Shuster, C. W., and Doudoroff, M. (1962), *J. Biol. Chem. 237*, 603.
Stadtman, E. R. (1957), *Methods Enzymol. 4*, 2280.
Stern, J. R. (1957), *Biochim. Biophys. Acta 26*, 448.
Stern, J. R., and Del Campillo, A. (1956), *J. Biol. Chem. 218*, 985.
Stern, J. R., Del Campillo, A., and Lehninger, A. L. (1955), *J. Am. Chem. Soc. 77*, 1073.
Stern, J. R., Del Campillo, A., and Raw, I. (1956), *J. Biol. Chem. 218*, 971.
Wakil, S. J., and Bressler, R. (1962), *J. Biol. Chem. 237*, 687.
Wieland, T., and Rueff, L. (1953), *Angew. Chem. 65*, 186.

The Amino Acid Sequence of Bovine Carboxypeptidase A. I. Preparation and Properties of the Fragments Obtained by Cyanogen Bromide Cleavage*

Masao Nomoto,† N. G. Srinivasan,‡ Ralph A. Bradshaw,§ Roger D. Wade, and Hans Neurath

ABSTRACT: Bovine carboxypeptidase A_γ has been subjected to cleavage by cyanogen bromide and the fragments were produced isolated by gel filtration on Sephadex G-75 equilibrated in 0.1 M propionic acid. In addition to the amino- and carboxyl-terminal peptides previously isolated and characterized, two large fragments, designated F_I and F_{III}, have been obtained. The first of these, F_I, occurs mainly as a high molecular weight aggregate in 0.1 M propionic acid while F_{III} shows little or no aggregation properties but exhibits nonideal behavior. Sedimentation equilibrium analyses in 6 M guanidine hydrochloride show F_I to possess a molecular weight about 23,000 and F_{III} a molecular weight of about 9000. Amino acid analyses of each fragment indicate that F_I and F_{III} are composed of 198 and 81 amino acid residues, respectively. Each fragment contains a single aspartyl residue as amino terminus. The relationship of the four cyanogen bromide fragments of bovine carboxypeptidase A_γ to the primary structure of the enzyme is discussed.

Bovine pancreatic carboxypeptidase A_γ is composed of a single polypeptide chain containing 300 amino acids. Extensive chemical and physical studies recently reviewed by Neurath *et al.* (1968, 1969), and Vallee and Riordan (1968) have disclosed many of the structural and mechanistic features of this enzyme. In addition, X-ray analysis of the crystal structure has been undertaken (Lipscomb *et al.*, 1968). In view of this concerted approach to the structure and function of this enzyme, it was deemed of importance to determine the complete amino acid sequence of carboxypeptidase A by chemical methods.

Earlier studies from this laboratory (Bargetzi *et al.*, 1964) suggested that a promising route to the solution of this structure was cleavage of the polypeptide chain at each of three methionyl residues by the CNBr method of Gross and Witkop 1961). This method has already been successfully applied to a number of sequence problems (Hofmann, 1964; Edmundson, 1963; Delaney and Hill, 1968), and indeed both the amino- (Sampath Kumar *et al.*, 1964) and carboxyl-terminal peptides (F_N and F_C)[1] of carboxypeptidase A composed of 15 and 6 amino acids, respectively, have been isolated (Bargetzi *et al.*, 1964). Complete amino acid sequence analysis of these fragments has supplied the first primary structural data for the protein.

In continuation of these studies, the preparation and characterization of the other two major fragments has been carried out. One of these fragments contains 81 residues and the other one 198. Amino acid analyses of these two fragments yield compositions which, together with the amino acid content of the N- and C-terminal peptides, account satisfactorily for the composition of carboxypeptidase A_γ determined from acid hydrolysates.

Experimental Procedure

Materials. Carboxypeptidase A_γ was obtained from Worthington Biochemical Corp. as an aqueous suspension containing 40 mg/ml of twice-crystallized protein. The specific activity of these preparations toward the ester substrate hippuryl-DL-β-phenyllactate was of the magnitude described by Bargetzi *et al.* (1963).

* From the Department of Biochemistry, University of Washington, Seattle, Washington 98105. *Received April 8, 1969.* This work was supported by the National Institutes of Health (GM 04617 and GM 15731-01), the American Cancer Society (P-79K), the Office of Naval Research (NONR 477-35), and the National Science Foundation (GB 4990).

† Present address: Seikagaku Kogyo Co., Ltd., 9, 2-Chome, Nihombashi-Honcho, Chuo-Ku, Tokyo, Japan.
‡ Present address: Cancer Research Institute, Madras, India.
§ Portions of this work were completed during the tenure of U. S. Public Health Service Postdoctoral Fellowship 1-F2-GM-23, 968-02. Present address: Department of Biological Chemistry, Washington University School of Medicine, St. Louis, Mo. 63110.

[1] F_C, F_N, F_{III}, and F_I, CNBr fragments of carboxypeptidase A.

Cyanogen bromide was obtained from Eastman Organic Chemicals and used without further treatment. The reagent was stored *in vacuo* in sealed desiccators at 4° when not in use.

Trifluoroacetic acid, obtained from Matheson Coleman and Bell Co., was distilled and the fraction boiling between 71 and 73° used.

Sephadex was purchased from Pharmacia Fine Chemicals.

Methods. Acid hydrolysis was carried out by the method of Moore and Stein (1963) and amino acid analyses were performed in Spinco 120 amino acid analyzers (Spackman *et al.*, 1958).

Tryptophan was determined spectrophotometrically by the method of Bencze and Schmid (1957).

Amino-terminal groups were determined by the fluorodinitrobenzene method of Sanger (1945) as described by Levy (1955) with slight modifications, and by the cyanate method of Stark and Smyth (1963). In the former system a solution of 4 M urea containing 1 M NaCl and 0.35 M Tris-HCl buffer (pH 8.0) was used to dissolve the fragment. DNP-amino acids were identified by two-dimensional chromatography utilizing *t*-amyl alcohol saturated with 3% ammonia and 1.5 M phosphate buffer (pH 5.0). Detection of the water-soluble DNP-amino acids was carried out by high-voltage paper electrophoresis at pH 2.1 at 2000 V for 90 min using formic acid–pyridine–water (40:3:960, v/v) buffer.

Sedimentation analyses were carried out with a Spinco Model E ultracentrifuge at room temperature. Molecular weights were determined by the meniscus depletion method of Yphantis (1964). Computations were done on an IBM 7094 computer using a program written by Teller *et al.* (1969).

Reaction with CNBr. Two methods for the cleavage of carboxypeptidase A with CNBr have been used. In the first, conditions essentially identical with those described by Bargetzi *et al.* (1964) and Sampath Kumar *et al.* (1964) were employed. In a typical experiment 25 ml of aqueous suspension of carboxypeptidase A containing a total of 1 g of enzyme was centrifuged at 2000 rpm at 0° for 20 min. The packed pellet was washed at least twice with distilled, deionized water. The washed crystals were dissolved in 60 ml of trifluoroacetic acid, 20 ml of water, and followed by 2.2 g of CNBr. The solution was clarified by centrifugation, and the supernatant solution was allowed to stand at room temperature for 24 hr. The reaction was terminated by slowly pouring the digest into a 2-l. beaker containing 400 ml of water. After 1-hr stirring, the mixture was centrifuged at 8000 rpm at 0° for 1 hr. The precipitate was suspended in 500 ml of 5% trifluoroacetic acid, stirred for 1 hr, and centrifuged again at 7000 rpm at 0° for 20 min. The supernatant and washings contained the soluble amino- and carboxyl-terminal peptides. After an additional wash the precipitate containing the two large fragments was resuspended in water and collected by lyophilization.

In an alternate procedure the washed crystals were dissolved in formic acid and sufficient water was added to make the final solvent concentration 70%. CNBr was added in the same proportions as described above and the reaction was allowed to proceed overnight at room temperature in the dark. The reaction was terminated by the addition of 15 volumes of water/volume of reaction mixture and lyophilized directly. Although this procedure does not remove the soluble amino- and carboxyl-terminal fragments, subsequent gel filtration (*vide infra*) does. In addition, the second procedure was found to be preferable since a purple color was often formed during the reaction in trifluoroacetic acid, a phenomenon not noted in the 70% formic acid medium. However, no detectable difference in cleavage pattern or in the content of acid-stable amino acids was noted between the fragments prepared by the two methods. The purple color has been arbitrarily attributed to some tryptophan destruction.

Isolation of Fragments. The lyophilized CNBr reaction mixture (300 mg) was suspended in 10 ml of 100% propionic acid and diluted to 100 ml of water. Although the fragments were insoluble when treated directly with dilute acid, almost clear solutions were obtained by the above procedure. In fact, solubilization and subsequent gel filtration can be accomplished on samples dissolved in 10–33% propionic acid. The sample was centrifuged to remove the slight cloudiness and the clear supernatant was applied to a 2.5 × 117 cm column of Sephadex G-75 equilibrated in 0.1 M propionic acid. Prior to construction of the column the absorbant was swollen in degassed 0.1 M propionic acid for 24 hr. The column was monitored by reading the absorbance at 280 and 220 mμ in cuvets of 1-mm path length in a Zeiss PMQII spectrophotometer. The peptide fragments were recovered by either lyophilization or alkali precipitation and centrifugation.

Alkylation of the Half-cystinyl Residues. Modification of the half-cystinyl residues was carried out in the following manner. Guanidine hydrochloride (5 g; Mann Ultra Pure) and Tris (121 mg; Sigma Chemical Co.) were dissolved in water to make a solution of 10 ml. Dithioerythritol (Calbiochem) was added to make the final concentration 20:1 moles/mole relative to the fragment F_I added (assuming a molecular weight of F_I = 22,000). The F_I (usually about 300 mg) was dissolved in this solution and the pH was adjusted to 9.2–9.5. Nitrogen gas was blown over the reaction mixture and the vessel was sealed. The reaction was allowed to proceed for 4 hr at room temperature. At the end of the reduction the solution was adjusted to pH 8.0 with 1 N HCl and a preadjusted solution of iodoacetate was added to make the final concentration 0.2 M. The alkylation was allowed to proceed for 5 min at which time 200 μl of 2-mercaptoethanol was added and the pH was held constant until no further change (~5 min) was observed. The reaction mixture was dialyzed exhaustively against water and the modified F_I was recovered by lyophilization. The amount of *S*-carboxymethylcysteine observed averaged 1.70–1.90 residues/molecule uncorrected for destruction. No other carboxymethylamino acids were observed.

Essentially the same procedure was followed for aminoethylation except that the ethyleneimine was added in aliquots and the pH was held constant with 6 N HCl.

Results

Isolation of the Fragments. The separation of the two large CNBr fragments of carboxypeptidase A_γ by gel filtration on Sephadex G-75 is shown in Figure 1. Contrary to expectation three distinct components were detected by ultraviolet absorbance measurements. A fourth peak, eluting in the final position, was also observed. However, subsequent analysis revealed that this latter pool contained the soluble F_N and F_C fragments whose presence was found to be contingent on the prior treatment of the sample. The high 220-mμ absorbance in this pool is due to the more concentrated propionic acid used to dissolve the reaction mixture. The first three pools were consequently designated F_I, F_{II}, and F_{III} corresponding to their

order of elution. F_I and F_{III} were routinely pooled according to the solid bars (see Figure 1) and collected after lyophilization or by centrifugation after alkali precipitation. F_{II}, which was only collected for analytical measurements, was pooled from tubes 83 to 95.

Amino acid analysis of each of the three pools revealed that F_I and F_{II} contained half-cystine but that F_{III} did not. In addition the compositions of F_I and F_{II} were strikingly similar while that of F_{III} was unique. Furthermore, each pool contained detectable quantities of homoserine and its lactone, while methionine was totally absent. On the basis of these preliminary data it was concluded that F_{III} is one of the two remaining fragments while F_I and F_{II} are probably related to each other and represent the other expected fragment.

These conclusions were substantiated by several additional experiments. Rechromatography of F_I, collected by alkaline precipitation, showed predominantly the same peak with a trace of F_{II}. However, rechromatography of F_I, collected after lyophilization, showed a marked departure from this behavior. Instead of the single major peak appearing at the breakthrough, two peaks were observed, one appearing in the position of F_I on the original chromatogram and the other moving in the position of F_{II}. Although F_I was still predominant, as much as 45% of the F_I peak could be converted into F_{II} by lyophilization and rechromatography. This result suggests that F_I and F_{II} represent the components of an aggregation–disaggregation phenomenon, in equilibrium with one another.

To further test this hypothesis, sedimentation equilibrium and velocity experiments were carried out on fragments F_I and F_{II}. An aliquot containing about 2 mg/ml of F_I was removed from the peak tubes (see Figure 1) and dialyzed overnight against 500 ml of 0.1 M propionic acid. The sample was divided and subjected to sedimentation equilibrium and sedimentation velocity analysis. Sedimentation velocity analysis gave an uncorrected s_{20} value of 7.45 S while the equilibrium run showed a weight-average molecular weight of over 300,000 (*vide infra*). Since native carboxypeptidase A_γ possesses a molecular weight of 34,440 (Smith and Stockell, 1954), it is evident that F_I, in this solvent, exists as an aggregate of one or more of the CNBr fragments of carboxypeptidase A.

Further substantiation of the polymeric nature of F_I in 0.1 M propionic acid was obtained by sedimentation analysis of this fragment in 6 M guanidine hydrochloride. An aliquot was drawn from the peak tube after separation on Sephadex G-75 (see Figure 1) and dialyzed for 16 hr against three changes of 6 M guanidine hydrochloride in water (pH 5.5). A precipitate was formed in the bag during dialysis but had spontaneously solubilized at the end of 16 hr. Sedimentation velocity analysis of F_I in this solvent gave an uncorrected s_{20} = 0.4 S. This low value clearly indicates disaggregation of this fraction by the action of 6 M guanidine hydrochloride. Sedimentation equilibrium analyses in the same solvent (*vide infra*) supported this conclusion.

Since both residues of half-cystine are found in F_I (and F_{II}), it was of interest to determine whether they might participate in either covalent or noncovalent bonding leading to the stable polymer. Consequently, modification of these residues by two different means was carried out and the effect of these modifications on the mobility of F_I on Sephadex G-75 tested. In the first case, F_I, which was reacted with ethyleneimine to produce a product which contained two residues of

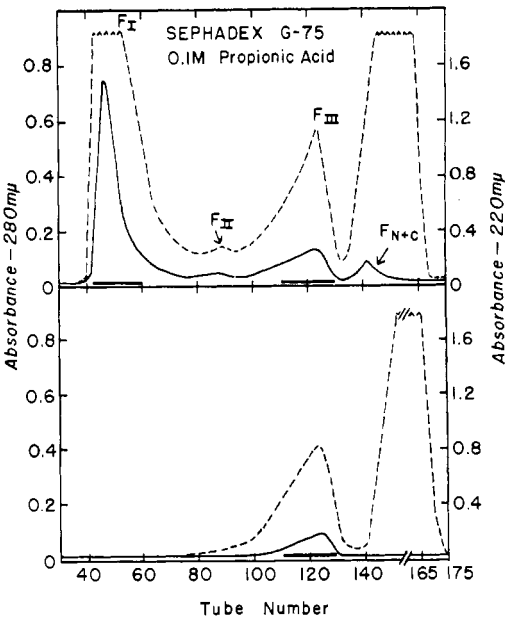

FIGURE 1: Elution profiles of the separation of the CNBr fragments of carboxypeptidase Aγ on a column (2.5 × 117 cm) of Sephadex G-75 equilibrated with 0.1 M propionic acid. The column was developed at 30 ml/hr and monitored at 280 (solid line) and 220 mμ (dashed line). Fraction volume = 4.0 ml. Upper, whole digest; lower, refractionation of F_{III}.

S-aminoethylcysteine, was chromatographed on Sephadex G-75 in 0.1 M propionic acid. The sample was dissolved in 33% propionic acid and run as described (*vide supra*). All of the protein appeared as a single peak in the column void volume in an analogous fashion to the untreated pool. Sedimentation velocity measurements on S-aminoethylated F_I, previously dialyzed against 0.1 M propionic acid, gave a value of s_{20} = 6.97 S in good agreement with the unmodified material. Similarly, alkylation of F_I with iodoacetate showed no differences in chromatographic behavior from the unmodified material. Therefore, it may be concluded that the half-cystinyl residues present in F_I do not contribute to the aggregation phenomenon.

These data, along with the amino acid analyses, suggest that F_I represents an aggregate of one of the large CNBr fragments of carboxypeptidase A_γ. Although rigorous evidence is lacking to establish that F_{II} is truly the monomeric form of this fragment, it is clear that both pools are composed of the same polypeptide chain. The chromatographic behavior on G-75 suggests that it is the monomeric form.

Rechromatography of F_{III} gave a single peak corresponding to the position of the initial pool as shown in the lower portion of Figure 1. The leading edge observed with this pool can be attributed either to interactions with the absorbant or to weak aggregation properties.

Characterization of Fragment F_{III}. The third fragment, F_{III}, eluted from the Sephadex G-75 column, was judged to be a single fragment by electrophoresis, sedimentation analysis, and amino-terminal analysis.

Electrophoretic analysis of this pool, utilizing a Beckman Microzone apparatus, at pH 2.1 and 6.5, showed a single band with ponceau red stain. These data suggest that this pool contains one of the two larger cyanogen bromide fragments expected from cleavage of carboxypeptidase A_γ.

TABLE I: Amino-Terminal Residues of Fragment F_{III}.[a]

Amino Acid	End-Group Residues/9100 g	Control	End-Group Control
Aspartic acid	*1.13*	0.02	*1.11*
Threonine	0.02	0.02	0
Serine	0.05	0.02	0.03
Glutamic acid	0.08	0.03	0.05
Glycine	0.20	0.03	0.17
Alanine	0.05	0.01	0.04
Isoleucine	0.02		0.02
Leucine	0.14		0.14

[a] Results of duplicate analyses.

TABLE II: Amino Acid Composition of F_{III}.

Amino Acid	22-hr Hydrolysis	48-hr Hydrolysis	72-hr Hydrolysis	Av	Integral Value
Lysine	3.79	4.00	4.00	3.93	4
Histidine	1.70	2.08	2.06	1.95	2
Arginine	3.63	3.99	3.97	3.86	4
Aspartic acid	6.29	6.27	6.33	6.30	6
Threonine	5.12	4.97	4.75	5.25[a]	5
Serine	6.54	6.10	5.45	7.00[a]	7
Glutamic acid	7.89	8.19	8.06	8.05	8
Proline	3.75	3.87	3.78	3.80	4
Glycine	6.82	6.81	6.84	6.82	7
Alanine	5.00	5.00	5.00	5.00	5
Valine	3.90	4.09	4.03	4.03[b]	4
Isoleucine	6.18	6.64	6.60	6.60[b]	7
Leucine	6.67	6.81	6.74	6.74	7
Tyrosine	3.15	3.18	3.19	3.17	3
Phenylalanine	4.06	4.06	4 14	4.09	4
Tryptophan					3[c]
Homoserine + lactone	0.83	0.96	0.85	0.88	1
Total					81

[a] Extrapolated to zero time of hydrolysis. [b] 72-hr value. [c] By alkaline spectra.

Sedimentation equilibrium analyses of F_{III} in 6 M guanidine hydrochloride gave values for M_n, M_w, M_z, and M_{z+1} of 8,800, 9,700, 11,800, and 17,500. The distribution of the various averages indicates the presence of only a very small amount of contaminant. In 0.1 M propionic acid F_{III} exhibits nonideal behavior indicating protein solvent interaction. However, in both systems the apparent molecular weight is of the order of 9000, in agreement with the amino acid analysis data (*vide infra*).

Amino-terminal analyses by the fluorodinitrobenzene method of Sanger (1945) and the cyanate method of Stark and Smyth (1963) show primarily a single aspartic acid residue. In the fluorodinitrobenzene analysis, DNP-aspartic acid was the only detectable end group in the ether-soluble fraction. The recovery was 0.46 residue/molecule, determined spectrophotometrically, based on a molecular weight of 9100 for F_{III}. No DNP-arginine could be found in the water-soluble fraction of the acid hydrolysate when examined by high-voltage paper electrophoresis at pH 2.1.

Analysis by the cyanate procedure yielded 1.11 residues of aspartic acid after correction, as shown in Table I. Glycine and leucine were the only additional amino acids found in significant amounts. The high levels of glycine may be explained, in part, by the generation of this residue from side reactions of the cyanate reagent (Stark and Smyth, 1963), which are not corrected for in the control. The increased levels of leucine must arise from contamination either in the protein sample or material generated in the analyses.

The amino acid composition of F_{III}, as deduced from acid hydrolysates, is given in Table II. The stoichiometry was calculated on the basis of 5.00 residues of alanine/molecule, which gave a recovery of 95% of the weight of the sample in milligrams of amino acid.

The data in each column are the average of triplicate analyses with the final integral values given in the last column. Serine and threonine were deduced by linear extrapolation. Tryptophan was estimated to be present as three residues based on the tryptophan to tyrosine ratio of 1.00:0.96 determined spectrophotometrically. The total content of tryptophan plus tyrosine computed by the method of Bencze and Schmid (1957) on the basis of a molecular weight of 9100 was 6.8-g residues as compared with a total of 6 expected from the data given in Table II. The total number of amino acids in F_{III} has been calculated to be 81 and the molecular weight equal to 9142.

Characterization of Fragment F_I. In view of the apparent chemical identity of fragments F_I and F_{II}, characterization and chemical sequence studies[2] have been carried out with the F_I pool because of the much larger quantities present in this fraction. The purity of F_I has been harder to establish due primarily to the insolubility of this fragment. Consequently, satisfactory electrophoretic patterns have not been obtained on any media. However, sedimentation equilibrium analyses in 0.1 M propionic acid and 6 M guanidine hydrochloride, summarized in Table III, indicate that in the latter solvent, the fragment possesses a molecular weight of the order anticipated, *i.e.*, 23,000, for the fourth fragment expected from CNBr cleavage. In contrast, the molecular weights observed in 0.1 M propionic acid show a complex aggregation system. These data are consistent with a mixture composed of a wide distribution of molecular forms, bound by noncovalent forces, and indicate the reason for the behavior of this fragment on Sephadex gel filtration. These values are somewhat higher than the value of 160,000 reported previously (Neurath *et al.*, 1968), perhaps because the samples used in these experiments were equilibrated by gel filtration rather than by dialysis. It may be expected that such a complex aggregation system would be extremely sensitive to equilibration procedures and hence the differences observed are not significant since both

[2] R. A. Bradshaw, L. H. Ericsson, K. A. Walsh, and H. Neurath, in preparation.

TABLE III: Sedimentation Equilibrium Analysis of Fragment F_I in 0.1 M Propionic Acid and 6 M Guanidine Hydrochloride.

Solvent	M_n	M_w	M_z	M_{z+1}
Propionic acid	264,000	322,000	406,000	543,000
Guanidine hydrochloride	23,400	26,500	29,600	32,500

sets of data are only of value in indicating the polymeric nature of F_I in 0.1 M propionic acid.

It is also of interest to note that the molecular distribution in 6 M guanidine hydrochloride also indicates some polymeric character. These data, however, are insufficient to characterize this phenomenon to any greater degree.

The homogeneity of this fragment is substantiated by the amino-terminal analyses by the cyanate method, summarized in Table IV. Clearly, F_I contains one residue of aspartic acid as amino terminus. The only other significant end group observed was glutamic acid, present to the extent of 0.14 residue/molecule. The extraneous glutamic acid most probably arises from incomplete removal of pyrrolidonecarboxylic acid (Stark and Smyth, 1963).

The amino acid composition of fragment F_I presented unique problems because of the presence of allotypic variants in carboxypeptidase A prepared from pooled glands[3] (Bargetzi *et al.*, 1964; Pétra and Neurath, 1969). Analyses of F_I prepared from unfractionated carboxypeptidase A_γ consistently yielded less satisfactory results as compared with the purified species. For this reason, fragment F_I was prepared from carboxypeptidase A^{Val} and A^{Leu},[4] reduced and alkylated, and subjected to quantitative amino acid analysis after acid hydrolysis. The results of these experiments are summarized in Tables V and VI. In each case the average of two analyses for 24, 48, and 72 hr are given. The fourth column is the average of the three sets of primary data and the last column shows the assumed integral values. The values for serine and threonine were obtained by linear extrapolation to zero time of hydrolysis and the values of valine and isoleucine were taken to be equivalent to the 72-hr value, as the longest time of hydrolysis can be expected to give the greatest release of these amino acids. Homoserine lactone was present in small amounts but quantitation of this residue or of homoserine was not possible. Consequently their sum has been assumed to be one residue. Tryptophan was measured spectrophotometrically and consistently gave values close to 4.5 residues. However, independent measurements using colorimetric procedures, to be reported as part of another study,[5] gave values much closer to four residues per molecule, suggesting the lower values reported in Tables V and VI as being correct. The values of glutamic acid in both samples can be expected to be somewhat high due to the presence of homoserine, which chromatographs in the same position as glutamic acid under the conditions employed for these analyses. Accordingly, a lower value than that apparent from the primary data has

[3] P. H. Pétra, R. A. Bradshaw, K. A. Walsh, and H. Neurath, in preparation.

[4] The designations A_γ^{Val} and A_γ^{Leu} have been assigned to the two purified forms of carboxypeptidase A_γ and refer to allotypic replacement present in the antepenultimate position of the polypeptide chain of the whole protein (Pétra and Neurath, 1969). The authors are grateful to Dr. Philip H. Pétra for providing the purified proteins used in these studies.

[5] T. M. Radhakrishnan, R. A. Bradshaw, D. A. Deranleau, K. A. Walsh, and H. Neurath, unpublished experiments.

TABLE IV: Amino-Terminal Residues of Fragment F_I.

Amino Acid	End-Group Residues/ 22,000 g	Control	End-Group Control
Aspartic acid	*0.87*	0.02	*0.85*
Serine	0.03	0.02	0.01
Glutamic acid	0.17	0.03	0.14
Glycine	0.05	0.03	0.02
Alanine	0.03	0.01	0.02
Isoleucine	0.03		0.03
Leucine	0.05		0.05
Tyrosine	0.03		0.03

TABLE V: The Amino Acid Composition of *S*-Carboxymethyl Fragment F_I from Bovine Carboxypeptidase A_γ^{Val}.[a]

Amino Acid	24 hr	48 hr	72 hr	Av	Integral Value
Lysine	10.85	10.75	10.90	10.83	11
Histidine	4.45	4.30	4.25	4.33	4
Arginine	6.26	6.31	6.00	6.19	6
S-CM-cysteine	1.76	1.82	1.89	1.82	2
Aspartic acid	16.46	16.90	17.11	16.82	17
Threonine	15.35	15.03	14.61	15.90[b]	16
Serine	22.21	20.35	19.10	23.70[b]	24
Glutamic acid	15.90	15.94	15.90	15.91	15[c]
Proline	6.28	5.91	6.21	6.13	6
Glycine	16.01	16.16	16.10	16.09	16
Alanine	14.00	14.00	14.00	14.00	14
Valine	10.45	10.74	10.92	10.92[d]	11
Isoleucine	10.76	12.26	12.59	12.59[d]	13
Leucine	14.94	15.26	15.27	15.15	15
Tyrosine	12.51	12.80	12.69	12.66	13
Phenylalanine	9.44	10.03	9.82	9.76	10
Tryptophan					4[e]
Homoserine + lactone					(1)[f]
Total					198

[a] The data for each hydrolysis time is the average of two analyses. Calculations were made on the basis of 14.00 residues of Ala/mole of fragment assuming the molecular weight of F_I to be 23,000. [b] Extrapolated to zero time. [c] Taken to be 15 because of the contamination of homoserine. [d] Equivalent to 72-hr value. [e] Determined spectrophotometrically. [f] Homoserine lactone identified in small amounts but quantitation of homoserine and the lactone not possible.

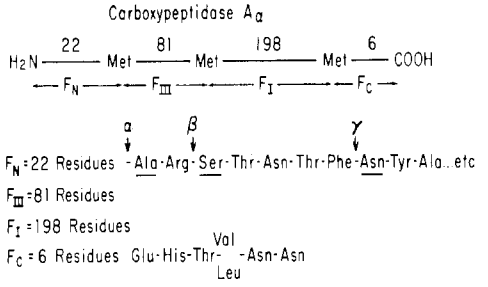

FIGURE 2: Schematic representation of the linear arrangement of the CNBr fragments of carboxypeptidase A_α. The Greek symbols refer to the alternate sites of cleavage to produce the activation forms.

been chosen. Significantly, the total number of amino acids in each sample is 198. The molecular weight of F_I, calculated from these data and taken as an equimolar mixture of the two allotypic forms, is 22,758.

The data presented indicate that fragment F_I contains two more allotypic variants involving glutamic acid, alanine, isoleucine, and valine. These results have been verified and their distribution ascertained by peptide analyses.[8]

Discussion

On the basis of the three residues of methionine present in carboxypeptidase A, a total of four fragments should be formed after treatment of the protein with CNBr. Two fragments were readily isolated from the reaction mixture on the basis of solubility and have been shown to correspond to the amino- and carboxyl-terminal portions of the enzyme. Since these fragments comprise only 21 of the 300 residues present in carboxypeptidase A_γ, the remaining two fragments must comprise the major portion of the molecule.

Accordingly, Sephadex gel filtration on G-75 in 0.1 M propionic acid was applied to the separation of the insoluble portion of the CNBr reaction mixture in an attempt to isolate the remaining two fragments. Two major and two minor fractions were obtained by this route. One of the minor fractions was the last peak to be eluted from the column and its presence was found to be contingent on whether the soluble F_N and F_C fragments were removed prior to the separation. Consequently, the content of this pool has been attributed to these previously characterized pieces. The other minor fraction, designated F_{II} (see Figure 1), has been identified as being the same as F_I. Both amino acid analysis and sedimentation equilibrium analysis indicate that F_I and F_{II} are related to each other in a polymer–monomer relationship. Hence, these two pools account for one of the remaining fragments.

The fourth fragment was isolated in pure form from pool F_{III}. Sedimentation, electrophoretic, amino-terminal, and amino acid analysis all support the homogeneity of this fragment. Thus, the remaining two CNBr fragments of carboxypeptidase have been isolated by gel filtration and shall subsequently be referred to as F_I and F_{III}, their chromatographic designations.

The purity of F_I has been harder to establish than that of the more soluble F_{III}. Its elution on the void volume of the G-75 column does not afford any opportunity for removal of any large contaminants. The unusual early elution position for a fragment of this size can be explained by the sedimentation

TABLE VI: The Amino Acid Composition of *S*-Carboxymethyl Fragment F_I from Bovine Carboxypeptidase A_γ^{Leu}.[a]

Amino Acid	24 hr	48 hr	72 hr	Av	Integral Value
Lysine	10.93	10.90	10.71	10.84	11
Histidine	4.22	4.06	4.17	4.15	4
Arginine	6.12	6.12	6.22	6.15	6
S-CM-cysteine	1.72	1.73	1.78	1.74	2
Aspartic acid	16.60	16.45	16.56	16.53	17
Threonine	15.41	14.70	14.50	15.80[f]	16
Serine	21.97	20.11	18.97	23.60[f]	24
Glutamic acid	16.65	16.46	16.56	16.55	16[b]
Proline	6.40	6.44	6.27	6.37	6
Glycine	16.01	15.68	15.49	15.73	16
Alanine	13.03	12.77	12.67	12.82	13
Valine	10.55	11.13	11.66	11.66[c]	12
Isoleucine	9.81	11.36	11.75	11.75[c]	12
Leucine	15.00	15.00	15.00	15.00	15
Tyrosine	12.86	12.81	12.89	12.85	13
Phenylalanine	9.94	10.28	9.94	10.05	10
Tryptophan					4[d]
Homoserine + lactone					(1)[e]
Total					198

[a] The data for each hydrolysis time is the average of two analyses. Calculations were made on the basis of 15.00 residues of leucine/mole of fragment assuming the molecular weight of F_I to be 23,000. [b] Taken to be 16 because of the contamination of homoserine. [c] Equivalent to 72-hr value. [d] Determined spectrophotometrically. [e] Homoserine lactone identified in small amounts but quantitation of homoserine and the lactone not possible. [f] Extrapolated to zero time.

data. In 0.1 M propionic acid, the molecular weight of F_I is over 300,000 which probably represents an average of many different polymeric species having a monomer molecular weight of about 23,000.

The approximate molecular weight values for the two fragments obtained by sedimentation equilibrium in 6 M guanidine hydrochloride are in reasonable agreement with the values anticipated from the measured molecular weight of carboxypeptidase A. On the basis of composition, molecular weights of 2,658, 9,148, 22,758, and 704 for F_N, F_{III}, F_I, and F_C have been calculated yielding a total molecular weight of 35,268 for carboxypeptidase A_α. The shorter carboxypeptidase A_γ, whose F_N has a molecular weight of 1879 rather than 2658, has a calculated weight of 34,489, which is in excellent agreement with the molecular weight of carboxypeptidase A_γ of 34,440 (Smith and Stockell, 1954).

The amino acid composition data supply even more striking evidence that the four fragments account satisfactorily for the entire carboxypeptidase A_γ molecule. A comparison of the amino acid composition of the four fragments with the amino acid composition of carboxypeptidase A_α as determined from acid hydrolysates is given in Table VII. The comparison has been made to carboxypeptidase A_α as the acid hydrolysis data

TABLE VII: Comparison of the Sum of the Amino Acid Compositions of the CNBr Fragments of Bovine Carboxypeptidase A_α with the Composition of the Native Enzyme.

Amino Acid	F_N[a]	F_C[a,b]	F_{III}	F_I[b]	Total Fragments	Carboxypeptidase A_α[c]
Lysine	0	0	4	11	15	15
Histidine	1	1	2	4	8	8
Arginine	1	0	4	6	11	11
Aspartic acid	4	2	6	17	29	28
Threonine	4	1	5	16	26	28
Serine	1	0	7	24	32	33
Glutamic acid	1	1	8	15.5	25.5	25
Proline	0	0	4	6	10	10
Glycine	0	0	7	16	23	22.5
Alanine	2	0	5	13.5	20.5	20
Half-cystine	0	0	0	2	2	2
Valine	0	0.5	4	11.5	16	16
Methionine	1	0	1	1	3	3
Isoleucine	1	0	7	12.5	20.5	20
Leucine	1	0.5	7	15	23.5	23
Tyrosine	3	0	3	13	19	19
Phenylalanine	2	0	4	10	16	16
Tryptophan	0	0	3	4	7	8
Total	22	6	81	198	307	307–308

[a] Taken from Sampath Kumar *et al.* (1964). [b] Calculated as an equimolar mixture of the Leu and Val allotypes (Pétra and Neurath, 1969). [c] Taken from Bargetzi *et al.* (1963).

are more detailed for this form of the enzyme (Bargetzi *et al.*, 1963). The differences between carboxypeptidase A_α and A_γ have been rigorously documented (Sampath Kumar *et al.*, 1964). Clearly, the agreement between the sum of the compositions of the four fragments and the composition of the whole protein is excellent. With the exception of the four residues present in fractional amounts, due to the allotypic forms (Pétra and Neurath, 1969), only aspartic acid, threonine, serine, and tryptophan are not in exact agreement. In the case of the serine and threonine, the source of error is readily located since nonlinear extrapolations were applied to the correction of these residues in the hydrolysates of the whole protein. In fact, if linear extrapolations are applied to these data (Bargetzi *et al.*, 1963), values of 26 and 32 are obtained for threonine and serine, respectively, in exact agreement with the values obtained from the sum of the fragments. The disagreement in the aspartic acid content is not readily explained, although the error may be reflected in the low values obtained for this residue, particularly in fragment F_I and carboxypeptidase A_γ^{Leu}.

The discrepancy in the tryptophan content is also without ready explanation. The value of 4 for each of the F_I samples was calculated from ratios yielding values of 4.4 to 4.6. The lower value has been favored because of independent measurements utilizing 2,4-dinitrobenzyl bromide (Barman and Koshland, 1967), which yield values much closer to 4.[5] However, the exact tryptophan content will probably only be ascertained after complete sequence analysis of the protein.

TABLE VIII: Peptic Peptides from Carboxypeptidase A_γ Supplying Overlap Data for the Alignment of the CNBr Fragments F_N, F_{III}, F_I, and F_C.

(1) $\underrightarrow{\text{Thr}}\ \underrightarrow{\text{Ile}}\ \underrightarrow{\text{Met}}\ \underrightarrow{\text{Glu}}\ \underrightarrow{\text{His}}\left(\text{Thr}, \begin{matrix}\text{Val}\\ \text{Leu}\end{matrix}, \text{Asn}, \text{Asn}\right)$

(2) $\underrightarrow{\text{Asp}}\ \underrightarrow{\text{Ser}}\ \underrightarrow{\text{Met}}\ \underrightarrow{\text{Asp}}$

The amino end-group data indicate that each fragment possesses an amino-terminal aspartic acid residue. Although neither the fluorodinitrobenzene nor the cyanate procedure can distinguish the chemical nature of the side chain, *i.e.*, aspartic acid or asparagine, independent sequence analysis was relied upon to establish both these residues to be aspartic acid.[2,6] Thus the amino termini of the four CNBr fragments from carboxypeptidase A_γ have been established as Asn, Asp, Asp, and Glu (Bargetzi *et al.*, 1964). A consequence of these assignments is the observation that mutually contaminating quantities of F_N, F_{III}, and F_I cannot be detected by amino end-group analysis.

The linear arrangement of the four fragments in the molecule can be made on the basis of two overlapping peptides isolated from peptic digestion of whole carboxypeptidase A_γ. These peptides are summarized in Table VIII. Peptide 1 clearly supplies the overlap with F_C since this fragment possesses the only amino-terminal glutamic acid residues. Peptide 2, which contains a Met-Asp sequence, must overlap either fragment F_I or F_{III}. Since the carboxyl-terminal sequence of F_N is known to be Asp-Phe-Met (Sampath Kumar *et al.*, 1964), this peptide must contain the carboxyl-terminal sequence of either F_I or F_{III}. Further information concerning the carboxyl-terminal sequence of both F_I and F_{III} is available, however. In the case of F_{III}, the carboxyl-terminal sequence has been shown to be Asp-Ser-Met (Neurath *et al.*, 1968, 1969)[2] and in F_I, Thr-Ile-Met.[6] Thus, on this basis, the peptides given in Table VIII supply sufficient data to align the CNBr fragments in the order F_N-F_{III}-F_I-F_C. These data are summarized in Figure 2, which shows the alignment of the four fragments of carboxypeptidase A_α and its relationship to carboxypeptidase A_β and carboxypeptidase A_γ.

The isolation and characterization of the two remaining CNBr fragments of carboxypeptidase A supplies a feasible route for the determination of the complete amino acid sequence of this protein. These studies are now in progress.

Acknowledgments

We are grateful to Mr. Richard R. Granberg and Mr. Thomas Oas for their assistance with the amino acid analyses. The expert technical assistance of Mr. C. V. Shih throughout this work is duly appreciated. We thank Drs. Kenneth A. Walsh and Philip H. Pétra for many helpful discussions.

[6] R. A. Bradshaw, D. R. Babin, M. Nomoto, N. G. Srinivasan, L. H. Ericsson, H. Neurath, and K. A. Walsh, in preparation.

References

Bargetzi, J. P., Sampath Kumar, K. S. V., Cox, D. J., Walsh, K. A., and Neurath, H. (1963), *Biochemistry 2*, 1468.
Bargetzi, J. P., Thompson, E. O. P., Sampath Kumar, K. S. V., Walsh, K. A., and Neurath, H. (1964), *J. Biol. Chem. 239*, 3767.
Barman, T. E., and Koshland, D. E., Jr. (1967), *J. Biol. Chem. 242*, 5771.
Bencze, W. L., and Schmid, K. (1957), *Anal. Chem. 29*, 1193.
Delaney, R., and Hill, R. L. (1968), *J. Biol. Chem. 243*, 4206.
Edmundson, A. B. (1963), *Nature 198*, 354.
Gross, E., and Witkop, B. (1961), *J. Am. Chem. Soc. 83*, 1510.
Hofmann, T. (1964), *Biochemistry 3*, 356.
Levy, A. L. (1955), *Methods Biochem. Anal. 2*, 360.
Lipscomb, W. N., Hartsuck, J. A., Reeke, G. N., Jr., Quiocho, F. A., Bethge, P. H., Ludwid, M. L., Steitz, T. A., Muirhead, H., and Coppola, J. C. (1968), *Brookhaven Symp. Biol. 21*, 24.
Moore, S., and Stein, W. H. (1963), *Methods Enzymol. 6*, 819.
Neurath, H., Bradshaw, R. A. Ericsson, L. H., Babin, D. R., Pétra, P. H., and Walsh, K. A. (1968), *Brookhaven Symp. Biol. 21*, 1.
Neurath, H., Bradshaw, R. A., Pétra, P. H., and Walsh, K. A. (1969), *Proc. Roy. Soc. (London) B* (in press).
Pétra, P. H., and Neurath, H. (1969), *Biochemistry* (in press).
Sampath Kumar, K. S. V., Clegg, J. B., and Walsh, K. A. (1964), *Biochemistry 3*, 1728.
Sanger, F. (1945), *Biochem. J. 39*, 507.
Smith, E. L., and Stockell, A. (1954), *J. Biol. Chem. 207*, 501.
Spackman, D. H., Stein, W. H., and Moore, S. (1958), *Anal. Chem. 30*, 1190.
Stark, G. R., and Smyth, D. C. (1963), *J. Biol. Chem. 238*, 214.
Teller, D. C., Horbett, T. A., Richards, E. G., and Schachman, H. K. (1969), *Ann. N. Y. Acad. Sci.* (in press).
Vallee, B. L., and Riordan, J. F. (1968), *Brookhaven Symp. Biol. 21*, 91.
Yphantis, D. A. (1964), *Biochemistry 3*, 297.

Identification of the Amino Acid Replacements Characterizing the Allotypic Forms of Bovine Carboxypeptidase A*

Philip H. Pétra, Ralph A. Bradshaw,† Kenneth A. Walsh, and Hans Neurath

ABSTRACT: The two allotypic forms of chromatographically purified bovine carboxypeptidase A have been examined in detail in order to ascertain the nature of the amino acid replacements characterizing each species. In addition to the valine-leucine replacement already identified in the carboxyl-terminal fragment, two other replacements, located in the F_I fragment, have been identified. The location and distribution of these interchanges have been confirmed by means of tryptic and thermolytic digestion followed by isolation of the appropriate peptides. One form of the enzyme contains isoleucine, alanine, and valine at positions 179, 228, and 305, while the other possesses valine, glutamic acid, and leucine at the same loci.

These replacements supply a rational explanation for the differences in chromatographic behavior and heat stability observed for the two allotypic variants. The maintenance of these two sets of linked triple mutations is considered indicative of the evolutionary development of this enzyme.

Studies on the relationship of structure to function in an enzyme require a detailed knowledge of primary and three-dimensional structure of the protein molecule. These structural features, which are generally regarded as unique to each individual enzyme, are also a reflection of their genetic origin and, hence, of the evolutionary pathway from which this structure evolved. However, improved technology in the field of protein chemistry has supplied increasing evidence that the "uniqueness" of structure associated with individual enzymes may be an over-simplification of the true genetic and chemical character of these molecules and that several if not all protein molecules maintain various vestiges of the mutational events that are the basis of their evolution.

Bovine carboxypeptidase A (CPA)[1] appears to be a clear

* From the Department of Biochemistry, University of Washington, Seattle, Washington 98105. *Received April 15, 1969.* This work was supported by the National Institutes of Health (GM 04617 and GM 15731-01), the American Cancer Society (P-79K), the Office of Naval Research (NONR 477-35), and the National Science Foundation (GB 4990).

† Portions of this work were completed during the tenure of U. S. Public Health Service Postdoctoral Fellowship 1-F2-GM-23, 968-02. Present address: Department of Biological Chemistry, Washington University School of Medicine, St. Louis, Mo. 63110.

[1] Abbreviations used: CPA carboxypeptidase A, subscripts α, β, and γ refer to the various activation forms (Sampath Kumar *et al.*, 1964a), while superscripts Val and Leu refer to the allotypic forms (Bargetzi *et al.*, 1964; Pétra and Neurath, 1969); F_I, F_{III}, F_N, and F_C, cyanogen bromide fragments of carboxypeptidase A;[2] Tp, tryptic digest or peptide; and Th, thermolytic digest or peptide.

[2] R. A. Bradshaw, D. R. Babin, M. Nomoto, N. G. Srinivasan, L. H. Ericsson, H. Neurath, and K. A. Walsh, in preparation.