

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

THE SEQUENCES OF THE TRYPTIC PEPTIDES
FROM ACTINIDIN

A thesis presented in partial fulfillment
for the degree of Doctor of Philosophy
at Massey University.

Alan Carne

December, 1976

ABSTRACT

Actinidin is a plant thiol protease which is isolated from the fruit of Actinidia chinensis, the chinese gooseberry. Determination of the primary amino acid sequence of actinidin was undertaken to extend the limited structural information available on this group of enzymes, and therefore enable a better understanding of their physical and chemical properties.

The order of arrangement of the 220 amino acid residues in the primary sequence of actinidin was determined from the sequences of the tryptic peptides. S-carboxy[$^{14}\text{C}_2$]methyl actinidin was digested with trypsin, and the twelve tryptic peptides produced were initially separated into seven fractions by gel chromatography on Sephadex G-50. The first four fractions contained tryptic peptides that were purified by DEAE-cellulose chromatography. The last three fractions contained peptides that were sufficiently small to enable purification by paper techniques, and these peptides were sequenced directly by the dansyl-Edman method.

Further degradation of the tryptic peptides purified on DEAE-cellulose with either chymotrypsin, thermolysin, pepsin or Staphylococcus aureus V8 protease was necessary to provide smaller peptides that could be sequenced by the dansyl-Edman method. S. aureus V8 protease was particularly useful in the determination of amide residues, because of the enzyme specificity for the carboxyl groups of glutamic acid.

The fourth tryptic peptide in the sequence of actinidin could not be located in the tryptic peptide elution profiles of either the Sephadex G-50 or DEAE-cellulose columns. The sequence of this peptide was determined from a tryptic peptide obtained by digestion of maleylated carboxymethyl actinidin.

The N-terminal of actinidin was determined by the dansyl Edman method, and the C-terminal by analysis of cyanogen bromide fragments, and by digestion with carboxypeptidase A.

Radioactively labelling the active site cysteine residue with iodo[$^{14}\text{C}_2$]acetic acid, and subsequent purification of the radioactive tryptic digest peptide, enabled the isolation of the tryptic peptide containing the active site cysteine residue. Further digestion of this peptide with chymotrypsin and determination of the sequence of the smaller radioactive peptide, provided the sequence about the active site cysteine residue.

Alignment of the tryptic peptides to reconstruct the primary sequence of actinidin was accomplished with information from cyanogen bromide fragments, information from tryptic peptides of maleylated carboxymethyl actinidin, and information from the three dimensional X-ray crystallographic structure of actinidin determined by Dr E.N. Baker.

The low proportion of basic residues and high proportion of acidic residues in actinidin are in agreement with the enzyme being an acidic protein. Colorimetric analysis of the tryptophan residue content, using 2-nitrophenylsulphenyl chloride, confirmed the presence of six tryptophan residues in the sequence of actinidin.

The amino acid sequences about the seven cysteine residues and the single histidine residue in actinidin were very similar to the analogous sequences in papain and other plant thiol proteases. Furthermore, comparison of the primary sequence of actinidin with that of papain, and the fragments of sequence available for other plant thiol proteases, indicated a considerable homology throughout the sequences of these proteins.

ACKNOWLEDGEMENTS

I wish to thank my supervisor Dr C.H. Moore for his continued interest in the project, and for the benefit of his experience in the form of helpful suggestions, throughout the course of this work.

The useful discussions with Dr G.G. Midwinter and Dr E.N. Baker during the project are much appreciated, as is the assistance of Mr P.R. Macdonald and Dr I.M. Morrison with proof reading of the thesis.

I wish also to acknowledge the encouragement given by my parents and Pauline during the course of this work.

CONTENTS

	Page
Abstract	i
Acknowledgements	iii
Contents	iv
List of Figures	viii
List of Tables	xi
Abbreviations	xiv
CHAPTER I	INTRODUCTION
1.1.	Actinidin and its relation to other proteases. 1.
1.2.	Determination of a protein primary sequence. 7.
CHAPTER II	MATERIALS AND METHODS
	MATERIALS 14.
	METHODS 16.
Section 2.1.	Protein purification. 16.
2.1.1.	Enzyme assay. 17.
2.1.2.	Gel electrophoresis of purified actinidin. 18.
Section 2.2.	Chemical modification reactions.
2.2.1.	Reduction and S-carboxymethylation of actinidin. 18.
2.2.2.	Reaction of cyanogen bromide with actinidin. 19.
2.2.3.	Reaction of maleic anhydride with actinidin. 19.
Section 2.3.	Enzyme digestion of actinidin.
2.3.1.	Trypsin. 20.
2.3.2.	Chymotrypsin, thermolysin and <u>S. aureus</u> V8 protease. 21.
2.3.3.	Pepsin. 21.
2.3.4.	Carboxypeptidase A. 21.
Section 2.4.	Column chromatography.
2.4.1.	Sephadex gel filtration. 22.
2.4.2.	DEAE-cellulose chromatography. 23.
Section 2.5.	Peptide purification by paper techniques.
2.5.1.	High voltage paper electrophoresis. 24.

CONTENTS (contd.)

	Page
2.5.2. Paper chromatography.	25.
2.5.3. Peptide staining on paper.	25.
2.5.3.1. Ninhydrin-cadmium reagent.	26.
2.5.3.2. Chlorine-tolidine reagent.	26.
2.5.3.3. Specific amino acid staining on paper.	27.
(a) Ehrlich reagent for tryptophan.	
(b) Arginine reagent.	
(c) Pauly reaction for histidine and tyrosine.	
2.5.4. Peptide elution from paper.	28.
Section 2.6. Amino acid analysis.	
2.6.1. Determination of the amino acid composition of peptide and protein samples.	28.
2.6.2. Determination of tryptophan.	29.
Section 2.7. Determination of radioactivity.	
2.7.1. Liquid scintillation spectrometry.	29.
2.7.2. Radioautography.	30.
Section 2.8. Peptide sequencing.	
2.8.1. Dansyl N-terminal procedure for peptides.	30.
2.8.2. Dansyl-Edman procedure for peptide sequencing.	31.
2.8.3. Rapid Edman degradation.	32.
2.8.4. Determination of amide groups.	33.
Section 2.9. Determination of the active site cysteine residue in actinidin.	33.
 CHAPTER III	
RESULTS	
Section 3.1. Purification of actinidin.	35.
Section 3.2. Amino acid composition of actinidin.	38.
Section 3.3. Purification of the tryptic peptides of actinidin.	40.
Section 3.4. The sequences of the tryptic peptides.	47.
3.4.1. Peptide T ₁ .	49.
3.4.2. Peptide T ₂ .	50.
3.4.3. Peptide T ₃ .	51.
3.4.4. Peptide T ₄ .	59.
3.4.5. Peptide T ₅ .	67.
3.4.6. Peptide T ₆ .	68.

	Page
3.4.7. Peptide T ₇ .	78.
3.4.8. Peptide T ₈ .	86.
3.4.9. Peptide T ₉ .	92.
3.4.10. Peptide T ₁₀ .	94.
3.4.11. Peptide T ₁₁ .	95.
3.4.12. Peptide T ₁₂ .	98.
Section 3.5. Determination of the C-terminal region of actinidin.	
3.5.1. Determination of the C-terminal region of actinidin from the cyanogen bromide fragments.	99.
3.5.2. Confirmation of the C-terminal region of actinidin with carboxypeptidase A.	102.
Section 3.6. Determination of the amino acid sequence about the active site cysteine residue in actinidin.	103.
CHAPTER IV DISCUSSION	
Section 4.1. The purification of actinidin.	104.
Section 4.2. The amino acid composition of actinidin.	104.
Section 4.3. Purification of the tryptic peptides of actinidin.	105.
Section 4.4. Determination of the amino acid sequence of each tryptic peptide from actinidin.	107.
Peptide T ₁ .	108.
Peptide T ₂ .	108.
Peptide T ₃ .	109.
Peptide T ₄ .	110.
Peptide T ₅ .	113.
Peptide T ₆ .	113.
Peptide T ₇ .	116.
Peptide T ₈ .	118.
Peptide T ₉ .	119.
Peptide T ₁₀ .	120.
Peptide T ₁₁ .	121.
Peptide T ₁₂ .	122.
Section 4.5. The C-terminal region of actinidin.	122.

CONTENTS (contd.)

	Page
Section 4.6.	124.
Section 4.7.	128.
Section 4.8.	131.
REFERENCES	135.

LIST OF FIGURES

Figure		Page
1.1.	Proteolytic enzyme groups.	1.
1.2.	Amino acid sequences containing the essential cysteine in some thiol proteolytic enzymes.	4.
1.3.	General active site structure of the plant thiol protease enzymes.	5.
3.1.	DEAE-cellulose chromatography of actinidin.	36,37.
3.2.	Sephadex G-50 chromatography of the tryptic peptides of carboxymethyl actinidin.	41,42.
3.3.	Composite peptide maps of the electrophoretically mobile tryptic peptides of carboxymethyl actinidin.	43,44.
3.4.	Composite DEAE-cellulose chromatography of the tryptic peptides of carboxymethyl actinidin.	45,46.
3.5.	Chymotrypsin digest peptide maps of peptide T ₃ .	53.
3.6.	Thermolysin digest peptide maps of peptide T ₃ .	55.
3.7.	Reconstruction of peptide T ₃ from chymotrypsin and thermolysin digest peptides.	56.
3.8.	Determination of the amide and acid residues in peptide T ₃ ChTa by the Offord mobility procedure.	57,58.
3.9.	Sephadex G-75 chromatography of the tryptic peptides of maleylated carboxymethyl actinidin.	60.
3.10.	Maleyl diagonal peptide map at pH 6.5 of a chymotrypsin digest of the maleyl tryptic peptide containing peptide T ₄ .	61.
3.11.	Reconstruction of the maleyl tryptic peptide obtained from fraction A of the Sephadex G-75 elution profile from chymotrypsin and <u>S. aureus</u> V8 protease digest peptides to provide the sequence of peptide T ₄ .	66.
3.12.	Chymotrypsin digest peptide maps of peptide T ₆ .	70.

LIST OF FIGURES (contd.)

Figure		Page
3.13.	Sephadex G-25 chromatography of peptides obtained by digestion of peptide T ₆ with <u>S. aureus</u> V8 protease.	71.
3.14.	Composite peptide maps of the chymotrypsin digest peptides from peptides GEa and GEb from peptide T ₆ .	74.
3.15.	Reconstruction of peptide T ₆ from chymotrypsin and <u>S. aureus</u> V8 protease digest peptides.	75.
3.16.	Determination of the amide and acid residues in peptide T ₆ GEbChTa ₂ by the Offord mobility procedure.	76,77.
3.17.	Chymotrypsin digest peptide maps of peptide T ₇ .	80.
3.18.	Sephadex G-25 chromatography of peptides obtained by digestion of peptide T ₇ with <u>S. aureus</u> V8 protease.	81.
3.19.	Composite peptide maps of <u>S. aureus</u> V8 protease and pepsin digest peptides from peptide T ₇ .	84.
3.20.	Reconstruction of peptide T ₇ from chymotrypsin, pepsin and <u>S. aureus</u> V8 protease digest peptides.	85.
3.21.	Chymotrypsin digest peptide maps of peptide T ₈ .	88.
3.22.	Sephadex G-25 chromatography of peptides obtained by digestion of peptide T ₈ with <u>S. aureus</u> V8 protease.	89.
3.23.	Reconstruction of peptide T ₈ from chymotrypsin, pepsin and <u>S. aureus</u> V8 protease digest peptides.	91.
3.24.	Thermolysin digest peptide maps of peptide T ₁₁ .	97.
3.25.	Sephadex G-50 chromatography of the cyanogen bromide fragments from carboxymethyl actinidin.	100.
3.26	Sephadex G-25 chromatography of the cyanogen bromide Sephadex G-50 fraction CB-II+III.	100.
4.1.	Alignment of the tryptic peptides of actinidin.	127.

LIST OF FIGURES (contd.)

Figure		Page
4.2.	Amino acid sequences containing the essential cysteine in some thiol proteolytic enzymes.	128.
4.3.	Amino acid sequences about the histidine residue in the active site of some plant thiol proteases.	130.
4.4.	Homologies among the available sequences of some plant thiol proteolytic enzymes.	132,133.

LIST OF TABLES

Table		Page
3.1.	Purification of actinidin.	35.
3.2.	Amino acid composition of actinidin.	39.
3.3.	Peptide isolation data and amino acid composition of peptide T ₁ .	49.
3.4.	Peptide isolation data and amino acid composition of peptide T ₂ .	50.
3.5.	Amino acid composition of peptide T ₃ .	51.
3.6.	Peptide isolation data and amino acid compositions of the chymotrypsin digest peptides of peptide T ₃ .	52.
3.7.	Peptide isolation data and amino acid compositions of the thermolysin digest peptides from peptide T ₃ .	54.
3.8.	Peptide isolation data and amino acid compositions of the chymotrypsin digest peptides from the maleyl tryptic peptide containing peptide T ₄ .	62.
3.9.	The peptide isolation data and amino acid compositions of the <u>S. aureus</u> V8 protease digest peptides of peptide ChTA that was obtained by chymotrypsin digestion of the maleyl tryptic peptide containing peptide T ₄ .	64.
3.10.	Amino acid composition of peptide GE/Y obtained by <u>S. aureus</u> V8 protease digestion of the maleyl tryptic peptide containing peptide T ₄ .	65.
3.11.	Peptide isolation data and amino acid composition of peptide T ₅ .	67.
3.12.	Amino acid composition of peptide T ₆ .	68.
3.13.	Peptide isolation data and the amino acid compositions of the chymotrypsin digest peptides from peptide T ₆ .	69.
3.14.	Amino acid compositions of fractions GEa and GEb obtained by <u>S. aureus</u> V8 protease digestion of peptide T ₆ .	72.

LIST OF TABLES (contd.)

Table		Page
3.15.	Peptide isolation data and amino acid compositions of the chymotryptic peptides from peptides GEa and GEb that were obtained by <u>S. aureus</u> V8 protease digestion of peptide T ₆ .	73.
3.16.	Amino acid composition of peptide T ₇ .	78.
3.17.	Peptide isolation data and amino acid compositions of the chymotrypsin digest peptides from peptide T ₇ .	79.
3.18.	Peptide isolation data and amino acid compositions of the pepsin digest peptides of peptide GE ₁ that was obtained by <u>S. aureus</u> V8 protease digestion of peptide T ₇ .	82.
3.19.	Peptide isolation data and amino acid compositions of peptides GE ₂ and GE ₃ that were obtained by digestion of peptide T ₇ with <u>S. aureus</u> V8 protease.	83.
3.20.	Amino acid composition of peptide T ₈ .	86.
3.21.	Peptide isolation data and amino acid compositions of the chymotrypsin digest peptides from peptide T ₈ .	87.
3.22.	Peptide isolation data and amino acid compositions of the pepsin digest peptides from peptide GE ₁ that was obtained from peptide T ₈ by digestion with <u>S. aureus</u> V8 protease.	90.
3.23.	Peptide isolation data and amino acid composition of peptide GE ₂ that was obtained from peptide T ₈ by digestion with <u>S. aureus</u> V8 protease.	90.
3.24.	Peptide isolation data and amino acid composition of peptide T ₉ .	92.
3.25.	Peptide isolation data and amino acid compositions of the chymotrypsin digest peptides from peptide T ₉ .	93.
3.26.	Peptide isolation data and amino acid composition of peptide T ₁₀ .	94.
3.27.	Peptide isolation data and amino acid composition of peptide T ₁₁ .	95.

LIST OF TABLES (contd.)

Table		Page
3.28.	Peptide isolation data and amino acid compositions of the thermolysin digest peptides from peptide T ₁₁ .	96.
3.29.	Peptide isolation data and amino acid composition of peptide T ₁₂ .	98.
3.30.	Peptide isolation data and amino acid compositions of cyanogen bromide fragments from actinidin.	99.
3.31.	Peptide isolation data and amino acid compositions of the chymotrypsin digest peptides from cyanogen bromide fragment CB-IIa ₂ .	101.
4.1.	Amino acid composition of some plant thiol proteases.	106.
4.2.	Comparison of the amino acid composition of actinidin with the amino acid compositions of the tryptic peptides as determined from the amino acid sequence of each peptide.	126.

ABBREVIATIONS

BAW	Butanol/glacial acetic acid/ distilled water.
BAWP	Butanol/glacial acetic acid/distilled water/pyridine.
N- α -CBZ-lys-pNP	N- α -carboboxy-L-lysine-p-nitrophenyl ester.
Dansyl, DNS	1-dimethylaminonaphthalene-5-sulphonyl
↗	Symbol used to indicate direct dansyl-Edman sequencing.
DEAE	Diethylaminoethyl.
DTT	Dithiothreitol.
E.U.	Enzyme unit (μmol substrate transformed.min ⁻¹).
FDNB	1-fluoro-2,4-dinitrobenzene.
NpS	2-nitrophenylsulphenyl.
%	All percentages are weight/volume (w/v) unless otherwise stated.
PITC	Phenylisothiocyanate.
POPOP	1,4-di[2-(5-phenyloxazolyl)]benzene.
PPO	2,5-diphenyloxazole.
<u>S. aureus</u>	<u>Staphylococcus aureus</u> .
TFA	Trifluoroacetic acid.
TPCK	L-(1-tosylamido-2-phenyl)ethylchloromethyl ketone.
Tris	Tris(hydroxymethyl)amino methane.

Amino acid abbreviations:

Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
Asx	Aspartyl or Asparaginyl
CMCys	Carboxymethyl cysteine
Cys	Cysteine
CySO ₃ H	Cysteic acid
Gln	Glutamine
Glu	Glutamic acid
Glx	Glutamyl or Glutaminyl
Gly	Glycine

ABBREVIATIONS (contd.)

His	Histidine
Hse	Homoserine
$\overbrace{\text{Hse}}$	Homoserine lactone
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Met	Methionine
Phe	Phenylalanine
Pro	Proline
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

Peptides obtained by enzymic digestion:

T	Tryptic
ChT	Chymotryptic
Th	Thermolytic
P	Peptic
GE	<u>S. aureus</u> V8 proteolytic
a	Peptide acidic charged at pH 6.5.
b	Peptide basic charged at pH 6.5.
N	Peptide neutral at pH 6.5

Enzyme EC numbers used are according to:

Recommendations (1972) of I.U.P.A.C. and the International Union of Biochemistry.

Elsevier Scientific Publishing Company, Holland (1975)
And Supplement I (corrections and additions (1975)) published in Biochim. Biophys. Acta, 429, 1-45 (1976).

General abbreviations:

Are used according to the "Policy of the Biochemical Journal"
Biochem. J., 153, 1-21 (1976).