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.

example of this phenomenon. Although this protein was considered for many years to be a chemically pure species (Neurath, 1960), it is now established (Pétra and Neurath, 1969)[3] that a heterogeneous product is obtained from enzymes prepared either by the methods of Anson (1937) or Cox *et al.* (1964). One source of heterogeneity residues in the site of proteolysis during activation of the zymogen, procarboxypeptidase A. Three defined forms of the enzyme, A_α, A_β, and $A\gamma$, each differing in the length of the amino-terminal portion of the polypeptide chain, have been isolated in a pure form (Sampath Kumar *et al.*, 1964a,b; Pétra and Neurath, 1969).[3]

In addition to this heterogeneity, due to activation, a variation was also discovered in the carboxyl-terminal portion of the molecule (Bargetzi *et al.*, 1964), consisting of a valine-leucine replacement in the antepenultimate position of the peptide chain. Further experiments with single animals (Walsh *et al.*, 1966) demonstrated that this trait followed Mendelian genetics indicating that two alleles for carboxypeptidase A exist in the bovine population. These forms have been subsequently designated as the Val and Leu types.

Chromatographic separation of the six forms of carboxypeptidase A,[4] *i.e.*, CPA_α^{Val}, CPA_α^{Leu}, CPA_β^{Val}, CPA_β^{Leu}, CPA_γ^{Val}, and CPA_γ^{Leu}, into homogeneous fractions has made it possible to examine these proteins for further replacement sites. These experiments have been carried out with the Val and Leu forms of carboxypeptidase A and have led to the identification of two additional replacement sites.

No analogous characterization of the α and β forms has been made although several lines of evidence suggest that they are identical with regard to the allotypic variation observed for the γ form. Each α and β pair is characterized by a Val and Leu form and recent evidence has been obtained that the α form can be converted into the β and γ forms by bacterial proteolysis. These products have identical chromatographic mobilities as the corresponding β and γ forms prepared by the method of Anson (1937), suggesting that they contain the same allotypic replacements.[5]

Experimental Section

Materials. Carboxypeptidase A was purchased in 5–10-g lots from Worthington Biochemicals as a twice-recrystallized product. Preparation of the A^{Val} and A^{Leu} forms was carried out by the method of Pétra and Neurath (1969).

CNBr, trifluoroacetic acid, and phenyl isothiocyanate were purchased from Eastman Organics Co. Thermolysin was obtained as a crystalline preparation from Daiwa Kasei K. K., Osaka, Japan. Pyridine and *N*-ethylmorpholine were redistilled from solid ninhydrin before use. All other materials were reagent grade.

Methods. CNBr cleavage of carboxypeptidase A was performed in 70% formic acid with a 2:1 ratio of CNBr to enzyme (w/w). The fragments produced were recovered by lyophilization and purified on Sephadex G-75 equilibrated in 0.1 M propionic acid (Nomoto *et al.*, 1969). Tryptic and thermolytic digestions of fragment F_I were performed in a pH-Stat at pH 8.8 and 8.0, respectively. Solutions of F_I were prepared by suspending the fragment in water (10 mg/ml) and titrating with 1 N NaOH until the F_I fragment was completely dissolved. The pH of the solution was then adjusted to the appropriate value for digestion. Temperature was maintained at 37° by a circulating water bath. Trypsin was added from a stock solution (10 mg/ml in 0.001 N HCl) to a final concentration of 1% (w/w) relative to F_I. Thermolysin was added from a stock solution (0.5 mg/ml in 0.001 M calcium acetate) to a final concentration of 0.5–1% (w/w). An additional aliquot of enzyme was added at the end of 2 hr and the digestion was allowed to proceed for 3–4 hr. The reaction was terminated in each case by the addition of 6 N HCl to pH 2.0. The insoluble material formed was removed by centrifugation and washed twice with 0.05 N pyridine acetate (pH 2.5).

The soluble tryptic peptides were separated on a 0.9 × 25 cm column and the thermolytic peptides on a 2.0 × 25 cm column of Dowex 50-X8 (Spinco, resin AA-15) at 55° with a double-linear gradient of pyridine acetate as described previously (Bradshaw *et al.*, 1969). Subfractionations were performed on 0.9 × 150 cm columns of Dowex 1-X2. All columns were monitored by ninhydrin analysis of aliquots previously hydrolyzed with 2.5 N NaOH.

Recovery of peptides, electrophoresis, amino acid analysis, and subtractive Edman degradations were performed as described previously (Bradshaw *et al.*, 1969; Shearer *et al.*, 1967).

Results

The treatment of carboxypeptidase A with cyanogen bromide results in the formation of four fragments (Sampath Kumar *et al.*, 1964b; Bargetzi *et al.*, 1964) which have been denoted F_N, F_C, F_{III}, and F_I. The alignment of these fragments in the order they occur in the whole enzyme, *i.e.*, F_N-F_{III}-F_I-F_C, has been deduced (Nomoto *et al.*, 1969). The examination of each of these fragments for the location of allotypic replacements has been carried out in order to describe fully all of the gene products of the carboxypeptidase A loci of the appropriate chromosome. The treatment of each of these fragments is described below.

Analysis of F_N, F_C, and F_{III}. In each of the experiments described, carboxypeptidase A_γ^{Val} and A_γ^{Leu} were separated by chromatography on DEAE-cellulose (Pétra and Neurath, 1969). Each form was subsequently treated with CNBr to produce the desired fragments. F_N and F_C were separated from F_I and F_{III} by acid precipitation and further purified by Sephadex gel filtration and paper electrophoresis. Amino acid analyses of these peptides revealed that F_N from each form of the enzyme possessed integral numbers of each amino acid previously established by sequence analysis (Sampath Kumar *et al.*, 1964a) and was apparently devoid of any allotypic variants. Furthermore, F_C, derived from the carboxyl-terminal portion of each allotypic form of the enzyme, contained either a valine or leucine in the antepenultimate position, as previously established (Bargetzi *et al.*, 1964), with no additional substitutions (Pétra and Neurath, 1969).

In the studies reported in this communication detailed examination of F_{III} and F_I has been performed. Previous amino acid analyses of F_{III} prepared from carboxypeptidase A(Anson) had shown consistent integral numbers for all amino acids,

[3] P. H. Pétra and H. Neurath, in preparation.

[4] Except CPA_α^{Leu} and CPA_β^{Val} which cannot be separated by these methods. These forms have been obtained in pure form from single animals homozygous for the appropriate trait.

[5] P. H. Pétra, R. W. Tye, R. Sande, and H. Neurath, in preparation.

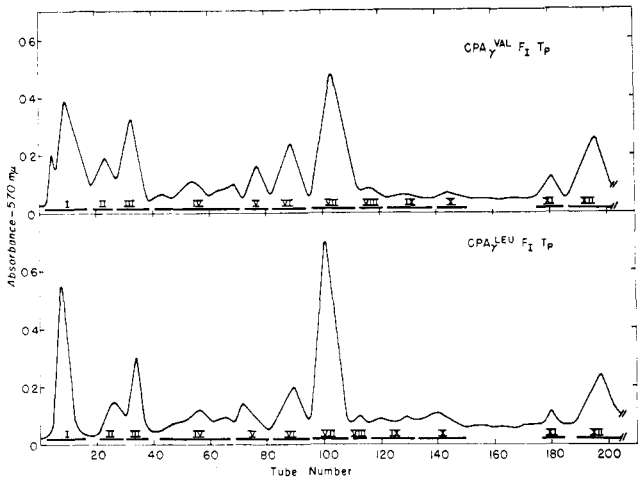

FIGURE 1: Initial portion of the elution profile of the soluble tryptic peptides from the two allotypic variants of carboxypeptidase $A_\gamma F_I$ on a column (0.9 × 25 cm) of Dowex 50-X8. The column was developed at 30 ml/hr with a double-linear gradient of pyridine acetate buffers as described in the text. Fractions of 2.0 ml were collected. The solid line is the absorbance at 570 mμ of aliquots, from every second tube, reacted with ninhydrin after alkaline hydrolysis.

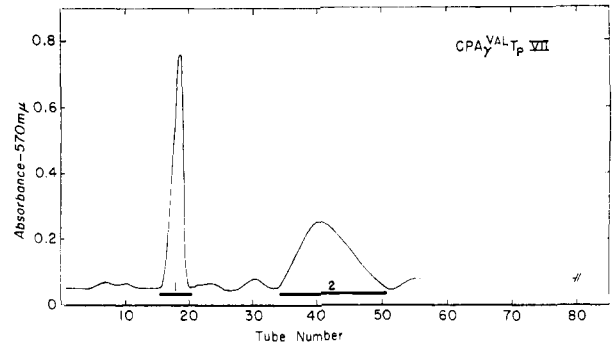

FIGURE 2: Elution profile of $CPA_\gamma^{Val}F_ITp$ VII on a 0.9 × 100 cm column of Dowex 1-X2. The column was developed at 30 ml/hr with a gradient of pyridine acetate. Fractions of 2.0 ml were collected.

suggesting that if any amino acid replacements were present in this fragment, they were compensated for by a reverse pair, *e.g.*, Ala for Gly linked with Gly for Ala.

F_{III}, prepared from both carboxypeptidase A_γ^{Leu} and carboxypeptidase A_γ^{Val}, was examined in this study for such replacements by amino acid analysis. In each case, the composition obtained was identical with that reported previously (Neurath *et al.*, 1968). These results suggest that fragment F_{III} does not possess any allotypic replacements. Additional evidence in support of this conclusion is supplied by the sequence analysis of F_{III} which has been completed.[2,6] Although these studies were carried out with carboxypeptidase A(Anson), no indication of any replacements was ever documented in these studies.

Analysis of F_I Tryptic Peptides. In contrast to the results obtained with F_{III}, several lines of evidence suggested that F_I contained at least one allotypic replacement. First, chromatography of the pooled enzyme on DEAE-cellulose which effectively separates the Val and Leu forms could not be rationalized on the basis of this replacement alone. Second, amino acid composition studies on the purified carboxypeptidase A_γ^{Val} and $_\gamma^{Leu}$ as well as F_I prepared from these proteins, indicated compositional differences in the alanyl, glutamyl, valyl, and isoleucyl content of these species (Nomoto *et al.*, 1969; Pétra and Neurath, 1969). Third, sequence analysis of F_I, from chromatographed enzyme, revealed the presence of peptides, in nearly equimolar mixtures, that varied by only a single residue in structure. These observations suggested that a detailed examination of the typtic peptides of F_I from the purified forms of carboxypeptidase A_γ was essential to establish whether the replacements indicated by composition were unique to the individual forms of the enzyme.

The separation of the soluble peptides derived from tryptic digestion of the F_I fragment from each form of carboxypeptidase A_γ is shown in Figure 1. Only the first half of the gradient is shown; the remaining peaks in each elution profile, not shown, occupy identical positions in the corresponding separations. These separations, which were carried out under essentially identical conditions, showed very similar patterns. The interpretation of these experiments was greatly facilitated by the detailed knowledge of the position of each soluble tryptic peptide from sequence analysis of F_I from CPA(Anson). In fact from these studies, particular attention was directed to pool VII, which had been observed to contain a mixture of peptides, two of which possessed very similar structures.

Each of the pools was subjected to electrophoresis at pH 3.75 and 6.5, followed by amino acid analysis. Only two differences were detected. In the first case, the peptide present in pool III of the $CPA_\gamma^{Leu}F_ITp$ separation was discovered to be in pool I of the $CPA_\gamma^{Val}F_ITp$ separation. Otherwise the distribution of peptides in the first three pools was identical. In each pool, the identity of the constituent peptides was checked against the corresponding peptides from unfractionated enzyme, and no anomalies were found.

In second case, pool VII, a distinct difference was discovered. High-voltage electrophoresis at pH 3.75 indicated that each pool contained two peptides. In each case, one spot gave a positive histidine test (Block, 1951) while the other was negative. However, electrophoresis at pH 6.5 gave a different pattern. Whereas pool VII from $CPA_\gamma^{Val}F_ITp$ again showed two spots, one neutral and one basic, pool VII from $CPA_\gamma^{Leu}F_ITp$ showed only a single neutral spot. The neutral spot in each sample was histidine positive. These results indicate that the histidine-negative peptides in each pool possess a different charge, the peptide in pool VII of $CPA_\gamma^{Leu}F_ITp$ containing an additional negative charge not fully expressed at pH 3.75.

Each of these pools was fractionated on Dowex 1-X2 and the results obtained are shown in Figures 2 and 3. The separation of pool VII of $CPA_\gamma^{Val}F_ITp$ (Figure 2) gave two pure peptides, $CPA_\gamma^{Val}F_ITp$ VII-1 and $CPA_\gamma^{Val}F_ITp$ VII-2. Electrophoresis at pH 6.5 indicated that pool VII-1 was basic and histidine negative, while pool Tp VII-2 was neutral and histidine positive. The separation $CPA_\gamma^{Leu}F_ITp$ VII required a longer column (150 cm as opposed to 100 cm used for $CPA_\gamma^{Val}F_ITp$ VII) to achieve resolution, but also yielded two pure peptides, $CPA_\gamma^{Leu}F_ITp$ VII-1 and $CPA_\gamma^{Leu}F_ITp$ VII-2. In contrast to the results obtained with pool VII of $CPA_\gamma^{Val}F_I$ Tp, both $CPA_\gamma^{Leu}F_ITp$ VII-1 and $CPA_\gamma^{Leu}F_ITp$ VII-2 were

[6] R. A. Bradshaw, in preparation.

TABLE I: Amino Acid Composition of the Tryptic Peptides of Carboxypeptidase $A_\gamma{}^{Val}F_I$ VII-1 and Carboxypeptidase $A_\gamma{}^{Leu}F_I$ VII-1.

Amino Acid	$CPA_\gamma{}^{Val}F_I$-Tp VII-1	$CPA_\gamma{}^{Leu}F_I$-Tp VII-1
Serine	0.97	0.96
Glutamic acid		*1.00*
Alanine	*3.04*	2.11
Valine	1.01	1.00
Leucine	1.00	1.00
Lysine	0.98	1.01
Total	7	7
% yield	63	75

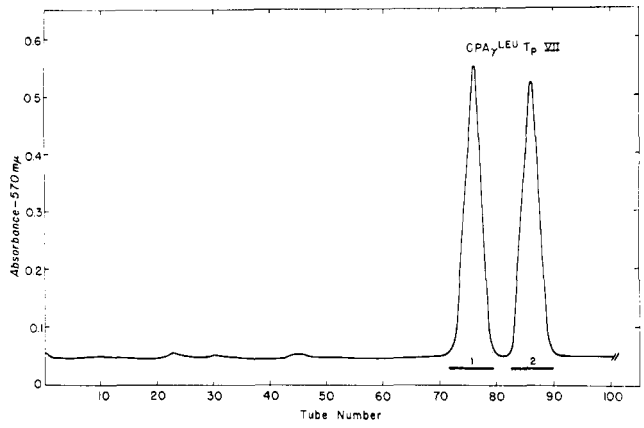

FIGURE 3: Elution profile of $CPA_\gamma{}^{Leu}F_ITp$ VII on a 0.9 × 150 cm column of Dowex 1-X2. Details are described in Figure 2.

neutral on electrophoresis at pH 6.5. Only pool Tp VII-2 was histidine positive.

Amino acid analysis of each of these four fractions gave the following results. Pool VII-2 from each separation yielded identical compositions, consistent with their electrophoretic behavior. However, the analyses of pool VII-1 from each separation, summarized in Table I, showed a single amino acid difference. $CPA_\gamma{}^{Val}F_ITp$ VII-1 contained three residues of alanine and no glutamic acid, while $CPA_\gamma{}^{Leu}F_ITp$ VII-1 contained two residues of alanine and one residue of glutamic acid. The charge distribution already noted on high-voltage electrophoresis at pH 6.5 indicated that the glutamic acid was in the acid form.

Sequence analysis by the subtractive Edman method and digestion by carboxypeptidase A and B gave the structures indicated in Table II. Clearly each peptide contains an identical structure with the exception of the fourth position, in which the $CPA_\gamma{}^{Val}F_ITp$ VII-1 peptide contains alanine while the $CPA_\gamma{}^{Leu}F_ITp$ VII-1 peptide has glutamic acid.

Examination of the remaining high yield peptides did not reveal any further differences. However, two factors suggested that the use of the tryptic peptides to examine allotypic replacements in F_I was insufficient. First, trypsin solubilizes only approximately 50% of the fragment, the remaining portion being found in a few large and insoluble pieces.[7] Second, low yields of several other peptides, including some previously identified in digests of F_I from CPA(Anson), which were compounded by the small amount of F_I available made identification of other allotypic replacements impossible. Consequently, a second digest of F_I from $CPA_\gamma{}^{Val}$ and $CPA_\gamma{}^{Leu}$ was carried out with thermolysin.

Thermolytic Peptides. The use of thermolysin was dictated by the pronounced success achieved with this enzyme in the sequence analysis of F_1[7] in which virtually all of the fragment was solubilized. However, the mixture produced was much more complex, at least in the early portion of the gradient, and required more detailed analysis of the fractions obtained. In this case, as with the tryptic peptides, the detailed knowledge of the separation carried out on the F_1 from CPA(Anson) allowed for the more ready identification of the peptides.

[7] R. A. Bradshaw, K. A. Walsh, and H. Neurath, in preparation.

The initial phase of the separation which contained all the identifiable replacements is shown in Figure 4. The separations observed were in good agreement with that obtained from the digest of F_I prepared from CPA(Anson).[7] In each sample, the fractions were pooled according to the solid bars, and subfractionation was carried out on Dowex 1-X2. The analysis of each purified fraction obtained in this manner was used to locate apparently similar peptides which differed by a single amino acid. Two pairs of such peptides were located, one in pool III and one in pool V. The purification of corresponding pools from $CPA_\gamma{}^{Val}F_ITh$ and $CPA_\gamma{}^{Leu}F_ITh$ for these two pairs is shown in Figures 5 and 6.

The separation of pool 3 from $CPA_\gamma{}^{Val}F_ITh$ is shown in the upper half of Figure 5. Only two peptide fractions were obtained which were consequently analyzed after high-voltage electrophoresis tests for purity. The lower half of Figure 5 shows the separation of the corresponding pool from $CPA_\gamma{}^{Leu}F_ITh$. Three pools were obtained in this case. Pool Th III-1 was found to be identical with pool Th III-2 of $CPA_\gamma{}^{Val}F_I$. Pool Th III-2, which was absent from the $CPA_\gamma{}^{Val}F_I$ separation, was readily identified as being present, in identical form,

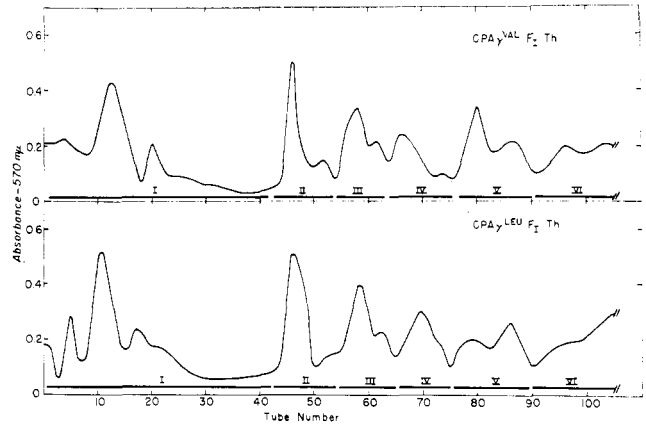

FIGURE 4: Initial portion of the elution profile of the soluble thermolytic peptides from the two allotypic forms of carboxypeptidase $A_\gamma F_I$ on a column (2.0 × 25 cm) of Dowex 50-X8. The column was developed at 80 ml/hr and collected in 6-ml fractions. Other details described in Figure 1.

TABLE II: The Structure of the Allotypic Tryptic Peptides from $CPA_\gamma^{Val}F_I$ and $CPA_\gamma^{Leu}F_I$.

Peptide $CPA_\gamma^{Val}F_ITp$ VII-1:	Ser-Ala-Val-Ala-Ala-Leu-Lys
Peptide $CPA_\gamma^{Leu}F_ITp$ VII-1:	Ser-Ala-Val-Glu-Ala-Leu-Lys

TABLE III: The Structure of the Allotypic Thermolytic Peptides from $CPA_\gamma^{Val}F_I$ and $CPA_\gamma^{Leu}F_I$.

Peptide $CPA_\gamma^{Val}F_ITh$ III-1:	Ala-Val-Ala-Ala
Peptide $CPA_\gamma^{Leu}F_ITh$ III-3:	Ala-Val-Glu-Ala
Peptide $CPA_\gamma^{Val}F_ITh$ V-3:	Ile-Val-Asp
Peptide $CPA_\gamma^{Leu}F_ITh$ V-3:	Val-Val-Asp

in an adjacent pool. Pool Th III-3, also a pure peptide, gave a composition identical with pool Th III-1 from the $CPA_\gamma^{Val}F_I$ separation, except that it contained one residue of glutamic acid in place of alanine. Sequence analysis, shown in the upper part of Table III, readily identified these tetrapeptides as part of the same portion of the molecule, already isolated in the tryptic peptides. Thus the replacement found in pool Th III is the same as that identified with the tryptic peptides, and supports that assignment.

The subfraction of pool Th V from each allotype is shown in Figure 6. The upper part of Figure 6 shows the separation of pool Th V from $CPA_\gamma^{Val}F_I$. Five peptide pools were obtained in contrast to the four that were obtained by fractionation of the same pool from $CPA_\gamma^{Leu}F_ITh$ shown in the bottom half of the figure. The pool missing in the $CPA_\gamma^{Leu}F_ITh$ separation, pool V-2, was isolated in the adjacent pool. Each of the peptides from the other pools which showed identical elution positions were found to possess identical compositions except Th V-3. Peptide $CPA_\gamma^{Val}F_ITh$ V-3 gave the composition of a tripeptide composed of isoleucine, valine, and aspartic acid, whereas $CPA_\gamma^{Leu}F_ITh$ V-3 was composed of two residues of valine and one of aspartic acid. Each sample required hydrolysis for 72 hr to give full release of the valine and isoleucine. Sequence analysis of these peptides gave the results shown in the lower half of Table III. Significantly, isoleucine and valine, the single difference, was found to be the initial residue in each peptide. The yield for peptide $CPA_\gamma^{Val}F_ITh$ V-3 was 80%, and for peptide $CPA_\gamma^{Leu}F_ITh$ V-3 was 69%.

Discussion

The initial observations of Bargetzi *et al.* (1964) that the carboxyl-terminal cyanogen bromide fragment of bovine carboxypeptidase A contained a single valine-leucine replacement was the first indication that this enzyme possessed heterogeneity in addition to that arising from different sites of activation of the zymogen, procarboxypeptidase A (Sampath Kumar *et al.*, 1964a). The detailed chromatographic analyses of Pétra and Neurath (1969) have substantially clarified this problem by indicating that the various activation forms, *i.e.*, α, β, and γ, can each be resolved into two forms[4] which are distinguished by the valine-leucine replacement in the antepenultimate position. Amino acid analyses of these purified proteins, as well as fragments derived therefrom, suggested that additional amino acid replacements were present. In order to establish the nature and distribution of these additional replacements, examination of each of the cyanogen bromide fragments (Nomoto *et al.*, 1969) was carried out. Neither F_N nor F_{III}, which together constitute the first 103 residues of carboxypeptidase A_α, showed any indication of heterogeneity from either amino acid analysis or sequence analysis on material prepared from unchromatographed enzyme. Consequently, it was concluded that any additional replacements must occur in the F_I segment of molecule covering residues 104–301.

Analysis of this fragment with both tryptic and thermolytic peptides derived from F_I prepared from each allotype of carboxypeptidase A_γ indicated two additional replacement sites. In the first case, the soluble tryptic peptides were examined and revealed a single alanine–glutamic acid interchange. Incorporation of these peptides into the chemical sequence[8] indicate that the site of replacement is position 228 in CPA_α.

The tryptic peptide analysis proved to be unsatisfactory for locating further replacements because of the large insoluble core and low yield of some of the peptides. This latter problem was compounded by the limited amount of F_I available from the purified species. Therefore, a second digest using thermolysin was prepared, since this enzyme solubilizes virtually all of the F_I fragment albeit that the peptide products are generally somewhat smaller than those obtained from the tryptic digest.[7] In this case, the alanine-glutamic acid replacement at position 228 was again documented, and, in addition, a second replacement involving an isoleucine–valine interchange was found. Although the site was demonstrated with a tripeptide, an examination of the complete sequence of carboxypeptidase[8] indicates that this tripeptide combination, Val-Val-Asp and Ile-Val-Asp, can only be incorporated at one place in the molecule. The site of the replacement is residue 179.

Considerable evidence, in addition to the complete sequence, is also available to support these assignments. First, tryptic, chymotryptic, and peptic peptides obtained from heterogeneous F_I, and possessing various overlaps to the adjacent sequences, confirm this location.[7] It is of interest to note that, whereas peptides from the alanine-glutamic site were always clearly separated, peptides from the isoleucine-valine site were obtained as an equimolar mixture regardless of the type of digest. Chromatographies on Dowex 50-X8, 50-X2, and 1-X2 were equally ineffective in separating the allotypic forms of these peptides. In fact, the peptides were only obtained in pure form when they were isolated from the purified species (*vide supra*). Second, the amino acid analyses of F_I from CPA_γ^{Val} and CPA_γ^{Leu} show a clear-cut distribution of alanine, glutamic acid, isoleucine, and valine to fit exactly the distribution deduced from the isolated peptides (Nomoto *et al.*, 1969).

It is significant that neither amino acid composition nor peptide analysis indicates any further replacements. Because of

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

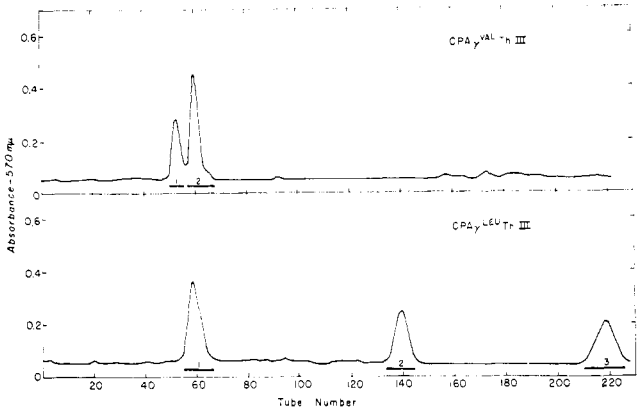

FIGURE 5: Elution profile of CPA$_\gamma^{Val}$ and CPA$_\gamma^{Leu}$F$_1$Th III on a 0.9 × 150 cm column of Dowex 1-X2. Details are described in Figure 2.

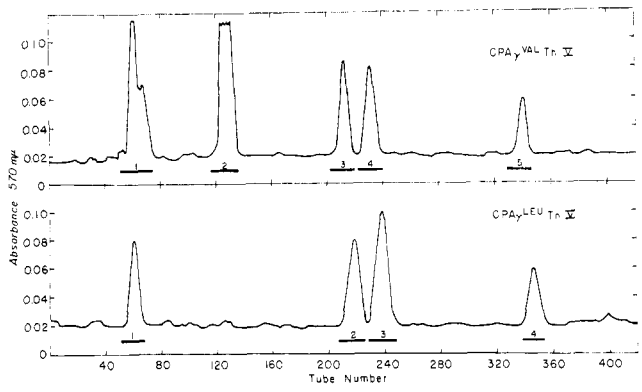

FIGURE 6: Elution profile of CPA$_\gamma^{Val}$ and CPA$_\gamma^{Leu}$ F$_1$Th V on a 0.9 × 150 cm column of Dowex 1-X2. Details are described in Figure 2.

the extreme difficulty in obtaining some of the portions of the F_1 fragment in high yield, it is not possible at present to rigorously exclude the possibility that other replacements do not exist. In fact, replacements of the type aspartic acid-asparagine and glutamic acid-glutamine cannot be reliably examined by present techniques because of the partial loss of some amides during peptide isolation. At best, it can be stated that no further evidence exists to suggest that there are any more replacements than those which have been demonstrated.

An examination of the three-dimensional model of carboxypeptidase A, determined from X-ray analysis (Lipscomb *et al.*, 1968), reveals several interesting features of the replacement sites. Residue 179 occupies a position in the interior portion of the molecule with a predominantly hydrophobic environment. On the other hand, residue 228 is a surface residue in which the side chain protrudes into the solvent surrounding the protein molecule, providing a possible explanation for the ion-exchange chromatographic separation of the two allotypic forms of the enzyme. The third site, 305, occurs in a helical segment with the side chain located in a manner to maximize hydrophobic contact with the main body of the molecule. In all three cases, the chemical nature of each pair of amino acids is in keeping with its location in the protein molecule.

It is of interest to note the distribution of the three replacement loci in the whole molecule. All three positions occur in the carboxyl-terminal half while the amino-terminal portion is constant. If these sites are considered in terms of the zymogen, procarboxypeptidase A, which is the whole gene product, approximately 60 residues must be added to the amino end (Neurath, 1960) thus displacing the replacement sites to the last third of the polypeptide chain.

An additional significance of these findings which deserves comment is the fact that these three replacements are linked; that is, each molecule of carboxypeptidase A possesses one set of three replacement amino acids and there is no evidence, at present, to suggest that there exist molecules which possess different mixtures of the replacement amino acids at the three loci. Consequently, it appears as if there are two forms of carboxypeptidase A which are characterized, at positions 179, 228, and 305, on the one hand, by the amino acids, isoleucine-alanine-valine, and on the other hand, by valine-glutamic acid-leucine.

The most significant aspect of this exclusive pairing resides in its origin. Although each replacement combination is genetically characterized by only a single base mutation, the origin of three such mutations, with no apparent intermediates remaining in the population, is harder to reconcile. This situation is worthy of additional comment.

Two alternate mechanisms suggest themselves. First, the appearances of two distinct gene products of the carboxypeptidase A loci of the chromosome may be the result of gene duplication. This hypothesis demands that each species possesses two genes for carboxypeptidase A and that one gene has undergone independent mutation, in the absence of selective pressure, to produce the second form of the enzyme now present in the population. This mechanism suffers from two criticisms. First, it is difficult to reconcile the apparent Mendelian character of this trait, as documented by Walsh *et al.* (1966) in which single animal experiments were carried out to show that any given animal was either homozygous for *either* form of the enzyme or was heterozygous in a 50:50 ratio for both forms. Gene duplication would not provide a simple route to produce an animal homozygous for the carboxypeptidase A resulting from the mutation of the second gene. Second, this mechanism does not explain the absence of the intermediate products resulting from the first and second mutations.

The alternative explanation is simpler in that it merely calls for a series of mutations for which the final product, with three replacements, is preserved along with one preceding gene product. This argument can be presented in one of two ways. Either the ancestor gene was different from either of the two present forms of the enzyme and was lost (Neurath *et al.*, 1969), or the product of the ancestral gene is, in fact, identical with one of the present forms, and the second form was derived from it.

Neither explanation is completely satisfactory. First, it is not clear why the missing intermediates, which must be considered lost due to the absence of selective pressure, were not functionally identical with the CPA$_\gamma^{Val}$ and CPA$_\gamma^{Leu}$ forms since no apparent functional differences between these enzymes can be demonstrated and second, why one form possessed sufficient selective advantage to become incorporated into the population to at least 40% (Pétra and Neurath, 1969).

A possible explanation of these problems may lie in the chronology of the events that introduced these allotypic forms into the population. If, in fact, the mutational events which

caused the formation of the second gene for CPA occurred either together or in rapid succession at a point very early in the evolutionary development of the bovine species, then the lack of intermediates and the 60–40 distrubution of the allotypes in the present population may be perceived without the necessity to invoke an argument for selective pressure. Development of the present population from the very small gene pool in which these mutational events occurred could result in the present distribution.

It may be countered that the lack of evidence concerning apparent selective pressure only reflects a lack of data concerning the interaction of the enzyme and its zymogen with its physiological environment. The veracity of this statement is obvious and only serves to indicate the magnitude of the problems confronting molecular biologists and protein chemists in their studies of evolution.

Acknowledgment

We thank Mr. Mark Kupperberg for assistance with the chromatography of carboxypeptidase and Mrs. Milagros Santarin, Mrs. Brita Moody, Mrs. Lynne Maston, Mrs. Merry Jarvis, and Mr. C. V. Shih for assistance with the peptide isolation and characterization. Mr. Richard R. Granberg and Mr. Thomas Oas performed the amino acid analyses.

References

Anson, M. L. (1937), *J. Gen. Physiol. 20*, 777.

Bargetzi, J. P., Sampath Kumar, K. S. V., Cox, D. J., Walsh, K. A., and Neurath, H. (1964), *J. Biol. Chem. 239*, 3767.

Block, R. J. (1951), *Arch. Biochem. Biophys. 31*, 266.

Bradshaw, R. A., Garner, W. H., and Gurd, F. R. N. (1969), *J. Biol. Chem. 244, 2149*.

Cox, D. J., Bovard, F. C., Bargetzi, J. P., Walsh, K. A., and Neurath, H. (1964), *Biochemistry 3*, 44.

Lipscomb, W. N., Hartsuck, J. A., Reeke, G. N., Jr., Quiocho, F. A., Bethge, P. H., Ludwig, M. L., Steitz, T. A., Muirhead, H., and Coppola, J. C. (1968), *Brookhaven Symp. Biol. 21*, 24.

Neurath, H. (1960), *Enzymes 4*, 11.

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) (in press).

Nomoto, M., Srinivasan, N. G., Bradshaw, R. A., Wade, R. D., and Neurath, H. (1969), *Biochemistry 8*, 2755. (this issue; preceding paper).

Pétra, P. H., and Neurath, H. (1969), *Biochemistry* (in press).

Sampath Kumar, K. S. V., Clegg, J. B., and Walsh, K. A. (1964a), *Biochemistry 3*, 1728.

Sampath Kumar, K. S. V., Walsh, K. A., and Neurath, H. (1964b), *Biochemistry 3*, 1726.

Shearer, W. T., Bradshaw, R. A., Gurd, F. R. N., and Peters, T., Jr. (1967), *J. Biol. Chem. 242*, 5451.

Walsh, K. A., Ericsson, L. H. and Neurath, H. (1966), *Proc. Natl. Acad. Sci. U. S. 56*, 1339.