ragi trypsin/α-amylase inhibitor. As can be seen, for most of the tryptic and chymotryptic peptides from the maize inhibitor, matching peptides can be found in the ragi trypsin/α-amylase inhibitor. It should be noted that the maize trypsin inhibitor peptides which match with the ragi trypsin/α-amylase

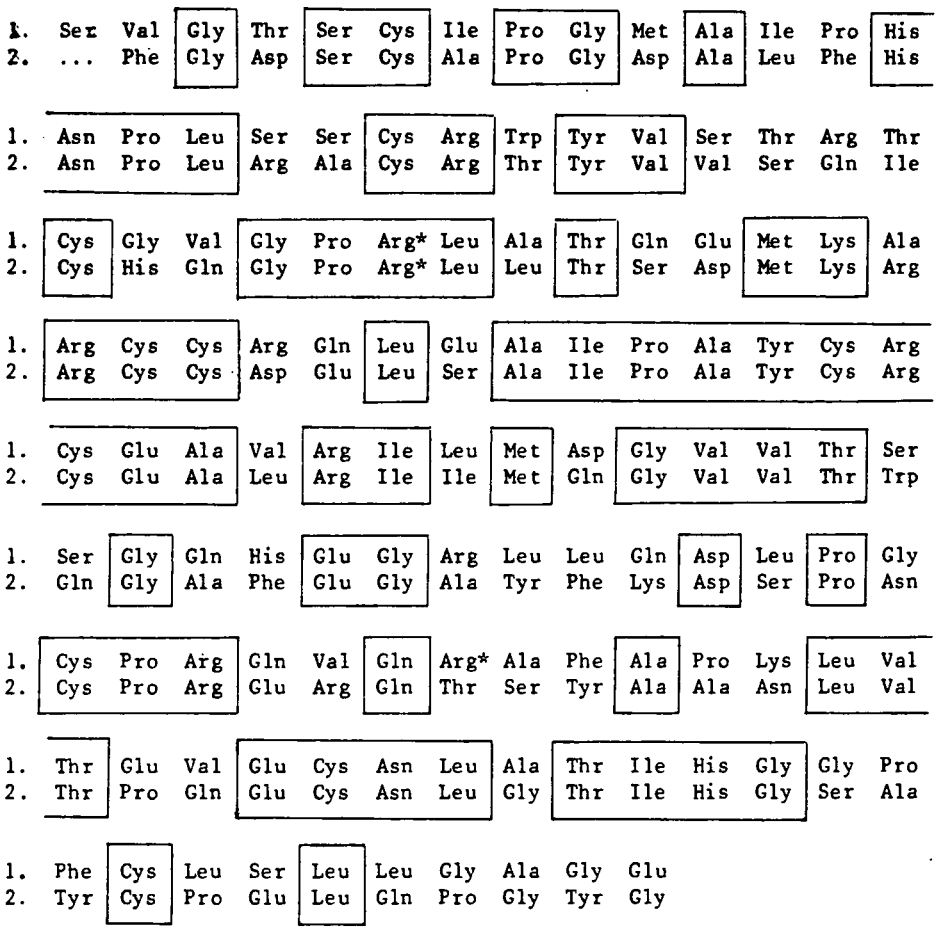


Figure 36 - Comparison of the primary structures of the ragi trypsin/ α-amylase inhibitor (1) and the barley trypsin inhibitor (2) (Odani et al., 1983a). The identical residues are enclosed in boxes; (...) indicates gap; (*) indicates trypsin inhibitory reactive sites. Parameter values: u=10, v=12, d=10. Distance measure=572. Similarity coefficient= 0.530; z= -21.453 (significant homology).

Ser Ala Gly Thr Ser Cys Val Pro Pro (Ser,Gly,Cys,Pro,His,Asx) Ala Ile
T-I

Leu Arg Thr Gly Ile Pro Gly Arg Leu Pro Pro Leu Glx Lys Thr Cys Gly Ile
T-II CT-IV CT-III

Gly Pro Arg Gln Val Gln Arg Leu Gln Asx Leu Pro Cys Pro Gly Arg Arg
CT-V CT-VI CT-I

Gln Leu Ala Asx Met Ile Ala Tyr Cys Pro Arg Cys Arg
CT-VII CT-II

Figure 37a - The amino acid sequence of the trypsin inhibitor from maize seeds (Hochstrasser et al.,1970). The nomenclature of the peptides is that used by the authors. T denotes a tryptic peptide and CT denotes a chymotryptic peptide obtained from the cleavage of a tryptic peptide.

1. Ser Val Gly Thr Ser Cys Ile Pro Gly Met Ala Ile Pro His Asn Pro Leu Ser
2. Ser Ala Gly Thr Ser Cys Val Pro Pro(Ser,Gly,Cys,Pro,His,Asx Ala Ile Leu

1. Ser Cys Arg Trp Tyr Val Ser Thr Arg Thr Cys Gly Val Gly Pro Arg Leu Ala
2. Arg Thr Cys Gly Ile Gly Pro Arg

1. Arg Cys Cys Arg Gln Leu Glu Ala Ile Pro ... Ala Tyr Cys ... Arg Cys Glu
2. Gln Leu ... Ala Asx Met Ile Ala Tyr Cys Pro Arg

1. Ala Val Arg Ile Leu Met Asp Gly Val Val Thr Ser Ser Gly Gln His Glu Gly

1. Arg Leu Leu Gln Asp Leu Pro Gly Cys Pro ... Arg Gln Val Gln Arg Ala Phe
2. Leu Gln Asx Leu Pro ... Cys Pro Gly Arg Gln Val Glx Arg

1. Ala Pro Lys Leu Val Thr Glu Val Glu Cys Asn Leu Ala Thr Ile His Gly Gly

1. Pro Phe Cys Leu Ser Leu Leu Gly Ala Gly Glu

Figure 37b - Comparison of the amino acid sequence of the ragi trypsin/α-amylase inhibitor (1) with fragments of the maize trypsin inhibitor (2) (Hochtrasser et al., 1970). Identical residues are enclosed in boxes. (...) indicates gap.

inhibitor in Figure 37b do not occur in the sequence order suggested by Hochtrasser et al (1970) but are scattered around the inhibitor molecule, and also not concentrated in any particular region. Inspection of the alignment of the ragi inhibitor sequence and the maize inhibitor peptides and consideration of the specificities of trypsin and chymotrypsin, strongly suggests that Hochtrasser et al (1970) did not recover all the possible tryptic and chymotryptic peptides from their digests. This implies that the maize inhibitor sequence in fact contains more than the 65 amino acids suggested by these workers. Also, as a matter of fact, in their report Hochtrasser et al (1970) do not present experimental evidence to substantiate the overlappings proposed. Thus, although the apparently fragmentary nature of the sequence data for the maize trypsin inhibitor does not permit a very meaningful comparison, the limited information available does suggest that homology exists with thr ragi trypsin/α-amylase inhibitor.

The alignment of positions 70 to 122 of the ragi trypsin/α-amylase inhibitor with the pancreatic secretory trypsin inhibitor (Kazal) and the alignment of positions 69 to 121 of the barley trypsin inhibitor (Odani et al., 1983b) with the pancreatic secretory trypsin inhibitor are presented in Figures 38a and 38b, respectively. Only the similarity in the second alignment is sufficient to be called homology. The values of z of both alignments, which determine the level of statistical significance of an alignment, have borderline values, but the invariant positions in both alignment are almost the same.

Ikenaka et al (1974) have shown that the region around the reactive site of the pancreatic secretory trypsin inhibitor

1. Asn Ile Leu Gly Arg Glu Ala Lys Cys Thr Asn Glu Val Asn Gly Cys Pro Arg
2. Trp Gln Gly Ala Phe Glu Gly Ala Tyr Phe Lys Asp Ser Pro Asn Cys Pro Arg

1. Ile Tyr Asn Pro Val Cys Gly Thr Asp Gly Val Thr Tyr Ser Asn Glu Cys Leu
2. Glu Arg Gln Thr Ser Tyr Ala Ala Asn Leu Val Thr Pro Gln ... Glu Cys Asn

1. Leu Cys Met Glu Asn Lys Glu Arg Gln Thr Pro Val Leu Ile Gln Lys Ser Gly
2. Leu Gly Thr Ile His Gly Ser Ala Tyr Cys Pro Glu Leu Gln Pro Gly Tyr Gly

1. Pro Cys
2. ... ...

Figure 38a - Comparison of the amino acid sequence of the pancreatic secretory trypsin inhibitor (Kazal) (1) (Greene and Bartelt, 1969) with segment 69 to 121 of the barley trypsin inhibitor (Odani et al., 1983a). Identical residues are enclosed in boxes. (...) indicates gap. Parameters values: u=10, v=12, d=10. Minimum distance measure=464. Similarity coefficient=0.189. z=-3.402 (significant homology).

1. Asn Ile Leu Gly Arg Glu Ala Lys Cys Thr Asn Glu Val Asn Gly Cys Pro Arg
2. Ser Ser Gly Gln His Glu Gly Arg Leu Leu Gln Asp Leu Pro Gly Cys Pro Arg

1. Ile Tyr Asn Pro Val Cys Gly Thr Asp Gly Val Thr Tyr Ser Asn Glu Cys Leu
2. Gln Val Gln Arg Ala Phe Ala Pro Lys Leu Val Thr Glu Val ... Glu Cys Asn

1. Leu Cys Met Glu Asn Lys Glu Arg Gln Thr Pro Val Leu Ile Gln Lys Ser Gly
2. Leu Ala Thr Ile His Gly Gly Pro Phe Cys Leu Ser Leu Leu Gly Ala ... Gly

1. Pro Cys
2. Glu ...

Figure 38b - Comparison of the amino acid sequence of the pancreatic secretory trypsin inhibitor (Kazal) (1) (Greene and Bartelt, 1969) with segment 70 to 122 of the ragi trypsin/α-amylase inhibitor (2). Identical residues are enclosed in box-s; (...) indicates gap. Parameters values: u=10, v=12, d=10. Minimum distance measure=474. Similarity coefficient=0.171; z= -2.627 (insignificant homology).

(Kazal) and the Streptomyces subtilisin inhibitor are similar, and Laskowski and Kato (1980) pointed out that this similarity is sufficient to be called homology. However, the remainder of the molecules appear to be unrelated. The similarities around the reactive site of both inhibitors were taken as an indication that the pancreatic secretory trypsin inhibitor and the Streptomyces subtilisin inhibitor families may have a distant divergent relationship (Laskowski and Kato, 1980). However, as the similarities between these inhibitors seem to be restricted to the regions around the reactive sites and the subsidiary enzyme-inhibitor binding site, and no details on the three dimensional structure of the inhibitors are known, a decision on the the precise evolutionary relationship between these inhibitor families cannot be made (Laskowski and Kato, 1980).

Subsequent to our report (Campos and Richardson, 1983) of the amino acid sequence of the ragi trypsin/α-amylase inhibitor and its homology with the wheat α-amylase inhibitor (0.28) and the barley trypsin inhibitor, Odani et al. (1983b) have discussed the homology between the ragi trypsin/α-amylase inhibitor, the barley trypsin inhibitor and the pancreatic secretory trypsin inhibitor. They suggested a tandem two domain structure for the ragi trypsin/α-amylase inhibitor and the barley trypsin inhibitor as follows:

H-(proteinase inhibitory domain)-(α-amylase inhibitory domain)-OH ,

and also suggested that the regions in the ragi trypsin/α-amylase inhibitor and barley trypsin inhibitor homologous to the pancreatic secretory trypsin inhibitor constituted the α-amylase inhibitory domain. Their suggestion imply an extraordinary divergent evolution of these inhibitors from a common ancestor. However, whilst the

regions around the reactive sites of the pancreatic secretory trypsin inhibitor and the Streptomyces subtilisin inhibitor are homologous, comparison of the amino acid sequence of the Streptomyces subtilisin inhibitor with the ragi trypsin/α-amylase inhibitor and barley trypsin inhibitor does not reveal any homology and furthermore, the invariant positions when the sequences of the pancreatic secretory trypsin inhibitor and Streptomyces subtilisin inhibitor are compared are different from the ones which are observed when the pancreatic secretory trypsin inhibitor is compared with the ragi trypsin/α-amylase inhibitor or with the barley trypsin inhibitor. It seems thus that taking into consideration sequence similarities, a divergent relationship between them seems to be unlikely, although a very distant divergent relationship between the pancreatic secretory trypsin inhibitor and the ragi trypsin/α-amylase inhibitor or between the pancreatic secretory trypsin inhibitor and the Streptomyces subtilisin inhibitor cannot be discarded.

The list of proteins which show a significant degree of similarity with the ragi trypsin/α-amylase inhibitor also includes the 0.53 (Maeda et al., 1983) and 0.28 (Kashlan and Richardson, 1981) α-amylase inhibitors from wheat. The two α-amylase inhibitors are highly homologous (Maeda et al., 1983), and when both of them are compared with the ragi trypsin/α-amylase inhibitor sequence (Figures 39a and 39b) the similarities observed are sufficient to be called homology. Also of interest is the fact that the invariant positions of both alignments are almost the same and the invariant positions in these alignments are also invariant in the alignment of the 0.53 and 0.28 inhibitors (see Figure 2), thus giving further indication of the divergent relationship between these proteins.

When the barley trypsin inhibitor and the wheat α-amylase inhibitors are compared the invariant positions obtained are the same as in the comparison of the wheat α-amylase inhibitors with the ragi trypsin/α-amylase inhibitor. While this work was in progress, Odani et al. (1983a) published a report in which they showed a homology between the barley trypsin inhibitor and the 0.28 α-amylase inhibitor from wheat. Their alignment is almost identical to the one presented here (Figures 39a and 39b), with the exception of some matches in the C-terminal region of both inhibitors. This difference is explained by the weak homology in this region.

In a report on the amino acid sequence of the castor bean (Ricinus communis) reserve protein, which is composed of a small and a large subunit joined by a disulphide bridge, Sharief and Li (1982) drew attention to an apparent homology between the small subunit and positions 116 to 149 of the sweet plant protein Thaumatin I from the fruit of Thaumatococcus daniellii. Also a homology was apparent between positions 1 to 25 of the large subunit and positions 11 to 38 of the Bowman-Birk type inhibitor from lima bean (Sharief and Li, 1982). Both similarities are sufficient to be called homology and indeed the homology between the large subunit and the lima bean trypsin inhibitor can be extended to all menbers of the Bowman-Birk inhibitor family and all the invariant positions are centred around the reactive site of the inhibitors. However, if one compares the whole sequence of the sweet plant protein thaumatin I with the small subunit of the castor bean protein the similarities then observed are not sufficient to be called homology. The same observation is found when the whole of the large subunit is compared with the whole lima bean trypsin inhibitor molecule.

When the whole sequence of both subunits of the castor

1. Ser Val Gly Thr Ser ... Cys Ile Pro Gly Met Ala Ile Pro His Asn Pro Leu Asp Ser Cys Arg Trp
2. Ser Gly Pro Trp Ser Trp Cys Asn Pro Ala Thr Gly Tyr Lys Val Ser Ala Leu Thr Gly Cys Arg Ala

1. Tyr Val Ser Thr Arg Thr Cys Cys Val Gly Pro Arg Leu Ala Leu Gln Glu Met Lys Ala Arg Cys Cys
2. Met Val ... ... Lys Leu Gln Cys Val Gly Ser Gln Val Pro Glu Ala ... Val Leu Arg Asp Cys Cys

1. Arg Gln Leu Glu Ala Ile ... Pro Ala Tyr Cys Arg Cys Glu Ala Val Arg Ile Leu Met Asp Gly Val
2. Gln Gln Leu Ala Asp Ile Asn Asn Glu Trp Cys Arg Cys Gly Asp Leu Ser Ser Met Leu Arg Ala Val

1. Val Thr Ser Ser Gly Gln His Glu Gly Arg Leu Leu Gln Asp Leu Pro Gly Cys Pro Arg Gln Val ...
2. Ala Gln Glu Leu Gly Val Arg Glu Gly ... ... Lys Glu Val Leu Pro Gly Cys Arg Lys Glu Val Met

1. ... ... Gln Arg Ala Phe Ala Pro Lys Leu Val Thr Glu Val Glu Cys Asn Leu Ala Thr Ile His Gly
2. Lys Leu Thr Ala Ala Ser Val Pro Glu Val Cys Lys Val Pro Ile Pro Asn Pro Ser Gly Asp Arg Ala

1. Gly Pro Phe Cys Leu Ser Leu Leu Gly ... Ala Gly Glu
2. Gly Val Cys Tyr Gly Asp Trp Ala Ala Tyr Pro Asp Val

Figure 39a - Comparison of the primary structures of ragi trypsin/α-amylase inhibitor (1) and the wheat 0.28 α-amylase inhibitor (2) (Kashlan and Richardson, 1981). The identical residues are enclosed in boxes. (...) indicates gap. Minimum distance measure= 1032; similarity coefficient= 0.169; z= -4.425 (significant homology). Parameters values: u=10, v=12, d=10.

1. Ser Val Gly Thr Ser Cys Ile Pro Gly Met Ala Ile Pro His Asn Pro Leu Asp Ser Cys Arg Trp Tyr
2. Ser Gly Pro Trp Met Cys Tyr Pro Gly Gln Ala Phe Gln Val Pro Ala Leu Pro Gly Cys Arg Pro Leu

1. Val Ser Thr Arg Thr Cys Gly Val Gly Pro Arg Leu Ala Thr Gln Glu Met Lys Ala Arg Cys Cys Arg
2. Leu Lys Leu Gln Cys Asn Gly Ser Gln Val Pro Glu Ala ... ... Val Leu Arg Asp Cys Cys Gln

1. Gln Leu Glu Ala Ile Pro Ala Tyr Cys Arg Cys Glu Ala Val Arg Ile Leu Met Asp Gly Val Val Thr
2. Gln Leu Ala Asp Ile Ser Glu Trp Pro Arg Cys Gly Ala Leu Tyr Ser Met Leu Asp Ser Met Tyr Lys

1. Ser Ser Gly Gln His Glu Gly ... Arg Leu Leu Gln Asp Leu Pro Gly Cys Pro Arg ... Gln Val ...
2. Glu His Gly Val Ser Glu Gly Gln Ala Cys Thr Gly Ala Phe Pro Ser Cys Arg Arg Glu Val Val Lys

1. ... Gln Arg Ala Phe Ala Pro Lys Leu Val Thr Glu Val Glu Cys Asn Leu Ala Thr Ile His Gly Gly
2. Leu Thr Ala Ala Ser Ile Thr Ala Val Cys Arg Leu Pro Ile Val Val Asp Ala Ser Gly Asp Gly Ala

1. Pro Phe Cys ... Leu Ser Leu Leu Gly Ala Gly Glu
2. Tyr Val Cys Lys Asp Val Ala Tyr Pro Asp Ala Ala

Figure 39b - Comparison of the primary structures of ragi trypsin/α-amylase inhibitor (1) and the wheat 0.53 α-amylase inhibitor (2) (Maeda et al., 1983). The identical residues are enclosed in boxes. (...) indicates gap. Minimum distance measure=1032; similarity coefficient= 0.187; z= -5.153 (significant homology). Parameters values: u=10, v=12, d=10.

bean reserve protein is compared with the ragi trypsin/α-amylase inhibitor, the similarities found are sufficient to be called homology. In the alignment obtained the homologous regions are clustered around the N-terminal end of the ragi trypsin/α-amylase inhibitor. A much strong similarity is found when residues 1 to 95 of the ragi trypsin/α-amylase inhibitor are compared with both subunits of the castor bean protein. As can be seen in Figure 40a the small and large subunits of the castor bean protein can be aligned with the residues 1 to 35 and 36 to 95, respectively of the ragi trypsin/α-amylase inhibitor. In the castor bean protein the two subunits are joined by disulphide bridge (Sharief and Li, 1982) and it is worth mentioning that when the large and small subunits of the castor bean protein are aligned with the barley trypsin inhibitor, the C-terminal of the small subunit and the N-terminal of the large subunit are situated in a position immediately adjacent to the reactive site of hte ragi trypsin/α-amylase inhibitor (Arg^{34}-Leu). This suggests the possibility that the castor bean protein may be synthesized as a single polypeptide precursor (Odani et al., 1983c) and that a post-translational modification may occur consisting of the cleavage of the peptide bond Arg-Gln which would give rise to the two subunits. In this context it is useful to recall the indication by Sharief and Li (1982) that the native castor bean reserve protein seems to be resistant to hydrolysis by proteolytic enzymes.

The castor bean reserve protein can also be aligned with the barley trypsin inhibitor (Figure 40b) and the invariant positions which are observed are almost the same as in the case of the alignment between both subunits of the castor bean protein and rhe ragi trypsin/α-amylase inhibitor. After our publication of a paper

1. Ser Val Gly Thr Ser Cys Ile Pro Gly Met Ala Ile Pro His Asn Pro Leu Asp Ser Cys Arg Trp Tyr
2. Pro Ser Gln Gln Gly Cys Arg Gly Gln Ile Gln Glu Gln Gln Asn ... Leu Arg Gln Cys Gln Glu Tyr

1. Val Ser Thr Arg Thr Cys Gly Val Gly Pro Arg* Leu Ala Thr Gln Glu Met Lys Ala Arg ... Cys Cys
2. Ile Lys Gln Gln Val Ser Gly Gln Gly Pro Arg Arg ... ... Gln Glu Arg Ser Leu Arg Gly Cys Cys

1. Arg Gln Leu Glu Ala Ile Pro Ala Tyr Cys Arg Cys Glu Ala Val Arg ... ... Ile Leu Met Asp Gly
2. Asp His Leu Lys Gln Met Gln Ser Gln Cys Arg Cys Glu Gly Leu Arg Gln Ala Ile Gln Gln Gln Gln

1. Val Val Thr Ser Ser Gly Gln His Glu Gly Arg Leu Leu Gln Asp Leu Pro Gly Cys Pro Arg Gln Val
2. Leu Gln Gly Gln Asn Val Phe Glu Ala Phe Arg Thr Ala Ala Asn Leu Pro Ser Met Cys Gly Val Ser

1. Gln Arg* Ala Phe Ala Pro
2. Pro Thr Gln Cys Arg Phe

Figure 40a - Comparison of the primary structure of the castor bean reserve protein (2) (Sharief and Li, 1982) with positions 1 to 95 of the ragi trypsin/α-amylase inhibitor (1). Identical residues are enclosed in boxes. (...) indicates gap. (*) indicates trypsin inhibitory reactive sites. Minimum distance measure=788; similarity coefficient=0.171; z= -4.069 (significant homology). Parameters values: u=10, v=12, d=10.

1. Phe Gly Asp Ser Cys Ala Pro Gly Asp Ala Leu Pro His Asn Pro Leu Arg Ala Cys Arg Thr Tyr Val
2. Pro Ser Gln Gln Gly Cys Arg Gly Gln Ile Gln Glu Gln Gln Asn Leu Arg Gln Cys Gln Glu Tyr Ile

1. Val Ser Gln Ile Cys His Gln Gly Pro Arg* Leu Leu Thr Ser Asp Met Lys Arg Arg Cys Cys Asp
2. Lys Gln Gln Val Ser Gly Gln Gly Pro Arg Arg Gln Glu Arg Ser Leu ... Arg Gly Cys Cys Asp

1. Glu Leu Ser Ala Ile Pro Ala Tyr Cys Arg Cys Glu Ala Leu Arg Ile ... Ile Met Gln Gly Val Val
2. His Leu Lys Gln Met Gln Ser Gln Cys Arg Cys Glu Gly Leu Arg Gln Ala Ile Gln Gln Gln Gln Leu

1. Thr Trp Gln Gly Ala Phe Glu Gly Ala Tyr Phe Lys Asp Ser Pro Asn Cys Pro Arg Glu Arg Gln Thr
2. Gln Gly Gln Asn Val Phe Glu Ala Phe Arg Thr Ala Ala Asn Leu Pro Ser Met Cys Gly Val Ser Pro

1. Ser Tyr Ala Ala Asn
2. Thr Gln Cys Arg Phe

Figure 40b - Comparison of the primary structure of the castor bean reserve protein (2) (Sharief and Li, 1982) with positions 1 to 95 of the barley trypsin inhibitor (1) (Odani et al., 1983a). Identical residues are enclosed in boxes. (...) indicates gap. (*) indicates trypsin inhibitory reactive site. Minimum distance measure= 724; similarity coefficient= 0.238; z= -8.048 (significant homology). Parameters values: u=10, v=12, d=10.

describing the amino acid sequence of the ragi trypsin/α-amylase inhibitor (Campos and Richardson, 1983) and while this work was still in progress, Odani et al (1983c), reported the homology between the castor bean reserve protein, the barley trypsin inhibitor and the ragi trypsin/α-amylase inhibitor. In this report they also agreed that the homology between both subunits of the castor bean reserve protein and the barley trypsin inhibitor and the ragi trypsin/α-amylase inhibitor is more conceivable than the one between its small subunit and the sweet plant protein thaumatin I or the one between the large subunit and the segments of the amino acid sequences of menbers of the Bowman-Birk inhibitor family.

The alignment of the ragi trypsin/α-amylase inhibitor with the apparently homologous protein barley trypsin inhibitor, castor bean reserve protein, the 0.28α-amylase inhibitor and the 0.53α-amylase inhibitor is shown in Figure 41. This alignment is highly tentative and it may not be the best alignment to show the relationships between these proteins. As summarized in Table XVII the similarities between each of the proteins is sufficient to be called homology.

It should be noted that in the alignment shown in Figure 41 some of the invariant positions which are obtained when, say, the castor bean reserve protein and the barley trypsin inhibitor are aligned using the computer method, are lost in the tentative alignment proposed in Figure 41. This is in part due to the sheer number of possible alignments between these five sequences, which makes the task of finding an optimal alignment virtually impossible. A computer porgram was developed to compute the alignment of an n number of amino acid sequences, based on the algorithm proposed by Gotoh (1982). Unfortunately

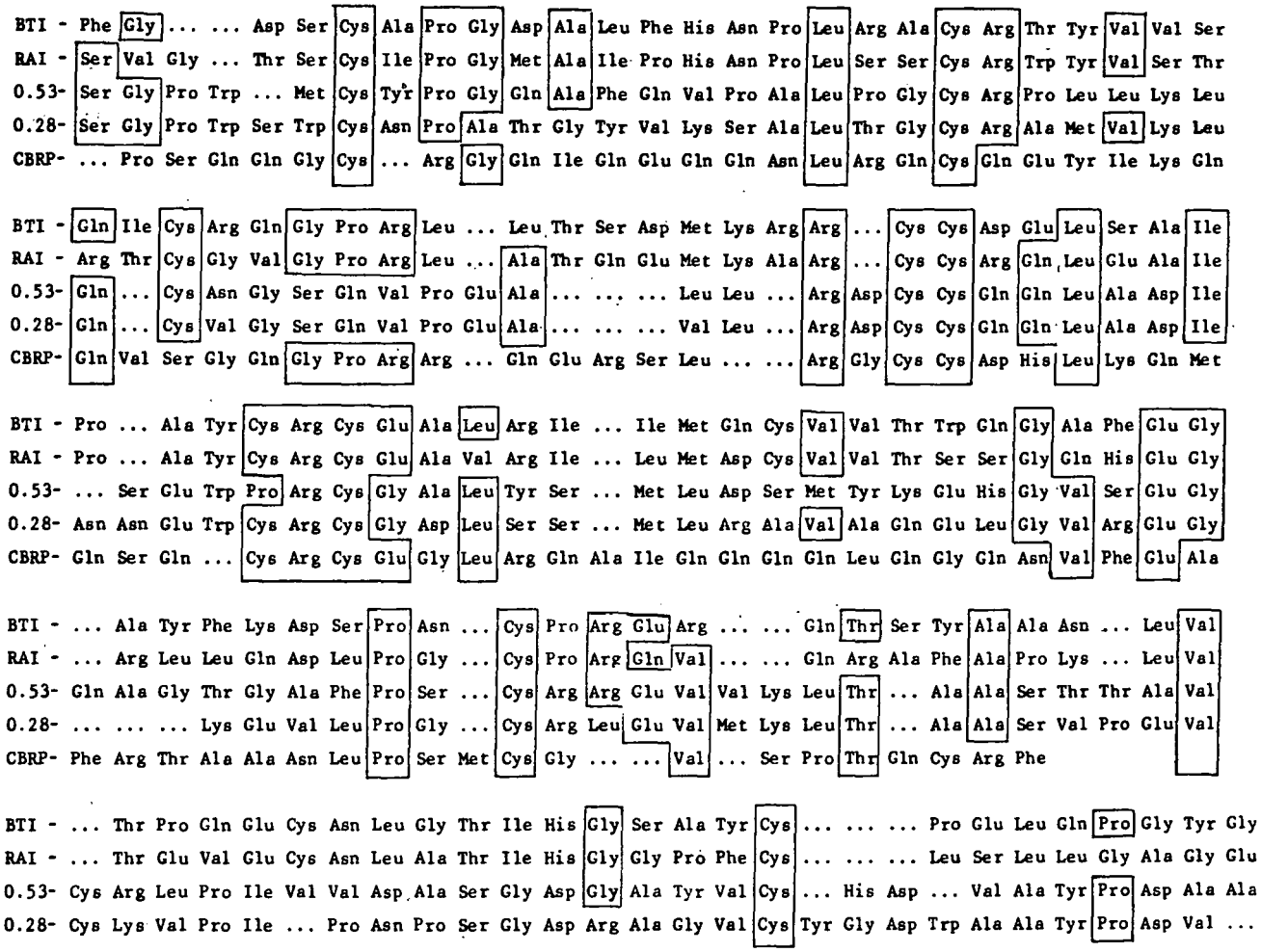


Figure 41 - Comparison of the amino acid sequences of barley trypsin inhibitor (BTI) (Odani et al., 1983a), ragi trypsin/α-amylase inhibitor (RAI), 0.53 wheat α-amylase inhibitor (0.53) (Maeda et al., 1983), 0.28 α-amylase inhibitor (0.28) (Kashlan and Richardson,1981) and castor bean reserve protein (Sharief and Li, 1982). Identical residues are enclosed in boxes. (...) indicates gap.

Table XVII - A matrix based on the amino acid sequence comparisons of the ragi trypsin/ α-amylase inhibitor (RAI), barley trypsin inhibitor (BTI) (Odani et al., 1983a), castor bean reserve protein (CBRP) (Sharief and Li, 1982), 0.28 wheat α-amylase inhibitor (0.28) (Kashlan and Richardson, 1981) and 0.53 α-amylase inhibitor (0.53) (Maeda et al., 1983). The values of z and of the minimum distance measure of each alignment are plotted in the lower half and upper half respectively.

	RAI	BTI	CBRP	0.28	0.53
RAI		572	788	1032	1018
BTI	-21.453		724	1038	1010
CBRP	- 4.069	-8.048		1054	1062
0.28	- 4.425	-4.116	-2.958		610
0.53	- 5.153	-5.496	-2.765	-21.211	

the complexity of this algorithm is exponential in time. This means that the algorithm is only viable if aplied to very short sequences. In practice the sequences are of considerable length and it could easily take several hundreds hours to compute the "optimal" alignment of the sequences.

The maize trypsin inhibitor is excluded from the alignment shown in Figure 41, due to the apparently fragmentary nature of its sequence data, although it is believed likely that a reassessement of its amino acid sequence will confirm the suggestions made about its homology with the ragi trypsin/α-amylase inhibitor family members. The pancreatic secretory trypsin inhibitor is also excluded on the grounds that the data available are not sufficient to allow a decision on whether the similarities with the barley trypsin inhibitor and the ragi trypsin inhibitor are indicative of a divergent or convergent relationship.

Taking into account the facts so far discussed, it is reasonable to argue that the ragi trypsin/α-amylase inhibitor and the homologous barley trypsin inhibitor do not fit into any of the inhibitor families proposed by Laskowski and Kato (1980) and that a new family should be created to accomodate them. This family should include the 0.28 and the 0.53 α-amylase inhibitors from wheat and the castor bean reserve protein. It remains to be seen whether the presence of α-amylase inhibitors and trypsin/α-amylase inhibitors is unique to this family, as the possibility that members of other inhibitor families may also have α-amylase inhibitory activity, apparently has not been investigated. In this connection, Mundy et al (1983) and Hejgaard et al (1983) have recently described the isolation and N-terminal sequence of a subtilisin/

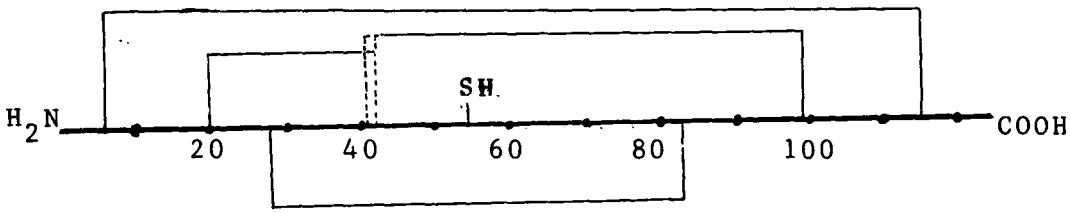
α-amylase inhibitor from barley seeds. It is interesting that the dual inhibitory activity in the protein molecule was recongnised due to the fact that while working with an α-amylase and a subtilisin inhibitor, they saw that the N-terminal sequence of both inhibitors were the same (Mundy, J., personal communication). Although data on the complete amino acid sequence of this subtilisin/α-amylase inhibitor has yet to be published, the N-terminal sequence of its 40 amino acids bears no homology either with members of the ragi trypsin/α-amylase inhibitor family or with members of other inhibitor families. However, this N-terminal sequence is strongly homologous to the N-terminal sequence of a subtilisin inhibitor from rice and thus they will probably constitute another new family.

The determination of the homology of a seed reserve protein with proteinase and α-amylase inhibitors may also serve as a starting point in a more precise comprehension of the genetic and physiological relationships between these proteins, although any homology does not necessarily confirm the hypothesis that proteinase inhibitors are a special kind of reserve protein. However, at least in the case of the ragi trypsin/α-amylase inhibitor, it has been shown to be hydrolysed *in vitro* by an endogenous proteinase from ragi seeds and its mobilisation during germination resembles the mobilisation of a reserve protein (Pattabiraman, T.N., personal communication).

It also remains to be clarified what is the exact evolutionary relationship between the pancreatic secretory trypsin inhibitor and the members of the ragi trypsin/α-amylase inhibitor family. As discussed before, the similarities between the pancreatic secretory trypsin inhibitor and the members of the ragi trypsin/α-amylase inhibitor family are sufficient to be called homology, while the same holds for

the similarities between the pancreatic secretory trypsin inhibitor and the Streptomyces subtilisin inhibitor, but there is no apparent homology between the Streptomyces subtilisin and the ragi trypsin/α-amylase inhibitor family members. The statistical analysis of the alignment ruled out the possibility that the similarities observed in the sequences were due to chance alone.

In making any comparisons of the amino acid sequences of the ragi trypsin/α-amylase inhibitor, the barley trypsin inhibitor and the 0.28 and 0.53 α-amylase inhibitors from wheat, it would be of interest to elucidate the positions of the intramolecular disulphide bridges in these inhibitors. At present the only one of these inhibitors for which details of the disulphide bridges are available is the 0.53 α-amylase inhibitor from wheat which was recently reported by Maeda et al. (1983b). In this inhibitor there are 4 disulphide bridges and one free cysteine residue, the locations of which are shown below:



It should be noted however that Maeda et al. (1983b) were unable to ascertain the exact disulphide bridge partners of residues 20 and 99, either of which could have been bonded

to cysteine residues 41 or 42.

In the comparison of the amino acid sequences of the 0.28 and 0.53 α-amylase inhibitors (Figure 2), the cysteine residues are distributed in almost identical positions, and so it would be reasonable to expect that the topological relationships between the disulphide bridges in these inhibitors might be very similar, except that the 0.28 inhibitor has the possibility of forming 5 disulphide bridges as it contains an additional cysteine residue (Kashlan and Richardson, 1981). However in the alignment of the ragi trypsin/ α-amylase inhibitor with either the 0.28 or the 0.53 wheat α-amylase inhibitors (Figures 39a and 39b), only 8 cysteine residues (0.28) and 7 cysteine residues (0.53) occupy identical positions, thus indicating that the topological relationships of the disulphide bridges in the wheat α-amylase inhibitors and in the trypsin (barley) and trypsin/ α-amylase (ragi) inhibitors belonging to the ragi trypsin/ α-amylase inhibitor family are likely to be different.

Laskowski and Kato (1980) have pointed out that the easiest family characteristics to appreciate at a glance are the topological relationships between the disulphide bridges. This does not appear to hold for the ragi trypsin/ α-amylase inhibitor family if the wheat α-amylase inhibitors are included. It should be said however that the suggestion of dividing the inhibitors into families was made primarily for inhibitors of proteolytic enzymes (Laskowski and Kato,

1980) and that in fact judging by the identical distribution of cysteine residues in the ragi trypsin/α-amylase inhibitor and in the barley trypsin inhibitor, the topological distribution of the disulphide bridges in these two inhibitors will probably be the same.

Finally, in conjunction with any further studies on the location of the disulphide bridges in these inhibitors it would also be desirable to determine the region(s) in the amino acid sequences of the ragi trypsin/α-amylase inhibitor and in the 0.28 and 0.53α-amylase inhibitors which are responsible for the inhibition of α-amylase, as it would provide further insight not only in the relationship between structure and inhibitory activity, but also into their evolutionary relationships.

Another line of possible future research in this area would be to examine as many of the extant proteinase inhibitors for the possesion of inhibitory activity towards amylase (or other enzymes), and the converse to see if any of the α-amylase inhibitors presently known have other inhibitory activities. This research is indicated since it seems extremely unlikely that the ragi (this work; Manjunath et al., 1983) and the barley (Mundy et al., 1983) bifunctional inhibitors are the only ones present in the plant kingdom.

3.3. Prediction of the Secondary Structure of the Trypsin/α-Amylase Inhibitor from Ragi (Indian Finger Millet, Eleusine coracana Gaertn.

The prediction of the secondary structure of the ragi trypsin/α-amylase inhibitor is presented in Figure 42. As no experimental data on the secondary structure contents of the inhibitor was available (e.g. from circular dichroism), a preliminary predictive run was made which indicated a 26% of α-helix content. From this percentage the optimal decision constants were chosen (Garnier et al., 1978) and an optimized prediction made, which gave to the inhibitor a 28% α-helix, 47% of extended chain, 12% of reverse turn and 13% of random coil.

In the secondary structure prediction presented in Figure 42, positions 25, 26, 28, 70 and 83 are occupied by serine, threonine, threonine, serine and proline, respectively. As shown before, there is a significant level of heterogeneity in all of these positions with alanine, lysine, alanine, proline and alanine also occupying positions 25, 26, 28, 70 and 83, respectively. In order to ascertain if these heterogeneities altered the predicted conformation of the region where they occur, predictive runs were made and in every case no alteration in predicted conformation was observed.

In a study on the conservation of chain reversal regions in proteins, Chou and Fasman (1979) found that in the C-peptides of proinsulin, in 7 proteinase inhibitors of the pancreatic secretory trypsin inhibitor (Kunitz) family and in 12 species of pancreatic ribonuclease, despite the relatively low sequence conservation in these sets of proteins, β-turns were predicted to be highly conserved. For example, in the proteinase inhibitor set, where the sequence

```
1. Ser Val Gly Thr Ser Cys Ile Pro Gly Met Ala Ile Pro His Asn Pro Leu Asp Ser Cys Arg Trp Tyr Val Ser Thr Arg Thr
   E   E   E   E   E   E   E   E   E   E   E   E   E   T   C   T   T   T   T   E   E   E   E   E   E   E   E   T

2.     Phe Gly Asp Ser Cys Ala Pro Gly Asp Ala Leu Pro His Asn Pro Leu Arg Ala Cys Arg Thr Tyr Val Val Ser Gln Ile
       T   T   T   T   T   C   T   T   T   C   C   T   T   C   T   T   T   T   T   E   T   T   T   E   E   E   E

1. Cys Gly Val Gly Pro Arg*Leu Ala Thr Gln Glu Met Lys Ala Arg Cys Cys Arg Gln Leu Gln Ala Ile Pro Ala Tyr Cys Arg
   T   T   G   C   C   E   H   H   H   H   H   H   H   H   H   H   H   H   E   E   E   E   T   T   H   H   H
2. Cys His Gln Gly Pro Arg*Leu Leu Thr Ser Asp Met Lys Arg Arg Cys Cys Asp Glu Leu Ser Ala Ile Pro Ala Tyr Cys Arg
   T   T   T   C   T   T   C   C   T   T   T   T   T   T   T   T   T   T   T   T   T   T   C   T   T   T   T   T

1. Cys Glu Ala Val Arg Ile Leu Met Asp Gly Val Val Thr Ser Ser Gly Gln His Glu Gly Arg Leu Leu Gln Asp Leu Pro Gly
   H   H   H   H   H   E   E   E   E   E   E   E   E   E   C   C   C   C   E   E   E   E   E   E   T   C   T   T
2. Cys Glu Ala Leu Arg Ile Ile Met Gln Gly Val Val Thr Trp Gln Gly Ala Phe Glu Gly Ala Tyr Phe Lys Asp Ser Pro Asn
   T   H   H   H   E   E   E   E   E   E   E   C   C   C   C   C   C   C   C   T   T   T   T   T   T   C   T   T

1. Cys Pro Arg Gln Val Gln Arg*Ala Phe Ala Pro Lys Leu Val Thr Glu Val Glu Cys Asn Leu Ala Thr Ile His Gly Gly Pro
   C   E   E   E   E   E   E   E   E   C   C   C   H   H   H   H   H   H   H   H   H   H   E   E   T   T   C   C
2. Cys Pro Arg Glu Arg Gln Thr Ser Tyr Ala Ala Asn Leu Val Thr Pro Gln Gln Cys Asn Leu Gly Thr Ile His Gly Ser Ala
   C   T   T   T   T   T   T   C   T   C   T   C   C   C   C   C   T   T   T   T   T   T   T   T   T   T   T   C

1. Phe Cys Leu Ser Leu Leu Gly Ala Gly Glu
   E   E   E   E   E   C   C   H   H   H
2. Tyr Cys Pro Glu Leu Gln Pro Gly Tyr Gly
   T   C   T   T   E   C   T   T   T   C
```

Figure 42 - Prediction of the secondary structures of the trypsin/α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) (1) and the trypsin inhibitor from barley (2) (Odani et al., 1983a).
(*): Trypsin inhibitory reactive site; T: β-turn; C: Random coil; E: Extended chain; H: α-helix.

conservation is 20%, they predicted a 85% β-turn conservation. They interpreted these results as an indication that chain reversal regions play an essential role in keeping intact the active structural domains in hormones and enzymes for their specific biological function. These studies, also on a predictive basis, were later supported by Tlomak and Nowak (1981). Their studies included the pancreatic secretory trypsin inhibitor (Kazal) family. Besides supporting the prediction of a conservation of chain reversal, they also suggested that in these two proteinase inhibitors families, there is a region conserved in the vicinity of the active site, composed of one β-turn with an adjacent extended chain structure. This suggestion in interesting, as it may explain why residues such as proline, which strongly favours the formation of β-turn (Chou and Fasman, 1978), are frequently found in the vicinity of the active site of proteinase inhibitors. In an attempt to check these suggestions, the prediction of the secondary structure of the barley trypsin inhibitor (Odani et al., 1983a) was done. As can be seen in Figure 42, the predicted secondary structures of the ragi trypsin/α-amylase inhibitor and the barley trypsin inhibitor are rather different, despite their strong sequence homology, and as far as these two inhibitors are concerned, do not give positive evidence for the suggestions of Chou and Fasman (1979) and Tlomak and Nowak (1981). In order to extend these observations to other inhibitor families, the secondary structure of some inhibitors of the Bowman-Birk family were also predicted and compared (Figure 43). In this family, the members are strongly related and their predicted secondary structures are also highly homologous, with a high content of reverse chain (around 78%). However, due to this high β-turn content a decision on the hypothesis of β-turn conservation cannot be made.

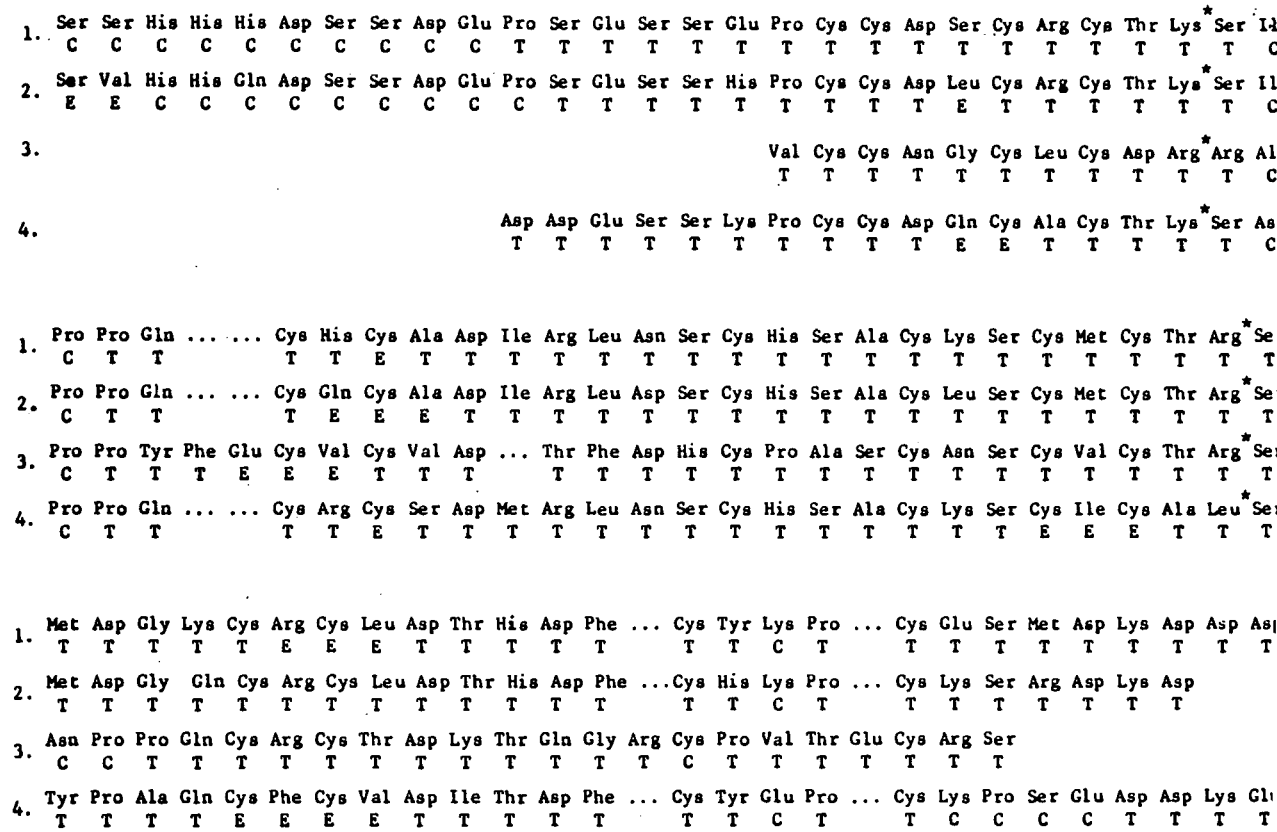


Figure 43 - Prediction of the secondary structures of the mung bean trypsin inhibitor (1) (Zhang et al., 1982), adzuki bean trypsin inhibitor (Yoshikawa et al., 1979) (2), peanuts trypsin inhibitor (3) (Norioka and Ikenaka, 1983), and the soybean Bowman-Birk inhibitor (4) (Odani and Ikenaka, 1973). (*) indicates reactive (inhibitory) site; T: β-turn; C: random coil; E: extended chain; H: α-helix.

An interesting feature in the predicted secondary structure of the ragi trypsin/α-amylase inhibitor, the barley trypsin inhibitor and the inhibitors from the Bowman-Birk family is their high content of β-turn, which may provide an additional factor for the remarkable thermal stability of these inhibitors (Ryan, 1973; Richardson, 1977, 1980, 1981; Liener and Kakade, 1980).

3.4. The Complete Amino Acid Sequence of the α-Amylase Inhibitor from Seeds of Ragi (Indian Finger Millet, Eleusine coracana Gaertn)

3.4.1. Amino Acid Analysis

The results of the amino acid analysis of the α-amylase inhibitor from ragi are presented in Table XIII. The values obtained are in good agreement with the amino acid composition calculated from the sequence except for the values of Asp, Ala and Val.

It should be noted that the amino acid composition of the ragi α-amylase inhibitor shown in Table XIII is remarkably different from that of the trypsin/α-amylase inhibitor in the same tissue (Table VI).

3.4.2. N-Terminal Analysis

The N-terminal sequence analysis of the reduced and carboxymethylated α-amylase inhibitor was performed by the DABITC-PITC double coupling method. The sequence obtained is given below:

Ala-Ile-Pro-Cys-Gly-Gln-Val-Ser-Ser-Ala-Ile-Gly-Pro

The identity of Ile^{2} and Ile^{11} was ascertained by comparison with peptides Ch1-4 and Ch3-4.

3.4.3. Tryspin Digestion

15 mg of the reduced and carboxymethylated inhibitor were subjected to hydrolysis with trypsin, under the conditions described

Table XVIII - The amino acid composition of the α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.).

Amino Acid	Analysis	Sequence
CM-Cys	6.59	7
Asp	8.29	7
Thr	1.90	2
Ser	15.22	16
Glu	2.53	2
Pro	ND	4
Gly	8.68	9
Ala	17.21	20
Val	4.03	5
Met	0.10	0
Ile	4.46	5
Leu	5.23	5
Tyr	1.68	2
Phe	0.14	0
Lys	2.03	2
His	0.21	0
Arg	8.66	9
Trp	ND	0

ND= Not determined.

in Section 2.2.6.3.1. Attempts were made to purify the fragments produced directly by RP-HPLC (Figure 44), but only a few of the peptides obtained were pure enough to allow direct sequencing on them (Table XIX). In order to obtain a better separation, the tryptic digest (12 mg) was fractioned on a Biogel P-4 column, equilibrated in 0.005M ammonium bicarbonate (Figure 45). Each of the pooled fractions was subjected to RP-HPLC; the HPLC chromatograms of the HPLC separations of peaks 2, 3 and 4 are shown in Figures 46, 47 and 48, respectively. No pure peptides were obtained from the RP-HPLC of peak T1. The sequences, elution times and positions in the sequence of each of the peptides purified are presented in Table XX.

Peptides T-9, T2-5 and T3-7 (positions 1-19)

Ala-Ile-Ser-Cys-Gly-Gln-Val-Ser-Ser-Ala-Ile-Gly-Pro-Cys-Leu-Ala-Tyr-Ala-Arg

The identity of Ile^{2} was ascertained by subjecting each peptide to a cycle of degradation by the "Dansyl-dman" method.

Peptides T2-1, T2-2, T2-3, T3-3 and T3-4 (positions 20-34)

Gly-Ala-Gly-Ala-Ala-Pro-Ser-Ala-Ser-Cys-Gln-Ser-Gly-Val-Arg

Peptides T-5 and T4-4 (positions 34-41)

Ser-Leu-Asn-Ala-Ala-Ala-Arg

The identity of Leu^{36} was ascertained by subjecting the peptides to a cycle of degradation by the "Dansyl-Edman" method, and by semi-quantitative amino acid analysis by the dansyl-chloride method.

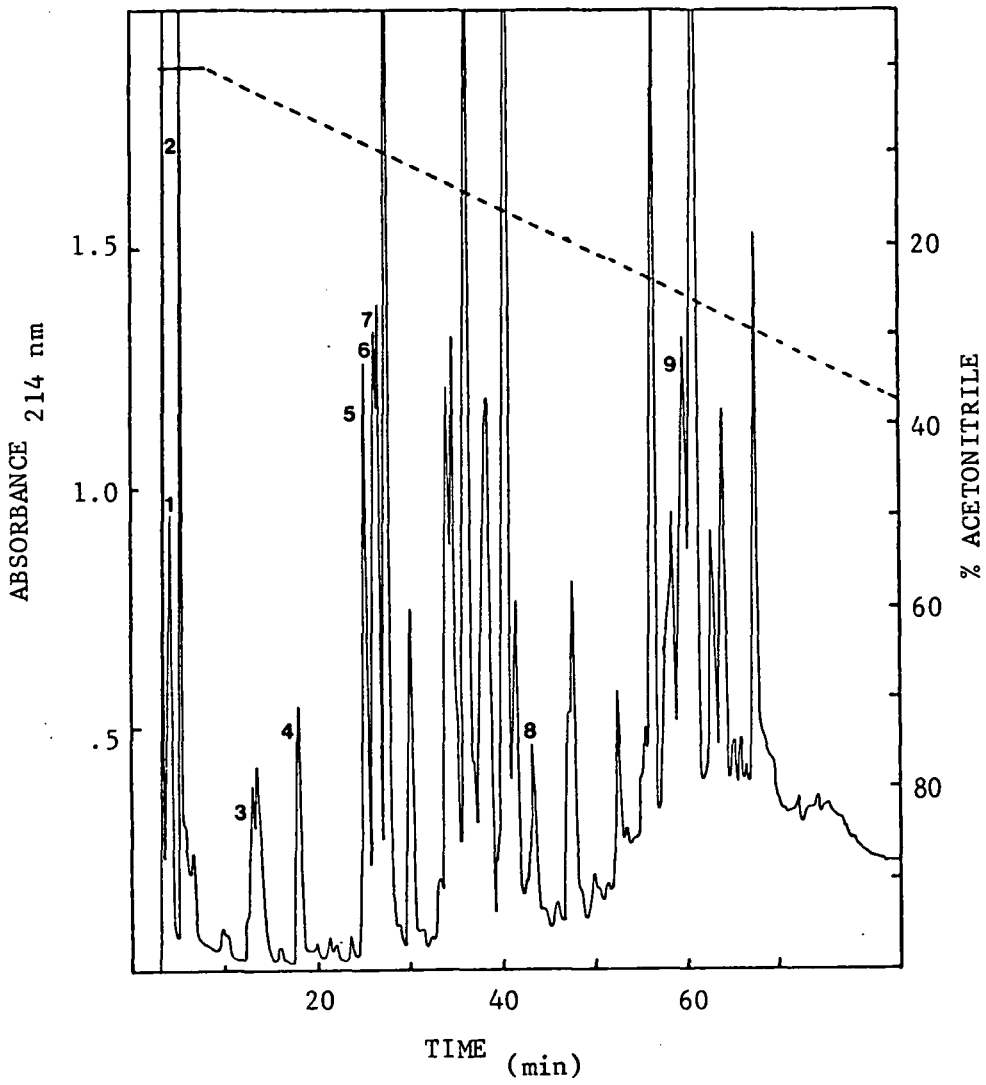


Figure 44 - RP-HPLC separation of the trypsin digest of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70 % (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table XIX - Peptides obtained by the digestion with trypsin of the α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.)

Peptide No. and Position in Sequence		Elution Time	Sequence
T-1	(93-95	4	Val-Asn-Asn
T-2	(56-60)	6	Ser-Ala-Ala-Ser-Arg
T-3	(42-47)	13	Thr-Thr-Ala-Asp-Arg-Arg
T-4	(76-79)	18	Cys-Gly-Val-Arg
T-5	(35-41)	25	Ser-Leu-Asn-Ala-Ala-Ala-Arg
T-6	(61-68)	26	Val-Ser-Gly-Leu-Asn-Ala-Gly-Lys
T-7	(69-75)	28	Ala-Ser-Ser-Ile-Pro-Gly-Arg
T-8	(77-82)	43	Gly-Val-Arg-Leu-Pro-Tyr
T-9	(1-19)	60	Ala-Ile-Ser-Cys-Gly-Gln-Val-Ser-Ser-Ala-Ile-Gly-Pro-Cys-Leu-Ala-Tyr-Ala-Arg

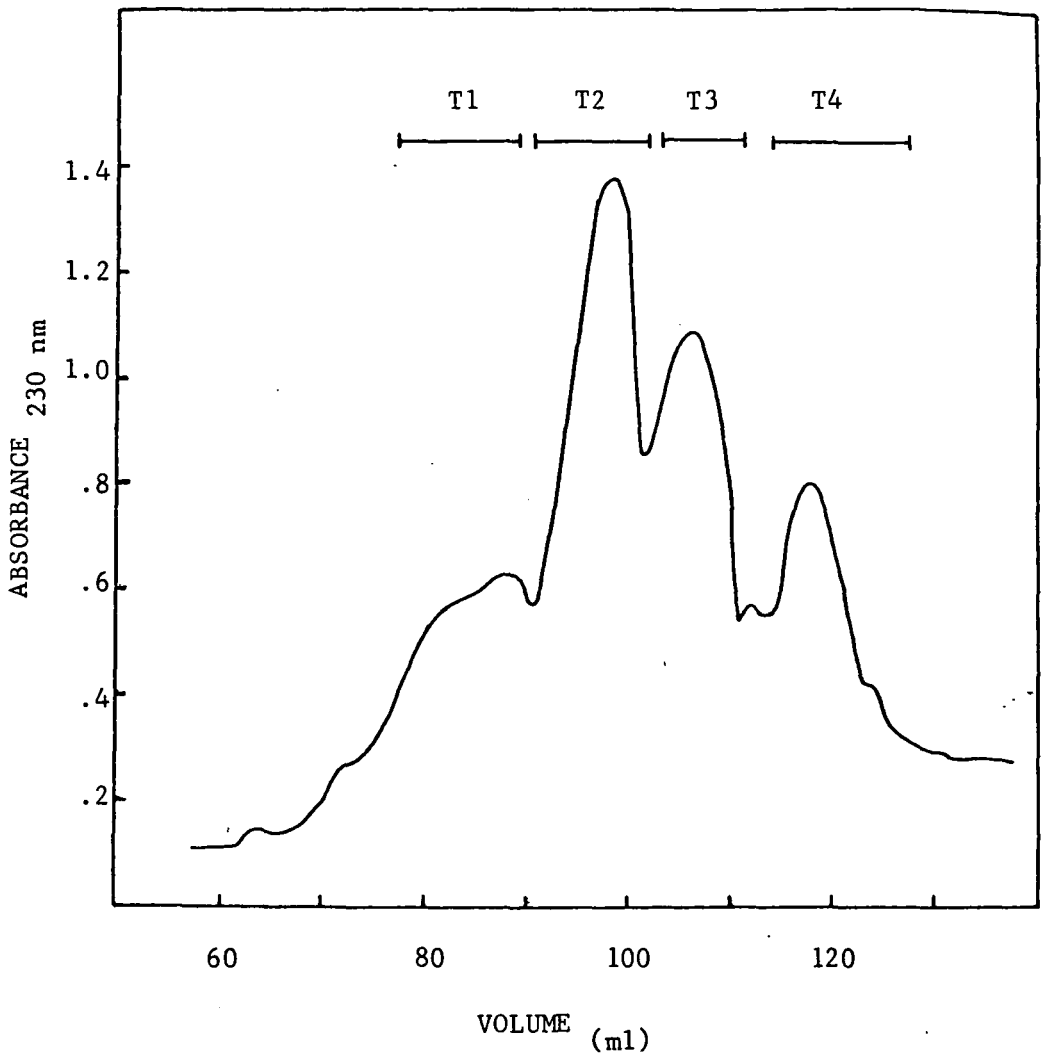


Figure 45 - Chromatography of the products of the digestion by trypsin of the α-amylase inhibitor from seeds of ragi (Indian Finger millet. Eleusine coracana Gaertn.) on a Biogel P-4 column (0.8 x 190cm) in 0.05M ammonium bicarbonate (flow rate 6 ml/h., fraction size 1.4 ml).

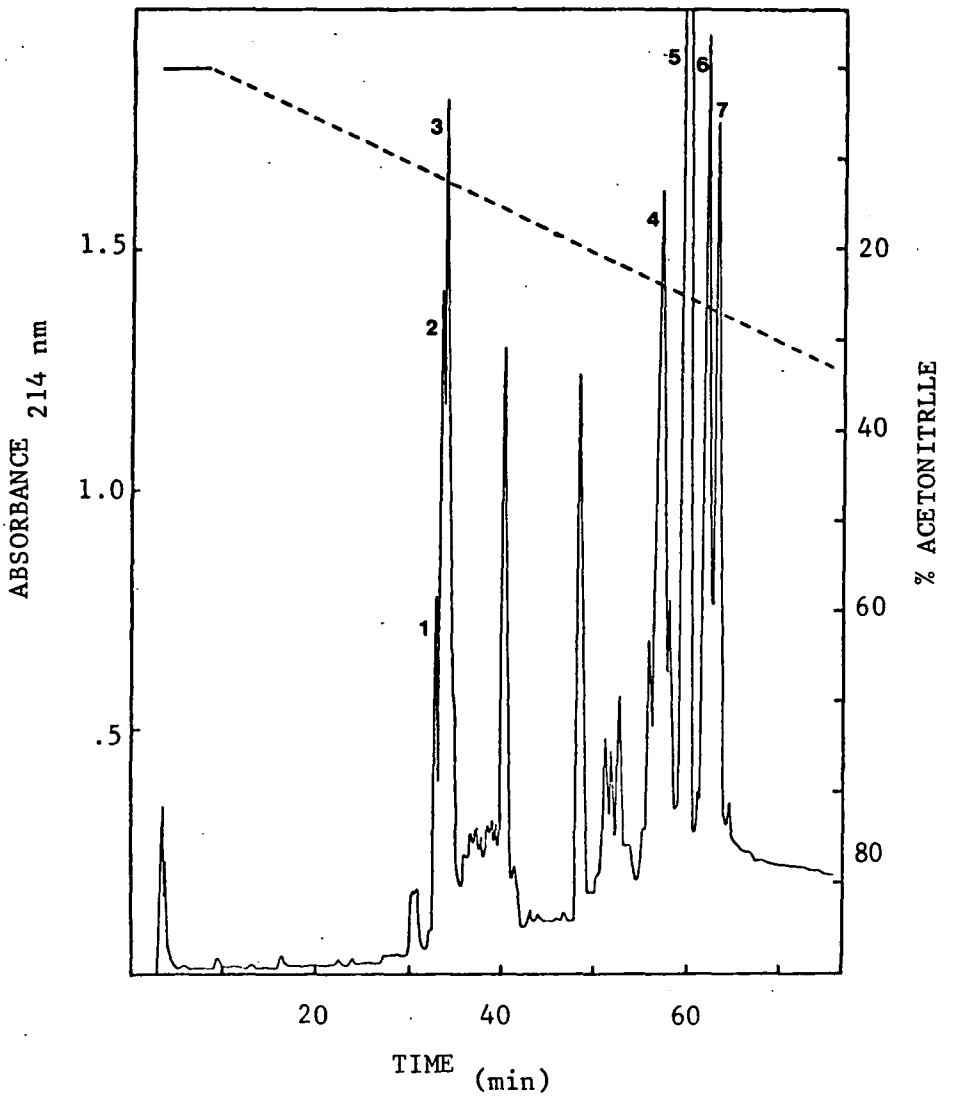


Figure 46 - RP-HPLC separation of fraction T-2 from the trypsin digestion of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

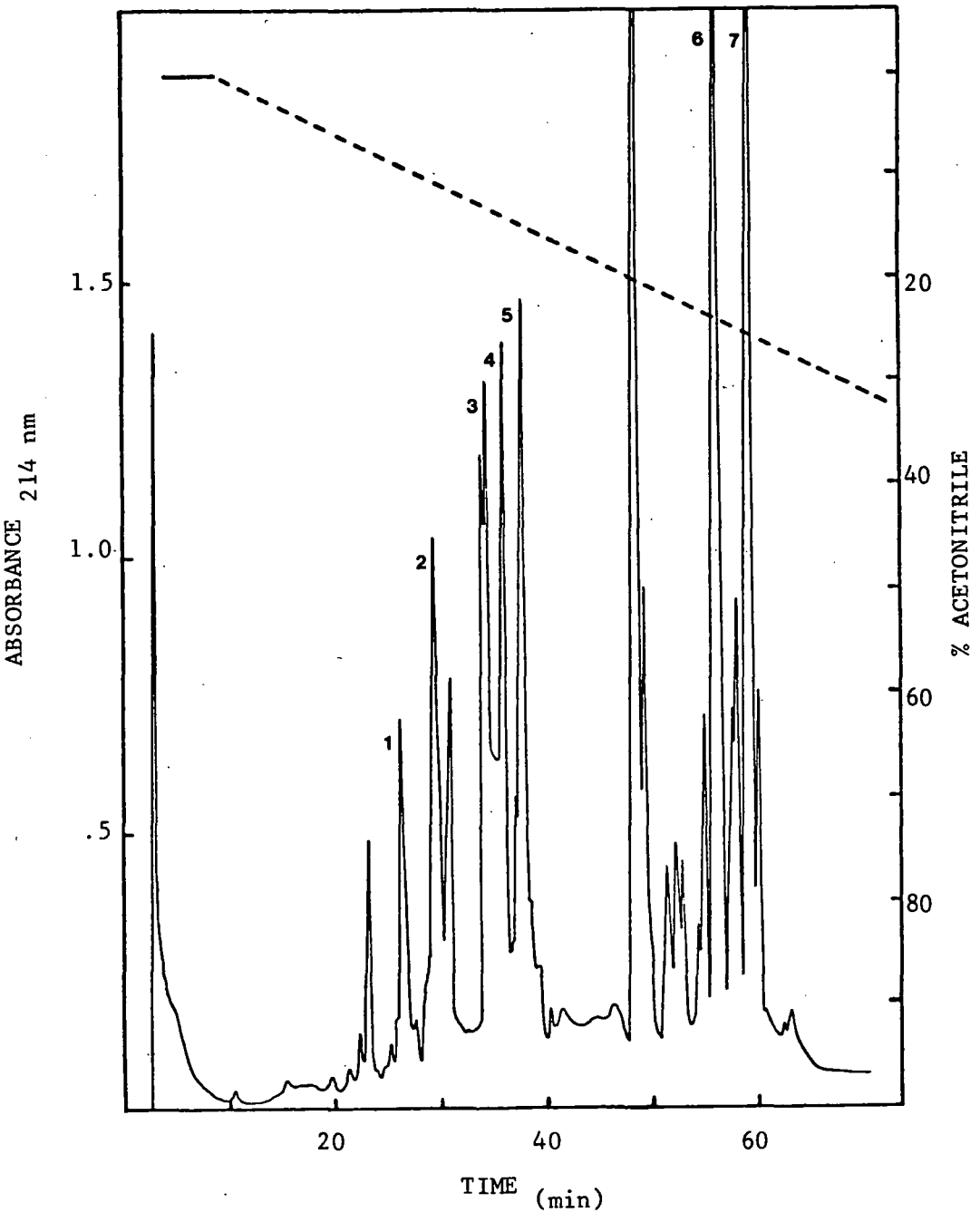


Figure 47 - RP-HPLC separation of fraction T-3 from the trypsin digestion of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracaca Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25 cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

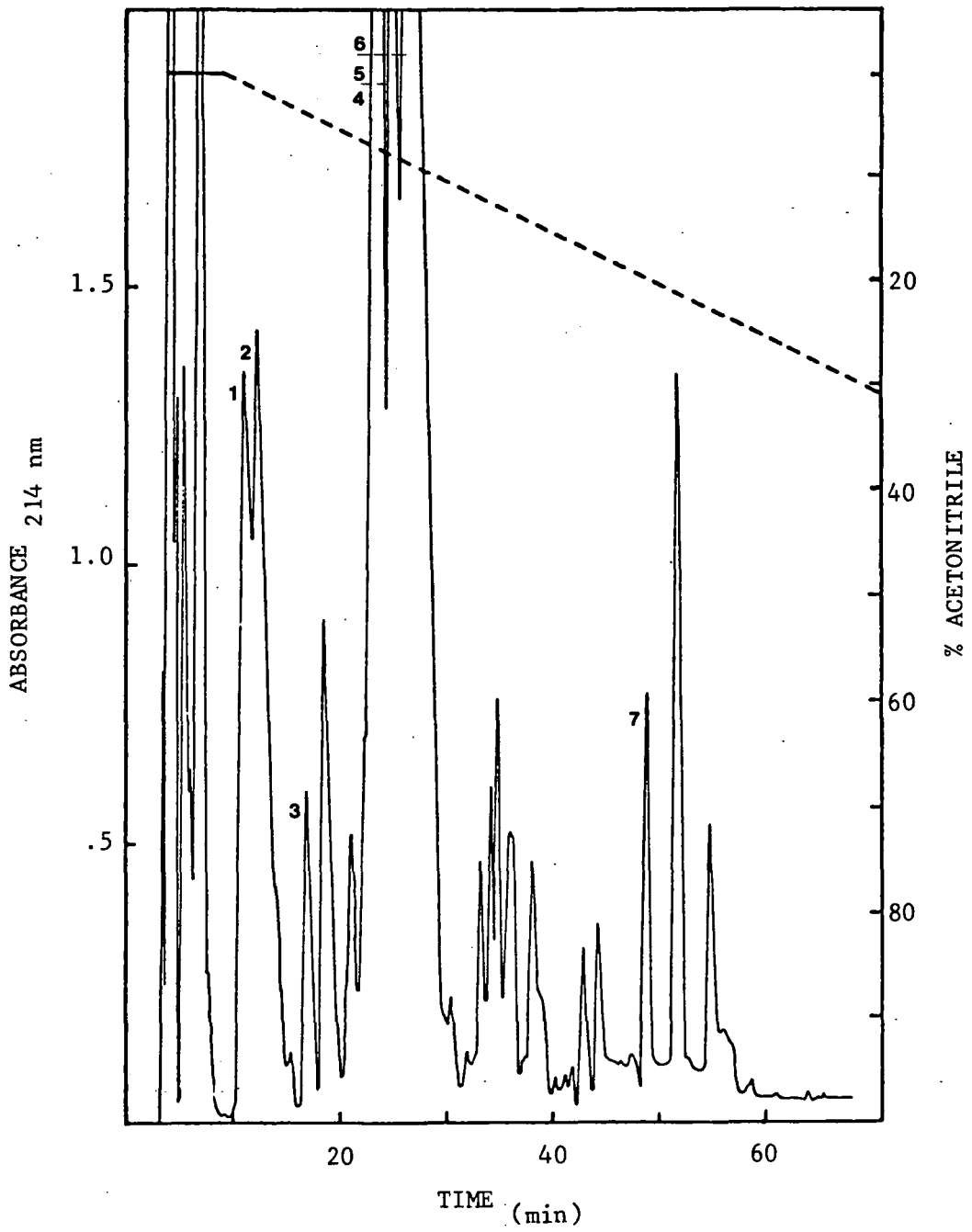


Figure 48 - RP-HPLC separation of fraction T-4 from the trypsin digestion of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rat of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table XX - Peptides obtained by the RP-HPLC separation of Biogel P-4 fractions T2 to T4 from trypsin digestion of the α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.)

Peptide No. and Position in Sequence	Elution Time	Sequence
T2-1 (20-34)	33	Gly-Ala-Gly-Ala-Ala-Pro-Ser-Ala-Ser-Cys-Gln-Ser-Gly-Val-Arg
T2-2 (20-34)	34	Gly-Ala-Gly-Ala-Ala-Pro-Ser-Ala-Ser-Cys-Gln-Ser-Gly-Val-Arg
T2-3 (20-34)	35	Gly-Ala-Gly-Ala-Ala-Pro-Ser-Ala-Ser-Cys-Gln-Ser-Gly-Val-Arg
T2-4 (80-92)	57	Leu-Pro-Tyr-Ala-Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg
T2-5 (1-19)	59	Ala-Ile-Ser-Cys-Gly-Gln-Val-Ser-Ser-Ala-Ile-Gly-Pro-Cys-Leu-Ala-Tyr-Ala-Arg
T2-6 (76-92)	62	Cys-Gly-Val-Arg-Leu-Pro-Tyr-Ala-Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg
T2-7 (76-92)	63	Cys-Gly-Val-Arg-Leu-Pro-Tyr-Ala-Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg
T3-1 (61-68)	27	Val-Ser-Gly-Leu-Asn-Ala-Gly-Lys
T3-2 (48-55)	30	Ala-Ala-Cys-Asn-Cys-Ser-Leu-Lys
T3-3 (20-34)	35	Gly-Ala-Gly-Ala-Ala-Pro-Ser-Ala-Ser-Cys-Gln-Ser-Gly-Val-Arg
T3-4 (20-34)	36	Gly-Ala-Gly-Ala-Ala-Pro-Ser-Ala-Ser-Cys-Gln-Ser-Gly-Val-Arg

Table XX - (Cont.)

Peptide No. and Position in Sequence		Elution Time	Sequence
T3-5	(83-95)	38	Ala-Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg-Val-Asn-Asn
T3-6	(80-92)	56	Leu-Pro-Tyr-Ala-Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg
T3-7	(1-19)	60	Ala-Ile-Ser-Cys-Gly-Gln-Val-Ser-Ser-Ala-Ile Gly-Pro-Cys-Leu-Ala-Tyr-Ala-Arg
T4-1	(42-47)	11	Thr-Thr-Ala-Asp-Arg-Arg
T4-2	(42-47)	12	Thr-Thr-Ala-Asp-Arg-Arg
T4-3	(76-79)	17	Cys-Gly-Val-Arg
T4-4	(35-41)	23	Ser-Leu-Asn-Ala-Ala-Ala-Arg
T4-5	(61-68)	24	Val-Ser-Gly-Leu-Asn-Ala-Gly-Lys
T4-6	(69-75)	26	Ala-Ser-Ser-Ile-Pro-Gly-Arg
T4-7	(80-92)	49	Leu-Pro-Tyr-Ala-Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg

Peptides T4-1, T4-2 and T-3 (positions 42-47)

Thr-Thr-Ala-Asp-Arg-Arg

Peptide T3-2 (positions 48-55)

Ala-Ala-Cys-Asn-Cys-Ser-Leu-Lys

The identity of Leu^{54} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by comparison with peptide Ch4-1.

Peptide T-2 (positions 56-60)

Ser-Ala-Ala-Ser-Arg

Peptides T-6, T3-1, and T4-5 (positions 61-68)

Val-Ser-Gly-Leu-Asn-Ala-Gly-Lys

The identity of Leu^{64} was ascertained by subjecting peptide T3-1 to 3 cycles of degradation by the "Dansyl-Edman" method, by semi-quantitative amino acid analysis by the dansyl-chloride method and by comparison with peptides Th-7, Th-11 and Th-12.

Peptides T-7 and T4-6 (positions 69-75)

Ala-Ser-Ser-Ile-Pro-Gly-Arg

The identity of Ile^{72} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by comparison with peptide $Ch4\text{-}6$.

Peptides T-4 and T4-3 (positions 76-79)

Cys-Gly-Val-Arg

Peptide T-8 (positions 77-82)

Gly-Val-Arg-Leu-Pro-Tyr

The identity of Leu^{80} was ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method and by comparison with peptides T2-4, T3-7 , T4-7 . Peptide T-8 arose due to the rather unusual cleavages of the peptide bonds Cys^{76}-Gly and Tyr^{82}-Ala, the later being a chymotrypsin cleavage site. Cleavage at the carboxyl side of cysteine by trypsin have been reported before in the bovine β-crystallin Bp chain by Driessen et al. (1983).

Peptides T2-6 and T2-7 (positions 76-92)

Cys-Gly-Val-Arg-Leu-Pro-Tyr-Ala-Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg

The identity of Leu^{80} was ascertained by comparison with peptides T2-4, T3-6 and T4-7 and by subjecting peptides T2-6 and T2-7 to four cycles of degradation by the "Dansyl-Edman" method.

In peptides T-8, T2-6 and T2-7 the peptide bond Arg^{79}-Leu was not cleaved by trypsin. This is probably due to the proximity of proline at position 81.

Peptides T2-4, T3-6 and T4-7 (positions 80-92)

Leu-Pro-Tyr-Ala-Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg

The identity of Leu^{80}, Ile^{84} and Ile^{88} were ascertained by subjecting the peptides to 5 cycles of degradation by the "Dansyl-Edman" method, by N-terminal analysis by the dansyl-chloride method and by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide T3-5 (positions 83-95)

Ala-Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg-Val-Asn-Asn

Peptide T3-5 is the product of an anomalous tryptic cleavage at the peptide bond Tyr^{82}-Ala. As the same peptide was also obtained in the chymotryptic cleavage of the inhibitor (see peptides Ch1-2 and Ch2-3), most probably its presence in the tryptic digest is due to a chymotrypsin contamination in the trypsin preparation utilised. The identity of Ile^{84} and Ile^{88} were ascertained by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide T-1 (positions 93-95)

Val-Asn-Asn

Peptide T-1 was identified as the C-terminal peptide, as it lacked arginine and lysine.

Apart from the cleavage of the peptide bonds Cys^{76}-Gly and Tyr^{82}-Ala, the observed cleavage sites in the digestion of the α-amylase inhibitor with trypsin were consistent with the specificity of trypsin. The existence of peptides T-8 (Gly^{77}-Val-Arg-Leu-Pro-Tyr) might suggest that the residue at position 76 (cysteine) was mistakenly identified, since neither trypsin or chymotrypsin (the usual contaminant of trypsin preparations) normally cleave at the carboxyl side of cysteine. In this case special attention was dedicated to this region of the inhibitor molecule and in all cases, the residue at position 76 was positively identified as cysteine (see peptides T-4, T2-6, T2-7, Ch4-7, Th-11 and Th-12).

3.4.4. Chymotrypsin Digestion

15 mg of the reduced and carboxymethylated inhibitor were subjected to chymotrypsin digestion, under the conditions described in Section 2.2.6.3.2. The fragments obtained were initially fractioned on a Sephadex G-50 column, equilibrated in 0.05M ammonium bicarbonate (Figure 49) and each of the pooled fractions was subsequently subjected to RP-HPLC. The chromatograms of the HPLC separation of peaks 1, 2, 3 and 4 are shown in Figures 50, 51, 52 and 53, respectively The sequences, elution times and positions in the sequence of each of the peptides obtained are presented in Table XXI.

Peptide Ch1-4 (positions 1-15)

Ala-Ile-Ser-Cys-Gly-Gln-Val-Ser-Ser-Ala-Ile-Gly-Pro-Cys-Leu

The identities of Ile^{2}, Ile^{11} and Leu^{17} were ascertained by subjecting the peptides to a cycle of degradation by the "Dansyl-Edman" method and by semi-quantitative amino acid analysis by the dansyl-chloride method. Also the specificity of chymotrypsin, as well as the comparison with peptides Ch3-4 and Ch4-5 helped to ascertain the identity of Leu^{17}.

Peptide Ch3-4 and Ch4-5 (positions 9-17)

Ser-Ala-Ile-Gly-Pro-Cys-Leu-Ala-Tyr

The cleavage of the peptide bond Ser^{8}-Ser, which gave rise to peptides Ch3-4 and Ch4-5, does not correspond to the usual specificity of chymotrypsin. Further examples of cleavage by chymotrypsin at the carboxyl side of serine will be given later (see peptides Ch1-1, Ch2-1, Ch4-2 and Ch4-7 below). The identity of Ile^{11} and Leu^{17} was ascertained

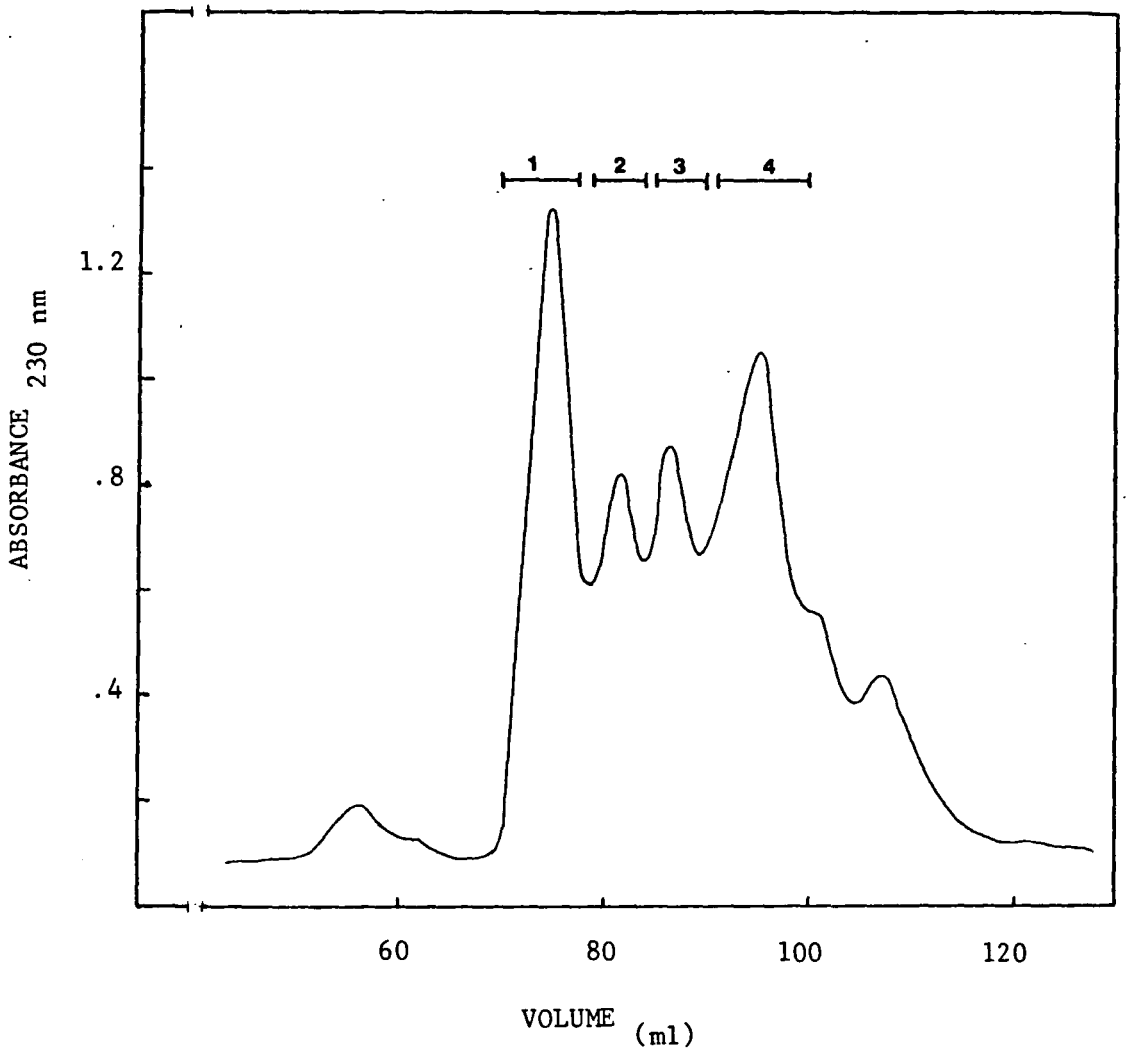


Figure 49 - Chromatography of the products of the digestion by chymotrypsin of the α-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) on a Sephadex G-50 column (0.8 x 190 cm) in 0.05M ammonium bicarbonate (flow rate 6 ml/h., fraction size 1.4 ml).

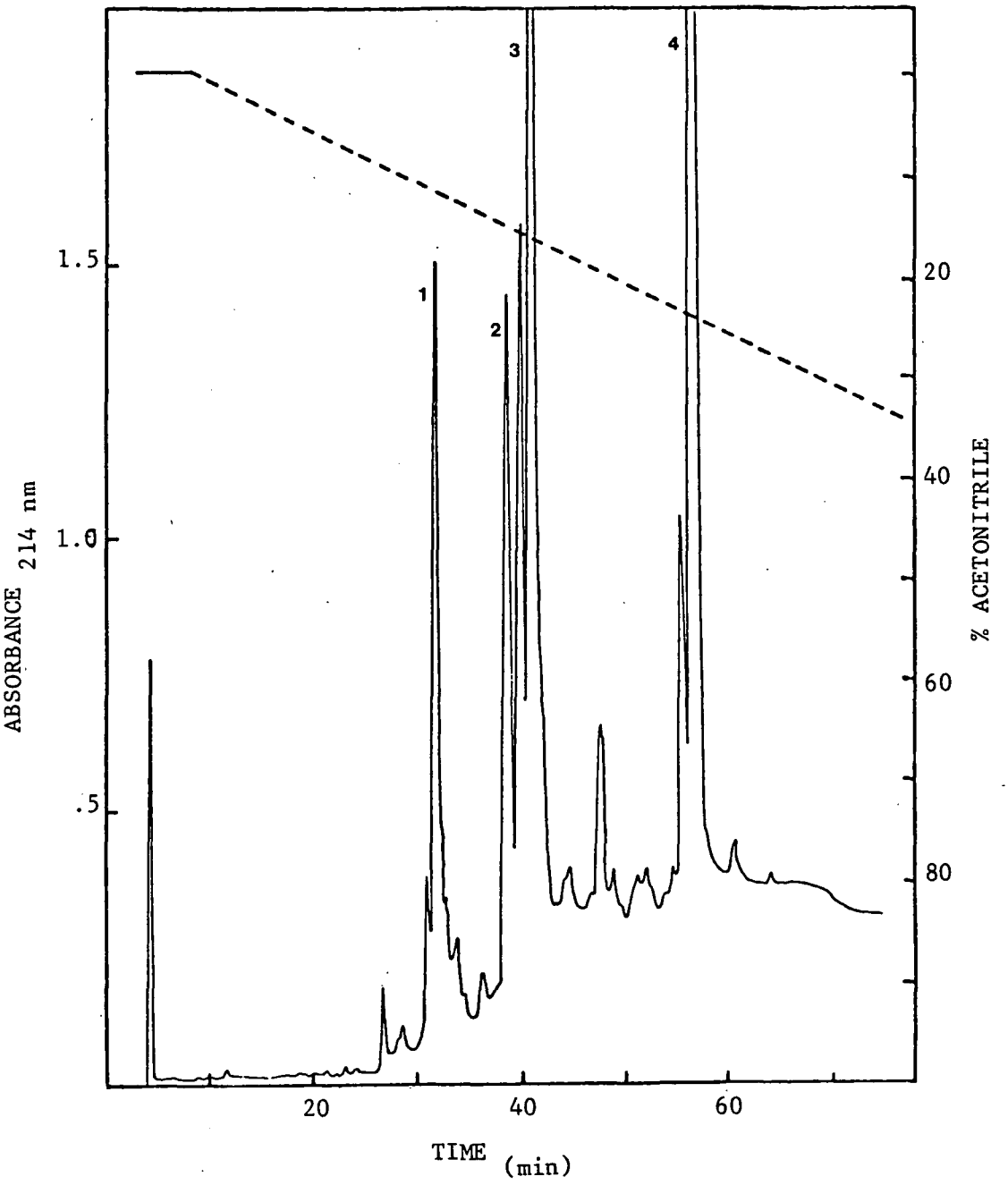


Figure 50 - RP-HPLC separation of fraction Ch-1 from the chymotrypsin digestion of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

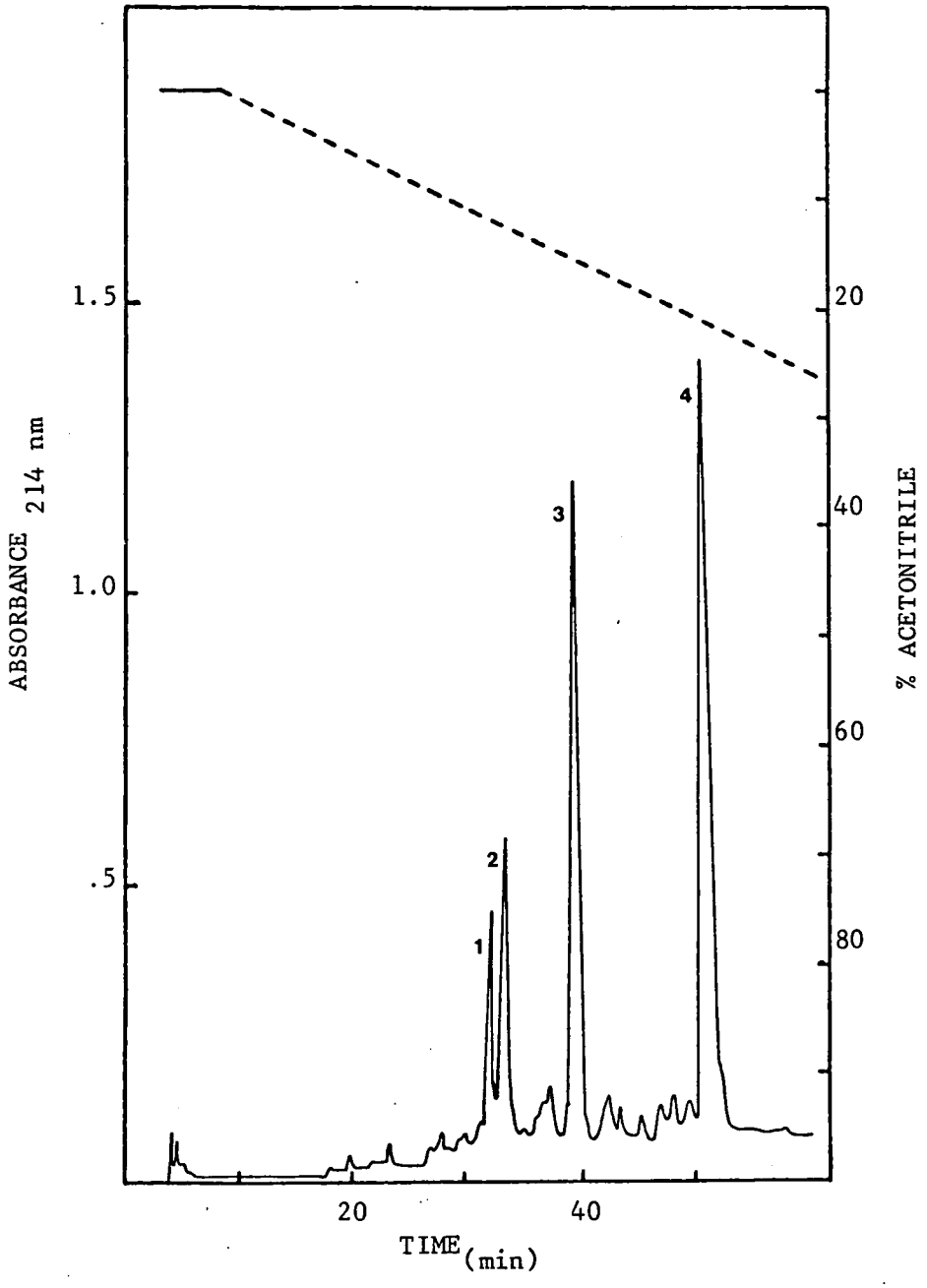


Figure 51 - RP-HPLC separation of fraction Ch-2 from the chymotrypsin digestion of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow ratw of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted lines indicates concentration of acetonitrile.

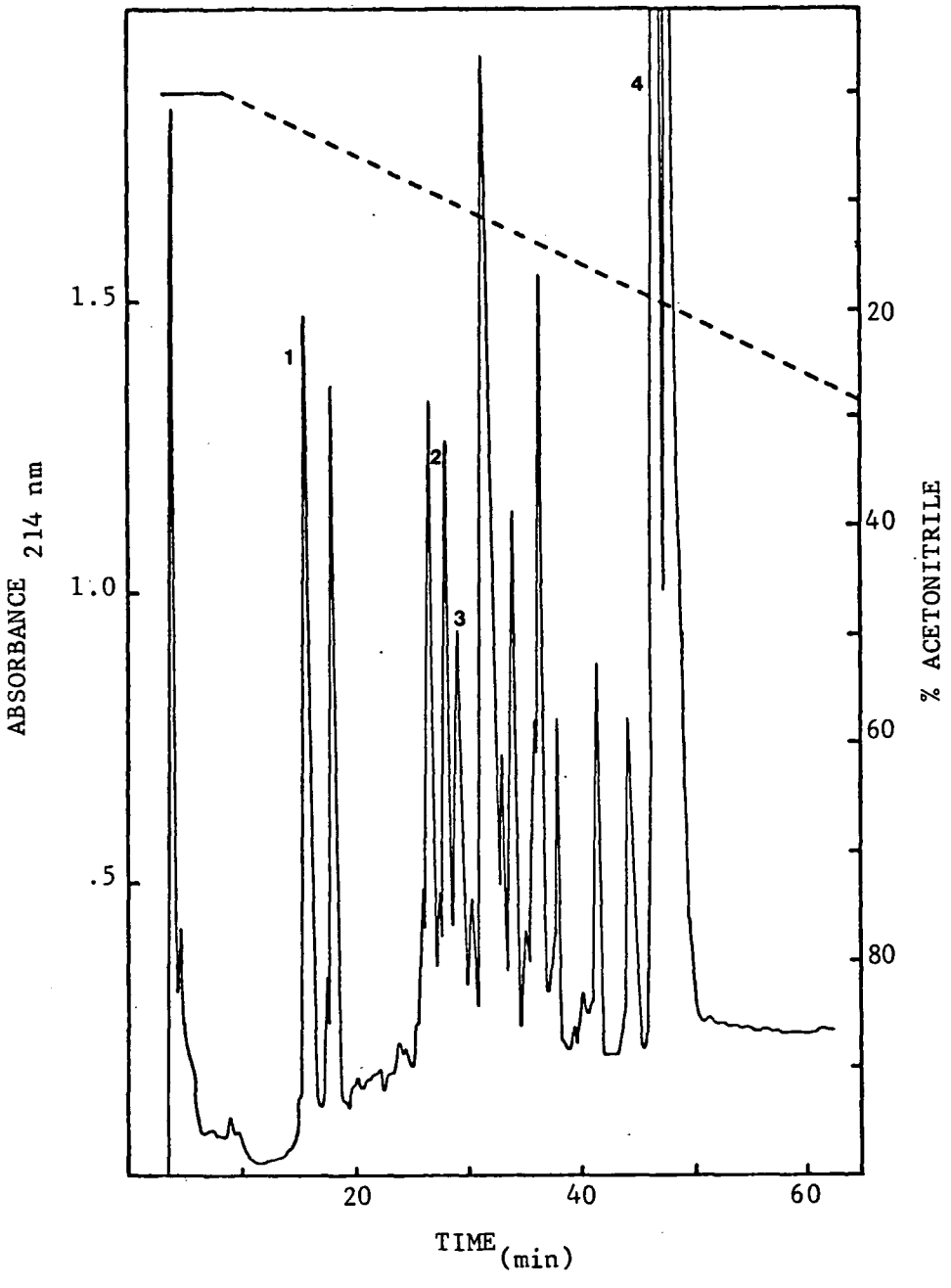


Figure 52 - RP-HPLC separation of fraction Ch-3 from the chymotrypsin digestion of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 u-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

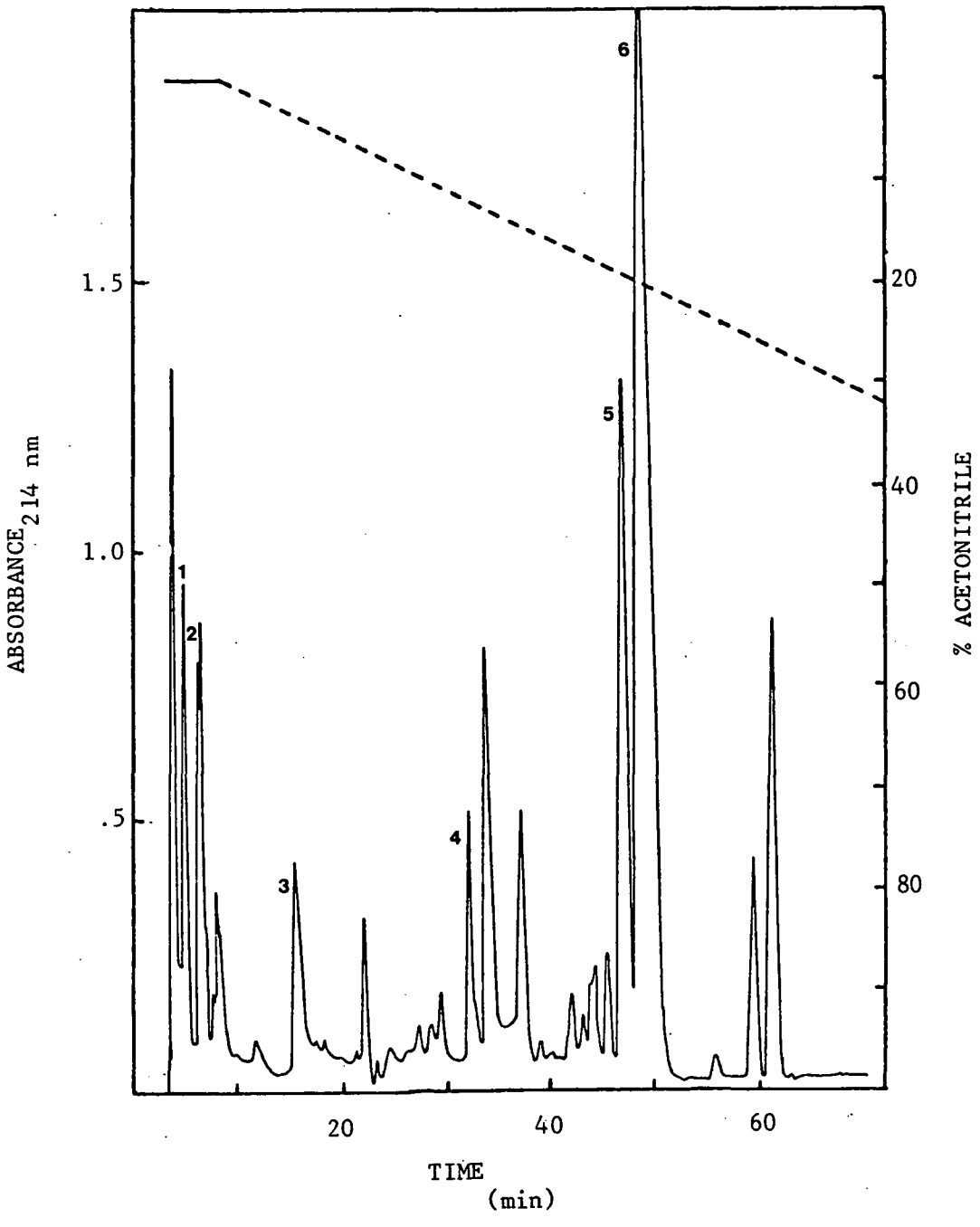


Figure 53 - RP-HPLC separation of fraction Ch-4 from the chymotrypsin digestion of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table XXI - Peptides obtained by the RP-HPLC separation of Sephadex G-50 fractions Ch1 to Ch4 from chymotrypsin digestion of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.)

Peptide No. and Position in Sequence		Elution Time	Sequence
Ch1-1	(18-26)	32	Ala-Arg-Gly-Ala-Gly-Ala-Ala-Pro-Ser
Ch1-2	(83-95)	38	Ala-Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg-Val-Asn-Asn
Ch1-3	(18-36)	41	Ala-Arg-Gly-Ala-Gly-Ala-Ala-Pro-Ser-Ala-Ser-Cys-Gln-Ser-Gly-Val-Arg-Ser-Leu
Ch1-4	(1-15)	56	Ala-Ile-Ser-Cys-Gly-Gln-Val-Ser-Ser-Ala-Ile-Gly-Pro-Cys-Leu
Ch2-1	(18-26)	32	Ala-Arg-Gly-Ala-Gly-Ala-Ala-Pro-Ser
Ch2-2	(43-54)	33	Thr-Ala-Asp-Arg-Arg-Ala-Ala-Cys-Asn-Cys-Ser-Leu
Ch2-3	(83-95)	39	Ala-Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg-Val Asn-Asn
Ch2-4	(65-82)	50	Asn-Ala-Gly-Lys-Ala-Ser-Ser-Ile-Pro-Gly-Arg-Cys-Gly-Val-Arg-Leu-Pro-Tyr
Ch3-1	(43-51)	15	Thr-Ala-Asp-Arg-Arg-Ala-Ala-Cys-Asn
Ch3-2	(57-64)	28	Ala-Ala-Ser-Arg-Val-Ser-Gly-Leu
Ch3-3	(37-42)	29	Asn-Ala-Ala-Ala-Arg-Thr
Ch3-4	(9-17)	46	Ser-Ala-Ile-Gly-Pro-Cys-Leu-Ala-Tyr
Ch4-1	(52-54)	5	Cys-Ser-Leu
Ch4-2	(65-70)	7	Asn-Ala-Gly-Lys-Ala-Ser
Ch4-3	(43-51)	15	Thr-Ala-Asp-Arg-Arg-Ala-Ala-Cys-Asn
Ch4-4	(55-64)	32	Lys-Ser-Ala-Ala-Ser-Arg-Val-Ser-Gly-Leu

Table XXI - (Cont.)

Peptide No. and Position in Sequence		Elution Time	Sequence
Ch4-5	(9-17)	47	Ser-Ala-Ile-Gly-Pro-Cys-Leu-Ala-Tyr
Ch4-6	(71-82)	49	Ser-Ile-Pro-Gly-Arg-Cys-Gly-Val-Ala-Arg-Leu-Pro-Tyr

by subjecting the peptides to 2 cycles of degradation by the "Dansyl-Edman" method and by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide Ch1-3 (positions 18 -36)

Ala-Arg-Gly-Ala-Gly-Ala-Ala-Pro-Ser-Ala-Ser-Cys-Gln-Ser-Gly-Val-Arg-Ser-Leu

The cleavage of the peptide bond Ser^{26}-Ala gave rise to peptides Ch1-1 (Ala-Arg-Gly-Ala-Gly-Ala-Ala-Pro-Ser) and Ch2-1 (Ala-Arg-Gly-Ala-Gly-Ala-Ala-Pro-Ser) . The data from semi-quantitative amino acid analysis by the dansyl-chloride method of peptides Ch1-1 and Ch2-1, provide support to the fact that Ser^{28} is the C-terminal residue of both peptides: both peptides lack cysteine, glutamine, arginine, valine and leucine. The peptide corresponding to positions 27 to 36 was not recovered. Peptides Ch1-1 and Ch2-1 are further examples in the present sequence, of cleavage at the carboxyl side of serine by chymotrypsin.

Peptide Ch3-3 (positions 37-42)

Asn-Ala-Ala-Ala-Arg-Thr

Peptide Ch2-2 (positions 43-54)

Thr-Ala-Asp-Arg-Arg-Ala-Ala-Cys-Asn-Cys-Ser-Leu

Peptides Ch3-1 and Ch4-3 (Thr-Ala-Asp-Arg-Arg-Ala-Ala-Cys-Asn) and Ch4-1 (Cys-Ser-Leu) resulted from the cleavage of the peptide bond Asn^{51}-Cys in peptide Ch2-2. Residue 54 was identified as leucine from the specificity of chymotrypsin. Semi-quantitative amino acid analysis by the dansyl-chloride method of peptides Ch2-2 and Ch4-1 provided

further confirmation of the identity of the residue at position 54.

Peptide Ch4-4 (positions 55-64)

Lys-Ser-Ala-Ala-Ser-Arg-Val-Ser-Gly-Leu

Peptide Ch3-2 (positions 57-64)

Ala-Ala-Ser-Arg-Val-Ser-Gly-Leu

Peptide Ch3-2 constitutes another example of the cleavage by chymotrypsin at the carboxyl side of serine.

Peptide Ch2-4 (positions 65-82)

Asn-Ala-Gly-Lys-Ala-Ser-Ser-Ile-Pro-Gly-Arg-Cys-Gly-Val-Arg-Leu-Pro-Tyr

The identity of Ile^{72} was ascertained by comparison with peptides Ch4-6. The cleavage of the peptide bond Ser^{70}-Ser gave rise to peptides Ch4-2 (Asn-Ala-Gly-Lys-Ala-Ser) and Ch4-6 (Ser-Ile-Pro-Gly-Arg-Cys-Gly-Val-Arg-Leu-Pro-Tyr). Again, this constitutes another example of the cleavage by chymotrypsin at the carboxyl side of serine. The identities Ile^{72} and Leu^{80} were ascertained by subjecting the peptides to 2 cycles of degradation by the "Dansyl-Edman" method, by semi-quantitative amino acid analysis by the dansyl-chloride method and by comparison with peptides T2-4, T3-6 and T4-7.

Peptides Ch1-2 and Ch2-3 (positions 83-95)

Ala-Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg-Val-Asn-Asn

The identities of Ile^{84} and Ile^{88} were ascertained by subjecting the peptides to 2 cycles of degradation by the "Dansyl-Edman" method and by semi-quantitative amino acid analysis by the dansyl-chloride method.

The treatment of the α-amylase inhibitor with chymotrypsin led to the expected cleavages at the carboxyl side of tyrosine, leucine and asparagine. In addition cleavages were also observed at the carboxyl side of serine and threonine. Cleavages at threonine residues were also observed in the amino acid sequence determination of the trypsin/α-amylase inhibitor from ragi seeds (this work), and have also been reported before by Croft (1980). Cleavage at the carboxyl side of serine residues have not been reported before.

3.4.5. Staphylococcus aureus (Strain V-8) Protease Digestion

10 mg of the reduced and carboxymethylated protein was used for digestion with the Staphylococcus aureus (Strain V-8) protease, under the conditions described in Section 2.2.6.3.4. The purification of the resulting peptides was performed by fractionating the digest on a Sephadex G-50 column equilibrated in 0.05M ammonium bicarbonate and subsequently subjecting each one of the pooled fractions (Figure 54) to RP-HPLC. The chromatograms of the HPLC separation of peaks 2 and 3 are shown in Figures 55 and 56, respectively. No pure peptides were recovered from the RP-HPLC separation of peak 1. The sequences, elution times and positions in the sequence of each of the peptides are presented in Table XXII.

Peptides V2-4 and V3-4 (Positions 1-20)

Ala-Ile-Ser-Cys-Gly-Gln-Val-Ser-Ser-Ala-Ile-Gly-Pro-Cys-Leu-Ala-Tyr-Ala-Arg-Gly

The identities of Ile^{2}, Ile^{11} and Leu^{15} were ascertained by subjecting the peptides to 2 cycles of degradation by the "Dansyl-Edman" method

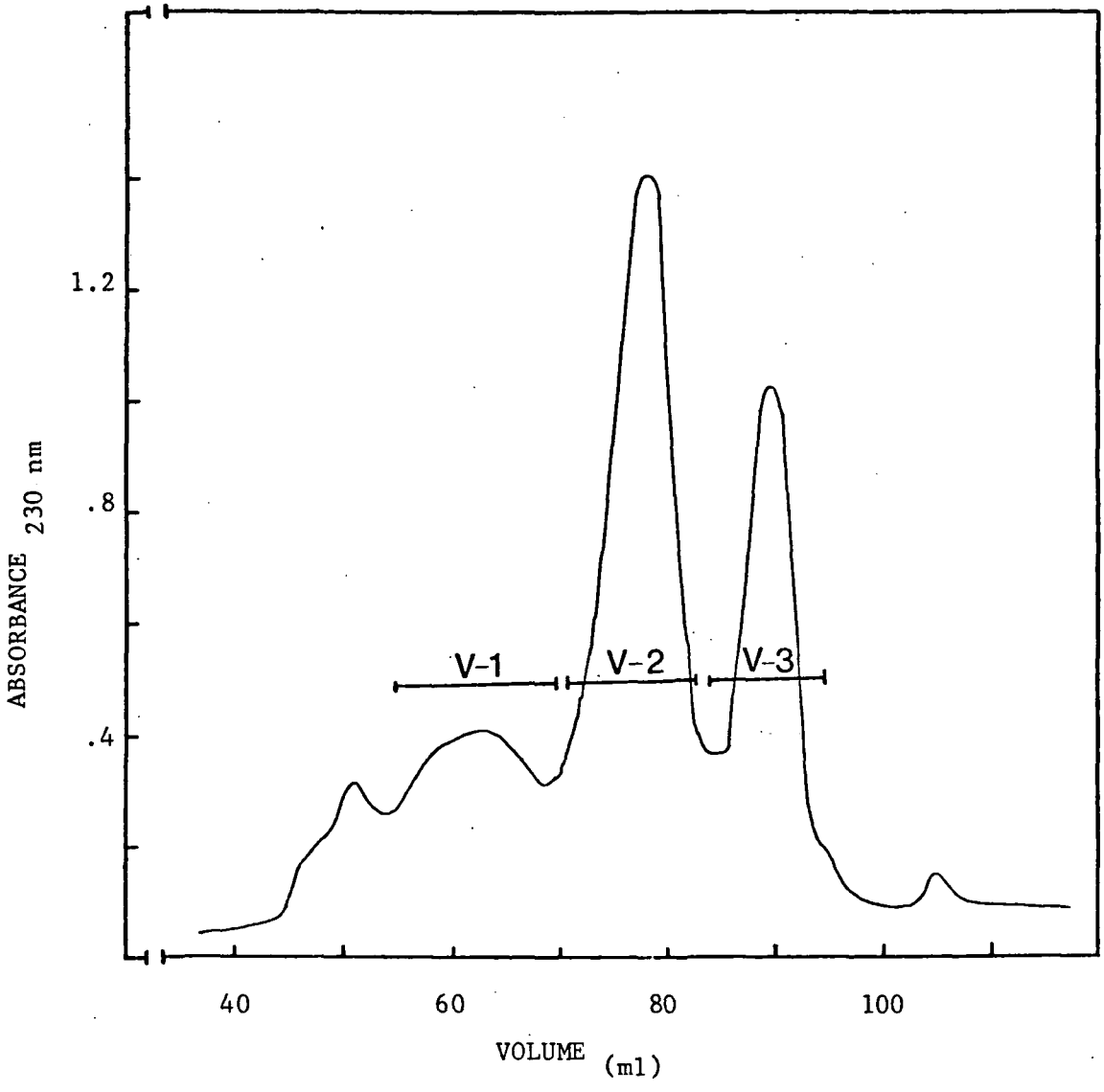


Figure 54 - Chromatography of the products of the digestion by Staphylococcus aureus (strain V-8) protease digestion of the ∝-amylase inhibitor from seeds of ragi (Indian finger millet, Eleusine coracana Gaertn.) on a Sephadex G-50 column (0.8 x 190 cm) in 0.05M ammonium bicarbonate (flow rate 6ml/h., fraction size 1.4 ml).

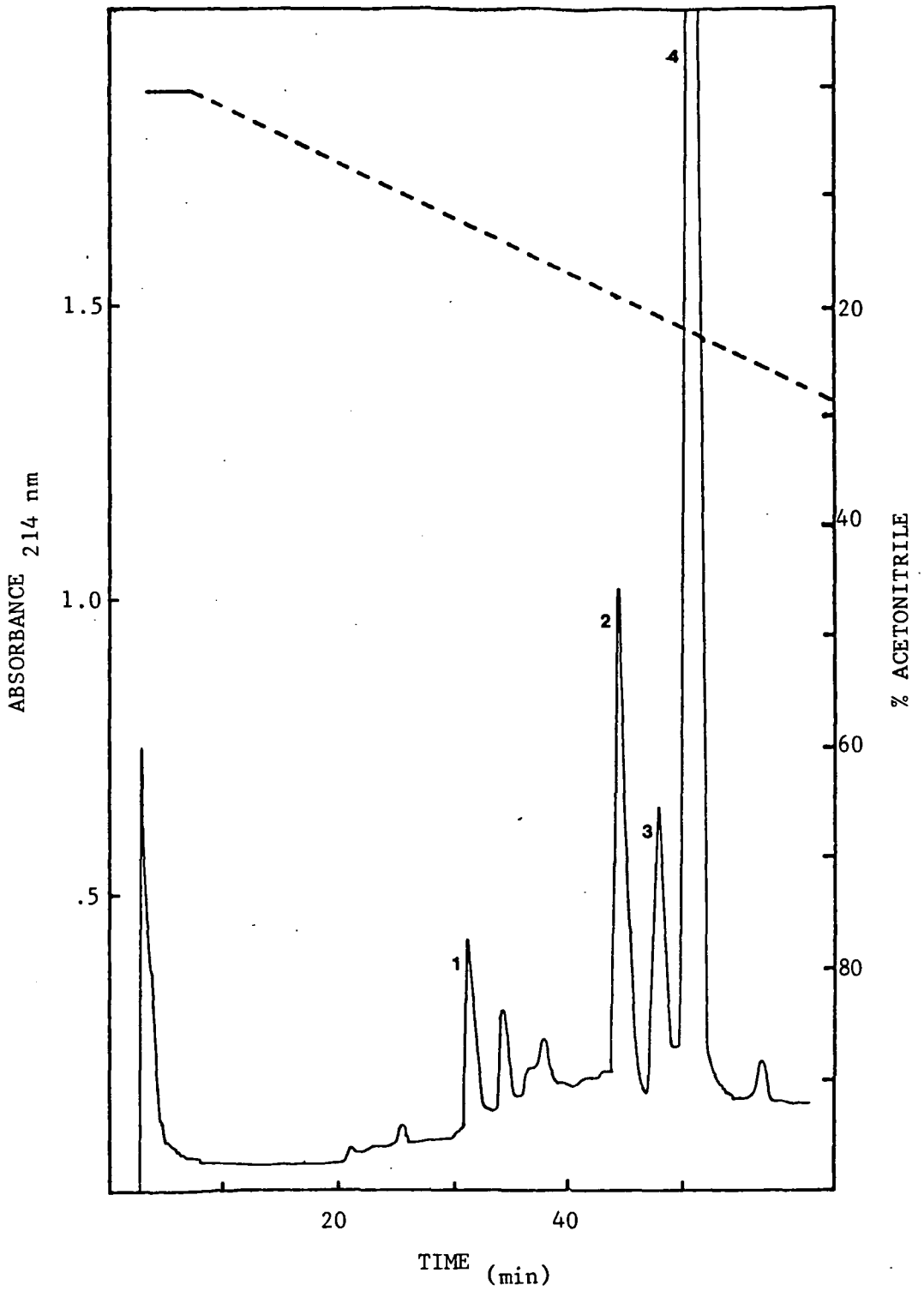


Figure 55 - RP-HPLC separation of fraction V-2 from the Staphylococcus aureus (strain V-8) protease of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions corresponding to peptides which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

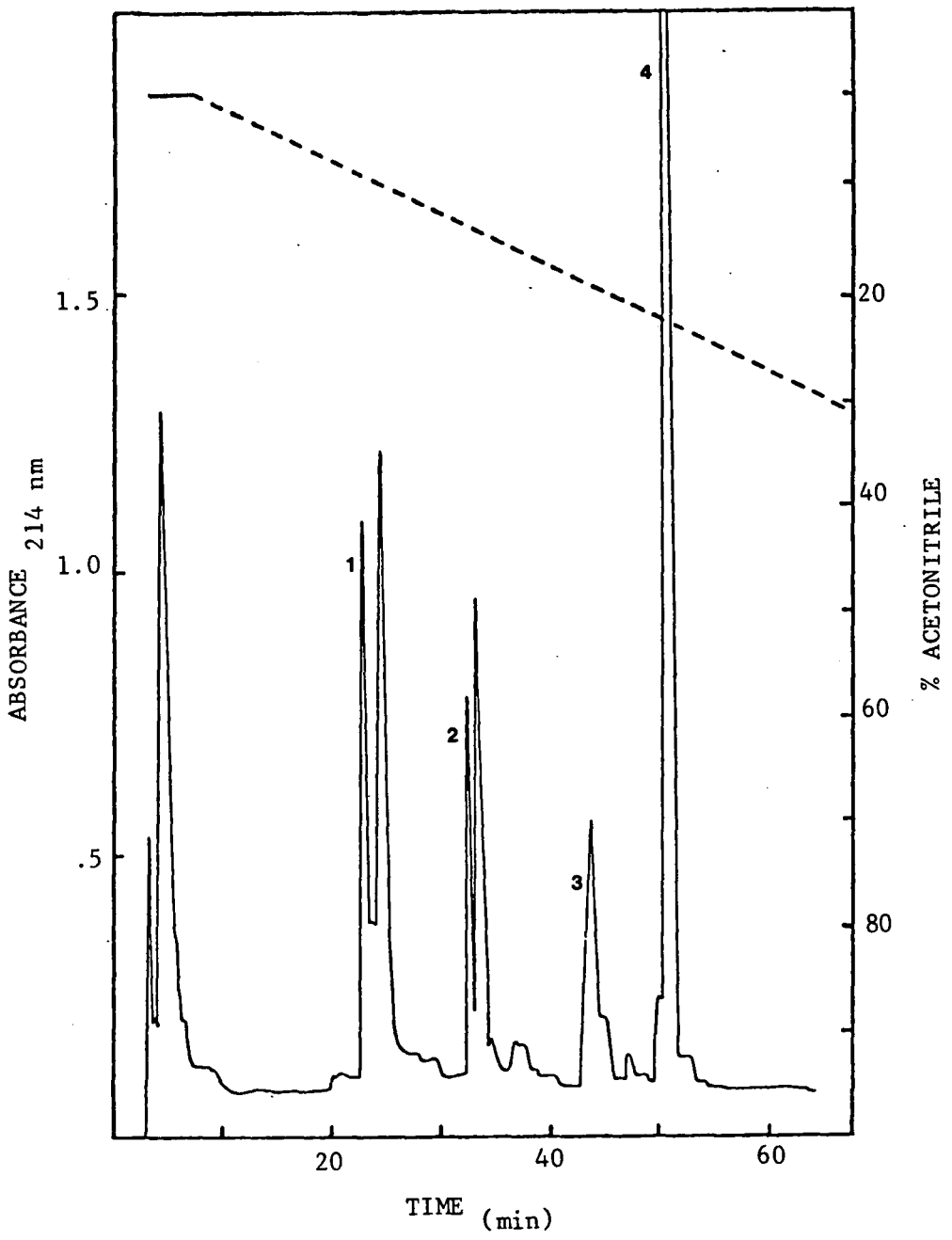


Figure 56 - RP-HPLC separation of fraction V-3 from the Staphylococcus aureus (strain V-8) protease of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) on a C-18 μ-Bondapak column (0.5 x 25 cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions corresponding to peptides which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table XXII - Peptides obtained by the RP-HPLC separation of Sephadex G-50 fractions V2 and V3 from Staphylococcus aureus (strain V-8) protease digestion of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.)

Peptide No. and Position in Sequence	Elution Time	Sequence
V2-1 (23-45)	33	Ala-Ala-Pro-Ser-Ala-Ser-Cys-Gln-Ser-Gly-Val-Arg-Ser-Leu-Asn-Ala-Ala-Ala-Arg-Thr-Thr-Ala-Asp
V2-2 (58-83)	47	Ala-Ser-Arg-Val-Ser-Gly-Leu-Asn-Ala-Gly-Lys-Ala-Ser-Ser-Ile-Pro-Gly-Arg-Cys-Gly-Val-Arg-Leu-Pro-Tyr-Ala
V2-3 (58-83)	50	Ala-Ser-Arg-Val-Ser-Gly-Leu-Asn-Ala-Gly-Lys Ala-Ser-Ser-Ile-Pro-Gly-Arg-Cys-Gly-Val-Arg Leu-Pro-Tyr-Ala
V2-4 (1-22)	52	Ala-Ile-Ser-Cys-Gly-Gln-Val-Ser-Ser-Ala-Ile-Gly-Pro-Cys-Leu-Ala-Tyr-Ala-Arg-Gly-Ala-Gly
V3-1 (86-95)	24	Ala-Ser-Ile-Asp-Cys-Ser-Arg-Val-Asn-Asn
V3-2 (84-95)	32	Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg-Val-Asn-Asn
V3-3 (9-21)	43	Ser-Ala-Ile-Gly-Pro-Cys-Leu-Ala-Tyr-Ala-Arg-Gly-Al
V3-4 (1-22)	51	Ala-Ile-Ser-Cys-Gly-Gln-Val-Ser-Ser-Ala-Ile-Gly-Pro-Cys-Leu-Ala-Tyr-Ala-Arg-Gly-Ala-Gly

and by comparison with peptides Ch3-4 and Ch4-6.

Peptide V3-3 (positions 9-21)

Ser-Ala-Ile-Gly-Pro-Cys-Leu -Ala-Tyr-Ala-Arg-Gly-Ala

The identities of Ile^{11} and Leu^{15} were ascertained by subjecting the peptide to 2 cycles of degradation by the "Dansyl-Edman" method, by semi-quantitative amino acid analysis by the dansyl-chloride method and by comparison with peptides Ch3-4 and Ch4-5.

Peptide V2-1 (positions 21-57)

Ala-Ala-Pro-Ser-Ala-Ser-Cys-Gln-Ser-Gly-Val-Arg-Ser-Leu-Asn-Ala-Ala-Ala-Arg-Thr-Thr-Ala-Asp

The identity of Leu^{36} was ascertained by comparison with peptides T-9, T6-12 and Th-6.

Peptides V2-2 and V2-3 (positions 58-83)

Ala-Ser-Arg-Val-Ser-Gly-Leu-Asn-Ala-Gly-Lys-Ala-Ser-Ser-Ile-Pro-Gly-Arg-Cys-Gly-Val-Arg-Leu-Pro-Tyr-Ala

The identity of Leu^{64} was ascertained by comparison with peptides Th-11, Th-12 and Th-3 and by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide V3-2 (positions 84-95)

Ile-Ser-Ala-Ser-Ile-Asp-Cys-Ser-Arg-Val-Asn-Asn

The identities of Ile^{84} and Ile^{88} were ascertained by N-terminal analysis by the dansyl-chloride method and by semi-quantitative amino acid

analysis by the dansyl-chloride method.

Peptide V3-1 (positions 86-95)

Ala-Ser-Ile-Asp-Cys-Ser-Arg-Val-Asn-Asn

The identity of Ile^{88} was ascertained by subjecting the peptide to 2 cycles of degradation by the "Dansyl-Edman" method.

The observed cleavage sites in the digestion of the α-amylase inhibitor with Staphylococcus aureus (strain V-8) protease, were not consistent with the specificity of the enzyme (Drapeau, 1977). Under the digestion conditions utilized here, only 3 fargments would be expected, as the molecule has only 2 aspartic acid residues and no glutamic acid residues. However, six fragments were recovered and none of them resulted from the cleavage of Asp-X peptide bonds. Instead, the fragments obtained were derived from the cleavage of Ser-Ser, Gly-Ala, Ala-Ala, Ala-Ile and Ser-Ala peptide bonds. Similar non-specific cleavages have previously been reported in the literature. For example, cleavage on the C-terminal side of some glycine and alanine residues was reported in Escherichia coli ribossomal protein L-11 (Dognin and Wittman-Liebold, 1977). Also, Richardson (1979) reported cleavage at the carboxyl side of glycine in the aubergine protease inhibitor. Cleavage at the carboxyl side of serine residues was reported in the β-crystallin Bp chain by Driessen et al (1983). It is not certain whether the non-recovery of the peptides expected to result from the cleavage of the Asp^{45}-Arg and Asp^{89}-Cys peptide bonds was due to the loss of the peptides during the purification of the digest by molecular exclusion chromatography or that the peptides were eluted together from the RP-HPLC column mixed with other peptides or, alternatively, that these peptide bonds were not

actually cleaved. However, in the analysis of the peptides resulting from the RP-HPLC separation of peak 3 (Figure 56), the fraction eluted at 35 minutes was shown to have a peptide with arginine at the N-terminal. Unfortunately this peak was composed of at least 3 different peptides and so their sequences could not be established.

3.4.6. Thermolysin Digestion

5 mg ot the reduced and carboxymethylated inhibitor was subjected to thermolysin hydrolysis, under the conditions described in Section 2.2.6.3.3. The purification of the resulting peptides was done by subjecting the hydrolysate directly to RP-HPLC without previous purification (Figure 5[7]). The sequences, elution times and positions in the sequence of the peptides purified are presented in Table XXIII.

Peptide Th5 (positions 1-6)

Ala-Ile-Ser-Cys-Gly-Gln

The identity of Ile^2 was ascertained by subjecting the peptide to a cycle of degradation by the "Dansyl-Edman" method and by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide Th9 (positions 1-9)

Ala-Ile-Ser-Cys-Gly-Gln-Val-Ser-Ser

The identity of Ile^2 was ascertained by subjecting the peptide to a cycle of degradation by the "Dansyl-Edman" method and by semi-quantitative amino acid analysis by the dansyl-chloride method.

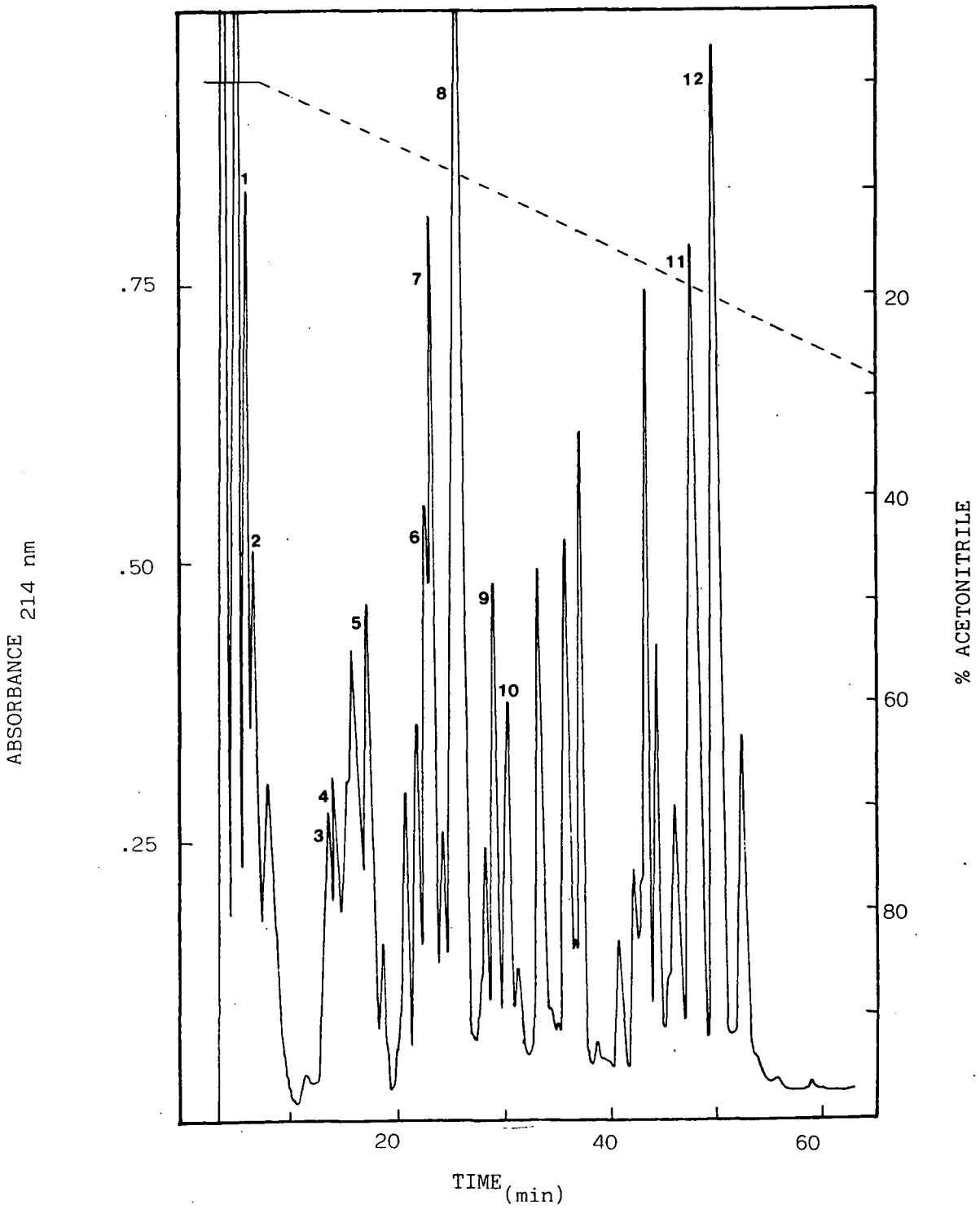


Figure 57 - RP-HPLC separation of peptides obtained by the digestion of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.) with thermolysin. The separation was performed on a C-18 μ-Bondapak column (0.5 x 25cm) with a linear gradient of 70% (v/v) acetonitrile in 0.1% (v/v) trifluoracetic acid at a flow rate of 1.0 ml/min. Fractions which were fully or partially sequenced are numbered. Dotted line indicates concentration of acetonitrile.

Table XXIII - Peptides obtained by the RP-HPLC separation of the thermolysin digest of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana)

Peptide No. and Position in Sequence		Elution Time	Sequence
Th-1	(48-53)	7	Ala-Ala-Cys-Asn-Cys-Ser
Th-2	(36-38)	8	Leu-Asn-Ala
Th-3	(64-68)	13	Leu-Asn-Ala-Gly-Lys
Th-4	(39-47)	14	Ala-Ala-Arg-Thr-Thr-Ala-Asp-Arg-Arg
Th-5	(1-6)	17	Ala-Ile-Ser-Cys-Gly-Gln
Th-6	(23-32)	22	Ala-Ala-Pro-Ser-Ala-Ser-Cys-Gln-Ser-Gly
Th-7	(11-17)	23	Ala-Gly-Pro-Cys-Leu-Ala-Tyr
Th-8	(10-17)	24	Ala-Ile-Gly-Pro-Cys-Leu-Ala-Tyr
Th-9	(1-9)	29	Ala-Ile-Ser-Cys-Gly-Gln-Val-Ser-Ser
Th-10	(18-26)	30	Ala-Arg-Gly-Ala-Gly-Ala-Gly-Ala-Ala-Pro-Ser
Th-11	(64-79)	47	Leu-Asn-Ala-Gly-Lys-Ala-Ser-Ser-Ile-Pro-Gly-Arg-Cys-Gly-Val-Arg
Th-12	(64-79)	50	Leu-Asn-Ala-Gly-Lys-Ala-Ser-Ser-Ile-Pro-Gly-Arg-Cys-Gly-Val-Arg

Peptide Th8 (positions 10-17)

Ala-Ile-Gly-Pro-Cys-Leu-Ala-Tyr

Peptide Th7 (positions 11-17)

Ile-Gly-Pro-Cys-Leu-Ala-Tyr

Peptide Th10 (positions 18-26)

Ala-Arg-Gly-Ala-Gly-Ala-Ala-Pro-Ser

Peptide Th6 (positions 23-32)

Ala-Ala-Pro-Ser-Ala-Ser-Cys-Gln-Ser-Gly

Peptide Th2 (positions 36-38)

Leu-Asn-Ala

The identity of Leu^{36} was ascertained by N-terminal analysis by the dansyl-chloride method and by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptide Th4 (positions 39-47)

Ala-Ala-Arg-Thr-Thr-Ala-Asp-Arg-Arg

Peptide Th1 (positons 48-53)

Ala-Ala-Cys-Asn-Cys-Ser

Peptide Th3 (positons 64-68)

Leu-Asn-Ala-Gly-Lys

The identity of Leu^{64} was ascertained by N-terminal analysis by the

dansyl-chloride method, and by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptides Th11 and Th12 (positions 64-79)

Lue-Asn-Ala-Gly-Lys-Ala-Ser-Ser-Ile-Pro-Gly-Arg-Cys-Gly-Val-Arg
64 72

The identities of Leu and Ile were ascertained by N-terminal amino acid analysis by the dansyl-chloride method and by semi-quantitative amino acid analysis by the dansyl-chloride method.

The observed cleavage sites in the digestion of the α-amylase inhibitor with thermolysin were consistent with the broad specificity of the enzyme. The direct fractionation of the digest by RP-HPLC without prior separation of the components by molecular exclusion chromatography, prevented the purification of many peptides, as some of the peaks collected were composed of more than one peptide (Figure 57.)). However, the peptides which were needed to establish consistent overlapping were recovered, namely Th4, Th10, Th11 and Th11.

The complete amino acid sequence of the ragi α-amylase inhibitor is shown in Figure 58. The only region in the sequence which lacks support from good overlapping segments (3 or more residues) of peptides is the sequence of residues 54 to 58. However, the sequence is entirely compatible with the results of the amino acid analysis of the inhibitor (Table XIII). The molecular weight of 9314 is in good agreement with the value calculated from SDS-polyacrylamide gel electrophoresis (Figure 5a).

The ragi α-amylase inhibitor has an odd number of cysteine residues (7) and in this aspect it resembles the 0./53 α-amylase

inhibitor from wheat (Maeda et al)., 1983) which also has an odd number of cysteine residues. Also notable is the high isoelectric point of this inhibitor (above 10) which reflects its high content of basic residues (9 arginine and 2 lysine residues) and the small number of strongly acidic residues (2 aspartic acid residues). Although the ragi α-amylase inhibitor was reported to possess a 10% content of carbohydrate (Shivaraj and Pattabiraman, 1981), no evidence was found to confirm it during this work. The presence of such a high content of carbohydrate in a protein, would be detected during acid hydrolysis of peptides derived from chemical or enzymatic hydrolysis of the protein.

Another noteworthy feature of the ragi α-amylase inhibitor is its unusually high content of alanine and serine residues, which is unique among trypsin and α-amylase inhibitors. The inhibitor is devoid of methionine, phenylalanine, histidine, and tryptophan residues. It is perhaps worth noting that the 0.28 inhibitor from wheat is also devoid of phenylalanine and histidine and that the 0.53 inhibitor contains only low levels of these amino acids.

3.4.7. Sequence Comparison

The ragi α-amylase inhibitor has no apparent homology with the trypsin/α-amylase inhibitor found in the same tissue or with other α-amylase inhibitors which have been sequenced so far (Aschauer et al., 1981; Kashlan and Richardson, 1981; Maeda et al., 1983; Hejgaard et al., 1983) or with any of the plant proteinase inhibitors (Ramshaw, 1982).

Gene duplication is a phenomenon which is thought to be the main process responsible for the evolution of the majority of the extant proteins and certainly most enzymes (Doolittle, 1981). The presence of internal sequence homology, which may constitute a sign of ancestral gene duplication, has been demonstrated to occur among all members of the Bowman-Birk inhibitor family (Richardson, 1977, 1980, 1981; Laskowski and Kato, 1980; Ryan, 1981) and among members of the pancreatic secretory trypsin inmhibitor (Kazal) family (Laskowski and Kato, 1980). Regarding the ragi α-amylase inhibitor one of the noteworthy features of its sequence is the presence of two regions of similar sequence in the molecule. As illustrated in Figure 59, the sequences between positions 8 and 47 and between 53 and 92 , can be aligned and out of 40 residues in these regions, there are ten pairs of matching residues. However, when the alignment was analysed to find out whether the similarity between these two regions is indicative of homology, i.e. common ancestry, the results were indecisive as the value of z(-2.893) for the alignment was just below the value where the hypothesis of homology is accepted. This does not necessarily mean that gene duplication has not occurred, but merely that the similarity between them is too weak to be statistically significant.

3.5 Prediction of the Secondary Structure of the α -Amylase Inhibitor from Ragi (Indian Finger Millet, Eleusine coracana Gaertn)

The predicted secondary structure of the ragi α-amylase inhibitor is presented in Figure 60. The values of the decision constants (see Section 2.2.8) were chosen after a preliminary predictive run where the decision constants were set to zero. The predicted secondary

dansyl-chloride method, and by semi-quantitative amino acid analysis by the dansyl-chloride method.

Peptides Th11 and Th12 (positions 64-79)

Leu-Asn-Ala-Gly-Lys-Ala-Ser-Ser-Ile-Pro-Gly-Arg-Cys-Gly-Val-Arg

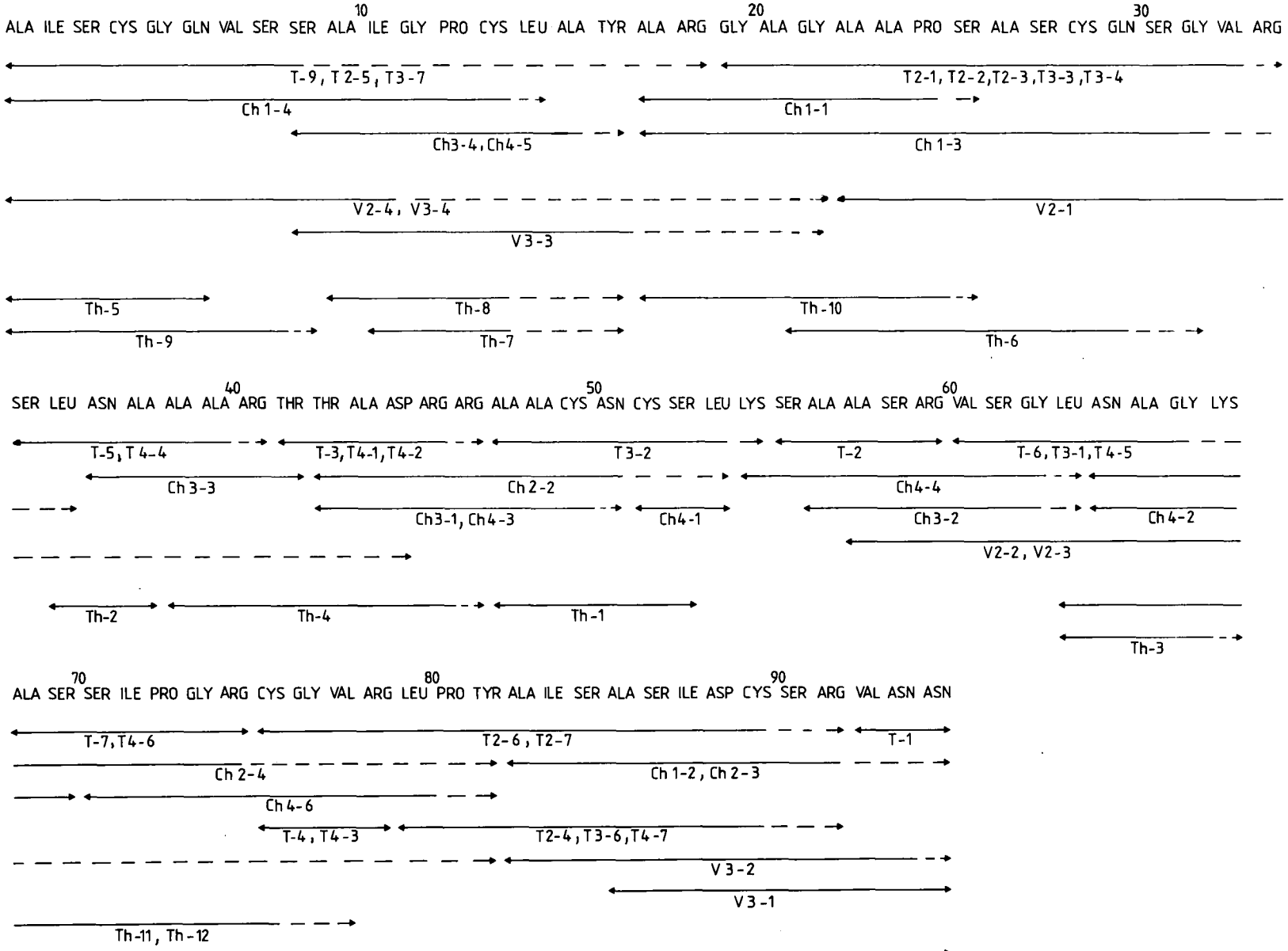
The identities of Leu^{64} and Ile^{72} were ascertained by N-terminal amino acid analysis by the dansyl-chloride method and by semi-quantitative amino acid analysis by the dansyl-chloride method.

The observed cleavage sites in the digestion of the α-amylase inhibitor with thermolysin were consistent with the broad specificity of the enzyme. The direct fractionation of the digest by RP-HPLC without prior separation of the components by molecular exclusion chromatography, prevented the purification of many peptides, as some of the peaks collected were composed of more than one peptide (Figure 57). However, the peptides which were needed to establish consistent overlapping were recovered, namely Th4, Th10, Th11 and Th11.

The complete amino acid sequence of the ragi α-amylase inhibitor is shown in Figure 58. The only region in the sequence which lacks support from good overlapping segments (3 or more residues) of peptides is the sequence of residues 54 to 58. However, the sequence is entirely compatible with the results of the amino acid analysis of the inhibitor (Table XIII). The molecular weight of 9314 is in good agreement with the value calculated from SDS-polyacrylamide gel electrophoresis (Figure 59)

The ragi α-amylase inhibitor has an odd number of cysteine residues (7) and in this aspect it resembles the 0.53 α-amylase

Figure 58 - The complete amino acid sequence of the α-amylase inhibitor. T = tryptic peptides, C = chymotryptic peptides, Th = thermolysin peptides, V = Staphylococcus aureus (strain V-8) peptides. (⟷) peptide sequenced by the DABITC-PITC double coupling method and/or the dansyl-Edman method; (← — →) regions of peptides which were not sequenced or gave unsatisfactory results


10
20
30
ALA ILE SER CYS GLY GLN VAL SER SER ALA ILE GLY PRO CYS LEU ALA TYR ALA ARG GLY ALA GLY ALA ALA PRO SER ALA SER CYS GLN SER GLY VAL ARG
T-9, T2-5, T3-7
T2-1, T2-2, T2-3, T3-3, T3-4
Ch1-4
Ch1-1
Ch3-4, Ch4-5
Ch1-3
V2-4, V3-4
V2-1
V3-3
Th-5
Th-8
Th-10
Th-9
Th-7
Th-6
40
50
60
SER LEU ASN ALA ALA ALA ARG THR THR ALA ASP ARG ARG ALA ALA CYS ASN CYS SER LEU LYS SER ALA ALA SER ARG VAL SER GLY LEU ASN ALA GLY LYS
T-5, T4-4
T-3, T4-1, T4-2
T3-2
T-2
T-6, T3-1, T4-5
Ch3-3
Ch2-2
Ch4-4
Ch3-1, Ch4-3
Ch4-1
Ch3-2
Ch4-2
V2-2, V2-3
Th-2
Th-4
Th-1
Th-3
70
80
90
ALA SER SER ILE PRO GLY ARG CYS GLY VAL ARG LEU PRO TYR ALA ILE SER ALA SER ILE ASP CYS SER ARG VAL ASN ASN
T-7, T4-6
T2-6, T2-7
T-1
Ch2-4
Ch1-2, Ch2-3
Ch4-6
T-4, T4-3
T2-4, T3-6, T4-7
V3-2
V3-1
Th-11, Th-12

inhibitor from wheat (Maeda et al., 1983) which also has an odd number of cysteine residues. Also notable is the high isoelectric point of this inhibitor (above 10) which reflects its high content of basic residues (9 arginine and 2 lysine residues) and the small number of strongly acidic residues (2 aspartic acid residues). Although the ragi α-amylase inhibitor was reported to possess a 10% content of carbohydrate (Shivaraj and Pattabiraman, 1981), no evidence was found to confirm it during this work. The presence of such a high content of carbohydrate in a protein, would be detected during acid hydrolysis of peptides derived from chemical or enzymatic hydrolysis of the protein.

Another noteworthy feature of the ragi α-amylase inhibitor is its unusually high content of alanine and serine residues, which is unique among trypsin and α-amylase inhibitors. The inhibitor is devoid of methionine, phenylalanine, histidine, and tryptophan residues. It is perhaps worth noting that the 0.28 inhibitor from wheat is also devoid of phenylalanine and histidine and that the 0.53 inhibitor contains only low levels of these amino acids.

3.4.7. Sequence Comparison

The ragi α-amylase inhibitor has no apparent homology with the trypsin/ α-amylase inhibitor found in the same tissue or with other α-amylase inhibitors which have been sequenced so far (Aschauer et al., 1981; Kashlan and Richardson, 1981; Maeda et al., 1983; Hejgaard et al., 1983) or with any of the plant proteinase inhibitors (Ramshaw, 1982).

Gene duplication is a phenomenon which is thought to be the main process responsible for the evolution of the majority of the extant proteins and certainly most enzymes (Doolittle, 1981). The presence of internal sequence homology, which may constitutes a sign of ancestral gene duplication, has been demonstrated to occur among all members of the Bowman-Birk inhibitor family (Richardson, 1977, 1980; 1981; Laskowski and Kato, 1980; Ryan, 1981) and among menbers of the pancreatic secretory trypsin inhibitor (Kazal) family (Laskowski and Kato, 1980). Regarding the ragi α-amylase inhibitor one of the noteworthy features of its sequence is the presence of two regions of similar sequence in the molecule. As illustrated in Figure 59 , the sequences between positions 8 and 47 and between 53 and 92, can be aligned and out of 40 residues in these regions, there are ten pairs of matching residues. However, when the alignment was analysed to find out whether the similarity between these two regions is indicative of homology, i.e. common ancestry, the results were indecisive as the value of z (-2.893) for the alignment was just below the value where the hypothesis of homology is accepted. This does not necessarily means that gene duplication has not occurred, but merely that the similarity between them is too weak to be statistically significant.

3.5. Prediction of the Secondary Structure of the α-Amylase Inhibitor from Ragi (Indian Finger Millet, Eleusine coracana Gaertn)

The predicted secondary structure of the ragi α-amylase inhibitor is presented in Figure 60. The values of the decision constants (see Section 2.2.8.) were chosen after a preliminary predictive run where the decision constants were set to zero. The predicted secondary

Ser	Ser	Ala	Ile	Gly	Pro	Cys	Leu	Ala	Tyr	Ala	Arg	Gly	Ala	Gly	Ala
Ser	Leu	Lys	Ser	Ala	Ala	Ser	Arg	Val	Ser	Gly	Leu	Asn	Ala	Gly	Lys

Ala	Pro	Ser	Ala	Ser	Cys	Gln	Ser	Gly	Val	Arg	Ser	Leu	Asn	Ala	Ala
Ala	Ser	Ser	Ile	Pro	Gly	Arg	Cys	Gly	Val	Arg	Leu	Pro	Tyr	Ala	Ile

Ala	Arg	Thr	Thr	Ala	Asp	Arg	Arg
Ser	Ala	Ser	Ile	Asp	Ser	Cys	Arg

Figure 59 - Homologous regions in the amino acid sequence of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.). Identical residues are enclosed in boxes. Minimum distance measure= 34.5; similarity coefficient= 0.25; z= -2.893 (insignificant homology). Parameters values: v=1, u=1, d=1.

Ala	Ile	Ser	Cys	Gly	Gln	Val	Ser	Ser	Ala	Ile	Gly	Pro	Cys	Leu	Ala	Tyr
E	E	E	T	T	T	T	T	T	T	C	C	T	E	E	E	E

Ala	Arg	Gly	Ala	Gly	Ala	Ala	Pro	Ser	Ala	Ser	Cys	Gln	Ser	Gly	Val	Arg
E	T	T	T	C	C	C	C	T	T	T	T	T	T	T	C	E

Ser	Leu	Asn	Ala	Ala	Ala	Arg	Thr	Thr	Ala	Asp	Arg	Arg	Ala	Ala	Cys	Asn
C	C	C	C	C	C	E	C	T	T	T	T	T	T	T	T	T

Cys	Ser	Leu	Cys	Ser	Ala	Ala	Ser	Arg	Val	Ser	Gly	Leu	Asn	Ala	Gly	Lys
T	T	T	T	C	E	E	E	E	E	C	C	C	C	C	T	T

Ala	Ser	Ser	Ile	Pro	Gly	Arg	Cys	Gly	Val	Arg	Leu	Pro	Tyr	Ala	Ile	Ser
C	T	E	E	T	T	T	T	T	E	E	E	E	T	T	E	E

Ala	Ser	Ile	Asp	Cys	Ser	Arg	Val	Asn	Asn
E	E	E	T	T	T	T	T	T	T

Figure 60 - Prediction of the secondary structure of the α-amylase inhibitor from ragi seeds (Indian finger millet, Eleusine coracana Gaertn.). (T): β-turn; (C): random coil; (E): extended chain; (H): α-helix.

structure of the ragi α-amylase inhibitor has no similarities either with the ragi trypsin/ α-amylase inhibitor or with other α-amylase inhibitors so far sequenced. However, notable in the predicted secondary structure is the absence of α-helix and the high content (50%) of β-turn, which at least in part, accounts for the thermal stability of the inhibitor noted by Shivaraj and Pattabiraman (1980, 1981).

3.6. The Complete Amino Acid Sequence of the Protease Inhibitor PI-I from the Tracy Cultivar of Soybean (Glycine max L. Merr)

The seeds of soybean are a very rich source of protein inhibitors of proteinases. On average these proteins comprise 6% of the proteins of dehulled, defatted, soybean meal (David et al., 1977). Two of these inhibitors, the so called Bowman-Birk (Birk et al., 1963) and Kunitz (Kunitz, 1946) inhibitors, were the first proteinase inhibitors ever to be isolated from plants. The Bowman-Birk inhibitor was initially sequenced by Odani et al (1971), but subsequently Odani and Ikenaka (1977a) showed that five proteinase inhibitors (named A, B, C-II, D-II, and E-I) could be isolated from an unnamed soybean cultivar, using CM-cellulose chromatography. Further studies (Odani and Ikenaka, 1976, 1978) showed that the inhibitors A and B were identical with the Bowman-Birk inhibitor (inhibitor B merely being a deamidated form of inhibitor A); whilst inhibitors C-II and D-II were shown to have sequences which were unique but highly homologous to the Bowman-Birk inhibitor. Inhibitor E-1 is a shortened version of D-II (Odani and Ikenaka, 1977b, 1978), lacking the first 9 N-terminal residues (Figure 61).

Other workers have also reported a similar multiplicity of closely related proteinase inhibitors in soybean seeds. For example, David et al. (1977) reported that during a search for proteins with a high content of methionine and half-cystine, they isolated five different but closely related inhibitors of proteinases from seeds of the Tracy cultivar of soybean. These inhibitors were named PI-I, PI-II, PI-III, PI-IV and PI-V. Immunological tests indicated that inhibitors

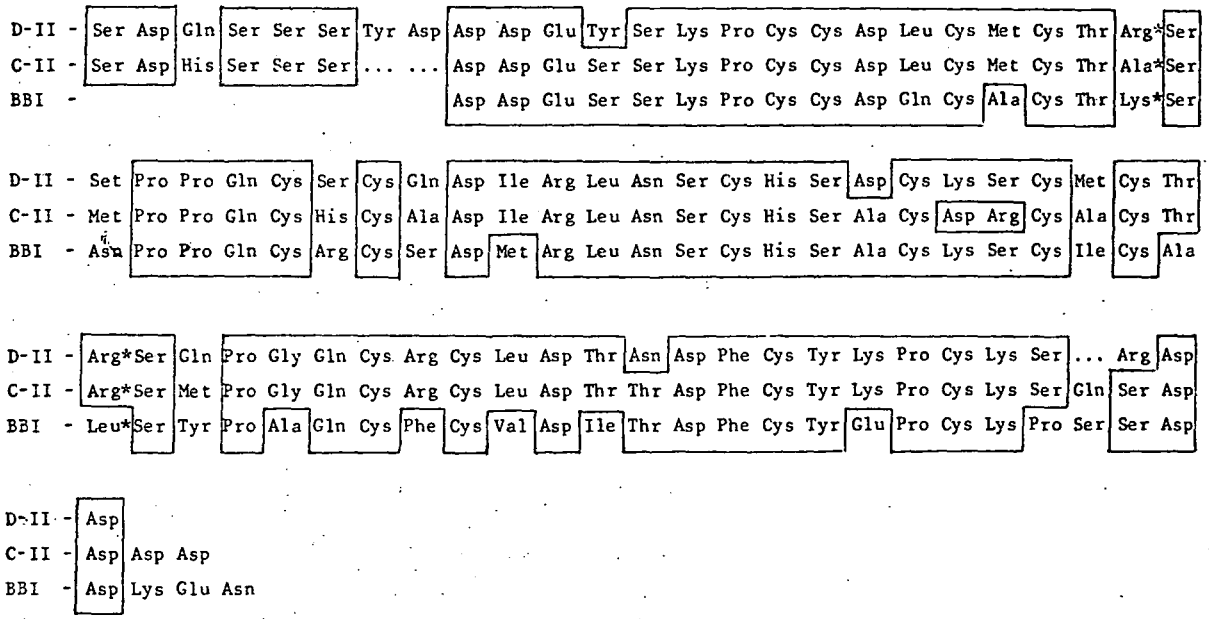


Figure 6 1- Comparison of the amino acid sequences of the soybean (Glycine max) double headed proteinase inhibitors (Odani and Ikenaka, 1976). The reactive sites are shown by asterisks. Identical residues are enclosed in boxes.

PI-I through IV were similar in antigenic properties, but little cross reaction occurred between PI-V and the group PI-I - IV. Both group of inhibitors were double-headed, but only PI-V inhibited trypsin and chymotrypsin. Inhibitors PI-I - IV inhibited only trypsin at two different reactive sites. Furthermore, sequence analysis demonstrated that the N-terminal sequence of the PI-IV inhibitor was the same as the Bowman-Birk inhibitor from the unnamed japanese variety of soybean, and PI-II was shown to have the same N-terminal sequence as inhibitor E-1 (David et al., 1977). Following these studies, a collaborative effort was set up between the laboratories of Dr. Richardson and Dr. Foard, with the aim of establishing the amino acid sequence of inhibitors PI-I, PI-II and PI-IV. Inhibitor PI-IV was shown to have the same sequence as inhibitor D-II by Kashlan (1980), and here the results of the sequencing of inhibitor PI-I are presented.

3.6.1. N-Terminal Analysis

The N-terminal amino acid was determined by the dansyl-chloride method. Two N-terminal amino acids, aspartic acid and glutamic acid, were found. However with the determination of the N-terminal sequence of the native form of the inhibitor by the DABITC-PITC double coupling method, two sequences differing only at the N-terminal amino acid were obtained:

Asp-Glu-Tyr-Ser-Lys-Pro and

Glu-Tyr-Ser-Lys-Pro .

The analysis of the fragments obtained by the trypsin digestion of the reduced and carboxymethylated inhibitor (see below) confirmed that the preparation was contaminated with the inhibitor PI-II (these two

inhibitors differ only in that inhibitor PI-II has an additional amino acid at the N-terminal - see Figure 64).

3.6.2. Tryptic Digestion

The fragments produced by the cleavage of the reduced and carboxymethylated inhibitor (20 mg) with trypsin were separated in a Biogel P-4 column, equilibrated in 0.05M pyridine-acetate buffer, pH 5.7 (Figure 62). Only fraction V could be directly sequenced without further purification, and it was shown to be composed of the peptides

Ser-Arg-Asp-Asp ,

Ser-Arg and

Asp-Asp

The remaining fractions (I - IV) were further purified by high voltage paper electrophoresis at pH 6.5 and pH 1.9. The sequencing of the peptides thus obtained (Table XXIV) was sufficient to establish the almost complete identity of inhibitor PI-I with inhibitors PI-II and D-II (Figures 63 and 61). PI-I differ from PI-II and D-II only by having a deletion at the N-terminal and by a microheterogeneity at position 30 (Ser/Pro) .

In the cultivar Tracy of soybean, the Bowman-Birk inhibitor constitutes 4% and the remaining structurally related inhibitors PI-I to PI-IV constitute about 2% of total protein (Foard et al., 1982). Inhibitors PI-I and PI-II have almost completely homologous sequences; PI-II and PI-IV differ from one another only in that PI-IV has an additional aspartic acid at the N-terminus and a microheterogeneity

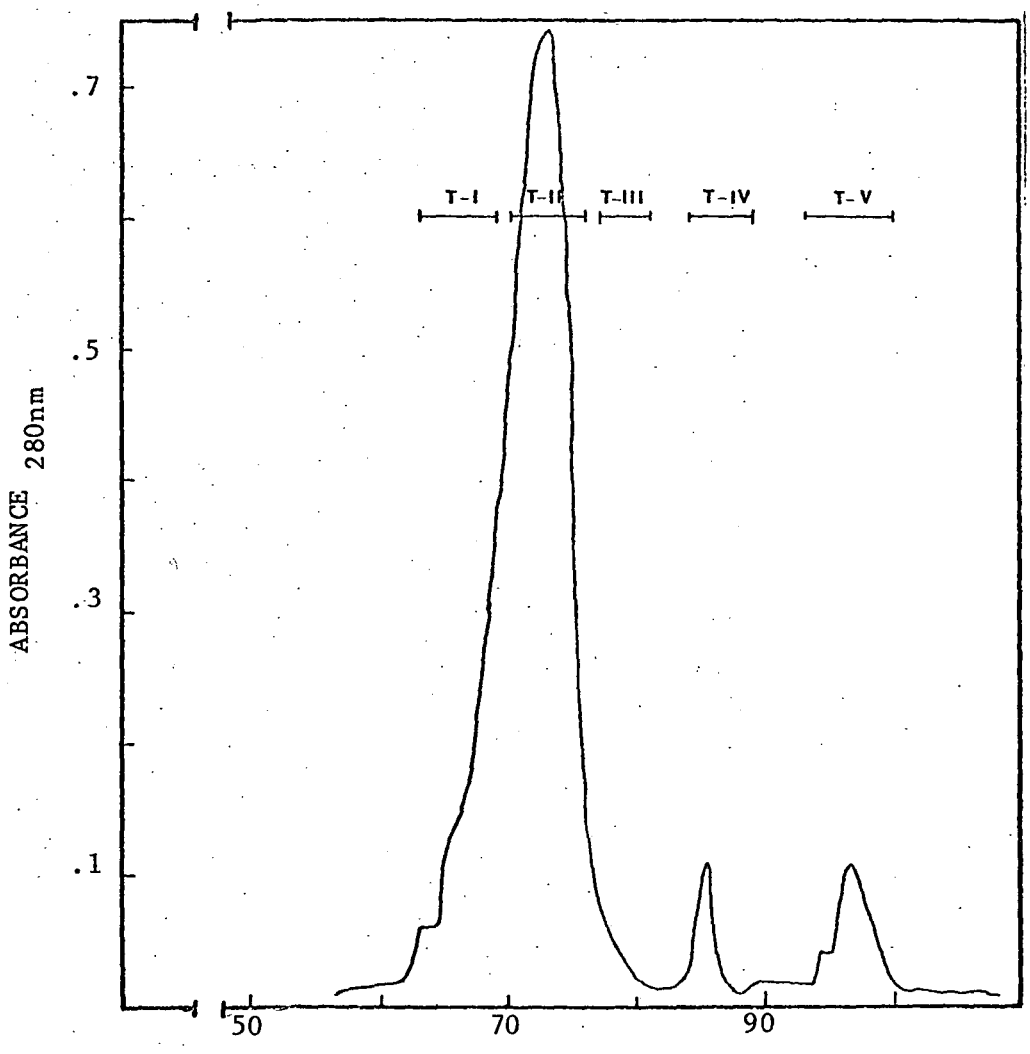


Figure 62 - Chromatography of the tryptic digest of the proteinase inhibitor PI-I from soybean (Glycine max) cultivar Tracy, on a Biogel P-4 column (0.8 x 190cm) in 0.05M pyridine acetate buffer, pH 5.4 (flow rate 6ml/h., fraction size 1.4ml). Those fractions indicated by bars, representing peaks I to V) were pooled and freeze-dried.

Peptide Number	Fraction Number	Mobility pH 6.5	Mobility pH 1.9	Sequence
I	PI-T1	- 0.87	-	Asp Glu Tyr Ser Lys Pro Cys Cys Asp Leu Cys Met Cys Thr Arg
II	PI-T2	- 0.82	-	Glu Tyr Ser Lys Pro Cys Cys Asp Leu Cys Met Cys Thr Arg
III	PI-T1	- 1.86	0.50	Ser Met Pro Pro Gln Cys Ser Cys Glu Asp Ile Arg
IV	PI-T2	- 0.73	-	Leu Asn Pro/Ser Cys His Ser Asp Cys Lys
V	PI-T4	- 0.41	-	Ser Cys Met Cys Thr Arg
VI	PI-T4	0.0	-	Ser Glu Pro Gly Gln Cys Arg
VII	PI-T2	- 1.80	0.57	Cys Leu Asp Thr Asn Asp Phe Cys Tyr Lys Pro Cys Lys
VIII	PI-T5	- 0.21	-	Ser Arg Asp Asp

Table XXIV - Peptides resulting from trypsin digest of soybean (Glycine max) cv. Tracy protease inhibitor PI-I and purified by high voltage paper electrophoresis.

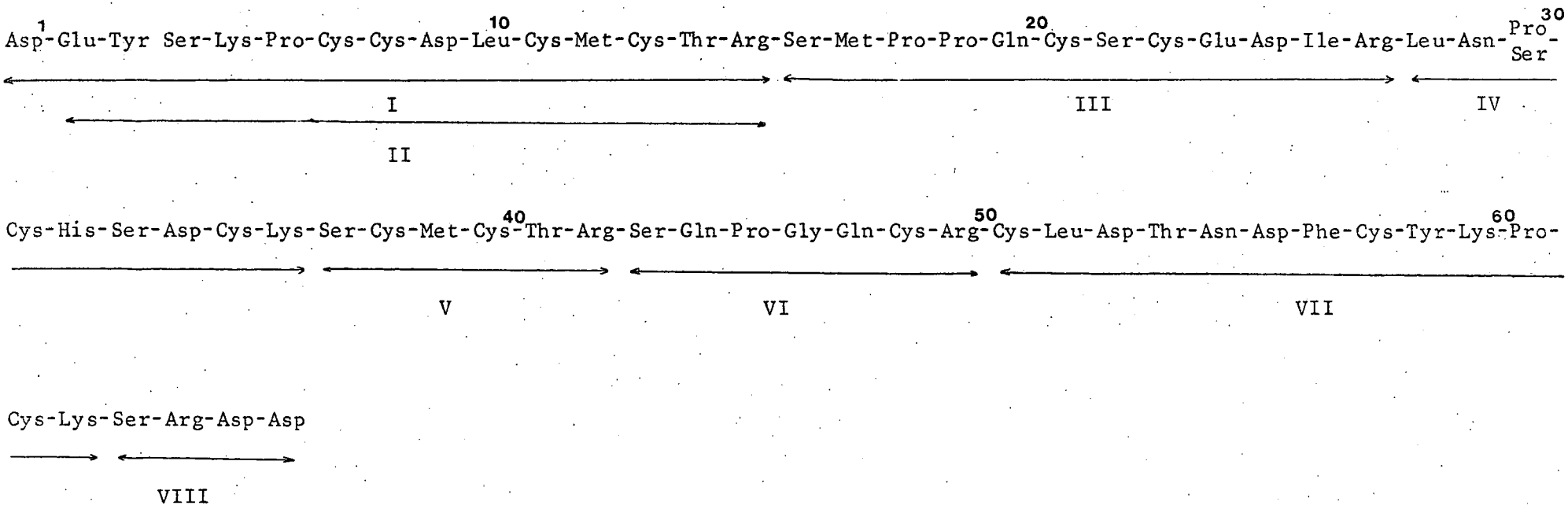


Figure 63 - The complete amino acid sequence of the soybean (Glycine max) cv. Tracy protease inhibitor PI-I.

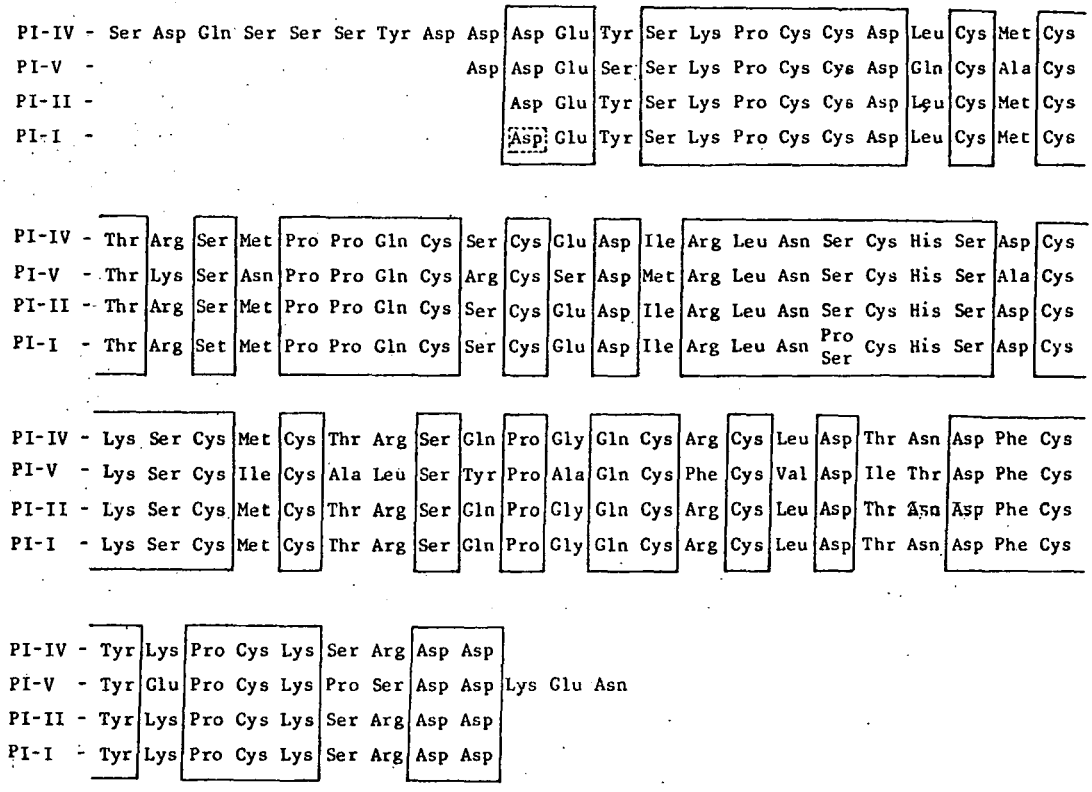


Figure 6 - Comparison of the amino acid sequence of inhibitors PI-IV (Kashlan, 1981), PI-V (David et al., 1977), PI-II (David et al., 1977) and PI-I. Identical residues are enclosed in boxes.

at position 30 (Ser/Pro) (Figure 63). Although the primary structure of PI-III is not yet known, its amino acid composition and the inhibition specificity (David et al., 1977) closely resembles inhibitors PI-I, PI-II and PI-IV).

The possibility of interconversion of these inhibitors through the loss of N-terminal amino acids is evident and indeed the existence of PI-I and PI-II differing only by one amino acid at the N-terminal is a positive indication. Supporting this idea is the fact that in order to get homogeneous samples of inhibitors PI-I to PI-IV for sequencing, it was frequently necessary to recromatograph them when several minor inhibitor bands for each inhibitor were obtained (D.E. Foard - personnal communication). Similar facts were reported by Odani and Ikenaka (1976) for the isolation of the four double headed inhibitors of the japanese cultivar of soybean. Similar indications of interconversion of inhibitors through the loss of N-terminal amino acids was established for inhibitors I and I' and II and II' from Phaseolus vulgaris (Yoshikawa et al., 1979; Kiyohara et al., 1981). However, it remains uncertain whether these interconversions are artifacts of the purification procedure or post-translational modifications or if the different forms of these inhibitors are products of different genes. The presence of multiple forms of the same inhibitor is also common in animals, but at least in the case of the α-1-proteinase inhibitor from rat serum, these multiple forms are derived from the sequential addition of sialic acid to the parental molecule (Ikehara et al., 1981).

Qualitative differences in proteinase inhibitors seem to be common among soybean cultivars. Working with the Fiskerby V cultivar, Wilson and Wilson (1982) and Wilson et al. (1982) found a different

pattern of low molecular weight proteinase inhibitors, despite the fact that they used the same procedure of purification as Odani and Ikenaka (1977) with the japanese cultivar of soybean. As the former workers assayed their inhibitors only through inhibition of trypsin, and immunologically, using antisera raised against the Bowman-Birk inhibitor, it is difficult to decide whether the content of proteinase inhibitors in the cultivar Tracy and FiskerbyV are similar or not. However, despite the general electrophoresis patterns being quite similar in both cultivars (Wilson et al., 1982) the Fiskerby V inhibitors react strongly with antisera raised against the Bowman-Birk inhibitor, while the inhibitors PI-I to PI-IV form a fully cross-reactive group that does not cross react with the Bowman-Birk inhibitor (David et al., 1977; Lin et al., 1980).

In the cultivar Tracy an inhibitor with similar inhibitory specificity as inhibitor C-II is apparently absent. Although inhibitors PI-I to PI-IV are double-headed, their active sites are specific for trypsin only, while C-II can inhibit trypsin, chymotrypsin and elastase; the trypsin and chymotrypsin reactive sites are in different positions in the molecule, but the elastase and chymotrypsin reactive sites are the same (Figure 6). The absence of this type of inhibitor, which may be functionally different from the others inhibitors found in the same seeds, as it has different specificity from them, deserves some comment. Studies of legume proteinase inhibitors in the resting and in the germinating seed have been hindered by the multiplicity of homologous inhibitors in these plants (Lorensen et al., 1981). Furthermore, the uncritical use of the term "isoinhibitor", has led to some confusion as sometimes inhibitors of close structural homology (e.g. C-II, D-II

and the Bowman-Birk inhibitor) but with different specificities towards proteases are called isoinhibitors. The use of this term refering to them is inadequate for two different reasons. Firstly , although they are homologous, their sequences are different and this raises the possibility that their cleavage by proteinases may also be different, and this might allow their differential degradation. Secondly, if their biological function is associated with specific inhibitory activity towards one or more proteinases, the use of the term "isoinhibitor" refering to them, disregards what may be a fundamental difference between them. This multiplicity of inhibitors may allow the precise control of proteinases present in a certain stage of development. This matter is specially relevant since at present the release (Wilson, 1980), degradation (Wilson et al., 1982) and appearence of new forms of inhibitors during germination (Wilson and Chen, 1983) is well established. As the ability of these proteins inhibiting proteinases is thought to be associated with their biological role, the answer of precisely what kind of inhibitory activity is released, degraded or formed is of decisive importance.

4. REFERENCES

Abe, M., Arai, S., Kato, H., and Fujimaki, M. (1980)
Agric. Biol. Chem. 44, 685-686.

Adams, C.A. and Rinne, R.W. (1981)
New Phytol. 89, 1-14.

Ako, H., Foster., R.J., and Ryan, C.A. (1974)
Biochemistry 13 132-138

Allen, G. (1981)
In "Laboratory Techniques in Biochemistry and Molecular Biology". (Work, T.S. and Burden, R.H., eds.) Volume 9, North-Holland Publishing Company, Amsterdan, Oxford and New York.

Anderson, R.S., Ewing, E.E. and Senesac, A.H. (1980)
Plant Physiol. 66, 451-456.

Applebaum, S.W. (1964)
J. Insect Physiol. 10, 897-901.

Archer, B.L. (1983)
Phytochem. 22, 633-639.

Aschauer, H., Vertesy, L., and Braunitzer, G. (1981)
Hoppe-Seyler's Z. Physiol. Chem. 362, 465-467.

Austen, B.M., and Smith, E.L. (1976)
Biochem. Biophys. Res. Comm. 72, 411-417.

Balls, A.K., and Ryan, C.A. (1963)
J. Biol. Chem. 233, 1355-1358.

Barrett, A.J., and Starkey, P.M. (1973)
Biochem J. 133, 709-724.

Baumgartner, B., and Chrispeels, M.J. (1976)
Plant Physiol. 58, 1-6.

Belozerskii, M.A., Dunaevskii, Y.E., and Voskoboinikova, N.E. (1982)
Dokladi Akademii Nauk SSSR 624, 991-993.

Berchtold, P., and Kiesselbach, N.H.K. (1981)
In "Regulators of Intestinal Absorption in Obesity, Diabetes and Nutrition" (Berchtold,P., Cairella, M., Jacobelli,A., and Silano, V., eds.), pp 181-200, Societa Editrice Universo, Roma.

Bernfeld, P. (1965)
Methods Enzymol. I, 149-154.

Bieth, J. (1974)
In "Bayer Symposium V - Proteinase Inhibitors" (Fritz, H., Greene, L.J., and Truscheit, E., eds.) pp 463-469, Springer Verlag, Berlin and New York.

Bieth, J. (1980)
Bull. Europ. Physiopath. Resp. 16, 183-195

Billet, E.E., Wallace, W., and Smith, H. (1978)
Biochim. Biophys. Acta 524, 219-224.

Birk, Y., Gertler, A., and Khalei, J. (1963)
Biochim. Biophys. Acta 67, 326-332.

Blanco-Labra, A., and Iturbe-Chinas, F.A. (1981)
J. Food Biochem. 5, 1-17.

Blumenthal, K.M., Moon, K., and Smith, E.L. (1975)
J. Biol. Chem. 250, 3644-3654.

Bock, W.M., Dongowski, G., and Krause, M. (1975)
Nahrung 19, 411-413

Bolognesi, M., Coda, A., Guarneri, M., and Menegatti, E. (1981)
J. Mol. Biol. 145, 603-605.

Borchers, R., and Ackerson, C.W. (1947)
Arch. Biochem. 13, 291-293.

Brooks, J.C., and Senior, A.E. (1971)
Arch. Biochem. Biophys. 147, 467-472.

Busetta, B., and Hospital, M. (1982)
Biochim. Biophys. Acta, 701, 111-118.

Buonocore, V., Petrucci, T., and Silano, V. (1977)
Phytochemistry 16, 811-820.

Buonocore, V., Gramenzi, F., Pace, W., Petrucci, T., Poerio,E., and Silano, V. (1980)
Biochem. J. 187, 637-645.

Buonocore, V., Poerio, E., and Silano, V. (1981)
In "Regulators of Intestinal Absorption in Obesity, Diabetes and Nutrition" (Berchtold, P., Cairella, Mn., and Silano, V., eds.), pp 87-101, Societa Editrice Universo, Roma.

Campos, F.A.P., and Richardson, M. (1983)
FEBS Lett. 152, 300-304.

Carasco, J.F., and Xavier-Filho, J. (1981)
Ann. Bot. 47, 259-266.

Carlson, G.L., Li, B.U.K., Bass, P., and Olsen, W.A. (1983)
Science, 219, 393-395.

Chandrasekher, G., Raju, D.S., and Pattabiraman, T.N. (1981)
J. Sci. Food Agric. 32, 9-16.

Chang, J.Y., Greaser, E.H., and Bentley, K.W. (1976)
Biochem. J. 153, 607-612.

Chang, J.Y. (1977)
Biochem. J. 163, 517-525.

Chang, J.Y., Brauer, D., and Wittman-Liebold, B. (1978)
FEBS Lett. 93, 205-214.

Chase, T., and Shaw, E. (1970)
Method Enzymol. 19, 20-27.

Chou, P.Y., and Fasman, G. D. (1974)
Biochemistry 13, 222-235.

Chou, P.Y., and Fasman, G.D. (1978)
In "Advances in Enzymology" (Meister, A., ed.), volume 47, pp 45-148, John Wiley and Sons.

Chou, P.Y., and Fasman, G.D. (1979)
Biophys. J. 26, 385-400.

Chrispeels, M.J., and Baumgartner, R. (1978)
Plant Physiol. 61, 617-623.

Chrzaszcz, T., and Janicki, J. (1933)
Biochem J. 26, 354-368.

Chrzaszcz, T., and Janicki, J. (1934)
Biochem. J. 28, 296-304.

Cleveland, T.E., and Black, L.L. (1982)
Plant Physiol. 69, 537-542.

Crestfield, A.M., Moore, S., and Stein, W.H. (1963)
J. Biol. Chem. 238, 622-631.

Croft, L.R. (1980)

Handbook of Protein Sequence Analysis, 2nd. Ed., John Wiley and Sons, Chechester, New York, Brisbane and Toronto.

David, L.R., Hwang, K.T., Lin, D., Yang, W.K., and Foard, D.E. (1971)

Biochim. Biophys. Acta 495, 369-382.

Dayhoff, M.F. (1979)

Atlas of Protein Sequence and Structure, Vol. 5, Suppl. 3, National Biomedical Research Foundation, Washington, D.C.

Dognin, M.J., and Wittman-Liebold, B. (1977)

FEBS Lett. 84, 342-346.

Doolitle, R.F.(1979)

In "The Proteins", Vol. IV, 3rd Ed. (Neurath, H., Hill, R.L., and Barder, eds.) pp 1-118, Acedemic Press, New York.

Doolittle, R.F. (1981)

Science, 214, 149-159.

Drapeau, G.R. (1977)

Methods Enzymol. 47, 189-191

Driessen, H.P.C., Herbrink, P., Bloemendal, H., and Jong, W.W. (1983)

Eur. J. Biochem. 121, 83-91.

Dulloo, R.M. (1980)

Indian J. Exp. Biol. 18, 295-299.

Edman, P. (1950)

Acta Chem. Scand. 4, 283-293.

Edman, P. (1953)

Acta Chem. Scand. 7, 700-701.

Edman, P. (1956)

Acta Chem. Scand. 10, 761-768.

Edman, P., and Begg, G. (1967)

Eur. J. Biochem. 2, 80-91.

Elliot, B.B., and Leopold, A.C. (1953)

Physiol. Plantarum 6, 65-71.

Emmens, H., Welling, G.W., and Beintema, J.J. (1976)

Biochem. J. 157, 317-323.

Erlanger, B.F., Kokowski, N., and Cohen, W. (1961)

Arch. Biochem. Biophys. 95, 271-275.

Fisher, M.L., Anderson, A.J., and Albersheim, P. (1973)

Plant Physiol. 51, 489-493.

Foard, D.E., Gutay, P.A., Ladin, B., Beachy, R.N., and Larkins, S. (1982)

Plant Mol. Biol. 1, 227-243.

Fredericq, L. (1878)

Bull. de l'Acad. Royale des Sciences, des Lettres et des Beau-Arts de Belgique, 47eme annee, 2. Serie, 46, 213-228.

Freed, R.C., and Ryan, D.S. (1978)

F. Food Sci. 43, 1316-1321.

Frerichs, H., Daweke, H., Gries, F.A., Gruneklee, D., Hessing, I.,Jahnke, K., Keup, U., Miss, H., Puls,W., Schmidt, D.P., and Zumfeld, C. (1973)

Diabetologia 9, 68-83

Gaier, J.R., Tulinskii, A., and Liener, I.E. (1981)

J. Biol. Chem. 256, 11417-11419.

Garnier, J., Osguthorpe, D.J., and Robson, B. (1978)

J. Mol. Biol. 120, 97-120.

Gatehouse, A.M.R., Gatehouse, J.A., Dobie, P., Kilminster, A.M., and Boulter, D. (1979)

J. Sci. Food Agric. 30, 948-958.

Gatehouse, A.M.R., Gatehouse, J.A., and Boulter, D. (1980)

Phytochem. 19, 751-756.

Gatehouse, A.M.R., and Boulter, D. (1983)

J. Sci. Food Agric. 34, 345-350.

Goldstein, Z., Trop, M., and Birk, Y. (1973)

Nature, New Biol. 246, 29-31.

Gotoh, O. (1982)

J. Mol. Biol. 162, 705-708.

Graham, J.S., and Ryan, C.A. (1981)

Biochem. Biophys. Res. Comm. 101, 1154-1170.

Granum, P.E. (1978)

J. Food Biochem. 2, 103-120.

Granum, P.E., and Whitaker, J.R. (1977)

J. Food Biochem. 1, 385-401.

Gray, W.R. (1972)

Methods Enzymol. 25, 121-138.

Gray, W.R.,and Hartley, B.S. (1963)

Biochem. J. 89, 379-380.

Green,N.M., and Work, E. (1953)

Biochem. J. 54, 347-352.

Green, T.R., and Ryan, C.A. (1972)

Science 175,776-777.

Greene, L.J., and Bartelt, D.C. (1969)

J. Biol. Chem. 244, 2646-2657.

Ham, W.E., and Sandstedt, R.M. (1944)

J.Biol. Chem. 154, 505-506.

Han, K.K., Richard, C., and Biserte, G. (1983)

Int. J. Biochem. 13, 873-884.

Hartley, B.S. (1960)

Ann. Rev. Biochem. 29, 45-63.

Hass, G.M., and Ryan, C.A. (1980a)

Phytochemistry 19, 1329-1333.

Hass, G.M., and Ryan, C.A. (1980b)

Biochem. Biophys. Res. Comm. 97, 1481-1486.

Hass, G.M., Nau, H., Biemann, K., Grahu, D.T., Ericsson, L.H., and Neurath, L. H. (1975)

Biochemistry 14, 1334-1340.

Heilmann, J., Barollier, J., and Watzke, E. (1957)

Hoppe-Seiler's Z. Physiol. Chem. 357, 1751-1770.

Hejgaard, J., Svendsen, I.B., and Mundy, J. (1983)

Carsberg Res. Comm. 48, 91-94.

Hill, A.F. (1937)

Economic Botany - A Textbook of Useful Plants and Plants Products.

McGraw-Hill Book Company, Inc., New York and London.

Hochtrasser, K., Ilchmann, K., and Werle, E. (1970)

Hoppe-Seyler's Z. Physiol. Chem. 351, 721-728.

Hoff, J., Jones, C.M., Sosa, M.P., and Rodis, P. (1972)

Biochem. Biophys. Res. Comm . 49, 1525-1529.

Horiguchi, T., and Kitagishi, K. (1971)

Plant and Cell Physiol. 12, 907-915.

Ikehara, Y., Miyasato, M., Ogata, S., and Kimimitsu, O. (1981)

Eur. J. Biochem. 115, 253-260.

Ikenaka, T., Odani, S., Sakai, M., Nabeshima, Y., Sato, S., and Murao, S. (1974)

J. Biochem. 76, 1191-1209.

Ilse, D., and Edman, P. (1963)

Aust. J. Chem. 16, 411-416.

Inagami, T., and Sturtvant, J.M. (1960)

J. Biol. Chem. 235, 1019-1023.

Jacob, R.T., Bhat, P.G., and Pattabiraman, T.N. (1983)

Biochem. J. 209, 91-97.

Jaffe, W.G., Moreno, R., and Wallis, V. (1973)

Nutr. Rep. Intern. 7, 169-174.

Janin, J., Sweet, R.M., and Blow, D. (1974)

In "Bayer Symposium V - Proteinse Inhibitors" (Fritz, H., Tschesche, H., Greene, L.J., and Truscheit, E., eds.) pp 513-519, Springer-Verlag, Berlin and New York.

Janzen, D.H. (1980)

J. Ecol. 68, 929-952

Janzen, D.H. (1981a)

Recent Advances in Phytochemistry 15, 211-258

Janzen, D.H. (1981b)

In "Advances in Legume Systematics" (Polhill, R.M., and Raven, P.H., eds.).

Janzen, D.H., Juster, H.B., and Bell, E.A. (1977)

Phytochemistry 16, 223-229.

Jering, H., and Tschesche, H. (1976)

Eur. J. Biochem. 61, 453-457.

Jolly, S.O., and Tolbert, N.E. (1978)

Plant Physiol. 62, 197-201.

Jones, T. D-D., and Gupta, M. (1981)

Biochem. Biophys. Res. Comm. 102, 1310-1316.

Johnston, B.F. (1958)

Staple Food Economies of Western Tropical Africa. Stanford University Press.

Joubert, F.J. (1981)

Hoppe-Seyler's Z. Physiol. Chem. 362, 1515-1521.

Joubert, F.J. (1982)

South African J. Chem. 35, 67-71.

Joubert, F.J. (1983)

Phytochemistry 22, 53-57.

Kabsch, W., and Sander, C. (1983)

FEBS Lett. 155, 179-182.

Kanamori, M., Ibuki, F., Tashiro, M., Yamada, M., and Miyoshi, M. (1975)

J. Nutr. Sci. Vitaminol. 21, 421-428.

Kanchisa, M.I. (1982)

Nucl. Acid Res. 10, 183-196.

Kashlan, N. (1980)

Studies on the Structure and Function of some Enzyme Inhibitors of plants. Ph.D Thesis, University of Durham, England.

Kashlan, N., and Richardson, M. (1981)

Phytochemistry 20,1781-1784.

Kaspar, C.B. (1970)

In "Protein Sequence Determination" (Needleman, S.B., ed.), pp 114-161, Springer-Verlag, Berlin, Heidelberg, New York.

Kassel, P. (1978)

In "Chemistry and Biochemistry of amino acids,peptides and proteins" (Weinstein, B., ed.), Vol. 5, pp 1-29, Marcel Dekker Inc., New York and Basel.

Keilova, H., and Tomasek, V. (1976)

Collect. Czech. Chem. Comm. 41, 489-497.

Kim, B.M., Shimoda, T., and Tunatsu, M. (1977)

J. Fac. Agric. Kyusha Univ. 21, 1-6.

Kirsi, M., and Mikola, J. (1971)

Planta, 96, 281-291.

Kirsi, M., and Mikola, J. (1977)

Physiol. Plant. 39, 110-114.

Kiyohara, T., Yokota, K., Masaki, Y., Matsui, O., Iwasaki, T., and Yoshikawa, M. (1981).

J. Biochem. 90, 721-728.

Kneen, E., and Sandstedt, R.M. (1946)

Arch. Biochem. 9, 235-249.

Kortt, A.A., and Jernyn, M.A. (1981)

Eur.J. Biochem. 115, 551-557.

Kortt, A.A., Tao, Z.J., and Rubira, M. (1983)

Phytochemistry 22, 767-768.

Kowalski, D., Leary, T., McKee, R.E., Sealock, R.W., Wang, D., and Laskowski, M. (1974)

In "Bayer Symposium V - Proteinase Inhibitors" (Fritz, H., Greene, L.J., and Truscheit, E., eds.), pp. 311-324, Springer-Verlag, Berlin and New York.

Kunitz, M. (1945)

Science 101, 668-669

Kunitz, M. (1946)

J. Gen. Physiol. 29, 149-154.

Laemli, U. K. (1970)

Nature 227, 680-685.

Lai, Y.K., and Scandalios, J.G. (1982)

Plant Sci. Lett. 27, 7-20

Lang, J.A., Chang-Hum, L.E., Reyess, P.S., and Briggs, G.M. (1974)

Fed. Proc. 33, 718-721.

Laskowski, M., and Kato, I. (1980)

Ann. Rev. Biochem. 49, 593-626.

Laskowski, M., and Sealock, R.W. (1971)

In "The Enzymes" (Boyer, P., ed.), vol. 3, pp 375-389, Academic Press, New York.

Leary, T.R., and Laskowski, M. (1973)

Fed. Proc. 32, 465-470.

Liener, I.E., and Kakade, M.L. (1969)

In "Toxic Constituents of Plant Foodstuffs" (Liener, I.E., ed.), pp. 7-68, Academic Press, New York.

Liener, I.E., and Kakade, M.L. (1980)

In "Toxic Constituents of Plants Foodstuffs" (Liener, I.E., ed.), 2nd. ed., pp. 7-71, Academic Press, New York.

Lin, K.D., Hwang, D.L., and Foard, D.E. (1980)

J. Chromat. 195, 385-391.

Lorensen, E., Prevosto, R., and Wilson, H.A. (1981)

Plant Physiol. 68, 88-92.

Maeda, K., Hase, T., and Matsubara, H. (1983)

Biochim. Biophys. Acta 743. 52-57.

Maeda, K., Wakabayashi, S., and Matsubara, H. (1983b)
J. Biochem. 94, 865-870.

Maier, K., Muller, H., Tesch, R., Trolp, R., Witt, I., and Holzer, H. (1979)

J. Biol. Chem. 254, 12555-12562.

Majunath, N.H., Veerabhadrappa, P.S. and Virupaksha, T.K. (1983)

Phytochemistry 22, 2348-2358.

Majunder, A.L., and Biswas, B.B. (1973)

Phytochemistry 12, 321-325.

Malhotra, O.P., and Rani, I. (1978)

Indian J. Biochem. Biophys. 15, 229-236.

Marshall, J.J. (1975)

ACS Symposium Series 16, 244-266

Marshall, J.J., and Lauda, C.M. (1975)

J. Biol. Chem. 250, 8030-8037.

Matoo, A.K., and Modi, V.V. (1970)

Enzymologia 39, 237-247.

Matsubara, H. (1970)

Methods Enzymol. 19, 642-651.

McFarland, D.D., and Ryan, C.A. (1974)

Plant Physiol. 54, 706-710.

McPhallen, C.A., Evans, C., Kayakawa, K., Jonassen, I., Svendsen, I., and James, M.N.G. (1983)

J. Mol. Biol. 168, 449-450.

Mikola, J. (1983)

In "Seed Proteins" (Daussant, J., Mosse, J., and Vaughan, J., eds.) Academic Press, New York and London.

Mikola, J., and Andenara, T.M. (1970)

J. Inst. Brew. 76, 182-188.

Mikola, J., and Kirsi, M. (1972)

Acta Che. Scand. 26, 787-795.

Mundy, J., Svendsen, I.B., and Hejgaard, J. (1983)

Carlsberg Res. Comm. 48, 81-80.

Nelson, C.E., Walker-Simmons, M., and Ryan, C.A. (1981)

Plant Physiol.67, 841-844.

Norioka, S., and Ikenaka, T. (1973)

J. Biochem. 93, 479-485.

Odani, S., and Ikenaka, T. (1972)

J. Biochem. 71, 839-848.

Odani, S., and Ikenaka, T. (1973)

J. Biochem. 74, 697-715.

Odani, S., and Ikenaka, T. (1976)

J. Biochem. 80, 641-643.

Odani, S., and Ikenaka, T. (1977a)

J. Biochem. 82, 1513-1522.

Odani, S., and Ikenaka, T. (1977b)

J. Biochem. 82, 1523-1531.

Odani, S., and Ikenaka, T. (1978a)

J. Biochem. 83, 737-745.

Odani, S., and Ikenaka,T. (1978)

J. Biochem. 84, 1-9.

Odani, S., Koide, T., and Ikenaka, T. (1971)

Proc. Japan. Acad. 47, 621-629.

Odani, S., Koide, T., and Ono, T. (1983a)

J. Biol. Chem. 258, 7998-8003.

Odani, S., Koide, T., and Ono, T. (1983b)

J. Biochem. 93, 1701-1704.

Odani, S., Koide, T., Ono, T., and Ohmishi, K. (1983c)

Biochem. J. 213, 543-545.

Odani, S., Ono, T., and Ikenaka, T. (1980)

J. Biochem. 88, 297-301.

O'Donnel, M.D., Fitzgerald, O., and McGeeney, M. (1977)

Clin. Chem. 23, 560-565.

Orland, A.R., Ade, P., DiMaggio, D., Fanelli,C., and Vitozzi, L. (1983)

Biochem. J. 209, 561-564.

Osawa, M., and Laskowski, M. (1966)

J. Biol. Chem. 241, 3955-3961.

Ozols, J., and Gerard, C. (1977)

J. Biol. Chem. 252, 5986-5989.

Puls, W., Keup, U., Krause, H.P., O'Dea, K., and Sitt, R. (1981)
In "Regulators of Intestinal Absorptium in Obesity, Diabetes and Nutritio
(Berchtold, P., Cairella, M., Jacobelli, A., and Silano, V., eds.) pp.
153-180, Societa Editrice Universo, Roma.

Pusztai, A. (1967)
Nutr. Abst. Rev. 37, 1-9.

Pusztai, A. (1968)
Eur. J. Biochem. 5, 252-259.

Pusztai, A. (1972)
Planta 107, 121-127.

Ramshaw, J.A.M. (1982)
In "Nucleic Acids and Proteins in Plants. Part I. Structure, Biochemistry
and Physiology of Proteins" (Boulter, D., and Parthier, B., eds.), pp.
229-290, Springer-Verlag, Berlin, Heidelberg, New York.

Ramshaw, J.A.M., Thompson, E.W., and Boulter, D. (1970)
Biochem. J. 119, 535-541.

Rao, M.N., Shurpalekar, K.S., and Sundarsvalli, D.E. (1967)
Indian J. Biochem. 4, 185-189.

Read, J.W., and Hass, L.W. (1938)
Cereal Chem. 15, 59-68.

Reddy, M.N., Keim, P.S., Heinrickson, R.L., and Kezdy, F.J. (1975)
J. Biol. Chem. 250, 1741-1750.

Reddy, N.R., and Salunkhe, K. (1980)
J. Food Biochem. 4, 273-276.

Reed, C., and Penner, D. (1978)
Agronomy J. 70, 337-339.

Ozumi, T., Ishino, K., Bepper, T., and Arima, K. (1976)
J. Biol. Chem. 251, 2808-2812.

Pace, W., Parlamenti, R., Ron, A., Silano, V., and Vitozzi, L. (1978)
Cereal Chem. 55, 244-252.

Peng, J.H., and Black, L.L. (1976)
Phytopathology 66, 958-963.

Petrucci, T., Sannia, G., Parlamenti, R., and Silano, V. (1978)
Biochem. J. 173, 229-235.

Phillips, D.C., Sternberg, M.J.E., and Sutton, B.J. (1983)
In "Evolution from Molecules to Man" (Bendall, D.S., ed.), pp. 146-173,
Cambridge University Press.

Plunkett, G., and Ryan, C.A. (1980)
J. Biol. Chem. 255, 2752-2758.

Plunkett, G., Senior, D.F., Zuroska, G., and Ryan, C.A. (1982)
Arch. Biochem. Biophys. 213, 463-472.

Podstolski, A. (1981)
Physiol. Plant. 52, 407-410.

Powers, J.R., and Culberston, J.D. (1982)
Journal of Food Protection 45, 655-657.

Powers, J.R., and Whitaker, J.R. (1977)
J. Food Biochem. 1, 217-221.

Preston, K., and Kruger, M. (1976)
Can. J. Plant Sci. 56, 217-221.

Puls, W., and Keup, U. (1973)
Diabetologia 9, 97-101.

Rees, D.C., and Lipscomb, W.N. (1980)

Proc. Natl. Acad. Sci. USA 77, 4633-4637.

Rees, D.C., and Lipscomb, W.N. (1982)

J. Mol. Biol. 160, 475-498.

Richardson, M. (1974)

Biochem. J. 137, 101-112.

Richardson, M. (1977)

Phytochemistry 16, 159-169.

Richardson, M. (1979)

FEBS Lett. 104, 322-326.

Richardson, M. (1980)

Food Chem. 6, 235-253.

Richardson, M. (1-81)

J. Biol. Educ. 51, 178-182

Richardson, M., and Cossins, L. (1974)

FEBS Lett. 45, 11-13.

Richardson, M., and Kashlan, N. (1981)

In "Regulators of Intestinal Absorptium in Obesity, Diabetes and Nutrition" (Berchtold, P., Cairella, M., Jacobelli, A., and Silano, V., eds.), pp. 103-126, Societa Editrice Universo, Roma.

Robson, B. (1974)

Biochem. J. 141, 853-867

Robson, B., and Suzuki, E. (1976)

J. Mol. Biol. 107, 326-356.

Royer, A. (1975)

In "Les Proteines des Grains" (Miege, J., ed.) pp. 169-172, Editions du conservatoire Botanique, Geneva.

Royer, A., Miege, M.N., Grange, A., Miege, J., and Mascherpa, J.M. (1974)
Planta 119, 1-16.

Ruhlman, A., Kukla, D., Schwager, P., Bartels, K., and Huber, R. (1973)
J. Mol. Biol. 77, 417-436.

Ryan, C.A. (1973)
Ann. Rev. Biochem. 24, 173-196.

Ryan, C.A. (1979)
In "Herbivores: Their Interactions with Metabolites" (Rosenthal, G.A., and Janzen, D.H., eds.) pp. 599-618, Academic Press, New York.

Ryan, C.A. (1981)
In "Biochemistry of Plants, Volume 6 - Proteins and Nucleic Acids" (Marcus, E., ed.), pp. 351-370, Academic Press, New York.

Ryan, C.A., and Walker-Simmons, M. (1981)
In "Biochemistry of Plants, Vol. 6 - Protein and Nucleic Acids" (Marcus, E., ed.), pp. 321-347, Academic Press, New York.

Ryan, C.A., Hass, G.M., and Kuhn, R.W. (1974)
J. Biol. Chem. 249, 5495-5499.

Salmia, M.A. (1981)
Physiol. Plant. 51, 253-258.

Salmia, M.A., and Mikola, J. (1980)
Physiol. Plant. 48, 126-130.

Sakato, K., Tanaka, H., and Masawa, M. (1975)
Eur. J. Biochem. 55, 211-219.

Sandstedt, R.M., and Beckford, O.C. (1946)
Cereal Chem. 23, 548-555.

Saunders, R.M., and Lang, J.A. (1973)

Phytochemistry 12, 1237-1240.

Savaiano, D.A., Powers, J.R., Costello, M.J., Whitaker, J.R., and Clifford, A.J. (1977)

Nutr. Reports Int. 15, 443-449.

Scherchter, I., and Berger, A. (1967)

Biochem. Biophys. Res. Comm. 27, 151-162.

Sealock, R.W., and Laskowski, M. (1969)

Biochemistry, 8, 3703-3710.

Seemuller, U., Eulitz, M., Fritz, H., and Strobe, A. (1980)

Hoppe-Seyler's Z. Physiol. Chem. 361, 1841-1846.

Shain, Y., and Mayer, A.M. (1968)

Phytochemistry 7, 1491-1498.

Shainkin, R., and Birk, Y. (1970)

Biochim. Biophys. Acta 221, 502-513.

Sharief, F.S., and Li, S.S.L. (1982)

J. Biol. Chem. 257, 14573-14579.

Sharma, K.K., and Pattabiraman, T.N. (1980)

J. Sci. Food Agric. 31, 981-991.

Sharma, K.K., and Pattabiraman, T.N. (1982)

J. Sci. Food Agric. 33, 255-262.

Shivaraj, B., and Pattabiraman, T.N. (1976)

Arogya - J. Health Sci. 2, 131-135.

Shivaraj, B., and Pattabiraman, T.N. (1980)

Indian J. Biochem. Biophys. 17, 181-185

Shivaraj, B., and Pattabiraman, T.N. (1981)
Biochem J. 193, 29-36.

Shivaraj, B., Rao, N., and Pattabiraman, T.N. (1982)
J. Sci. Food Agric. 33, 1080-1091.

Silano, V. (1978)
Cereal Chem. 55, 722-731.

Silano, V., Pocchiari, F., and Kasarda, D.D. (1973)
Biochem. Biophys. Acta 317, 139-148.

Silano, V., Poerio, E., and Buonocore, V. (1975)
Mol. Cell. Biochem. 18, 87-90.

Smith, I. (1958)
Chromatographic and Electrophoretic Techniques 1, 7-35.

Smith, T.F., Waterman, M.S., and Fitch, W.W. (1981)
J. Mol. Evol. 18, 38-46.

Sorenson, J.C., and Scandalios, J.G. (1980)
Plant Physiol. 66, 688-691.

Steers, E., Craven, G.R., and Anfinsen, C.R. (1965)
J. Biol. Chem. 240, 2478-2485.

Steffens, R., Fox, F.R., and Kassel, B. (1978)
J. Agric. Food Chem. 26, 170-175.

Subbulakshmi, G., Kumar, K.B., and Venkataraman, L.V. (1976)
Nutr. Rep. Intern. 3, 23-29.

Sumathi, R., and Pattabiraman, T.N. (1975)
Indian J. Biochem. Biophys. 12, 383-385.

Suguira, M. (1974)

Japanese Patent, No. 7437,275; cited in Chem. Abstr. 82, 134930b. (1975).

Surendranathan, K.K., and Nair, P.M. (1978)

Plant Sci. Lett. 12, 169-176.

Svendsen, I., and Jonassen, I.B. (1980)

Carlsberg Res. Comm. 45, 389-395.

Svendsen, I., Boisen, S., and Hejgaard, J. (1982)

Carlsberg Res. Comm. 47, 45-53,

Szopa, J., and Wagner, K.G. (1980)

Eur. J. Biochem. 111, 211-215.

Thompson, E.W., Laylock, M.W., Ramshaw, J.A., and Boulter, D. (1970)

Biochem. J. 117, 183-188.

Tlomak, P., and Nowak, K. (1981)

Acta Biochimica Polonica 28, 241-251.

Travis, J., and Salvesen, G.J. (1983)

Ann. Rev. Biochem. 52, 655-709.

Vanlelyveld, L.J., and Bester, A.I. (1978)

Phytopathol. Z. 93, 69-73.

Vartak, H.G., Rele, M.V., and Jagannatham, V. (1980)

Arch. Biochem. Biophys. 204, 134-140.

Vesely, D.L., Graves, W.R., Lo, T.M., Fletcher, M.A. and Levey, G.S. (1977)

Biochem. Biophys. Res. Comm. 77, 1295-1299.

Walker-Simmons, M., and Ryan, C.A. (1977)

Plant Physiol. 60, 61-63.

Walsh, K (1975) In "Proteinases and Biological Control" (E. Reich, D.B. Rifkin, and E. Shaw, Eds) pp. 1-11, Cold Spring Harbor, New York.

Warchalewski, J.R. (1977)

Bull. Acad. Polon. Sci., Ser. Sci. Biol. 25, 725-730.

Warchalewski, J.R. (1983)

Die Nahrung 27, 103-117

Waterman, M.S., Smith, T.F., and Beyer, W.A. (1976)

Adv. Math. 20, 367-387

Weiel, J., and Hapner, K.D. (1976)

Phytochemistry 15, 1885-1887.

Weselake, R.J., MacGregor, W., and Hill, R.D. (1983)

Plant Physiol. 72, 809-812.

Whitaker, J.R. (1981)

In "Impact of Toxicology on Food Processing" (Ayres, J.C., and Kirschman, J.C., eds.) pp. 57-104, Avi Publishing Company Ins., Westport, Connecticut.

Wilson, A.L.T., and Wilson, K.A. (1982)

Phytochemistry 21, 1547-1551.

Wilson, A.L.T., Rightmire, B.R., and Wilson, K.A. (1982)

Plant Physiol. 70, 493-497.

Wilson, K.A. (1980)

Phytochemistry 19, 25-17-2519.

Wilson, K.A., and Chen, J.C. (1983)

Plant Physiol. 71, 493-497

Wilusz, T., Wieczorek, M., Polanowski, A., Denton, A., Cook, J., and Laskowski, M. (1983)

Hoppe-Seyler's Physiol. Z. Physiol. Chem. 364, 93-95

Wrigley, C.W. (1976)

In "Isoelectric Focusing" (Catsimpolas, N., ed.) pp. 93-117, Academic Press, New York.

Xavier-Filho, J. (1973)

Physiol. Plant. 28, 149-155.

Xavier-Filho, J. (1974)

J. Food Sci. 39, 422-426.

Xavier-Filho, J., and Ainouz, I.L. (1977)

Biol. Plant. 19, 183-189.

Xavier-Filho, J., and Negreiros, A.N.M. (1979)

Zeitschrift fur Pflanzenphysiologie 95, 423-430.

Yetter, M.A., Saunders, R.M., and Boles, H.P. (1979)

Cereal Chem. 56, 243-244.

Yoshikawa, M., Kiyohara, T., Iwasaki,T., Ishi, Y., and Kimura, N. (1979)

Agric. Biol. Chem. 43, 787-796.

Yoshikawa, M., Iwasaki, T., Fujii, M., and Ogaki, M. (1976)

J. Biochem. 79, 765-773.

Yoshimoto, T., Walter, R., and Tsuru, D. (1980)

J. Biol. Chem. 255, 4786-4792.

Yung, B.Y.K., and Trowbridge, C.G. (1980)

J. Biol. Chem. 255, 9724-9730.

Zhang, Y., Luo, S., Tan, F., Chi, C., Xu, L., and Zhang, A. (1982) Scientia Sinica 25, 268-276.