The Amino Acid Sequence of a Carboxypeptidase Inhibitor from Potatoes†

G. Michael Hass,‡ H. Nau, K. Biemann, Donald T. Grahn, Lowell H. Ericsson, and Hans Neurath*

ABSTRACT: The carboxypeptidase inhibitor from Russet Burbank potatoes (C. A. Ryan et al. (1974b), *J. Biol. Chem. 249,* 5495) is a mixture of approximately equal amounts of two polypeptide chains containing 38 and 39 amino acid residues, respectively. The chains differ in their amino terminal sequence only, one beginning with <Glu-His-Ala . . . and the other with <Glu-Gln-His-Ala . . . . . Specific cleavage procedures utilized in determining the complete amino acid sequence of the inhibitor included acid cleavage of the aspartyl–proline bond and tryptic and chymotryptic digestion. Mass spectrometry, automatic Edman degradation, and subtractive Edman degradation were employed in sequencing the resulting peptide fragments.

The purification and characterization of a naturally occurring inhibitor of the pancreatic carboxypeptidases A and B have recently been reported (Ryan et al., 1974b). The inhibitor is a single polypeptide chain containing approximately 38 amino acid residues, three disulfide bonds, a blocked amino terminus, and glycine as the carboxyl terminus (Ryan et al., 1974b).

The present report describes the determination of the amino acid sequence of this carboxypeptidase inhibitor. This information should provide a basis for subsequent chemical modifications designed to elucidate the interaction of this inhibitor with the carboxypeptidases.

Experimental Section

Materials. The carboxypeptidase inhibitor was prepared from Russet Burbank potatoes by the method of Ryan et al. (1974b). As judged by electrophoretic, chromatographic, and chemical analyses previously described (Ryan et al., 1974b) the inhibitor is considered to be a homogeneous protein.

Trypsin (treated with L-1-tosylamido-2-phenylethyl chloromethyl ketone, lot (TRTPCK-2B), and α-chymotrypsin (3X crystallized, lot CDI-8JA) were purchased from Worthington Biochemical Corporation. Dowex 1-X2 (Bio-Rad Laboratories) and Sephadex G-25F (Pharmacia Fine Chemicals) were used according to the manufacturers' specifications. 4-Vinylpyridine (Eastman) was vacuum distilled prior to use. [1-^{14}C]Bromoacetic acid (New England Nuclear) was diluted with unlabeled bromoacetic acid to give a specific radioactivity 0.02 Ci/mol. Pyridine was distilled from ninhydrin prior to use in chromatographic and electrophoretic systems (Hill and Delaney, 1967).

All other chemicals were reagent grade or better and were used without further purification.

Detection of Acetyl and Formyl Groups. Acetyl and formyl groups were identified as the corresponding 1-acyl-2-dansylhydrazines as described by Schmer and Kreil (1969). Suitable quantities (150 nmol) of each, inhibitor, horse cytochrome *c* (which contains an N-acetylated amino-terminal residue), acetyl-L-alanine, and formyl-L-phenylalanine, were treated with hydrazine and then with dansyl chloride. The 1-acyl-2-dansylhydrazines were extracted and subjected to thin-layer chromatography on Kieselgel G plates using chloroform–1-butanol–acetic acid (6:3:1) as solvent. The dansyl derivatives were detected by their fluorescence upon irradiation with ultraviolet light.

Preparation of Inhibitor for Sequence Analysis. The *S*-pyridylethyl[1] and *S*-carboxymethyl derivatives of inhibitor were prepared by reduction and alkylation according to the procedures of Cavins and Friedman (1970) and Crestfield et al. (1963), respectively. The modified peptide derivatives were separated from excess reagent and reaction by-products by gel filtration chromatography on a 1.5 × 80 cm column of Sephadex G-25 equilibrated and eluted with 0.1 *M* propionic acid.

Cleavage of the Asp–Pro Bond and Separation of Peptides. The aspartyl–proline bond of *S*-Cm-inhibitor was cleaved by incubation of 20 mg of protein in 4 ml of 10% acetic acid adjusted to pH 2.5 with pyridine (Fraser et al., 1972). After 4 days of incubation at 40°, the resulting fragments were separated by gel filtration on a 1.5 × 80 cm column of Sephadex G-25 equilibrated and eluted with 9% formic acid. Approximately 1-ml fractions were collected and the elution was monitored by absorbance at 280 nm and by ninhydrin analysis following alkaline hydrolysis of 0.05-ml aliquots from alternate fractions.

Two forms of peptide fragment 1 (AP-I) were separated by preparative paper electrophoresis at pH 6.5 for 2 hr at 2 kV using prewashed Whatman 3MM paper. A guide strip was stained with ninhydrin and three zones, which migrated

† From the Department of Biochemistry, University of Washington, Seattle, Washington 98195 (G.M.H., D.T.G., L.E., and H. Neurath), and the Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 (H. Nau and K.B.). *Received September 25, 1974.* This work was supported by grants from the National Institutes of Health (GM 15731), and the American Cancer Society (BC-91P) to the University of Washington and by a grant from the National Institutes of Health (GM 05472) to M.I.T. Preliminary accounts of this work have been presented (Ryan et al., 1974a; Hass et al., 1974).

‡ Investigator of the Howard Hughes Medical Institute. Present address: Department of Bacteriology and Biochemistry, University of Idaho, Moscow, Idaho 83843.

[1] Abbreviations used are: *S*-Cm-, *S*-carboxymethyl-; *S*-Pe-, *S*-pyridylethyl-; PTH, phenylthiohydantoin.

4.4, 6.5, and 11.0 cm toward the anode, were excised and eluted with 20% acetic acid. The first two zones, AP-I-A and AP-I-B, contained peptide and the third was used as a control for contaminating amino acids.

Partial Acid Hydrolysis of S-Cm inhibitor and Fragments AP-I-A,B. Samples of the inhibitor (0.5 μmol) and the mixture of peptides (1.0 μmol) were hydrolyzed in 6 *N* HCl at 110° for 8 and 20 min, respectively. The hydrolysates were then evaporated to dryness and prepared for mass spectral analysis (vide infra).

Citraconylation of the Lysine Residues and Hydrolysis with Trypsin. *S*-[^{14}CM]-Inhibitor was dissolved in 2.5 ml of 0.1 *M* potassium phosphate (pH 8.3). Citraconic anhydride (0.12 ml) was added over a 2-hr period in 0.01-ml portions. The temperature of the reaction was 2° and the pH was maintained at pH 8.3 by the addition of 10 *N* NaOH (Habeeb and Atassi, 1970).

After the final addition of the reagent, the temperature was raised to 37° and trypsin (1% by weight of inhibitor) was added to initiate hydrolysis of the arginyl bonds. After 1 hr, trypsin (1% by weight) was again added and after an additional hour digestion was terminated by the addition of 0.05 ml of a 1 *M* solution of diisopropyl phosphorofluoridate in 2-propanol. The pH of the reaction mixture was adjusted to pH 2.5 with 88% formic acid. After overnight incubation at pH 2.5 to deblock the amino groups, the solution was lyophilized. The peptides were dissolved in 2.5 ml of 25% formic acid and separated on a 1.5 × 80 cm column of Sephadex G-25. Peptide elution was monitored by absorbance at 280 nm and by radioactivity.

Peptide T-1, which contained both lysine residues and the single arginine residue of the inhibitor, was digested with trypsin (2% by weight) in 0.2 *M* ammonium bicarbonate at 25°. After 3-hr incubation the reaction mixture was acidified with formic acid and lyophilized. The tryptic peptides (T-1-A and T-1-B) derived from peptide T-1 were separated by chromatography on Sephadex G-25. The conditions of chromatography and of monitoring peptide elution were identical with those described for the separation of peptides T-1 and T-2 (vide supra).

Hydrolysis of S-Pe-Inhibitor with Trypsin. The modified inhibitor (20 mg) was suspended in 3 ml of 0.01 *M* Tris-HCl (pH 7.5) containing 0.01 *M* $CaCl_2$. Trypsin (5% by weight) was added and hydrolysis was carried out at pH 7.5 and 37° for 6 hr. The tryptic peptides were separated from trypsin and salts on a 1.5 × 80 cm column of Sephadex G-25 equilibrated and eluted with 9% formic acid. Elution of peptides was monitored by ninhydrin analysis after alkaline hydrolysis of 0.05-ml aliquots from alternate fractions. Fractions containing the peptides were combined and lyophilized.

Hydrolysis with Chymotrypsin. *S*-[^{14}Cm]-Inhibitor (12 mg) was treated with α-chymotrypsin (2% by weight) at 37° for 2 hr in 2.5 ml of 0.01 *M* Tris-HCl maintained at pH 8.0 by automatic addition of 0.1 *M* NaOH. Digestion was terminated by lyophilization.

The lyophilized digest was dissolved in 3% pyridine and the chymotryptic peptides were fractionated on a 0.6 × 55 cm column of Dowex 1-X2 at 37°, using a continuous gradient comprised of 100 ml each (1) 3% pyridine, (2) 0.5 *M* pyridine-acetic acid (pH 3.75), (3) 1.0 *M* pyridine-acetic acid (pH 5.5), and (4) 2.0 *M* pyridine-acetic acid (pH 5.0) (Bradshaw et al., 1969). Residual peptides were eluted with 50% acetic acid. The flow rate was 20 ml/hr and fractions of approximately 2 ml were collected. Elution of peptides was monitored by automatic alkaline hydrolysis and ninhydrin analysis (Hill and Delaney, 1967), of 15% of the eluate. Aliquots (0.05 ml) were taken from alternate tubes for radioactivity measurements.

Analysis of Peptides

Amino Acid Analysis. Aliquots of digested inhibitor and of each purified peptide were hydrolyzed with 6 *N* HCl at 110° for 18 hr (Moore and Stein, 1963). Amino acid analyses were performed by the method of Spackman et al. (1958) with a Beckman Model 120B amino acid analyzer. Reported peptide yields are corrected for aliquots removed for monitoring the elution.

Paper Electrophoresis. High-voltage paper electrophoresis was performed on each purified peptide using pyridine–acetic acid buffers at pH 3.6 or pH 6.5 (Bennett, 1967) for 1.25 hr at 2 kV. Peptides were detected by ninhydrin (Bennett, 1967) or by the chlorine–toluidine reagent (Smith, 1953). Tryptophan-containing peptides were identified by staining the electropherograms with *p*-dimethylaminobenzaldehyde (Smith, 1953).

Edman Degradation. Sequential determinations of fragment AP-2 derived from *S*-Cm- and *S*-Pe- inhibitor, and of the mixture of tryptic peptides from *S*-Pe inhibitor were made with a Beckman sequencer using the method of Edman and Begg (1967) as modified by Hermodson et al. (1972). The criteria adopted by this laboratory for the identification of amino acid residues in sequence by this technique have been described in detail previously (Hermodson et al., 1972). The amino acid sequences of peptide T-2 (residues 32–39) and of the first three residues of peptide C-3 were determined by Edman degradation (Edman, 1950) using the subtractive procedure of Konigsberg and Hill (1962). Peptide C-3 was also subjected to direct Edman degradation by the procedure of Peterson et al. (1972). The PTH-amino acid derivatives were identified as described by Hermodson et al. (1972).

Mass Spectrometry. The partial acid hydrolysate of peptide AP-1-A,B was esterified with 1 ml of 3 *N* methanolic HCl at room temperature for 45 min. The reagents were removed in vacuo. Methanol (2 ml) and methyl trifluoroacetate (1 ml) were added, and the pH was adjusted to 7.5–8.0 with triethylamine (Weygand et al., 1960). After the mixture was stirred for 8 hr at room temperature, the reagents were removed in vacuo and the mixture of *N*-trifluoroacetyl methyl esters was reduced to the corresponding mixtures of trifluorodideuterioethyl poly(amino alcohols) with 2 ml of 0.9 *N* $LiAlD_4$ in dimethoxyethane (Nau, 1974). These compounds were O-trimethylsilylated and 30% of the resulting mixture was coinjected into the gas chromatograph (Perkin-Elmer 990) with 1 μl of a mixture containing three *n*-alkanes ($C_{14}H_{30}$, $C_{22}H_{46}$, and $C_{32}H_{66}$). The gas chromatographic conditions were as follows: 3-ft glass column (2 mm i.d., 0.25 in. o.d.) filled with 10% OV-17 on Gas Chrom Q (Applied Science Laboratories); flow rate, 40 ml/min; temperature, linearly programmed from 80° to 330° at 12°/min. The gas chromatograph is directly coupled via a porous fritted glass helium separator to a low resolution mass spectrometer (Hitachi-Perkin-Elmer RMU-6L). An IBM 1800 computer is used on line for the continuous recording of the mass spectra (Hites and Biemann, 1968; Biller, 1972), for the generation of mass chromatograms (Hites and Biemann, 1970), and the assignment of retention indices through automatic location of the coinjected hydrocarbon standards (Nau and Biemann, 1974).

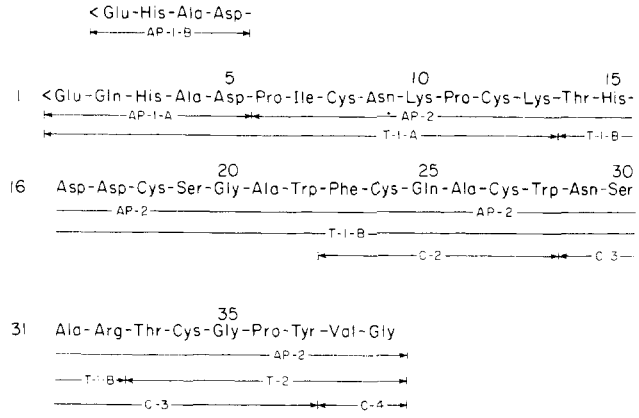

FIGURE 1: The amino acid sequence of carboxypeptidase inhibitor. One sequence begins with <Glu-Gln-His . . . and the other with <Glu-His . . . For details of the generation, purification, and characterization of the fragments, see text.

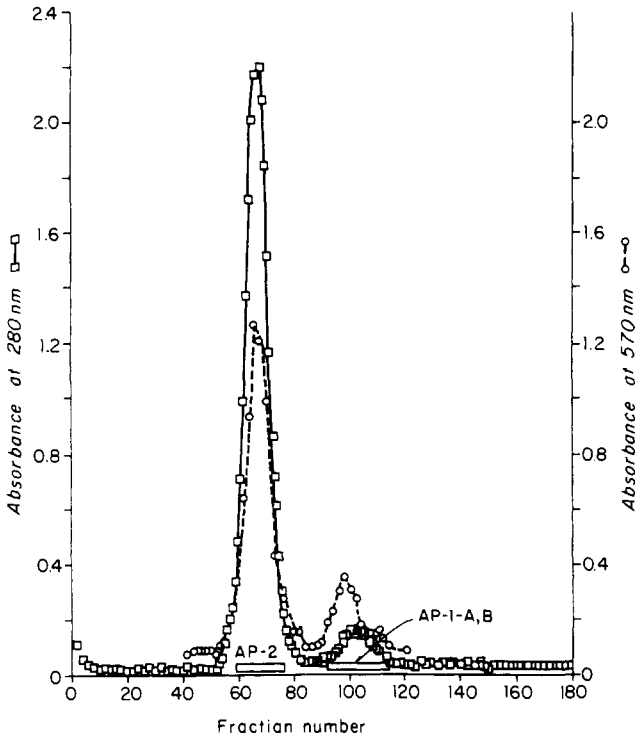

FIGURE 2: Gel filtration on a 1.5 × 80 cm column of Sephadex G-25 in 9% formic acid of the peptides derived from 12 mg of *S*-^{14}*Cm*-inhibitor by cleavage of the aspartyl-proline bond. Fractions of 1.0 ml were collected. Elution of peptides was monitored by ninhydrin analysis (O—O) after alkaline hydrolysis of 0.05-ml aliquots of alternate fractions as well as by absorbance at 280 nm (□—□).

The mixture of peptides generated by partial acid hydrolysis of *S*-Cm inhibitor was esterified as described above and acylated by heptafluorobutyric acid anhydride (Andersson, 1967). After evaporation, the residue was dissolved in 100 μl of a 1:1 mixture of ethyl acetate and chloroform. One-tenth of this solution was then subjected to analysis by gas chromatography–mass spectrometry analysis (vide supra).

Results

Identity of the Amino Terminal Residue. The inhibitor isolated by the procedure of Ryan et al. (1974b) has a blocked amino terminal residue. The most common blocking groups are pyrrolidonecarboxylic acid and *N*-acyl groups (acetyl, formyl, etc.). Since acetyl or formyl groups could not be detected by the method of Schmer and Kreil (1969) under conditions producing unequivocal identification of acetyl groups in acetyl-L-alanine and horse cytochrome *c* and of formyl groups in formyl-L-phenylalanine, the amino terminus of the inhibitor is presumed to be pyrrolidonecarboxylic acid. Mass spectral analyses (vide infra) of peptide fragments generated from the inhibitor by partial acid hydrolysis support this hypothesis.

Table I: Amino Acid Compositions of Peptides Resulting from Cleavage of Carboxypeptidase Inhibitor at the Aspartyl—Proline Bond.

Amino Acid	Inhibitor[a]	AP-I	AP-I-A	AP-I-B	AP-2
Lys	2				1.9 (2)
His	2	0.9 (1)[b]	1.0 (1)	1.0 (1)	1.0 (1)
Arg	1				0.8 (1)
CM-Cys	6	0.3 (0)			6.0 (6)
Asx	5	1.0 (1)	1.0 (1)	1.0 (1)	3.7 (4)
Thr	2				2.1 (2)
Ser	2				2.0 (2)
Glx	2.5	1.5 (1–2)	1.8 (2)	1.2 (1)	1.0 (1)
Pro	3				2.5 (2–3)
Gly	3	0.3 (0)			3.1 (3)
Ala	4	1.2 (1)	1.1 (1)	1.0 (1)	3.0 (3)
Val	1				0.9 (1)
Ile	1				1.1 (1)
Tyr	1				0.9 (1)
Phe	1				1.0 (1)
Trp[c]	2				+
Yield (%)		71	8	7	84
Residues	1.39	1–5 2–5	1–5	2–5	6–39

[a] From Ryan et al. (1974b). [b] Values in parentheses are approximations to nearest integer. [c] Positive test with *p*-dimethylaminobenzaldehyde (see text). Data for amino acids present at less than 0.2 residue/molecule were not included.

Cleavage at Asp$_5$–Pro$_6$. Evidence for Heterogeneity in the Amino Terminal Region, Identification of Residues 1–5, 6–16, and 20–29. The complete amino acid sequence of the inhibitor is presented in Figure 1. The heterogeneity in the amino terminal region of the inhibitor and a substantial amino acid sequence were elucidated by analysis of fragments generated by cleavage of the Asp$_5$–Pro$_6$ bond.

Treatment of *S*-Cm inhibitor under conditions which specifically cleave aspartyl-proline bonds (Fraser et al., 1972) produces two peptide fragments which were separated by gel filtration on Sephadex G-25 (Figure 2). The amino acid compositions of the fragments AP-1 and AP-2 are presented in Table I. Fragment AP-1 is heterogeneous containing an average of 1.5 glutamic acid residues in addition to single residues of histidine, alanine, and aspartic acid. Upon electrophoresis at pH 6.5 AP-1 separates into two ninhydrin-negative peptides labeled AP-1-A and AP-1-B having mobilities relative to aspartic acid of 0.18 and 0.25, respectively. The amino acid compositions of AP-1-A and AP-1-B and of unfractionated AP-1 suggest the following. (1) AP-1 is derived from the amino terminal region of the protein. This conclusion is based upon the presence of aspartic acid in this peptide fraction and upon the specificity of the cleavage procedure. (2) The inhibitor is a mixture of two polypeptide chains present in approximately the same amount and differing by a single glutamine residue.

Because fragments AP-1-A and AP-1-B were isolated by preparative paper electrophoresis in very low yield (Table

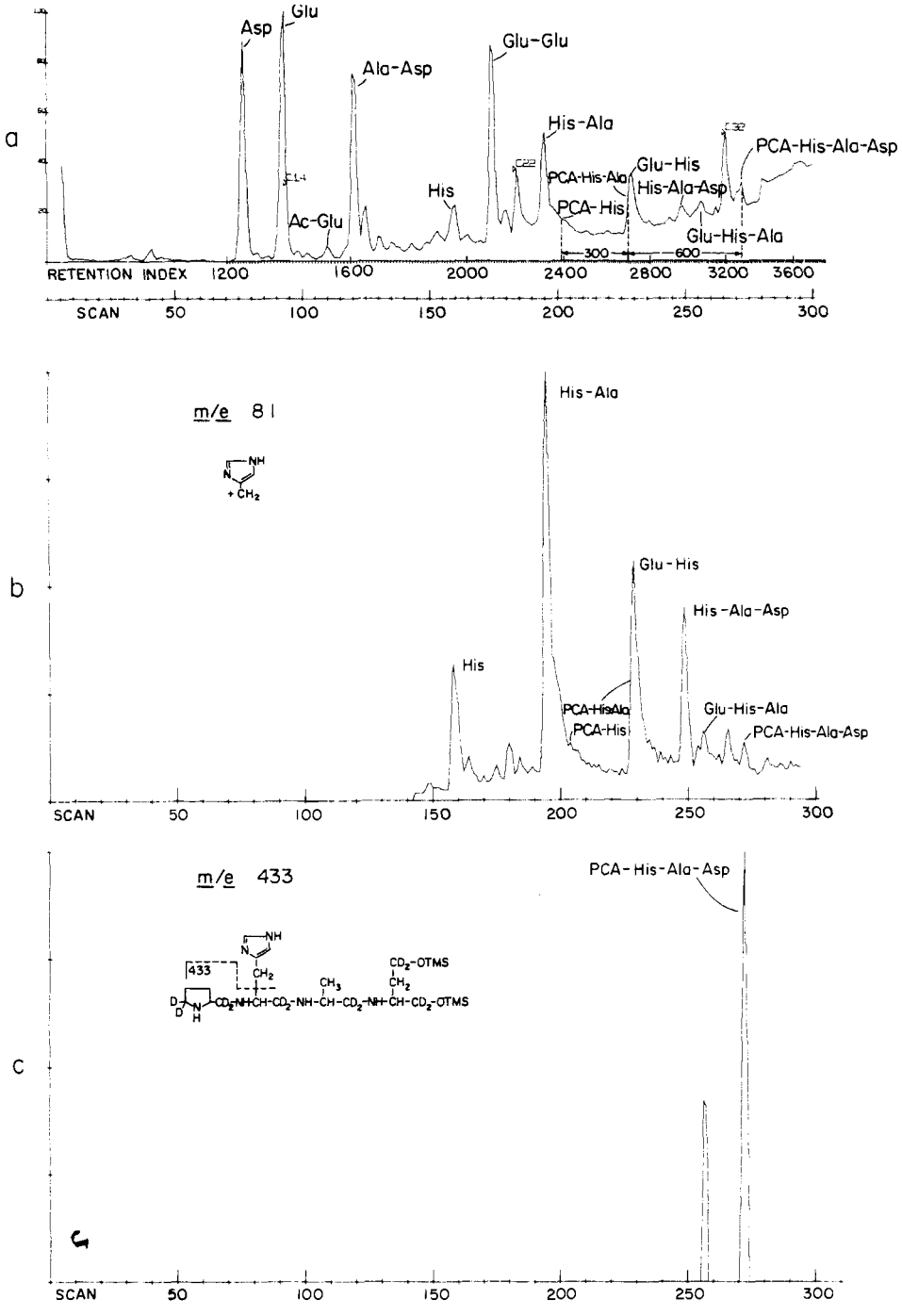

FIGURE 3: (a) Total ionization–retention index plot of a mixture of trifluorodideuterioethyl *O*-trimethylsilyl oligopeptide derivatives obtained by acid hydrolysis and subsequent derivatization. Gas chromatograph conditions: 3-ft glass column (2 mm i.d., 0.25 in. o.d.) filled with 10% OV-17 on Gas Chrom Q (Applied Science Lab.); flow rate, 40 ml/min; temperature, linearly programmed from 80 to 330° at 12°/min. The coinjected *n*-alkanes tetradecane, docosane, and dotriacontane were labeled by the computer as C (14), C (22), C (32). The identified derivatives (some structures are shown in Figure 3 and 4) are indicated by abbreviations of the corresponding underivatized compounds. (b) Mass chromatogram of the ion *m/e* 81. (c) Mass chromatogram of the ion *m/e* 433.

I), sequence analysis was performed on the mixture, AP-1. This mixture of peptides, which is too complex to be amenable to analysis by gas chromatography–mass spectrometry, using any derivatization technique, was subjected to a short acid hydrolysis to cleave the peptides (Kelley et al., 1975) without completely removing the blocking group from the amino terminus. This mixture was then derivatized to form the *N*-trifluorodideuterioethyl *O*-trimethylsilyl ethers, which have excellent gas chromatographic and mass spectrometric properties (Nau, 1974). Analysis of the mixture (Figure 3a) shows that extensive hydrolysis had occurred, in particular of bonds involving aspartic and glutamic acids, producing large amounts of these two amino acids. The other major fractions of this mixture (Figure 3a) were identified as derivatives of Ala-Asp, Glu-Glu (Figure 4a), His-Ala (Figure 4c), Glu-His (Figure 4b), His-Ala-Asp, and Glu-His-Ala. The fragments allow reassembly of the sequence I of one peptide.

The second residue of this peptide, AP-1-A, must be glutamine, since it is slightly less anionic than peptide AP-1-B (<Glu-His-Ala-Asp) at pH 6.5 (vide supra). The addition of a glutamic acid residue to AP-1-B would greatly increase

Sequence I

Glx-Glx
Glx-His
Glx-His-Ala
His-Ala
His-Ala-Asp
Ala-Asp

Glx-Glx-His-Ala-Asp

its mobility at this pH due to the additional negative charge, while the introduction of a glutamine would slightly decrease its mobility since the resulting peptide would be larger. The identification of glutamine in the second position of AP-1-A is also consistent with the observed incorporation of an average of 3.9 residues of leucine per molecule when the free carboxylates of the inhibitor are coupled to leucine methyl ester using water-soluble carbodiimide in 6 *M* guanidine hydrochloride.[2]

The first residue is likely to be either pyrrolidonecarboxylic acid or an *N*-acylglutamine. Since histidine is near the amino terminus (see sequence I above) it was expected that the ion *m/e* 81 corresponding to the side chain of histidine would indicate peptide derivatives containing the blocked N-terminus.

Indeed, the mass chromatogram of this ion (Figure 3b) shows, in addition to maxima indicating the derivatives of histidine, His-Ala, Glu-His, His-Ala-Asp, and Glu-His-Ala, a shoulder at scan 202, and a small but characteristic maximum at scan 272. The fraction emerging at scan 202 was identified by mass chromatograms of other sequence-indicating ions as <Glu-His and still another one emerging at scan 227 was <Glu-His-Ala in the front of the peak which has already been assigned to Glu-His. The maximum of the intensity of *m/e* 81 at scan 272 was shown to correspond to the derivative of <Glu-His-Ala-Asp by interpretation of additional mass chromatograms, including that of the M − 81 ion (*m/e* 443, see Figure 3c), an ion always abundant in histidine-containing peptide derivatives (see Figure 4b,c) and resulting from elimination of the side chain with charge retention on the backbone of the molecule (see structure in Figure 3c).

Retention indices were used for the confirmation of the results since it has been found that they can be predicted from values assigned to each amino acid residue (Nau et al., 1974). The retention indices of <Glu-His-Ala and <Glu-His differ by 300 units, those of <Glu-His-Ala-Asp and <Glu-His-Ala by 600 units, values close to those characteristic for an alanine and aspartic acid residue, respectively.

One (AP-1-B) of the two amino terminal peptides of the inhibitor has the sequence II.

Sequence II

<Glu-His-Ala-Asp

It should be noted that only small amounts of Glu containing peptides were present since most of the amino-terminal residue had been hydrolyzed and was present as glutamic acid and as Glu-Glu (Figure 4a). Sequence I as well as the amino acid composition of the peptide mixture suggested the presence of a second peptide containing an additional glutamine residue between the amino-terminal residue and histidine. Since a peptide containing the amino-terminal sequence <Glu-Glu, <Glu-Gln, *N*-Ac-Gln-Gln, etc.

[2] G. M. Hass in preparation.

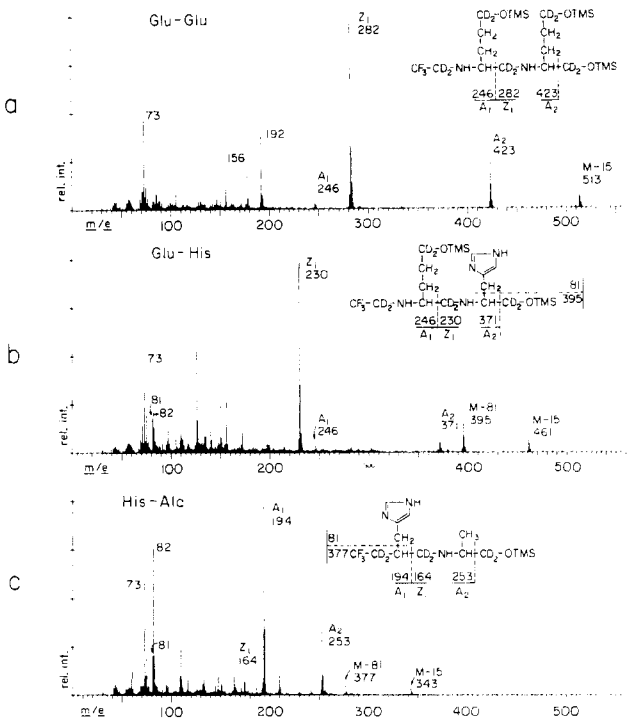

FIGURE 4: Mass spectra corresponding to (a) scan 174 (retention index 2100), (b) scan 229 (retention index 2700), and (c) scan 195 (retention index 2315) of the GC–MS–computer experiment shown in Figure 3. The structures of the identified derivatives are shown in the upper right corners; the original oligopeptides are denoted in the upper left corners.

was not detected, a sample of the entire inhibitor molecule was hydrolyzed under even milder conditions and the resulting mixture transformed to the corresponding heptafluorobutyryl methyl esters. A mass spectrometric analysis of this mixture revealed the presence of <Glu-Glu (retention index 3127) in addition to a number of *N*-heptafluorobutyryl oligopeptide methyl esters. Thus the second amino terminal peptide (AP-1-A) of the inhibitor has sequence III.

Sequence III

<Glu-Gln-His-Ala-Asp

During the search for possible blocking groups (acetyl, formyl, etc.) a compound eluting at scan 109 (retention index 1525) was identified as the derivative of *N*-acetylglutamic acid by ions *m/e* 192, 102, and 282, which are characteristic of compounds of this type (Nau et al., 1975). The origin of this blocking acetyl group, which may have been introduced during the isolation of the inhibitor or of the amino-terminal peptides, is under investigation.

The larger fragment (AP-2) contains all of the sites of tryptic and chymotryptic cleavage. Automatic Edman degradation of fragment AP-2 derived from *S*-Cm- or *S*-Pe-inhibitor provided both the identification of 20 of the 39 amino acid residues of the inhibitor (6–16 and 20–29) and the overlap between tryptic peptides T-1-A and T-1-B (vide infra). Unfortunately, identification of residues 17–19 was not accomplished by this approach primarily due to the overlap accumulated while sequencing through the proline residues in positions 6 and 11.

Tryptic Cleavage at Arginine. The Amino Acid Sequence of Residues 33–39 and Alignment of Tryptic Peptides. Tryptic digestion of *S*-Cm-inhibitor which had been treated with citraconic anhydride to reversibly block lysine residues produces two peptide fragments, T-1 and T-2 (Figure 5a)

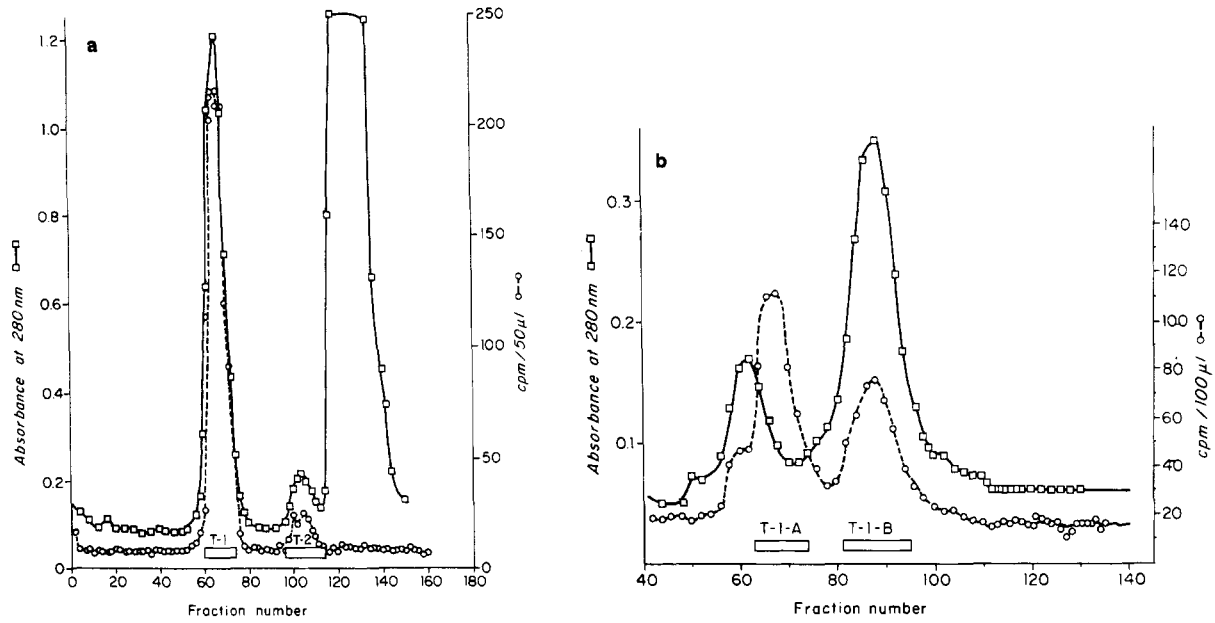

FIGURE 5: (a) Separation of the peptides derived by tryptic digestion of *S*-^{14}Cm-inhibitor which had been citraconylated to prevent cleavage of lysyl bonds. Gel filtration was performed on a 1.5 × 80 cm column of Sephadex G-25 in 9% formic acid. Fractions of 1.0 ml were collected. The column was monitored by absorbance at 280 nm (□—□) and radioactivity was determined in alternate fractions by counting 0.05-ml aliquots (○—○). High absorbance at 280 nm observed in fractions 120–140 is due to excess citraconic acid. (b) Separation of the tryptic peptides generated from T-1 after deblocking the lysine residues. Separation was accomplished on a 1.5 × 80 cm column of Sephadex G-25 in 9% formic acid. Fractions of 1.0 ml were collected. The column was monitored by absorbance at 280 nm (□—□). Radioactivity (○—○) was determined by counting 0.1-ml aliquots of alternate fractions. Absorbance at 280 nm observed in fractions 55–62 is due to uncleaved fragment T-1.

Table II: Amino Acid Compositions of Tryptic Peptides from Carboxypeptidase Inhibitor.

Amino Acid	Inhibitor[a]	T-1	T-I-A	T-1-B	T-2
Lys	2	2.3 (2)[b]	1.6 (2)		
His	2	2.2 (2)	0.9 (1)	1.0 (1)	
Arg	1	1.0 (1)		0.8 (1)	
CM-Cys	6	4.9 (5)	1.8 (2)	2.8 (3)	1.0 (1)
Asx	5	4.9 (5)	1.9 (2)	3.0 (3)	0.3 (0)
Thr	2	1.2 (1)		1.0 (1)	1.0 (1)
Ser	2	1.8 (2)		1.7 (2)	
Glx	2.5	2.7 (2–3)	1.3 (1–2)	1.1 (1)	
Pro	3	2.3 (2)	1.4 (1–2)		1.0 (1)
Gly	3	1.5 (1–2)		1.1 (1)	2.0 (2)
Ala	4	4.0 (4)	1.0 (1)	3.0 (3)	
Val	1				1.0 (1)
Ile	1	1.1 (1)	0.6 (1)		
Tyr	1				0.9 (1)
Phe	1	1.1 (1)		0.9 (1)	
Trp[c]	2	+ (2)		+ (2)	
Yield (%)		88	66	50	92
Residues	1–39	1–32	1–13	14–32	33–39

[a] From Ryan et al. (1974b). [b] Values in parentheses are approximations to nearest integer. Data for amino acids present at less than 0.2 residue/molecule were not included. [c] Positive test with *p*-dimethylaminobenzaldehyde (see text). Values for tryptophan in each peptide are based on the known amino acid sequence of the inhibitor.

Table III: Edman Degradation of Peptide T-2 (Residues 33—39) from *S*-Cm-Inhibitor.

Amino Acid[a]	Composition	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
CM-Cysteine	1.00	1.00	0.27	0.20	0.12	0.16	0.11
Threonine	0.95	0.18	0.09	0.09	0.19	0.12	0.08
Proline	1.15	1.06	1.02	1.00	0.34	0.31	0.21
Glycine	2.00	1.97	2.19	1.53	1.50	1.49	1.00
Valine	1.02	1.00	1.00	1.00	1.00	1.00	0.29
Tyrosine	0.94	0.68	0.74	0.65	0.35	0.10	0.07

[a] All other amino acids are present at less than 0.2 residue/Cm-cysteine. Boldfaced entries indicate the amino acid residue removed at each step.

(Table II). The amino acid sequence of T-2 was determined by subtractive Edman degradation (Table III). The apparent decrease in tyrosine content during step 4 presumably results from the relative instability of this amino acid during Edman degradation. Assignment of tyrosine to position 5 rather than position 4 in this peptide reflects the much greater percentage decrease of proline than tyrosine in step 4. This peptide is identified as the carboxyl-terminal segment of the protein since it lacks lysine and arginine and because T-2, like the intact inhibitor (Ryan et al., 1974b), has glycine in the carboxyl-terminal position.

The larger peptide, T-1, contains both lysine residues and the single arginine (Table II) and occupies positions 1–32. Treatment of T-1 with trypsin after deblocking of the lysine residues produces peptides T-1-A (residues 1–13) and T-1-B (residues 14–32; Figure 5b). Their amino acid compositions are consistent with the proposed sequence of the inhibitor (Table II). The Lys_{10}–Pro_{11} bond is resistant to tryptic digestion as anticipated from previous studies (Hirs et al., 1956).

Peptide T-1-A, derived from the amino terminal region of the inhibitor, is ninhydrin positive because of its lysine content but, like the intact inhibitor, is resistant to Edman degradation. Peptide T-1-A provides the overlap between fragments AP-1 and AP-2 as evidenced by its amino acid composition and the known amino acid sequences of peptides AP-1-A, AP-1-B, and the first nine residues of peptide AP-2.

Table IV: Amino Acid Compositions of Chymotryptic Peptides from Carboxypeptidase Inhibitor.

Amino Acid	Inhibitor[a]	C-1-A	C-1-B	C-2	C-3	C-4
Lys	2	2.1 (2)[b]	2.1 (2)			
His	2	2.2 (2)	2.0 (2)			
Arg	1				1.1 (1)	
CM-Cys	6	2.5 (2–3)	2.4 (2–3)	2.0 (2)	0.9 (1)	
Asx	5	4.0 (4)	3.6 (4)		1.1 (1)	
Thr	2	0.9 (1)	1.0 (1)		1.0 (1)	
Ser	2	1.0 (1)	1.3 (1)		1.0 (1)	
Glx	2.5	1.3 (1–2)	1.4 (1–2)	1.0 (1)		
Pro	3	1.9 (2)	1.4 (1–2)		1.0 (1)	
Gly	3	1.1 (1)	1.3 (1)		1.0 (1)	1.0 (1)
Ala	4	2.0 (2)	2.0 (2)	1.0 (1)	1.0 (1)	
Val	1	0.3 (0)			0.9 (1)	0.9 (1)
Ile	1	0.8 (1)	0.7 (1)			
Tyr	1				0.9 (1)	
Phe	1		0.6 (0–1)	0.8 (1)		
Trp[c]	2	+ (1)	+ (1)	+ (1)		
Yield (%)		33	17	64	61	86
Residues		1–22	1–23	23–28	29–37	38–39

[a] From Ryan et al. (1974b). [b] Values in parentheses are approximations to nearest integer. Data for amino acids present at less than 0.2 residue/molecule were not included. [c] Positive test with *p*-dimethylaminobenzaldehyde (see text). Values for tryptophan in each peptide are based on the known amino acid sequence of the inhibitor.

Table V: Edman Degradation of C-3 (Residues 29–37) from S-Cm-Inhibitor.

Amino Acid[a]	Composition	Step 1	Step 2	Step 3
CM-Cysteine	0.94	0.86	0.83	0.97
Aspartic acid	1.06	**0.15**	0.15	
Threonine	0.96	0.96	0.90	0.92
Serine	0.95	0.66	**0.57**	0.22
Proline	1.00	0.70	0.79	0.56
Glycine	1.02	1.00	1.00	1.00
Alanine	1.00	1.04	0.97	**0.68**
Tyrosine	0.90	0.57	0.68	0.63
Arginine	1.08	0.97	1.20	0.89

[a] All other amino acids were present at levels less than 0.2 residue/glycine. Values are expressed in terms of molar ratios relative to glycine. Boldfaced entries indicate the amino acid removed at each step.

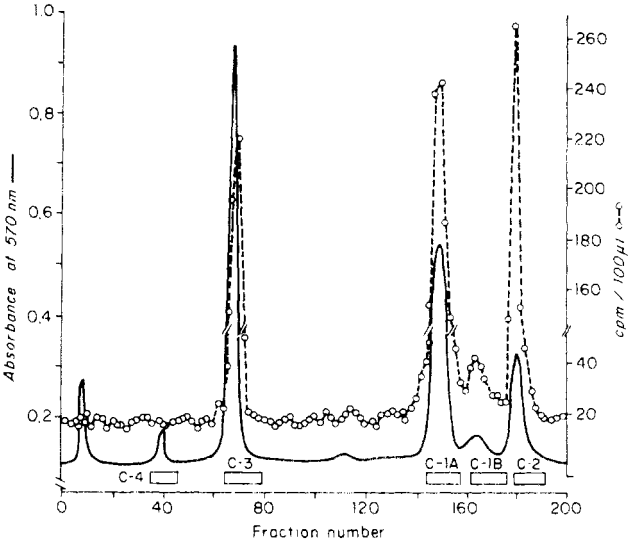

FIGURE 6: Chromatography of the chymotryptic peptides of *S*-^{14}Cm-inhibitor on a 0.6 × 55 cm column of Dowex 1-X2 at 37° using the continuous gradient of pyridine–acetic acid buffers described in the text. The flow rate was 20 ml/hr and fractions of 1.8 ml were collected. Elution of peptides was monitored by automatic ninhydrin analysis after alkaline hydrolysis (—) of 15% of the sample. Radioactivity (O—O) was determined by counting 0.05-ml aliquots of alternate fractions.

Since residues 17–19 could not be unambiguously assigned by automatic Edman degradation of fragment AP-2, presumably because of difficulties encountered at the proline residues in positions 6 and 11 (vide supra), automatic Edman degradation of peptide T-1-B, which adjoins these proline residues, was attempted. Unfortunately, peptide T-1-B was rapidly lost from the sequencer cup, due to its unfavorable solubility properties.

Tryptic Peptides of S-Pe-Inhibitor. Identification of Residues 17–19. *S*-Pyridylethyl derivatives of peptides are generally more amenable to automatic Edman degradation than are the corresponding *S*-carboxymethyl derivatives (Hermodson et al., 1972). The three tryptic peptides derived from *S*-Pe-inhibitor could not be readily separated, and thus the mixture was subjected to automatic Edman degradation. Peptide T-1-A, which is blocked, and T-2, which washed out of the sequencer cup after the first step of degradation, did not interfere with the unambiguous assignment of residues 17–19 in peptide T-1-B (Figure 1).

Chymotryptic Peptides. Assignment of Positions 29–31. The chromatographic elution profile (Dowex 1-X2) of the chymotryptic peptides derived from *S*-Cm-inhibitor is presented in Figure 6. Fractions labeled C-1A, C-1B, C-2, C-3, and C-4 are nearly pure peptides as judged by paper electrophoresis, while the other fractions exhibiting positive ninhydrin reaction are devoid of peptide material as determined by paper electrophoresis and amino acid analysis. The amino acid compositions of the chymotryptic peptides (Table IV) are consistent with the proposed amino acid sequence of the inhibitor (Figure 1).

Chymotryptic peptide C-3 provides both the overlap between tryptic peptides T-1-B and T-2 and the identification of positions 29–31 (Table V), which were required to complete the sequence of the inhibitor. The tentative assignment of asparagine at position 29 by automatic Edman degradation of AP-2 and of T-1-B derived from *S*-Pe-inhibitor

was confirmed (as Asx) by subtractive Edman degradation of peptide C-3 (Table V). The observation that C-3 is electrophoretically neutral at pH 6.5 demonstrates that position 29 is, in fact, asparagine rather than aspartic acid. The low levels of serine and tyrosine observed after step 1 of the Edman degradation of peptide C-3 are at least in part due to the instability observed for these types of residues. To confirm the sequence of residues 29–31, peptide C-3 was subjected to direct Edman degradation (Peterson et al., 1972). Asparagine was observed at step 1 of the degradation and alanine at step 3 with no PTH-amino acid derivatives appearing at step 2. These observations are consistent with the assignment of serine to position 2 of this peptide (vide supra), since serine is the only amino acid not detected by this method (Peterson et al., 1972).

Discussion

Naturally occurring protease inhibitors are of importance not only because they fulfill a critical physiological function in the regulation of proteolytic activity but also because protease inhibitor systems provide excellent models for elucidating mechanisms of enzyme action and for evaluating the forces involved in protein–protein interactions. The interactions of "serine proteases" with naturally occurring inhibitors from various sources have been extensively investigated (for a review see Laskowski and Sealock, 1971); however, few examples of protein inhibitors toward other classes of enzymes have been documented. The amino acid sequence of a carboxypeptidase inhibitor from potatoes was determined to provide a better understanding of this molecule for future investigation of enzyme–inhibitor interactions.

Properties of the inhibitor which particularly impeded determination of its amino acid sequence are a blocked amino terminus and heterogeneity of the preparations. The presence of amino terminal pyrrolidonecarboxylic acid is strongly suggested by mass spectrometric analysis of fragments generated by partial acid hydrolysis of *S*-CM-inhibitor and by the negative test for acetyl and formyl groups. The presence of a blocking group other than pyrrolidonecarboxylic acid cannot be excluded, however.

Heterogeneity in preparations of this inhibitor was suggested by amino acid analysis (Ryan et al., 1974b). Nearly integral values were observed for all amino acids except glutamic acid (2.5 residues/molecule). The isolation of approximately equal amounts of two peptides (AP-1-A and AP-1-B) derived from the amino-terminal region by cleavage of the aspartyl–proline bond and containing two and one residues of glutamic acid, respectively, suggests that the inhibitor is a mixture of two closely related types of polypeptide (Figure 1). Since the two forms of inhibitor differ only by a glutamine residue it is not surprising that separations based upon size or charge indicate that this protein behaves as a homogeneous peptide (Ryan et al., 1974b).

Determination of the amino acid sequence was accomplished by application of several analytical techniques to peptide fragments generated by a variety of procedures. Computerized mass spectrometric analyses of small peptides provided sequence information in the amino terminal region and identification of the blocking group. This approach was complemented by automatic and subtractive Edman degradation which allowed the assignment of the remainder of the amino acid sequence.

The presence of three disulfide bonds in such a small protein suggests an inordinately tight packing of the molecule. The location of these bonds remains to be determined.

With the knowledge of the amino acid sequence of the carboxypeptidase inhibitor the evolution and mechanism of action of this molecule may be investigated. None of the protease inhibitors whose amino acid sequences have been reported appears to be homologous to the carboxypeptidase inhibitor. However, the only other protease inhibitor from potatoes available for comparison at this time is chymotrypsin inhibitor I (Richardson, 1974).

Regarding mechanism of action, the carboxypeptidase inhibitor is particularly amenable to modification by group specific reagents because of its small size and stability to extremes of pH. Modification of the inhibitor by a broad spectrum of reagents and studies of the effects of modification on the strength of binding to the carboxypeptidases may allow the identification of specific regions of the inhibitor which interact with these enzymes (Hass et al., 1974).

Acknowledgments

The authors thank Dr. Harry M. Ako and Dr. Kenneth A. Walsh for helpful discussions during the course of this investigation, and Dr. R. W. Kuhn for his contribution to the initial phase of this work.

References

Andersson, B. A. (1967), *Acta Chem. Scand. 21,* 2906.
Bennett, J. C. (1967), *Methods Enzymol. 11,* 330.
Biller, J. E. (1972), Ph.D. Thesis, Massachusetts Institute of Technology.
Bradshaw, R. A., Garner, W. H., and Gurd, F. R. N. (1969), *J. Biol. Chem. 244,* 2149.
Cavins, J. F., and Friedman, M. (1970), *Anal. Biochem. 35,* 489.
Crestfield, A. M., Moore, S., and Stein, W. (1963), *J. Biol. Chem. 238,* 622.
Edman, P. (1950), *Acta Chem. Scand. 4,* 277, 283.
Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1,* 80.
Fraser, J. F., Poulsen, K., and Haber, E. (1972), *Biochemistry 11,* 4974.
Habeeb, A. F. S. A., and Atassi, M. Z. (1970), *Biochemistry 9,* 4939.
Hass, G. M., Grahn, D. T., Leary, T. R., Neurath, H., and Ryan, C. A. (1974), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 33,* 1529.
Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H., and Walsh, K. A. (1972), *Biochemistry 11,* 4493.
Hill, R. L., and Delaney, R. (1967), *Methods Enzymol. 11,* 339.
Hirs, C. H. W., Moore, S., and Stein, W. H. (1956), *J. Biol. Chem. 219,* 623.
Hites, R. A., and Biemann, K. (1968), *Anal. Chem. 40,* 1217.
Hites, R. A., and Biemann, K. (1970), *Anal. Chem. 42,* 855.
Kelley, J. A., Nau, H., Forster, H.-J., and Biemann, K. (1975), submitted to *Biochemistry.*
Konigsberg, W., and Hill, R. J. (1962), *J. Biol. Chem. 237,* 2547.
Laskowski, M. Jr., and Sealock, R. W. (1971), *Enzymes, 3rd Ed. 3,* 375.
Moore, S., and Stein, W. H. (1963), *Methods Enzymol. 6,* 819.
Nau, H. (1974), *Biochem. Biophys. Res. Commun. 59,* 1088.
Nau, H., and Biemann, K. (1974), *Anal. Chem. 46,* 426.

Nau, H., Kelley, J. A., Forster, H.-J., and Biemann, K. (1975), submitted to *Biochemistry*.
Peterson, J. D., Nehrlich, S., Oyer, P. E., and Steiner, D. F. (1972), *J. Biol. Chem. 247,* 4866.
Richardson, M. (1974), *Biochem. J. 137,* 101.
Ryan, C. A., Hass, G. M., and Kuhn, R. W. (1974a), 2nd International Research Conference on Proteinase Inhibitors.
Ryan, C. A., Hass, G. M., Kuhn, R. W., and Neurath, H. (1974b), *J. Biol. Chem. 249,* 5495.
Schmer, G., and Kreil, G. (1969), *Anal. Biochem. 29,* 186.
Smith, I. (1953), *Nature (London) 171,* 43.
Spackman, D. H., Stein, W. H., and Moore, S. (1958), *Anal. Chem. 30,* 1190.
Weygand, F., Kolb, B., Prox, A., Tilak, M. A., and Tomida, J. (1960), *Hoppe Seyler's Z. Physiol. Chem. 322,* 38.