Chymotryptic Inhibitor I from Potatoes

THE AMINO ACID SEQUENCE OF SUBUNIT A*

By M. RICHARDSON

Department of Botany, University of Durham, Durham DH1 3LE, U.K.

(Received 25 July 1973)

The amino acid sequence of subunit A of the potato chymotryptic inhibitor I was determined. The sequence was deduced from analysis of fragments and peptides derived from the protein by cleavage with cyanogen bromide, *N*-bromosuccinimide and dilute acid, and by digestion with trypsin, thermolysin, pepsin and papain. The molecule consists of a single polypeptide chain of 84 residues, which contains two homologous regions each of 13 amino acids. The protein does not appear to be homologous with any other known proteinase inhibitors.

There has been much recent interest in the numerous proteinase inhibitors found in plants (Fritz & Tschesche, 1971). The amino acid sequences of the inhibitors of chymotrypsin and trypsin from soya bean (Koide *et al.*, 1972; Odani & Ikenaka, 1972), lima bean (Tan & Stevens, 1971), maize and peanuts (Hochstrasser & Werle, 1971) have been reported. These data have proved useful in studies on the mechanisms of inhibition (Seidl & Liener, 1971; Krahn & Stevens, 1972).

Although a large number of different proteinase inhibitors have been isolated from potatoes (Balls & Ryan, 1963; Rancour & Ryan, 1968; Hochstrasser *et al.*, 1969; Iwasaki *et al.*, 1971; Hojima *et al.*, 1971; Kaiser & Belitz, 1972), there are as yet no details of the amino acid sequences of any of these inhibitors.

One proteinase inhibitor in particular from potatoes has been the subject of much investigation. This is the so-called chymotryptic inhibitor I (Ryan, 1966; Melville & Ryan, 1970), which is a potent inhibitor of both chymotrypsin and trypsin. This protein appears to be a tetramer composed of a heterogeneous mixture of protomers (subunits). Melville & Ryan (1972) have succeeded in separating these subunits into four main types (A, B, C and D) by ion-exchange chromatography on sulphoethyl-cellulose in 8*M*-urea. All four subunits were shown to have glutamic acid as their *N*-terminal amino acid, but to differ quantitatively from each other in their amino acid compositions, reactivity with chymotrypsin and trypsin, electrophoretic mobilities and susceptibility to digestion by pepsin.

The present paper reports the elucidation of the complete amino acid sequence of the subunit A of potato chymotryptic inhibitor I.

*Dedicated to the memory of Dr. B. T. Cromwell.

Experimental

Enzymes

Pepsin, trypsin ('tosylphenylalanine chloromethyl ketone'-treated) and papain (2×crystallized) were obtained from Worthington Biochemical Corp., Freehold, N.J., U.S.A. Thermolysin (crystalline) was obtained from Daiwa Kasei K.K., Osaka, Japan. Carboxypeptidase A (di-isopropyl phosphorofluoridate-treated) was obtained from Sigma (London) Chemical Co., London S.W.6, U.K.

Preparation of chymotryptic inhibitor I

Inhibitor I was purified from potato tubers (*Solanum tuberosum* L. var. 'Ulster Prince') by the method of Melville & Ryan (1972) except that a large (0.914m; 36in) rotary basket centrifuge (T. Broadbent and Sons Ltd., Huddersfield, U.K.) was used during stage I of extraction instead of a rack and cloth press. In addition the use of Celite as a filter aid during stage I was found to be unnecessary. The final yield of purified subunit A of inhibitor I was 148mg from 300kg fresh wt. of potato tubers.

The final preparation of subunit A yielded single bands when examined by polyacrylamide-gel disc electrophoresis at pH4.3 (Reisfeld *et al.*, 1962) and at pH8.3 (Clarke, 1964).

Samples of the protein were reduced and carboxymethylated as described by Crestfield *et al.* (1963).

Amino acid compositions

Amino acid compositions of peptides, large fragments and the whole protein were determined from analyses carried out on a Locarte amino acid

analyser. Freeze-dried samples of the native and the reduced and carboxymethylated protein (40nmol) were hydrolysed *in vacuo* at 110°C in 5.7M-HCl for 24, 48 and 72h. All hydrolyses of peptides (20–50nmol) were carried out under similar conditions for 48h. Tryptophan in the whole protein was determined by the spectrophotometric method of Beaven & Holliday (1952).

Sequence determination

Cleavage of the protein with cyanogen bromide. Cyanogen bromide (200mg) was added to the reduced and carboxymethylated protein (50mg) dissolved in 2ml of 70% (v/v) formic acid (Steers *et al.*, 1965). The reaction was allowed to proceed for 24h in the dark. The products of the reaction were freeze-dried and then separated by passage through a column (1cm×100cm) of Sephadex G-50 equilibrated with 70% (v/v) formic acid.

Cleavage of the protein with *N*-bromosuccinimide. *N*-Bromosuccinimide (3μmol) was added to the reduced and carboxymethylated protein (1μmol) in 70% (v/v) acetic acid. The reaction was allowed to proceed at room temperature for 30min in the dark before a further 3μmol of *N*-bromosuccinimide was added. After a further 30min of incubation the products of the reaction were separated by passage through a column (1cm×110cm) of Sephadex G-50 equilibrated with 70% (v/v) acetic acid.

Digestion of protein and fragments. The fragments which resulted from cleavage of the protein with cyanogen bromide were digested separately with trypsin, pepsin and thermolysin. The intact, reduced and carboxymethylated protein was digested with papain and also subjected to partial (dilute) acid hydrolysis. All digestions were carried out on approx. 1.5–2.0μmol.

The enzymic digestions with trypsin and thermolysin were carried out in 0.2M-pyridine–acetate buffer, pH8.0, at 37°C at an enzyme/substrate ratio of 1:50 (w/w) for 5h. The thermolysin was activated before use with 5mM-CaCl₂.

For digestion with pepsin, pH1.9 buffer (acetic acid–formic acid–water; 4:1:45, by vol.) was used with an enzyme/substrate ratio of 1:100 (w/w) for 1h and 5h at 37°C.

The digestion with papain [enzyme/substrate ratio 1:50 (w/w)] was carried out in 0.2M-pyridine–acetate buffer, pH6.5, containing 0.1% (v/v) 2,3-dimercapto-propan-1-ol as an activator for 5h at 37°C.

Partial acid digestion was performed *in vacuo* in 0.03M-HCl for 21h at 110°C (MacDowall & Smith, 1965). The dilute acid was flushed with N₂ before use.

Separation and purification of peptides. The peptides that resulted from the digestion of the cyanogen bromide fragments with trypsin were subjected to an initial fractionation on a column (1cm×110cm) of

Sephadex G-25 equilibrated with 70% (v/v) acetic acid. These and all other peptides were also purified by high-voltage paper electrophoresis at pH6.5 and 1.9 (Thompson *et al.*, 1970) and occasionally at pH4.7 (Michl, 1958), and by paper chromatography in BAWP (butan-1-ol–acetic acid–water–pyridine, 15:3:12:10, by vol.; Waley & Watson, 1953). All peptides were eluted from paper with 20% (v/v) pyridine and stored as frozen solutions. The yields of the various peptides after purification were only determined semi-quantitatively.

Edman degradations and *N*-terminal analyses. The amino acid sequences of the peptides were determined by using the dansyl-Edman method (Gray & Hartley, 1963) as described previously (Thompson *et al.*, 1970). The quantity of each peptide taken for Edman degradation was in the range 200–400nmol. After each round of degradation samples containing 10–20nmol of peptide were removed for dansylation. Dansyl-amino acids were identified by chromatography on polyamide sheets (Woods & Wang, 1967) by using the solvent systems described by Ramshaw *et al.* (1971).

***C*-Terminal analyses.** The *C*-terminal residues of the protein and certain peptides were determined by the use of carboxypeptidase A in 0.2M-*N*-ethylmorpholine–acetate buffer, pH8.5 (Ambler, 1972), in combination with the dansyl technique.

Peptide nomenclature. Peptides and fragments are numbered on the basis of their occurrence in the deduced sequence starting at the *N*-terminus. Peptides derived by further or minor cleavage are given by a suffix to the parent peptide. The following abbreviations are used: CNBr, peptides derived from cyanogen bromide cleavage; T, tryptic peptides; H, thermolysin peptides; P, peptic peptides; PA, papain peptides; A, peptides from dilute acid hydrolysis; NBS, peptides from *N*-bromosuccinimide cleavage.

Results

The amino acid compositions of subunit A of potato chymotryptic inhibitor I determined by analysis and calculated from the sequence are shown in Table 1. These values are in reasonable agreement. The low values obtained for valine are probably due to the high resistance to acid hydrolysis of the three Val-Val peptide bonds which occur in the sequence. Similarly the Ile-Ile, Ile-Leu, Leu-Ile and Ile-Leu-Leu sequences that occur in this protein may have contributed to the somewhat low values obtained for isoleucine and leucine.

Also shown, for comparative purposes, are the values obtained by other workers for this inhibitor.

Edman degradation of the whole protein revealed an *N*-terminal sequence of Glx-Phe-Glx-. Digestion with carboxypeptidase A for various periods released only valine and alanine in equal amounts.

Table 1. Amino acid composition of chymotryptic inhibitor I (subunit A) from potatoes (var. 'Ulster Prince')

The results are expressed as residues/molecule. The analysis values for the native protein and the reduced and carboxymethylated protein are average values obtained from two hydrolyses at 24, 48 and 72h. Optimum values for serine and threonine were calculated by extrapolation to zero time; and values for valine and isoleucine were calculated by extrapolation to infinite time. Tryptophan was determined as described in the Experimental section. Abbreviation: CmCys, carboxymethylcysteine. * Results of Melville & Ryan (1972) for subunit A from potato var. 'Russet Burbank'. † Results of Kiyoharā *et al.* (1973) for inhibitor I from potato var. 'Danshaku Imo'.

Amino acid	Native protein	Reduced and carboxymethylated protein	Sequence	Subunit A*	Inhibitor I†
CmCys/Cys	1.90	2.20	2	2.5	2.4
Asp	10.12	10.28	10	11.6	11.3
Thr	2.29	2.08	2	3.6	3.1
Ser	4.58	4.44	4	5.1	4.3
Glu	9.04	8.82	9	11.1	7.9
Pro	5.08	4.90	5	6.6	5.2
Gly	6.18	6.35	6	6.9	6.9
Ala	2.06	2.14	2	6.5	2.4
Val	9.80	9.60	11	9.6	9.0
Met	1.33	1.26	1	1.1	0
Ile	7.22	7.43	8	7.7	8.6
Leu	10.70	10.89	11	8.0	9.7
Tyr	0.60	0.37	—	2.3	1.0
Phe	3.13	3.15	3	3.1	2.7
His	0.93	0.94	1	0.4	1.2
Lys	4.05	4.27	4	5.2	4.9
Arg	3.99	3.88	4	3.9	4.3
Trp	1.00	1.00	1	—	1.2
Totals	84.00	84.00	84	95.2	86

Cleavage of the protein with cyanogen bromide yielded two major fragments (CNBr1 and CNBr3), and a minor fragment (CNBr2) which only occurred in very low yield. This result was consistent with the existence of a single methionine residue in subunit A, as indicated by the analysis, together with a low degree of microheterogeneity for methionine in another residue. Analysis of the end groups of fragments CNBr1 and CNBr3 revealed that the fragments were located at the *N*-terminus and the *C*-terminus of the protein respectively.

Preliminary experiments indicated that the reduced and carboxymethylated protein and the major large fragment derived by cleavage with cyanogen bromide (CNBr1) appeared to be still resistant to digestion with chymotrypsin. Since no quantitative studies have been carried out it is possible that the activity of the enzyme was prevented in both cases by traces of the protein which had failed to react with the modifying agent.

Both of the two major cyanogen bromide fragments, however, were readily digested with trypsin, thermolysin and pepsin. Table 2 shows the properties of the peptides that resulted from these digestions. Also shown are the peptides that resulted from the digestion of the whole protein with papain and dilute

acid. Many of the first 19 residues of the fragment CNBr3 were also established by a direct application of the dansyl-Edman method to two separate samples of the fragment. The first sample was sequenced in the normal manner, and the second sample was subjected to ten cycles of Edman degradation before any samples were removed for dansylation. Consideration of the results of the dansyl-Edman analyses on these peptides enabled the complete sequence of the protein to be deduced, as shown in Fig. 1. The amide content of the peptides was indicated by their electrophoretic mobilities at pH 6.5 (Offord, 1966) and also by the release of certain residues with carboxypeptidase A, as follows.

Glu-1, Glu-3 and Asp-5: all three residues were shown to be acids by the acidic electrophoretic mobilities of the peptides T1, H1, H1a, PA1a, PA1b and NBS-1.

Gln-9 and Glu-12: the neutral mobility of peptide T2 indicates that one residue is an amide and the other an acid. The acidic mobilities of peptides H2, H2a and P1a indicate that Glu-12 is the acid residue. Digestion of peptide H1 with carboxypeptidase A released glutamine (and leucine) confirming Gln-9.

Glu-23: the neutral mobility of peptides H4, H4a and P3e indicated that this residue must be an acid.

Table 2. *Properties of fragments and peptides derived from chymotryptic inhibitor I (subunit A) by chemical cleavage and digestion with various enzymes*

V_e/V_0 . Relative elution volume on gel filtration through Sephadex G-50 (a) or G-25 (b) equilibrated with 70% (v/v) formic acid or 70% (v/v) acetic acid respectively. m (pH 6.5), electrophoretic mobilities of the peptides on paper at pH 6.5 were measured relative to that of Dns-Arg-Arg (+ve) from the neutral amino acids. m (pH 1.9), electrophoretic mobilities of the peptides on paper at pH 1.9 relative to Dns-Arg from Dns-OH. $R_{Dns-Arg}$ (BAWP), mobilities of the peptides relative to Dns-Arg during paper chromatography in butan-1-ol-acetic acid-water-pyridine (15:3:12:10, by vol.). For details of peptide nomenclature see the Experimental section. The quantitative amino acid results are expressed as mol of each amino acid/mol of peptide. The results are uncorrected for destruction of serine and threonine or for incomplete cleavage at valyl and isoleucyl bonds. Trace contaminants (less than 0.2 mol/mol of peptide) are ignored.

Peptide	Purification			$R_{Dns-Arg}$ (BAWP)	Amino acid analysis																Position in sequence	
	V_e/V_0	m (pH6.5)	m (pH1.9)		Cys	Asp	Thr	Ser	Glu	Pro	Gly	Ala	Val	Met	Ile	Leu	Tyr	Phe	His	Lys		Arg
CNBr1	1.3 ^a				Not analysed (mixture of two large fragments)																{1-60 1-55	
CNBr2	2.4 ^a		0.47							0.9	1.2	2.0	0.7**									
CNBr3	1.8 ^a		0.62							1.0	1.3	1.0	3.9	1.9	2.0	0.2	1.7				3.8	
CNBr1	1.6 ^b	-2.30	0.66	0.17	0.8	4.2	0.9	1.2	1.0	1.1							0.9		1.1			
T1	1.6 ^b	0	0.48	0.64	0.9	1.2	1.2	0.3	1.8	2.0	1.0	1.0	1.0	0.8	1.8				1.1		+	
T2	1.6 ^b																		1.1			
T3	2.0 ^b	1.40																	1.0			
T4	2.0 ^b	-0.88	0.75	0.32				1.9											1.0			
T5	1.2 ^b				Not analysed (mixture of two large fragments)																{28-60 28-55	
CNBr3		0	1.00		1.0												0.8			0.9		
T7	-1.30				1.0	0.9														0.9		
T8	1.55																			0.9		
T9	-0.50		0.44		1.9	0.9	0.9	1.0	1.0	1.3	0.8	1.0	1.8	2.1	2.0	1.0				1.0		
T10		0	0.76																		0.9	
CNBr1			0.60		1.0	1.2	0.5	2.7	1.1											1.1		
H1a		-2.10	0.66	0.16	1.0	1.0		2.0	1.1											0.9		
H2	-2.30	-0.85	0.39	1.00	1.0	1.0		0.9	1.0													
H2a	-0.85	-0.85	0.39	0.80				1.0	1.0													
H3	0.90	0.90	0.76	0.51		0.8		1.1	0.8											1.2		
H3a	0.90	0.90	0.65	0.69		0.8		1.2	1.3											1.1		
H4	0	0	0.90	0.35				1.0												1.0		
H4a	0	0	1.16					0.9												0.9		
H5	0	0	0.66	0.24		1.0	0.8	2.0														
H5a	0	0	0.90	0.53				1.0														
H5b	0	0	0.55	0.20		1.1	0.9	0.9												1.1		
H6	0	0	0.44	0.65		1.1	0.9															
H6a	0	0	0.55	0.40		1.0	0.8															
H7	1.25		1.35																			
H8	0	0	0.70	1.15																0.9		
H9	0	0	0.35	0.55		1.2	0.9	0.9														
H9a	0	0	0.40	0.48		1.2	0.9															
H9b	0	0	0.70	0.56		Not analysed																
H10	-1.80		0.60	0.55	1.0																	
H11	-0.60		0.28	1.15	1.1		1.1	1.0	1.1													
H11a	-1.25		0.44	0.81	1.0																	

[illegible]

Table 2.—Continued

Peptide	Purification			<i>R</i> _{Phe-Arg} (BAWP)	Amino acid analysis																	Position in sequence	
	<i>V_e</i> / <i>V</i> ₀	<i>m</i> (pH 6.5)	<i>m</i> (pH 1.9)		Cys	Asp	Thr	Ser	Glu	Pro	Gly	Ala	Val	Met	Ile	Leu	Tyr	Phe	His	Lys	Arg		Trp
PA7a		-1.40	0.46		1.2			1.2				2.1											51-54
PA8		0	0.46	0.96					0.9	1.2		2.0				0.9							55-59
PA9		0	0.98	0.47	0.9	2.3						1.2	0.5				0.8				2.8		60-68
PA9*		0	0.86		1.0	2.1						0.9	0.7				0.6				3.1		60-68
PA9a		0	0.98	0.33			Not analysed																61-68
PA10		-1.40	0.30		2.1					1.0				1.0	1.8		1.0						69-75
PA11		0.65	0.61		Not analysed		Not analysed																76-84
PA11a		0.96	0.89		Not analysed		Not analysed																80-84
A1		-2.90			Not analysed		Not analysed																2-4
A1a		-1.75	0.70		Not analysed		Not analysed																2-3
A2		0	0.62	0.36			Not analysed																6-11
A3		0.98	0.86		Not analysed		Not analysed																16-20
A4		0	1.23		Not analysed		Not analysed																22-23
A5		0	0.92		Not analysed		Not analysed																24-27
A5a		-1.50	0.60		Not analysed		Not analysed																24-26
A6		0	0.50		Not analysed		Not analysed																30-33
A7		0.45	0.82		Not analysed		Not analysed					1.0		0.8	1.9		0.8						35-39
A8		0	0.37		Not analysed		Not analysed																41-46
A8a		0	0.43	0.54			Not analysed																42-46
A9		-1.75	0.55		Not analysed		Not analysed	(see also peptide A13)															47-50
A9a		0	0.62	1.00	Not analysed		Not analysed	(see also peptide A13a)															48-50
A10		-0.67	0.31		Not analysed		Not analysed																52-60
A10a		-0.85	0.43		Not analysed		Not analysed																52-58
A11		0	0.90	0.42	0.8													0.9			1.0		62-64
A11*		0	0.82	0.9													0.7	0.2			1.0		62-64
A12		1.45	1.08									1.0			0.8		0.9			2.2			66-70
A13		-1.75	0.55		Not analysed		Not analysed	(same peptide as A9)															72-75
A13a		0	0.62	1.00	Not analysed		Not analysed	(same peptide as A9a)															73-75
A14		0	0.62	0.80			0.7	1.1	1.1		0.9	3.2		1.0							1.0		76-84
A14a		0	0.70		Not analysed		Not analysed																77-84
NBS-1	1.7 ^b	-1.35	0.52		0.3	1.0	0.5	2.6		1.0					0.8		1.0		1.1				1-10
NBS-2	1.0 ^b						Not analysed	(mixture of unchanged protein and large <i>N</i> -bromosuccinimide-cleaved fragment)															1-84
																							{ 11-84

** As homoserine.

+. Tryptophan indicated by Ehrlich-positive zone and a faint blue fluorescence on paper after electrophoresis or chromatography, and a pink coloration at the trifluoroacetic acid stage of the Edman degradation.

‡. Peptides also purified by paper electrophoresis at pH 4.7.

T5

CNBr1

NBS-2

Fig. 1. Complete amino acid sequence of chymotryptic inhibitor I (subunit A) from potatoes

—, ———, Composition determined quantitatively; ———, composition determined qualitatively; —→, sequence determined by dansyl-Edman degradation; —→, standard results by dansyl-Edman procedure; → indicates that the last amino acid in the peptide after the removal of the remainder by Edman degradation was positively identified as the free amino acid; —→, sequence determined by using carboxypeptidase A. For details of peptide nomenclature see the Experimental section. For full details of the properties of the individual peptides see Table 2. Peptides marked * may be due to microheterogeneity or may have resulted from contamination of the protein with small amounts of the other subunits B and C.

Glu-26: the acidic mobility of peptide T4 confirms that Glu-23 is an acid and shows that Glu-26 must also be an acid.

Gln-28 and Asn-29: the neutral mobility of peptide H5 indicated that it contained only one acidic residue. This was shown above to be Glu-26. The neutral mobility of peptide H5*b* and the basic mobility of peptide P3*d* confirms that Gln-28 and Asn-29 must both be amides. However, traces of another neutral peptide corresponding to P3*d* were also found, showing that deamidation of either Gln-28 or Asn-29 may occur.

Asn-34: the neutral mobilities of the peptides H6 and P3*b*, and the slightly basic mobility of peptide PA4*a*, indicates that this residue must be an amide.

Asn-40: the neutral mobilities of the peptides H9, H9*a* and PA5 indicate that this residue is normally an amide, but a peptide PA5* was also found in good yield with an acidic mobility at pH6.5, indicating that Asn-40 may be subject to deamidation. This was confirmed by the finding of trace amounts of acidic peptides which corresponded to H9 and H9*a*. It is now well known (Ambler & Wynn, 1973) that the sequence -Asn-Gly- which occurs at positions 40-41 is very susceptible to deamidation.

Asp-47: the acidic nature of this residue was indicated by the acidic mobilities of the peptides H10, PA6, P4 and P4*b*.

Asp-51: the acidic mobilities of the peptides H11*a* and P5 show that this residue is an acid.

Gln-54: the neutral mobilities of the peptides H11*b*, P6 and P6*a* indicated that position 54 was occupied by an amide. The acidic mobilities of the peptides H11, PA7 and PA7*a* also indicated this. This was confirmed by digestion of peptide PA7*a* with carboxypeptidase A which released glutamine (and valine).

Asp-61 and Asp-65: the neutral mobility of the peptides T6, H13, P7, PA9 and PA9*a* indicated that both residues were acids. The acidic mobility of peptide T7 and the basic mobilities of peptides P7*a* and P7*b* confirmed this.

Asp-71 and Asp-72: the acidic mobilities of the peptides H15, H15*a*, P8 and PA10 indicated that both of these residues were acids.

Gln-79: the basic mobilities of peptides P9, H17 and PA11 indicated that this residue is an amide.

The possible existence of a disulphide bridge between Cys-4 and Cys-64 was not investigated.

Minor examples of microheterogeneity were observed at positions 9, 55 and 62. At position 9 glutamine was normally found, but small amounts of serine were also observed in this position in certain of the peptides (H1, P1, T2 and NBS-1). Although the amounts of dansyl-serine observed in this position were small relative to the other dansyl-amino acids found in the respective peptides, they were much greater than the traces of dansyl-serine which

are occasionally seen during this type of analysis. Further, the quantitative amino acid analyses of these peptides (Table 2) gave values for serine (0.3-0.5) which were greater than the amounts of serine observed as a contaminant in peptides not containing serine (values always <0.2). At position 55 leucine was normally found, but small amounts of methionine were observed in peptide PA8 and also indicated in this position by the recovery of the small cyanogen bromide peptide CNBr2 in very low yield relative to the other cyanogen bromide fragments. At position 62 phenylalanine was normally found, but a number of peptides (H13*, PA9* and A11*) were recovered in low yield which contained tyrosine instead of phenylalanine in this position. Traces of tyrosine were also observed in this position when samples of the fragment CNBr3 were sequenced by the dansyl-Edman method. It is possible that these examples of microheterogeneity resulted from the digestion of small amounts of the other subunits B and C which may have been present in trace amounts as contaminants. For example, work in this laboratory has shown that both subunits B and C contain tyrosine instead of phenylalanine in position 62 (M. Richardson & L. Cossins, unpublished work).

Discussion

The amino acid sequence of subunit A of potato chymotryptic inhibitor I does not appear to show much homology with any of the sequences determined for other plant proteinase inhibitors.

The amino acid composition of the subunit derived from analyses and also calculated from the sequence (Table 1) is similar in many respects to that previously reported for subunit A by Melville & Ryan (1972), with the exception of the values for alanine, serine, glutamic acid and tyrosine. However, it is perhaps difficult to make direct comparisons in this case, in view of the different sources of material used and also the fact that these workers appear to have based their analyses on a total of 95 residues in the subunit. The choice of this number is surprising in view of the estimates of a molecular weight of 9300-9800 obtained for the subunit by the same workers (Melville & Ryan, 1970). The molecular weight calculated from the sequence found during the present investigation is 9332.

The analysis agrees well with the results obtained by Kiyohara *et al.* (1973) for Inhibitor I. These workers estimated the minimum molecular weight at 9400 (85 or 86 residues) for the subunits of this protein.

The *N*-terminal acid detected both by Melville & Ryan (1972) and by Kiyohara *et al.* (1973) was glutamic acid (or glutamine), which is in agreement with the results of the present investigation. However,

Laskowski & Sealock (1971) for the active site of a trypsin inhibitor. However, these possibilities require further study.

In comparison with most of the other proteinase inhibitors of plant origin the chymotryptic inhibitor I from potatoes appears to be unusual in only containing two residues of cysteine. For example, the soya bean Bowman-Birk inhibitor (Odani & Ikenaka, 1972), the lima-bean inhibitor (Tan & Stevens, 1971) and the trypsin inhibitor from *Phaseolus vulgaris* (Wilson & Laskowski, 1973) all contain 14 residues of cysteine per molecule. Other proteinase inhibitors from the potato itself also appear to have a much higher content of cysteine (Hochstrasser *et al.*, 1969; Iwasaki *et al.*, 1972; Kaiser & Belitz, 1972).

It is perhaps noteworthy that the lysine and arginine residues of subunit A appear to be clustered together at opposite ends of the polypeptide chain. Thus the four lysine residues occur between positions 7 and 27, whereas the four arginine residues are located between positions 63 and 82.

Another striking feature of the amino acid sequence is the presence of two homologous regions. As illustrated in Fig. 2 there is a clear repetition or similarity of the sequence between residues 47 and 58 with that between residues 72 and 84. Similar repetitive sequences have been recorded previously in other proteins and are thought possibly to result from the extension of shorter polypeptide chains by a process of gene duplication (Dixon, 1966), although the possible significance of this as a mechanism of protein evolution has been disputed (McLachlan, 1972).

Similar homologous regions have also been recorded in the lima-bean proteinase inhibitor (Tan & Stevens, 1971) and in the Bowman-Birk inhibitor from soya beans (Odani & Ikenaka, 1972). In both of these inhibitors the anti-trypsin and anti-chymotrypsin sites are located in these homologous regions of the polypeptide chain (Krahn & Stevens, 1972; Odani & Ikenaka, 1972). It is therefore tempting to speculate that the inhibitory sites of the chymotryptic inhibitor I from potatoes could lie in the homologous sequences shown in Fig. 2. This is particularly so since residues 81-84 appear to obey the tentative rules proposed by

References

- Ambler, R. P. (1972) *Methods Enzymol.* **25**, 262-272
- Ambler, R. P. & Wynn, M. (1973) *Biochem. J.* **131**, 485-498
- Balls, A. K. & Ryan, C. A. (1963) *J. Biol. Chem.* **238**, 2976-2982
- Beaven, G. H. & Holliday, E. R. (1952) *Advan. Protein Chem.* **7**, 319-386
- Clarke, J. T. (1964) *Ann. N.Y. Acad. Sci.* **121**, 428-436
- Crestfield, A. M., Moore, S. & Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622-627
- Dixon, G. H. (1966) *Essays Biochem.* **2**, 128-204
- Fritz, H. & Tschesche, H. (1971) *Proc. Int. Res. Conf. Proteinase Inhibitors 1970, Munich*, Walter de Gruyter, Berlin
- Gray, W. R. & Hartley, B. S. (1963) *Biochem. J.* **89**, 378-380
- Hochstrasser, K. & Werle, E. (1971) *Proc. Int. Res. Conf. Proteinase Inhibitors 1970, Munich*, p. 169, Walter de Gruyter, Berlin
- Hochstrasser, K., Werle, E., Seigelmann, R. & Schwarz, S. (1969) *Hoppe-Seyler's Z. Physiol. Chem.* **350**, 897-902
- Hojima, Y., Moriya, H. & Moriwaki, C. (1971) *J. Biochem. (Tokyo)* **69**, 1019-1025
- Iwasaki, T., Kiyohara, T. & Yoshikawa, M. (1971) *J. Biochem. (Tokyo)* **70**, 817-826
- Iwasaki, T., Kiyohara, T. & Yoshikawa, M. (1972) *J. Biochem. (Tokyo)* **72**, 1029-1035
- Kaiser, K. P. & Belitz, H. D. (1972) *Chem. Mikrobiol. Technol. Lebensm.* **1**, 191-194
- Kiyohara, T., Iwasaki, T. & Yoshikawa, M. (1973) *J. Biochem. (Tokyo)* **73**, 89-95
- Koide, T., Tsunasawa, S. & Ikenaka, T. (1972) *J. Biochem. (Tokyo)* **71**, 165-167
- Krahn, J. & Stevens, F. C. (1972) *Biochemistry* **11**, 1804-1808

- Laskowski, M., Jr. & Sealock, R. W. (1971) *Enzymes*, 3rd edn., 3, 375-473
- MacDowall, M. A. & Smith, E. L. (1965) *J. Biol. Chem.* **240**, 281-289
- McLachlan, A. D. (1972) *J. Mol. Biol.* **64**, 417-437
- Melville, J. C. & Ryan, C. A. (1970) *Arch. Biochem. Biophys.* **138**, 700-702
- Melville, J. C. & Ryan, C. A. (1972) *J. Biol. Chem.* **247**, 3445-3453
- Michl, H. (1958) *J. Chromatogr.* **1**, 93-119
- Odani, S. & Ikenaka, T. (1972) *J. Biochem. (Tokyo)* **71**, 839-848
- Offord, R. E. (1966) *Nature (London)* **211**, 591-593
- Ramshaw, J. A. M., Richardson, M. & Boulter, D. (1971) *Eur. J. Biochem.* **23**, 475-483
- Rancour, J. M. & Ryan, C. A. (1968) *Arch. Biochem. Biophys.* **125**, 380-383
- Reisfeld, R. A., Lewis, U. J. & Williams, D. E. (1962) *Nature (London)* **195**, 281-283
- Ryan, C. A. (1966) *Biochemistry* **5**, 1592-1596
- Seidl, D. S. & Liener, I. E. (1971) *Biochim. Biophys. Acta* **251**, 83-89
- Steers, J. E., Craven, G. R., Anfinsen, C. B. & Bethune, J. L. (1965) *J. Biol. Chem.* **240**, 2478-2484
- Tan, C. G. L. & Stevens, F. C. (1971) *Eur. J. Biochem.* **18**, 515-523
- Thompson, E. W., Laycock, M. V., Ramshaw, J. A. M. & Boulter, D. (1970) *Biochem. J.* **117**, 183-192
- Waley, S. G. & Watson, J. (1953) *Biochem. J.* **55**, 328-337
- Wilson, K. A. & Laskowski, M. Sr. (1973) *J. Biol. Chem.* **248**, 756-762
- Woods, K. R. & Wang, K. T. (1967) *Biochim. Biophys. Acta* **133**, 369-370