Biochem. J. (1978) **173**, 73–83
Printed in Great Britain

The Amino Acid Sequence of the Tryptic Peptides from Actinidin, a Proteolytic Enzyme from the Fruit of *Actinidia chinensis*

By ALAN CARNE* and CHRISTOPHER H. MOORE
Department of Chemistry, Biochemistry and Biophysics, Massey University, Palmerston North, New Zealand

(*Received* 28 *September* 1977)

The amino acid sequences of the tryptic peptides of the thiol proteinase actinidin from *Actinidia chinensis* were determined by the manual dansyl–Edman procedure. There are 12 tryptic peptides, which give a polypeptide chain of 220 residues with a mol.wt. of 23 500. An alignment of the tryptic peptides was made by using the X-ray-crystallographic data of Baker [(1977) *J. Mol. Biol.* **115**, 263–277] determined at 0.28 nm resolution on crystalline actinidin. Detailed evidence for the amino acid sequences of the tryptic peptides has been deposited as Supplementary Publication SUP 50083 (14 pages) at the British Library Lending Division, Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K., from whom copies can be obtained on the terms indicated in *Biochem. J.* (1978) **169**, 5.

Actinidin is a proteolytic enzyme from the fruit of *Actinidia chinensis*, the Chinese gooseberry. Actinidin contains a thiol group that is essential for activity (Arcus, 1959; McDowall, 1970) and is grouped in the class of plant thiol proteinases. This class includes papain, ficin and stem bromelain and has been reviewed by Glazer & Smith (1971). There are substantial similarities in the chemical and physical properties of these enzymes. Papain, ficin and stem bromelain have similar sequences about the active-site cysteine and histidine residues, and kinetic studies indicate similar modes of action. However, variations are found in the amino acid compositions, molecular weights, isoelectric points and carbohydrate contents. For example, actinidin has a mol.wt. of 23 500, compared with 23 400, 25 500 and 33 500 for papain, ficin and stem bromelain respectively. The isoelectric point of actinidin is 3.1, which is very different from the values for papain (8.75), ficin (>9) and stem bromelain (9.55).

Papain is the only thiol proteinase for which the complete amino acid sequence (Light *et al.*, 1964; Mitchel *et al.*, 1970; Husain & Lowe, 1969, 1970*a*) and the three-dimensional structure (Drenth *et al.*, 1968*a*,*b*) are known. Structural information available for ficin and stem bromelain is limited to the amino acid sequences about the active-site cysteine and histidine residues (Wong & Liener, 1964; Husain & Lowe, 1970*b*,*c*), although partial amino acid-sequence data for bromelain were reported (Goto *et al.*, 1976). It has been suggested that the plant proteinases form a family of homologous proteins because of their structural and kinetic similarities.

* Present address: Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.

Determination of the amino acid sequences of the tryptic peptides from actinidin was undertaken in parallel with an X-ray-crystallographic analysis of the three-dimensional structure of actinidin (Baker, 1973, 1976, 1977) to determine the degree of homology with other plant thiol proteinases, and to provide a better understanding of their physical and chemical properties.

Experimental

Materials

Chinese gooseberries were obtained from the local fruit market. Trypsin, soya-bean trypsin inhibitor, pepsin and carboxypeptidase A were from Sigma Chemical Co., St. Louis, MO, U.S.A. Trypsin was treated with 1-chloro-4-phenyl-3-L-tosylamido-butan-2-one before use. Chymotrypsin was from BDH Chemicals, Poole, Dorset, U.K., *Staphylococcus aureus* V8 proteinase from Miles Laboratories, Slough, Bucks., U.K., and thermolysin from Calbiochem, La Jolla, CA, U.S.A. Iodo[2-^{14}C]acetic acid was from The Radiochemical Centre, Amersham, Bucks., U.K. Sodium tetrathionate was prepared from sodium thiosulphate by the method of Liu & Inglis (1972). Reagents for sequencing were purified as described by Edman & Begg (1967).

Preparation of actinidin

Actinidin was prepared as the *S*-sulphenyl thiosulphate derivative to prevent autodigestion during purification (Boland & Hardman, 1972). Fruit (1 kg) was blended in extraction buffer (1 litre) containing sodium tetrathionate (10 mM) and EDTA (1 mM), and

the enzyme extracted by stirring for 30min. The mixture was centrifuged and the enzyme precipitated from the supernatant with $(NH_4)_2SO_4$ to 70% saturation. The precipitate was resuspended in potassium phosphate buffer (0.1M, pH6.0) and dialysed against water for 16h. After dialysis the solution was clarified by adding $CaCl_2$ to 40mM followed by a solution of K_2HPO_4 to give a final concentration of 100mM; the calcium phosphate gel was removed by centrifugation. The clear supernatant was adjusted to 0.1M with KCl and loaded on to a DEAE-cellulose (Whatman DE-23) column (4.5cm × 20cm). The enzyme was eluted from the column with a linear gradient of 0.1–1.0M-KCl (2 litres) in potassium phosphate buffer (0.1M, pH6.0). Fractions containing enzyme with a specific activity of greater than 30units (μmol/min)/mg of protein were pooled. The enzyme was crystallized by dialysing a 1% (w/v) solution against 20%-sat. $(NH_4)_2SO_4$ in the potassium phosphate buffer. Enzymic activity was measured by its esterase activity on *N*-α-benzyloxycarbonyl-L-lysine *p*-nitrophenyl ester by the method of Boland & Hardman (1972).

Reduction and S-carboxymethylation

Actinidin (10μmol) was unfolded in a solution containing guanidine hydrochloride (6M), EDTA (1mM) and Tris/HCl (0.5M, pH8.0) and reduced at 4°C under N_2 for 16h by using a 5-fold molar excess of dithiothreitol over the total thiol groups. The protein solution was then dialysed against 2 litres of guanidine hydrochloride (6M) in Tris/HCl (0.5M, pH8.0). The protein was *S*-carboxymethylated by using a 3-fold molar excess of neutralized iodoacetic acid containing 100μCi of iodo[2-^{14}C]acetic acid (Hirs, 1967) and the reaction was stopped by the addition of 2ml of 2-mercaptoethanol. The protein was then dialysed against water. This caused it to be precipitated, but it redissolved on exhaustive dialysis over 2 days. The iodo[2-^{14}C]acetate was incorporated at 8.4×10^6 d.p.m./mol of actinidin (1.6×10^6 d.p.m./mol of carboxymethylcysteine).

Selective *S*-carboxy[^{14}C]methylation at the active-site cysteine residue was carried out on actinidin that had been activated with dithiothreitol in the absence of guanidine hydrochloride. The disulphide bridges of actinidin are not reduced by dithiothreitol at 4°C in the absence of a denaturing agent. The disulphide bonds were then reduced with dithiothreitol in the presence of guanidine hydrochloride and carboxymethylated with non-radioactive iodoacetate. The final product was labelled at 1.4×10^6 d.p.m./mol.

Maleylation

S-Carboxymethyl-actinidin that had been selectively labelled with ^{14}C at the active site was maleylated with maleic anhydride by the method of Butler & Hartley (1972). The maleyl groups were removed from maleylated peptides after the peptides had been separated by paper electrophoresis by suspending the paper in the vapour of pyridinium acetate [1% (v/v) pyridine/5% (v/v) acetic acid, pH3.5] at 60°C for 16h.

Enzymic digestions

A solution of carboxymethyl-actinidin (5mg/ml) was digested at 37°C under N_2 with trypsin at an enzyme/substrate ratio of 1:100 (w/w) for 1h. An equal amount of trypsin was then added and the digestion continued for a further 1h. The pH of the solution was maintained at pH8.0 by the addition of NaOH (0.05M). The digestion was stopped by the addition of soya-bean trypsin inhibitor (equimolar with trypsin) and the material was freeze-dried. Further degradations of the large tryptic peptides were accomplished by digestion with chymotrypsin, thermolysin or *S. aureus* V8 proteinase in 1% (w/v) NH_4HCO_3 or pepsin in 5% (v/v) formic acid. Digestions were made at 37°C for 2–3h with an enzyme/substrate ratio of 1:100 and were stopped by freeze-drying.

The *C*-terminal amino acids were determined with carboxypeptidase A. Carboxypeptidase A had no effect on native actinidin, but removed the *C*-terminal amino acids from actinidin that had been denatured with HCl (pH1.0) for 30min. Carboxypeptidase A (5μg), dissolved by the procedure of Harris (Ambler, 1972), was added to acid-denatured actinidin (300nmol) in *N*-ethylmorpholine buffer (0.2M) and left to react at 37°C for 2h. The reaction was stopped by freeze-drying and the amino acids released were measured directly on an amino acid analyser.

Peptide fractionation

The tryptic peptides from carboxymethyl-actinidin were fractionated initially on two 100cm × 2.5cm columns of Sephadex G-50 (superfine grade). A Sephadex G-75 column (100cm × 2.5cm) was used to separate the tryptic peptides from maleylated actinidin and Sephadex G-25 was used to fractionate smaller peptides produced by further enzymic digestion of the tryptic peptides. In all cases the eluent was NH_4HCO_3 (0.2M) with a flow rate of 20ml/h. The large tryptic peptides from the Sephadex G-50 fractionations were purified further on a Whatman DE-32 DEAE-cellulose column (100cm × 2.5cm) by using two linear gradients, one of 0.025–0.2M-NaCl (1 litre) followed by a second gradient of 0.2–0.5M-NaCl (0.5 litre), both in Tris/HCl buffer (0.05M) at pH8.0.

Small peptides were purified by high-voltage electrophoresis on paper (Michl, 1951) by using the

buffer systems at pH6.5, 3.5 and 2.1 of Tang & Hartley (1967) and by chromatography in butanol/acetic acid/water/pyridine (15:10:3:12, by vol.). The peptides were located with cadmium/ninhydrin reagent for amino groups (Heathcote & Haworth, 1969), Ehrlich's reagent for tryptophan (Smith, 1953), phenanthraquinone for arginine (Yamada & Itano, 1966) and Pauly reagent for histidine and tyrosine (Frank & Petersen, 1955). Peptides were eluted from paper with 0.02M-NH_3.

Amino acid analysis

Amino acid analyses were made on actinidin that had been oxidized with performic acid. Samples (1mg) were hydrolysed in evacuated tubes with 1ml of HCl [6M, containing 0.1% (w/v) phenol] for 24, 48 and 72h. Quantitative analyses of the hydrolysates were made with a Locarte single-column amino acid analyser. Tryptophan was determined colorimetrically by the method of Scoffone *et al.* (1968). Peptides isolated from enzymic digestions were hydrolysed for 16h.

Sequence determination

The manual dansyl–Edman procedure described by Hartley (1970) was used. The rapid Edman degradation procedure of Gray (1972) was used to obtain the sequence of the first three to five residues from the *N*-terminal end of the larger tryptic peptides. Tryptophan residues in sequences were found by staining samples, which had been spotted on paper, with Ehrlich's reagent after each Edman cycle.

Determination of amide groups

The electrophoretic mobilities of peptides at pH6.5 were determined relative to aspartic acid and overall charges were calculated on the basis of the observed mobility and the molecular weight estimated from the amino acid composition (Offord, 1966). In peptides that contained both amides and acids, the amide assignments were made by measuring the mobility of the peptide after each Edman degradation cycle. In some cases the cadmium/ninhydrin-stain colour difference between an *N*-terminal asparagine (yellow) and aspartic acid (red) provided supporting evidence. Additional evidence for confirming the positions of amide bonds was obtained by using *S. aureus* V8 proteinase, which is specific for the carboxy group of glutamic acid (Houmard & Drapeau, 1972; Emmens *et al.*, 1976).

Results and Discussion

Purification of actinidin

A 40-fold increase in the specific activity of actinidin was obtained during the purification (Table 1). The enzymic activity was associated with the trailing portion of the protein peak that was eluted from the DEAE-cellulose column (Fig. 1), in agreement with the results of Boland & Hardman (1972). Polyacrylamide-gel electrophoresis (Davis, 1964) of crystallized actinidin in 15% gels gave a single band with a mobility of 0.8 relative to Bromophenol Blue. Dansylation of crystallized actinidin gave a single *N*-terminal leucine. Table 2 shows the amino acid composition of crystallized actinidin, based on a mol.wt. of 23500.

Purification of the tryptic peptides of carboxymethyl-actinidin

The low proportion of basic amino acid residues in actinidin (six lysine and five arginine) prompted the digestion of actinidin with trypsin, since this would produce a relatively small number of peptides for purification. Sephadex G-50 chromatography of the tryptic digest of carboxy[^{14}C]methyl-actinidin separated the material into seven fractions (Fig. 2). Fractions I, II, III and IV were rechromatographed separately on DEAE-cellulose and a composite diagram of the separate runs is shown in Fig. 3. Fraction I contained the tryptic peptide T6, fraction II contained peptide T7, fraction III contained peptide T8 and fraction IV contained peptides T3 and T11. Fractions V, VI and VII contained peptides that could be purified by paper techniques. Fraction V contained peptide T2, fraction VI contained peptides T5 and T9, and fraction VII contained peptides T1, T10 and T12. The yields, electrophoretic mobilities and compositional data of the peptides are shown in Table 3.

Table 1. *Purification of actinidin from 1kg of fruit*

Procedure	Total protein (mg)	Total activity (units)	Specific activity (units/mg)
Centrifuged extract	14500	33350	2
Redissolved $(NH_4)_2SO_4$ precipitate	1700	27350	16
Supernatant after calcium phosphate treatments	1600	26400	16
DEAE-cellulose eluate	300	22500	75
Crystallized enzyme	230	18400	80

Table 2. *Amino acid composition of actinidin*

The number of residues of each amino acid was calculated from 24h, 48h and 72h hydrolysates, which were each an average of three analyses. The results for serine and threonine were taken from extrapolations to zero time and the values for valine, leucine and isoleucine were values at 72h. Cysteine and cystine were determined as cysteic acid, and methionine was measured as methionine sulphone. Tryptophan was determined separately by the method of Scoffone *et al.* (1968).

Amino acid	Content (residues/molecule)			Residue number per 23500 mol.wt.	Deduced sequence composition
Hydrolysis time ...	24h	48h	72h		
His	1.4	1.5	1.3	1.4	1
Lys	6.4	6.2	6.5	6.3	6
Arg	5.0	5.1	5.0	5.1	5
Cys	6.7	6.4	6.5	6.6	7
Asp	29.1	29.4	28.5	28.7	29
Thr	17.4	16.9	16.5	17.8	18
Ser	11.6	11.0	10.2	12.4	12
Glu	20.3	19.8	20.6	20.4	20
Pro	6.8	7.0	6.5	6.9	7
Gly	27.5	27.4	28.1	27.6	28
Ala	17.5	17.4	17.7	17.7	18
Val	16.1	16.5	16.7	16.7	17
Met	2.1	2.1	1.9	2.1	2
Ile	16.6	16.5	16.9	16.9	17
Leu	7.4	7.3	7.7	7.7	8
Tyr	14.4	14.5	14.0	14.3	14
Phe	5.6	4.9	5.2	5.3	5
Trp	—	—	—	5.8	6
					220

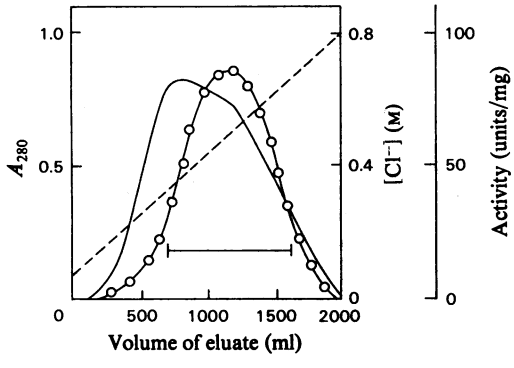


Fig. 1. *DEAE-cellulose chromatography of actinidin*

The dialysed solution of actinidin, which had been precipitated from the extraction medium with 70%-satd. $(NH_4)_2SO_4$, was loaded on to a DEAE-cellulose column (20cm×4.5cm diam.) equilibrated with potassium phosphate buffer (0.1M, pH6.0) containing 0.1M-KCl and then eluted from the column with a linear gradient (----) of 0.1M-KCl (1 litre) to 1.0M-KCl (1 litre) in the phosphate buffer. ——, A_{280}; ○, activity against benzyloxycarbonyl-lysine *p*-nitrophenyl ester. The bar represents the fraction taken for crystallization.

The relatively large amounts of actinidin available for each experiment (10μmol) enabled the purification procedures to be continued until the background contamination in the amino acid analyses was less than 0.2mol of amino acid/mol of peptide. These background values are not included in Table 3.

Supplementary information

Detailed evidence for the amino acid sequence of the protein has been deposited with the British Library Lending Division, Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K. (Supplementary Publication SUP 50083).

The information comprises the following.

(1) Tables showing the properties of the digestion fragments of peptides T3, T6, T7, T8 and T9 that were isolated, as shown in Fig. 4. The electrophoretic mobilities at pH6.5, pH3.5 and pH2.1, and chromatography mobilities are given where applicable. Relative overall yields of peptides, amino acid analyses, *N*-termini and the net charge on each peptide at pH6.5 are included in the Tables.

(2) Further information in the Tables giving details of peptide isolation, sequence and alignment, including the isolation of peptide T4.

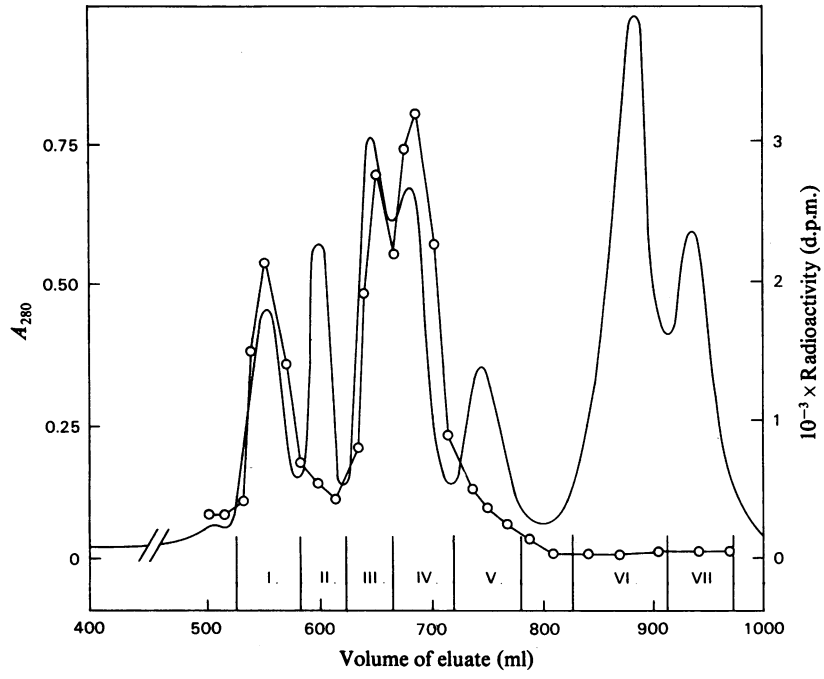


Fig. 2. *Sephadex G-50 chromatography of the tryptic peptides of carboxymethyl-actindin*
The tryptic digest (5 μmol) of carboxymethyl-actinidin was subjected to gel filtration on two Sephadex G-50 columns (100cm × 2.5cm diam.) connected in series, with 0.2M-NH_4HCO_3 as eluent with an ascending flow rate of 20ml/h. ——, A_{280}; ○, radioactivity. Fractions were pooled to give the seven major fractions I–VII.

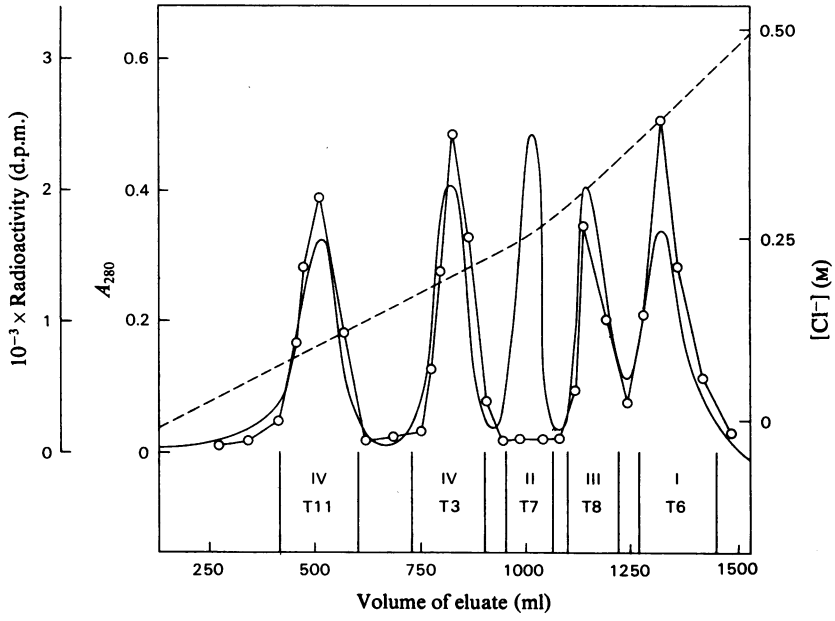


Fig. 3. *Composite diagram of DEAE-cellulose chromatography of the tryptic peptides of carboxymethyl-actinidin*
Fractions I, II, II and IV from the Sephadex G-50 separation were pooled separately and applied to a column of DEAE-cellulose (95cm × 2.5cm diam.). Each fraction was eluted with the same gradient of 0.025–0.2M-NaCl (500ml each in 0.05M-Tris/HCl, pH8.0) followed by a second gradient of 0.2–0.5M-NaCl (250ml each in the same buffer). The elution patterns from the four separate runs are combined in the diagram. ——, A_{280}; ○, radioactivity; ----, salt concentration of the gradient.

Table 3. *Peptides obtained from tryptic digestion of actinidin* (10 μmol)

The inclusion of mobility and R_F measurements indicates that the peptides were separated by paper electrophoresis followed by chromatography (see the text for details). Where no measurements are included the peptides were isolated by other methods described in the text. Mobility measurements were made relative to aspartic acid for pH6.5 and relative to dansyl-arginine for pH2.1. The negative sign indicates an acidic peptide. Amino acid residues with values less than 0.2 mol/mol of peptide are not included.

Peptide ...	T1	T2	T3	T5	T6	T7	T8	T9	T10	T11	T12
Mobility at pH6.5	0	0	—	0.37	—	—	—	−0.37	0.50	0	0
Mobility at pH2.1	0.60	0.73	—	—	—	—	—	—	—	0.47	0.50
R_F	0.86	0.58	—	0.26	—	—	—	0.52	0.82	0.44	0.48
N-Terminus	Leu	Ser	Ser	Thr	Gly	Tyr	Glx	Asn	Ile	Asn	Tyr
Charge at pH6.5	0	0	—	1	—	—	—	−2	1	0	0
10^{-6} × radioactivity (d.p.m/μmol)	—	—	2.5	—	2.2	—	1.1	—	—	1.3	—
Yield (% of peptide digested)	60	50	70	55	75	60	65	45	62	54	50
Carboxymethyl-cysteine	—	—	1.9	—	1.9	—	0.9	—	—	0.9	—
Aspartic acid	1.0	1.0	1.0	1.0	10.5	6.1	2.1	1.9	—	1.1	2.0
Threonine	—	—	1.0	1.8	2.8	3.8	2.8	1.9	—	1.8	—
Serine	0.9	0.8	1.8	—	—	1.0	1.0	1.0	—	0.9	—
Glutamic acid	—	—	3.0	1.1	5.2	4.2	2.0	2.0	—	—	—
Proline	1.0	—	—	—	0.9	1.9	0.9	—	—	1.9	—
Glycine	—	1.0	3.9	—	6.7	1.1	7.2	2.0	—	4.3	—
Alanine	—	2.0	3.4	—	1.9	5.9	2.9	—	—	2.3	—
Valine	0.9	1.8	1.0	—	1.0	4.8	4.8	—	—	2.0	—
Methionine	—	—	—	—	—	—	—	0.9	—	0.9	—
Isoleucine	—	1.0	2.0	—	3.9	1.0	4.0	—	1.0	1.2	—
Leucine	1.0	—	—	—	1.2	2.0	—	—	1.0	—	—
Tyrosine	1.0	—	—	—	3.0	4.0	2.9	1.1	—	1.3	1.0
Phenylalanine	—	—	0.9	—	2.0	1.0	1.0	—	—	—	—
Lysine	—	1.0	1.0	—	1.0	1.0	1.0	—	—	1.0	—
Histidine	—	—	—	—	—	—	1.0	—	—	—	—
Arginine	1.0	—	—	0.9	—	—	—	1.0	1.0	—	—
Tryptophan	+	—	+	—	—	+	+	+	—	—	—

(3) Figures providing evidence for the location in peptides of amide residues that could not be ascertained directly from the net charge on each of these peptides.

Sequence of the tryptic peptides

Peptides T1, T2, T5, T9, T10 and T12 were isolated by electrophoresis followed by chromatography on paper. The composition and chromatographic values of each peptide are shown in Table 3, and the sequences as determined by the dansyl–Edman method are given in Fig. 4. The amide residues were assigned unambiguously from the electrophoretic mobilities.

A chymotryptic digest of peptide T9 gave four peptides with compositions that identified the tryptophan residues, assigned the amides and confirmed the remaining sequence. Peptides T3, T6, T7 and T8 (Table 3) were too large to be completely sequenced by the dansyl–Edman technique. The first three to five residues of these peptides were determined by the rapid Edman procedure of Gray (1972) before digestion with other enzymes to give smaller peptides. These smaller peptides were aligned to give the sequences summarized in Fig. 4.

In some peptides the amide residues could not be assigned unambiguously on the basis of the net charge on the peptide. The glutamine residue at position 2 in peptide T3C1 was established from mobility measurements at pH6.5 on samples taken after each Edman cycle.

Peptide T6G1C4 contained both aspartic acid and asparagine. Digestion with thermolysin produced two peptides T6G1C4L1 and T6G1C4L2 and the amide residues in these two peptides were determined by the same method used for peptide T3C1.

In some secondary digests of purified tryptic peptides, all of the peptides to complete the sequence could not be isolated. For example, the *N*-terminal five-residue sequence of peptide T7G3 could not be found in the peptic digest of peptide T7; it is possible

that this sequence may have undergone partial enzyme digestion producing a number of small fragments in low yield.

However, in situations such as this, sufficient sequence information was obtained from different digests and other methods to give an unambiguous sequence for each tryptic peptide.

Peptide T4 presented a problem in that it was the only tryptic peptide of carboxymethyl-actinidin that could not be located in the Sephadex or DEAE-cellulose elution profiles. Comparison of the amino acid analysis of actinidin with the analyses of the 11 tryptic peptides indicated that peptide T4 should contain one cysteine and one arginine residue, but no aromatic residues. Further, results obtained from the high-resolution X-ray-crystallographic analysis of actinidin (Baker, 1977) indicated that peptide T4 was located between peptides T3 and T5. This location prompted the isolation of peptide T4 from a tryptic digest of maleylated carboxymethyl-actinidin. The lysine residues in peptides T2 and T3 would be blocked by the maleyl groups and not cleaved by trypsin, and peptide T4 could be isolated as the maleyl-peptide containing peptides T2, T3 and T4. Furthermore peptide T3 contained the active-site thiol group, which was specifically carboxymethylated (see the Experimental section). Maleylated carboxymethyl-actinidin (5μmol), radioactively labelled at the active site (1.4×10^6 d.p.m./mol), was digested with trypsin and the required peptide was separated as the only radioactive tryptic peptide on Sephadex G-75. Further digestion of this maleylated tryptic peptide with chymotrypsin and *S. aureus* proteinase enabled the sequence of peptide T4 to be determined.

The digestion of acid-denatured actinidin with carboxypeptidase A released 2mol of asparagine and 1mol each of tyrosine, lysine and valine per mol of actinidin.

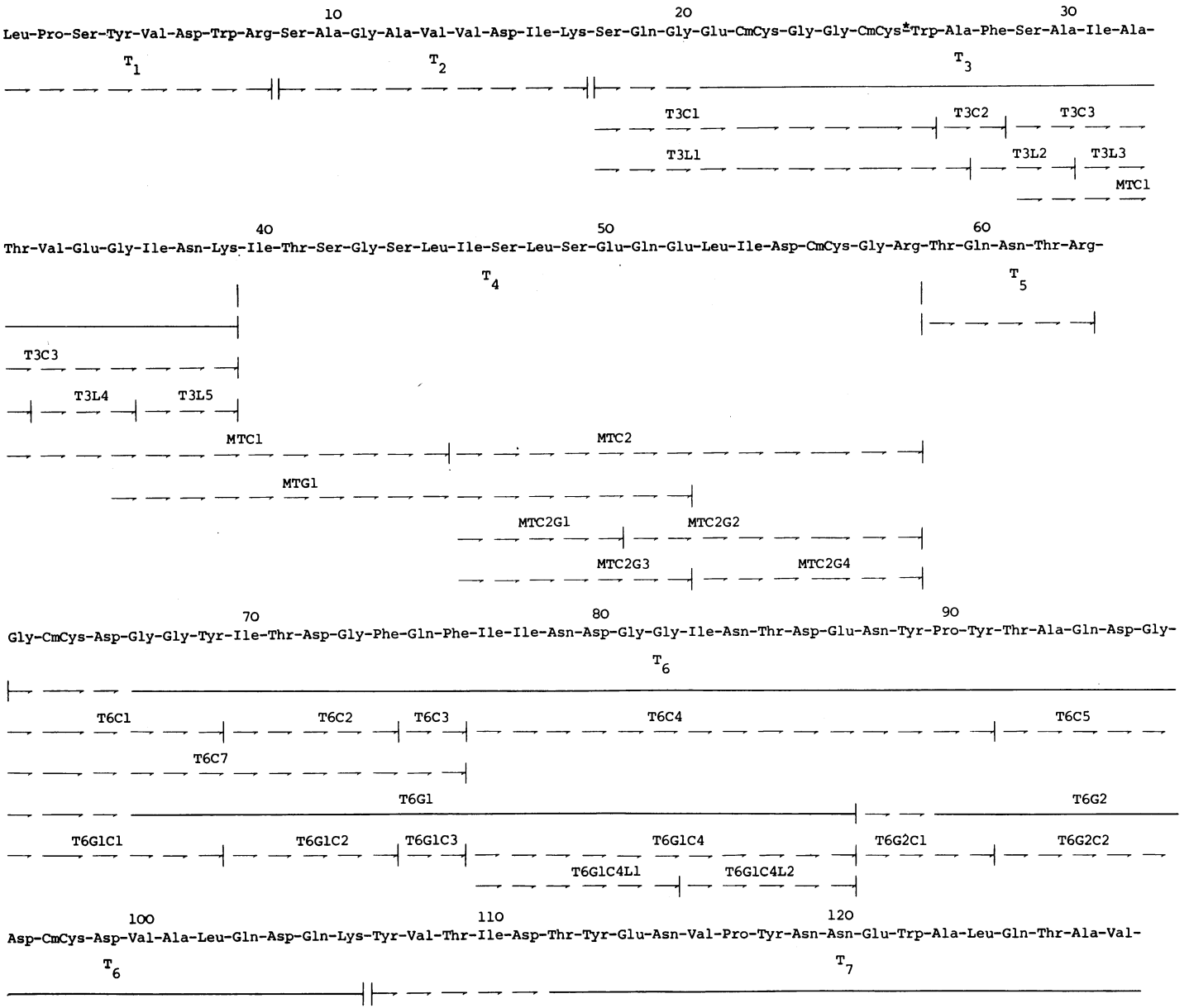
Tryptic-peptide alignment

The composition of the 12 tryptic peptides accounts for the complete amino acid analysis of actinidin. Ordering of the tryptic peptides was made possible by reference to the 0.28nm-resolution electron-density map obtained in the X-ray-crystallographic studies by Baker (1977). He showed that the course of the polypeptide chain could be followed easily with clear resolution of six tryptophan, five phenylalanine and 11 out of 14 tyrosine residues. Other residues that could usually be identified with confidence from the X-ray analysis without prior reference to sequence data included glycine, alanine and serine, and the branching of some side chains in the interior of the molecule distinguished them as leucine or isoleucine. This enabled the various tryptic peptides to be recognized along the course of the polypeptide chain and arranged in the correct order. They have therefore been numbered in order from the *N*-terminus, and the sequence T1 to T11 clearly fits the X-ray data.

Peptide T1 was the only tryptic peptide having the correct *N*-terminal residue for actinidin, and the tryptophan residue coincided with the location of the tryptophan at position 8 from the X-ray study. Peptides T2, T3 and T4 were overlapped during the isolation of peptide T4 and the location of the active-site cysteine residue at position 25 in the sequence was confirmed along with cysteine-22 and tryptophan-26 by the X-ray data. Peptide T5 was the only tryptic peptide that could fill the five-residue gap formed by the alignment of peptide T6. Cysteine-65, phenylalanine-74 and phenylalanine-76 positioned peptide T6 in the sequence, and cysteine-98 of peptide T6, coupled with cysteine-56 of peptide T4, explained the electron density between these two residues in the X-ray map. The aromatic residues in peptides T7 and T8 were adequate to enable the unambiguous alignment of these peptides. The sole histidine residue present in peptide T8, which was located in the active site at position 162 by the X-ray data, further confirmed this alignment along with cysteine-156 in the primary sequence. The two tryptophan residues in peptide T9 aligned with positions 184 and 188 in the polypeptide chain, as did the cysteine residue in peptide T11 with position 206. The gap of three residues between peptides T9 and T11 could only be filled by peptide T10. Peptide T12 was the only peptide that could not be ordered with certainty from the X-ray analysis, since the last two residues were not visible in the electron-density map. From studies with carboxypeptidase A, peptide T12 was shown clearly to be the *C*-terminal peptide, and it was the only tryptic peptide lacking a basic *C*-terminal residue. The location of the proline residues in the sequence coincided with bends in the well-defined polypeptide chain obtained from the X-ray analysis, further supporting the tryptic-peptide alignment.

Primary-sequence homology

Comparison of the primary sequence of actinidin with that of papain (Fig. 5) demonstrates the considerable degree of homology in their primary sequences, which parallels the close similarity in their three-dimensional structures. It is interesting to note the decrease in homology in the regions of residues 70–130 compared with the rest of the molecule. The X-ray data show that this region of the primary sequence is predominantly on the outside of the molecule, between two α-helical regions, and forms the main link between the two domains of the molecule, and changes in amino acid residues in this region would not alter the overall structure to any great extent.


10 20 30
Leu-Pro-Ser-Tyr-Val-Asp-Trp-Arg-Ser-Ala-Gly-Ala-Val-Val-Asp-Ile-Lys-Ser-Gln-Gly-Glu-CmCys-Gly-Gly-CmCys*Trp-Ala-Phe-Ser-Ala-Ile-Ala-
T1
T2
T3
T3C1
T3C2
T3C3
T3L1
T3L2
T3L3
MTC1
40 50 60
Thr-Val-Glu-Gly-Ile-Asn-Lys-Ile-Thr-Ser-Gly-Ser-Leu-Ile-Ser-Leu-Ser-Glu-Gln-Glu-Leu-Ile-Asp-CmCys-Gly-Arg-Thr-Gln-Asn-Thr-Arg-
T4
T5
T3C3
T3L4
T3L5
MTC1
MTC2
MTG1
MTC2G1
MTC2G2
MTC2G3
MTC2G4
70 80 90
Gly-CmCys-Asp-Gly-Gly-Tyr-Ile-Thr-Asp-Gly-Phe-Gln-Phe-Ile-Ile-Asn-Asp-Gly-Gly-Ile-Asn-Thr-Asp-Glu-Asn-Tyr-Pro-Tyr-Thr-Ala-Gln-Asp-Gly-
T6
T6C1
T6C2
T6C3
T6C4
T6C5
T6C7
T6G1
T6G2
T6G1C1
T6G1C2
T6G1C3
T6G1C4
T6G2C1
T6G2C2
T6G1C4L1
T6G1C4L2
100 110 120
Asp-CmCys-Asp-Val-Ala-Leu-Gln-Asp-Gln-Lys-Tyr-Val-Thr-Ile-Asp-Thr-Tyr-Glu-Asn-Val-Pro-Tyr-Asn-Asn-Glu-Trp-Ala-Leu-Gln-Thr-Ala-Val-
T6
T7

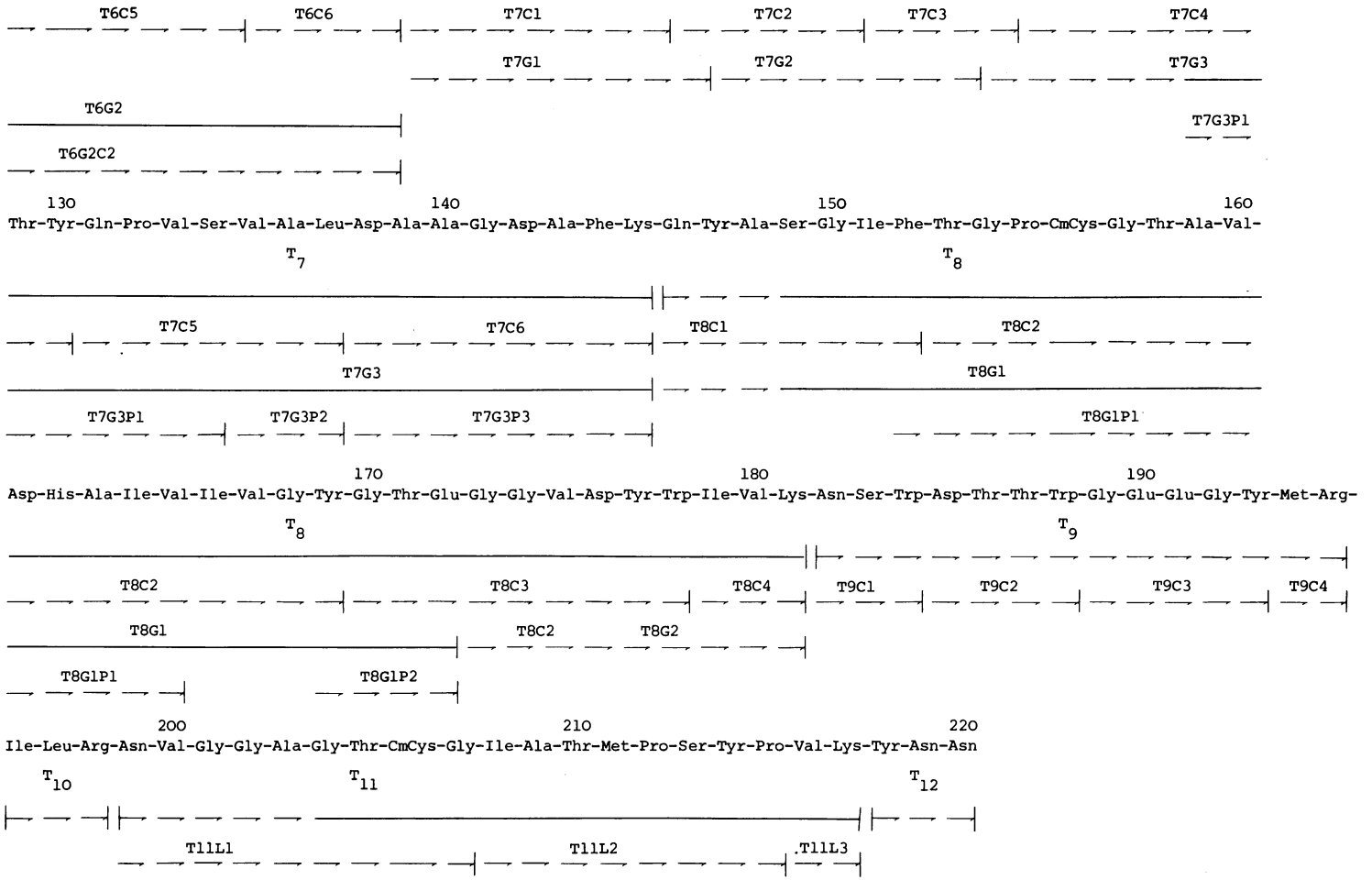


Fig. 4. *Sequences of the tryptic peptides and their alignment to give the primary sequence of actinidin*

The letters T, C, L and G indicate that the peptide was derived from a digestion by: T, trypsin; C, chymotrypsin; L, thermolysin; G, *S. aureus* proteinase. Additional letters indicate further digestion, e.g. TC indicates a chymotryptic peptide from a tryptic peptide. The letter M is used to show that the peptide was isolated as a maleylated derivative. ——›, Residue established by the dansyl–Edman procedure; ——, residues consistent with the amino acid analysis after hydrolysis. * Cysteine residue that was specifically labelled with iodo[2-^{14}C]acetic acid. CmCys, carboxymethylcysteine.

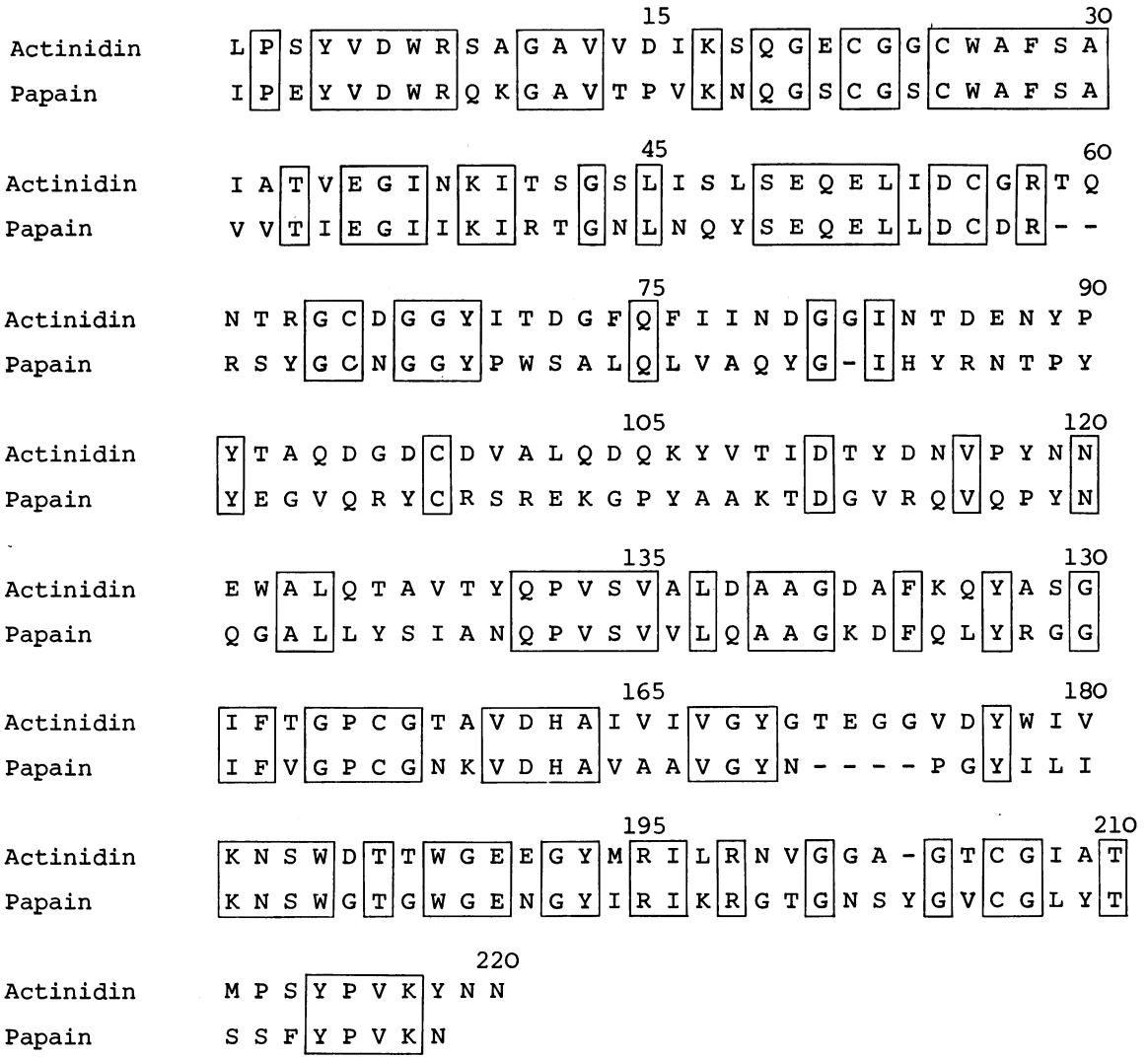


Fig. 5. *Sequences of actinidin and papain aligned to show homologies*

The dashes in the sequences indicate possible deletions or insertions of amino acids. The sequence numbering is for actinidin and the boxed areas indicate the identical residues. The following single-letter notation is used: A, alanine; C, cystine or cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine, I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.

References

Ambler, R. P. (1972) *Methods Enzymol.* **25**, 143–154
Arcus, A. C. (1959) *Biochim. Biophys. Acta* **33**, 242–244
Baker, E. N. (1973) *J. Mol. Biol.* **74**, 411–412
Baker, E. N. (1976) *J. Mol. Biol.* **101**, 185–196
Baker, E. N. (1977) *J. Mol. Biol.* **115**, 263–277
Boland, M. J. & Hardman, M. J. (1972) *FEBS Lett.* **27**, 282–284
Butler, P. J. G. & Hartley, B. S. (1972) *Methods Enzymol.* **25**, 191–199
Davis, B. J. (1964) *Ann. N.Y. Acad. Sci.* **121**, 404–427
Drenth, J., Hol, W. G. J., Visser, J. W. E. & Sluyterman, L. A. A. E. (1968*a*) *J. Mol. Biol.* **34**, 369–371
Drenth, J., Jansonius, J. N., Koekoek, R., Swen, H. M. & Wolthers, B. G. (1968*b*) *Nature* (*London*) **218**, 929–932
Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91
Emmens, M., Welling, G. W. & Beintema, J. J. (1976) *Biochem. J.* **157**, 317–323
Frank, H. & Petersen, H. (1955) *Hoppe-Seyler's Z. Physiol. Chem.* **299**, 1–5
Glazer, A. N. & Smith, E. L. (1971) *Enzymes*, 3*rd edn.* **3**, 501–546
Goto, K., Murachi, T. & Takahashi, N. (1976) *FEBS Lett.* **62**, 93–95
Gray, W. R. (1972) *Methods Enzymol.* **25**, 333–342
Hartley, B. S. (1970) *Biochem. J.* **119**, 805–822
Heathcote, J. G. & Haworth, C. (1969) *J. Chromatogr.* **43**, 84–92
Hirs, C. H. W. (1967) *Methods Enzymol.* **11**, 199–203
Houmard, J. & Drapeau, G. R. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3506–3509
Husain, S. S. & Lowe, G. (1969) *Biochem. J.* **114**, 279–288
Husain, S. S. & Lowe, G. (1970*a*) *Biochem. J.* **116**, 689–692
Husain, S. S. & Lowe, G. (1970*b*) *Biochem. J.* **117**, 333–340
Husain, S. S. & Lowe, G. (1970*c*) *Biochem. J.* **117**, 341–346
Light, A., Frater, R., Kimmel, J. R. & Smith, E. L. (1964) *Proc. Natl. Acad. Sci. U.S.A.* **52**, 1276–1283
Liu, T.-Y. & Inglis, A. S. (1972) *Methods Enzymol.* **25**, 57
McDowall, M. A. (1970) *Eur. J. Biochem.* **14**, 214–221
Michl, H. (1951) *Monatsh. Chem.* **82**, 489–493
Mitchel, R. E. J., Chaiken, I. M. & Smith, E. L. (1970) *J. Biol. Chem.* **245**, 3485–3492
Offord, R. E. (1966) *Nature* (*London*) **211**, 591–593
Scoffone, E., Fontana, A. & Rocchi, R. (1968) *Biochemistry* **7**, 971–979
Smith, I. (1953) *Nature* (*London*) **171**, 43–44
Tang, J. & Hartley, B. S. (1967) *Biochem. J.* **102**, 593–599
Wong, R. C. & Liener, I. E. (1964) *Biochem. Biophys. Res. Commun.* **17**, 470–474
Yamada, S. & Itano, H. A. (1966) *Biochim. Biophys. Acta* **130**, 538–540