PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Bovine Carboxypeptidase A -- Activation, Chemical Structure and Molecular Heterogeneity

H. Neurath, R. A. Bradshaw, P. H. Petra and K. A. Walsh

Phil. Trans. R. Soc. Lond. B 1970 **257**, 159-176
doi: 10.1098/rstb.1970.0019

Phil. Trans. Roy. Soc. Lond. B **257**, 159–176 (1970) [159]
Printed in Great Britain

III. CARBOXYPEPTIDASE

Bovine carboxypeptidase A—activation, chemical structure and molecular heterogeneity

By H. Neurath, R. A. Bradshaw, P. H. Pétra and K. A. Walsh
Department of Biochemistry, University of Washington, Seattle, Washington 98105, *U.S.A.*

The study of the structure and function of bovine pancreatic carboxypeptidase A has a long historical record which began in 1937 with the isolation of the crystalline enzyme by Anson (1937). Although the literature of this exopeptidase has been reviewed on several occasions (Neurath & Schwert 1950; Smith 1951; Neurath 1960; Vallee 1964; Neurath, Bradshaw, Ericsson, Babin, Petra & Walsh 1968; Vallee & Riordan 1968; Lipscomb, Hartsuck, Reeke, Quiocho, Ludwig, Steitz & Bethge 1968), nevertheless, it seems appropriate at the outset of a conference on carboxypeptidase to review in brief the major contributions that have been made over the years to the study of the structure and function of this enzyme, as summarized in table 1.

Table 1. History of carboxypeptidase A

isolation	Anson (1937); Putnam & Neurath (1946); Allan *et al.* (1964); Cox *et al.* (1964)
peptidase activity	Bergmann & Fruton (1941)
esterase activity	Snoke, Schwert & Neurath (1948)
kinetics and specificity	Neurath & Schwert (1950); Smith (1951)
competitive inhibitors	
β-phenylpropionate	Elkins-Kaufman & Neurath (1948)
indole acetate	Smith, Lumry & Polglase (1951)
molecular weight	Smith, Brown & Hanson (1949)
amino acid composition	Smith & Stockell (1954); Bargetzi *et al.* (1963)
Zn-metalloenzyme	Vallee & Neurath (1954); Vallee (1955)
apo-enzyme	Vallee, Rupley, Coombs & Neurath (1960)
-SH reactivity	Vallee, Coombs & Hoch (1960); Walsh *et al.* (1962)
modulation of specificity	
by metal replacement	Coleman & Vallee (1960)
by chemical modification	Simpson, Riordan & Vallee (1963); Simpson & Vallee (1966); Riordan, Sokolovsky & Vallee (1967); Sokolovsky & Vallee (1967)
zymogen activation	Keller, Cohen & Neurath (1956, 1958*a*); Brown *et al.* (1963); Freisheim, Walsh & Neurath (1967)
N-terminal variants	Sampath Kumar *et al.* (1964*a*)
C-terminal variants	Bargetzi *et al.* (1964); Walsh, Ericsson & Neurath (1966)
partial sequences	Sampath Kumar *et al.* (1964*b*); Bargetzi *et al.* (1964); Roholt & Pressman (1967); Neurath *et al.* (1968)
X-ray crystallography	Ludwig *et al.* (1963); Quiocho & Richards (1964); Lipscomb *et al.* (1968)

In brief, carboxypeptidase A catalyses the hydrolysis of the carboxyl-terminal peptide bond in peptides and proteins and the corresponding ester bond in model compounds, wherein the peptide nitrogen is replaced by an oxygen. Its specificity is directed primarily toward aromatic and hydrophobic side chains such as phenylalanine, tryptophan or leucine. In contrast, the action of carboxypeptidase B is preferentially directed toward the basic side chains of arginine or lysine. The earliest quantitative observations on the specificity of carboxypeptidase A are those of Bergman & Fruton (1941) for peptide substrates and Snoke, Schwert & Neurath (1948)

for ester substrates. The discovery that carboxypeptidase possessed both peptidase and esterase activity has clarified many aspects of enzyme mechanism, but it has also contributed its own problems, because detailed analysis has revealed that some of these synthetic substrates are bound to the enzyme in a complex manner. The most effective competitive inhibitors of carboxypeptidase A are derivatives of phenylalanine or tryptophan substrates such as β-phenylpropionate or indole acetate. Some of these, such as β-phenylpropionate, can bind to the enzyme at several sites (Bethune 1965). One of the major milestones in the study of carboxypeptidase is its identification as a zinc metalloenzyme. This finding has served to focus attention on the unique role of zinc and other transition elements of the IIB series in the activity and specificity of carboxypeptidase; it has generated a search for the nature of the ligands involved in zinc binding; and it has aided substantially in the functional and structural recognition of the active site. Similarly, the demonstrable modulation of enzymic specificity by chemical modification with site specific reagents (Vallee & Riordan 1968) has significantly broadened the range of approaches to the elucidation of the mechanism of action of carboxypeptidase A.

Despite the rather extensive characterization of the enzyme in terms of function (Vallee & Riordan 1968), our knowledge of the chemical structure of the enzyme is of more recent origin and is as yet incomplete. Thus neither the amino acid sequence (Neurath *et al.* 1968) nor the three-dimensional structure (Lipscomb *et al.* 1968) of this enzyme has been elucidated in full, though rapid advancement in both directions gives promise that these goals will soon be reached.

Detailed investigations of the chemical structure of carboxypeptidase A originated in our laboratory in 1964 as part of our effort to characterize the products arising from various methods of isolation of the enzyme generated under different conditions of zymogen activation. These investigations have led to the determination of the amino and carboxyl terminal sequences of carboxypeptidase A. Other regions of the molecule have been partially identified through efforts to differentiate between the two half cystinyl residues. A systematic and continuing investigation of the complete amino acid sequence of this enzyme met with difficulties at an early stage and has required new experimental approaches to the chemical characterization of this enzyme. These difficulties relate to (1) the apparent heterogeneity and complex structure of the zymogen, bovine pancreatic procarboxypeptidase A; (2) the multiple products of activation, and (3) genetic variations of the structure of the enzyme which probably arise from allelomorphism. While each of the contributing factors to be considered in this regard is not without precedent, together they have presented problems, since the usual criteria for the proof of structure were not applicable in this instance. The picture that finally emerged of this protein appears to be at variance with prevalent ideas of the uniqueness of the structure of enzymes.

Zymogen activation

Brown, Greenshields, Yamasaki & Neurath (1963) in our laboratory have shown that bovine pancreatic procarboxypeptidase A is an aggregate of subunits occurring either as a trimer, with a sedimentation coefficient of 6 Svedberg units, or as a dimer, with a sedimentation coefficient of 5 Svedberg units as shown in figure 1. The trimer is composed of subunit I, the immediate precursor of carboxypeptidase A, and subunits II and III, which are precursors of an endopeptidase similar in specificity to chymotrypsin. The dimer contains only subunits I and II. While the initial observations suggested that each of these forms is homogeneous, application of more recent chromatographic systems by R. W. Tye in our laboratory has indicated heterogeneity

of significant magnitude. Interaction between these subunits is so strong that it has not yet been possible to obtain subunit I, the immediate precursor of carboxypeptidase, in a native and viable form. As the closest approach, subunit I can be obtained by acylation of the complex with succinic anhydride, which causes spontaneous disaggregation into a mixture of components from which succinyl fraction I can be isolated by chromatography (Freisheim, Walsh & Neurath 1967).

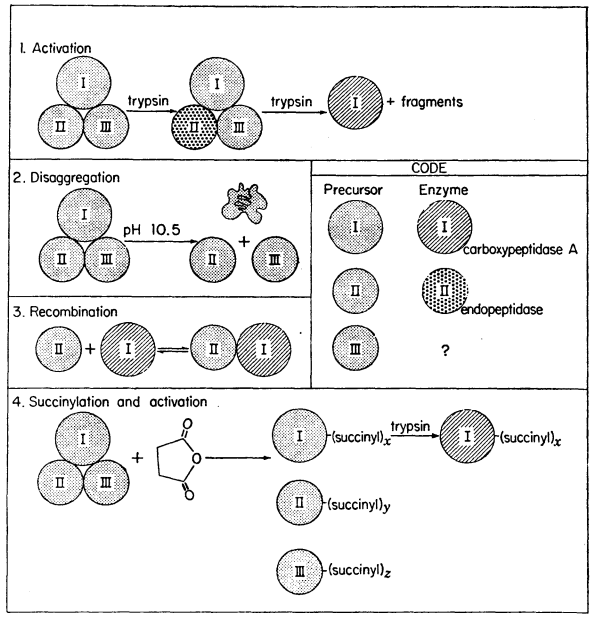

FIGURE 1. Schematic representation of the structure and reactions of the bovine procarboxypeptidase A complex. References are given in the text.

The generation of carboxypeptidase activity from the trimeric precursor complex, and hence the isolation of the crystalline enzyme, involves the participation of two enzymes (Keller, Cohen & Neurath 1956, 1958*a*, *b*). Trypsin at low concentration activates the endopeptidase, fraction II, which then in concert with additional trypsin generates carboxypeptidase activity. This reaction is slow, requiring several hours of incubation at 37 °C. In contrast, monomeric succinyl fraction I or by analogy other procarboxypeptidases which occur in monomeric form such as bovine B (Cox, Wintersberger & Neurath 1962), dogfish A (Lacko & Neurath 1967), or dogfish B (Prahl & Neurath 1966) are activated much more rapidly under the influence of but one enzyme such as trypsin. These observations suggest that the point of primary bond cleavage in the procarboxypeptidase unit is masked as a result of the quarternary structure of the complex.

The formation of the active carboxypeptidase is accompanied by the cleavage of a peptide bond releasing an N-terminal fragment of the enzyme precursor. This process is different in kind from the well known reactions involved in the activation of chymotrypsinogen or trypsinogen. The activation fragment is large containing some 60 amino acid residues in contrast to the activation peptide of trypsinogen which contains 6 residues. Moreover, the precise location of the bond being hydrolysed appears to be unimportant for the generation of enzymic activity. The products of activation arise from cleavages at different though closely related points in the zymogen molecule. There is evidence that some of the functional features of the enzyme are pre-existent in the enzyme precursor. Thus acetylation of the succinylated zymogen

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

with *N*-acetylimidazole prior to activation results in the same changes in enzymatic function that occur when the active enzyme, succinylcarboxypeptidase, is acetylated, i.e. the esterase activity increases as the peptidase activity decreases (Freisheim *et al.* 1967). In each case, these changes are prevented by β-phenylpropionate, a competitive inhibitor. The effects of metal replacements in the zymogen or in the active enzyme also have similar functional consequences. A full understanding of the structural relation between enzyme and zymogen must await a more detailed analysis of the chemical structure of procarboxypeptidase A.

Methods of preparation and amino acid composition of carboxypeptidase A

Three different procedures for the preparation of carboxypeptidase A have been described and are summarized in figure 2. Each preparation yields, as a major component, a form of the enzyme which differs chemically from the others. These differences relate to the locus of cleavage

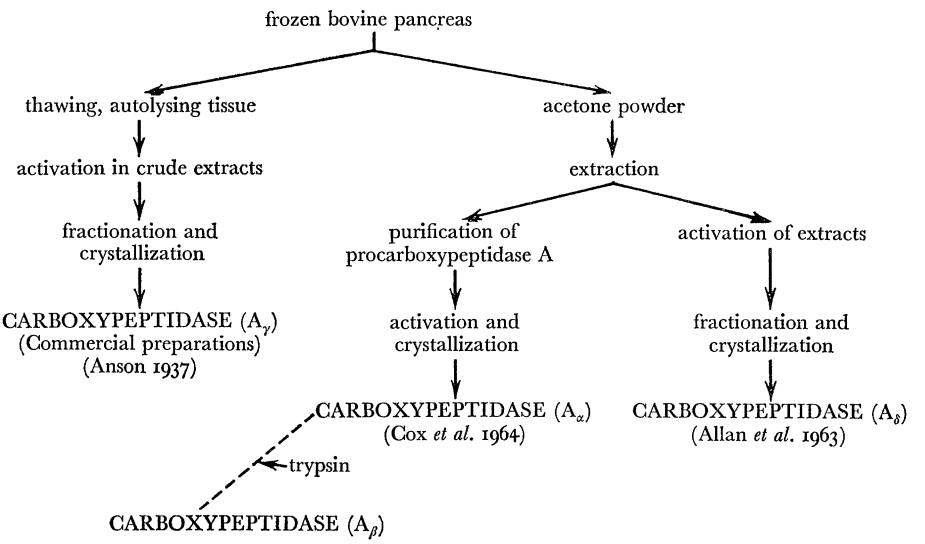

Figure 2. Flow diagram of the methods of preparation of carboxypeptidase A according to the procedures of Anson (1937), Allan *et al.* (1964), and Cox *et al.* (1964).

just mentioned during zymogen activation under the particular experimental conditions employed. The major product of the Anson (1937) preparation isolated from thawing and autolysing tissue is carboxypeptidase A_γ. The primary product of the preparation by Cox, Bovard, Bargetzi, Walsh & Neurath (1964) obtained from partially purified procarboxypeptidase A contains an additional N-terminal heptapeptide as compared to carboxypeptidase A_γ. Carboxypeptidase A_β is two amino acids shorter at the N-terminus than the α form and occurs as a minor component in all other preparations. The β enzyme can be generated from the α form by removing the N-terminal peptide alanylarginine with trypsin (Sampath Kumar, Clegg & Walsh 1964*a*). Until recently this form of the enzyme has not been available in a pure state. The third preparation was described by Allan, Keller & Neurath (1964) and is similar to the Anson enzyme in that it is derived by activation of the extract rather than the purified zymogen. It is isolated from an acetone powder rather than from frozen tissue and differs from the Anson enzyme also in the ease of reactivation of the apoenzyme by the addition of zinc.

The nature of the alternate products of activation in the several preparations are reflected in the amino acid composition of the N-terminal fragments as shown in figure 3. The N-terminal sequence corresponds, from left to right, to carboxypeptidases A_α, A_β and A_γ, respectively, and the different points of cleavage are indicated by the dotted lines. The products of these several activation processes are apparently nearly identical in their catlytic activities towards small substrates. Thus the catalytic apparatus is insensitive to variations in the N-terminal portion of the molecule.

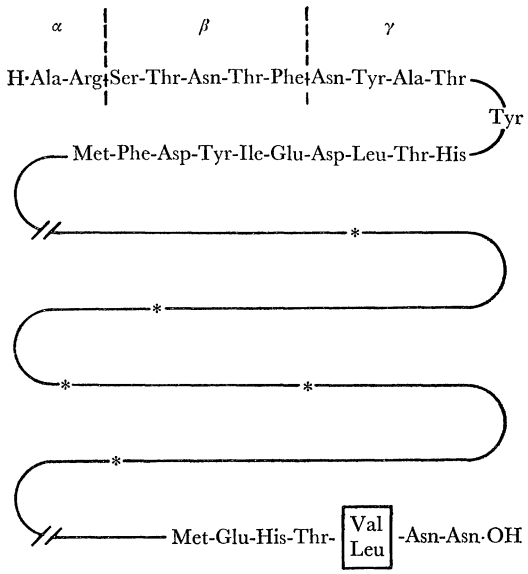

FIGURE 3. Schematic representation of bovine carboxypeptidase A illustrating the structure of the N- and C-terminal regions. The sites of activation are indicated by vertical dashed lines.

In addition to these variations in the N-terminal regions, two variant forms have been found which differ from each other in the C-terminal region by a Leu-Val replacement in the antepenultimate position. Two other amino acid replacements elsewhere in the molecule have been recently uncovered, and have thus introduced additional variations in the chemical structure and amino acid sequence of this protein.

RESOLUTION OF HETEROGENEITY OF CARBOXYPEPTIDASE A

The preparation of a molecularly homogeneous form of carboxypeptidase A thus appears a formidable task, requiring first pancreatic juice from an animal known to be homozygous with regard to the established amino acid replacements, and secondly the development of procedures that would yield a single product of activation. An alternative and much more suitable route to the preparation of a homogeneous enzyme involves the development of chromatographic procedures that would clearly separate all molecular species from each other. The development of such a chromatographic system proved difficult since the extreme insolubility of the enzyme had resisted all attempts at ion exchange chromatography. However, the inclusion of the competitive inhibitor β-phenylpropionate which, in addition to stabilizing the enzyme, also increases the protein solubility at low inhibitor concentrations, has solved the problem (Pétra

& Neurath 1968). The first system is shown in figure 4, which depicts the chromatography of commercial preparations of carboxypeptidase A_{Anson} isolated from pooled bovine pancreas glands. Five fractions, numbered consecutively in the order of elution, were separated from each other and each of the five fractions was crystallized. The breakthrough fraction showed only 10 % of the expected specific activity and probably represents a mixture of proteins. Proof of identity of these fractions was obtained by characterization of the N- and C-terminal fragments

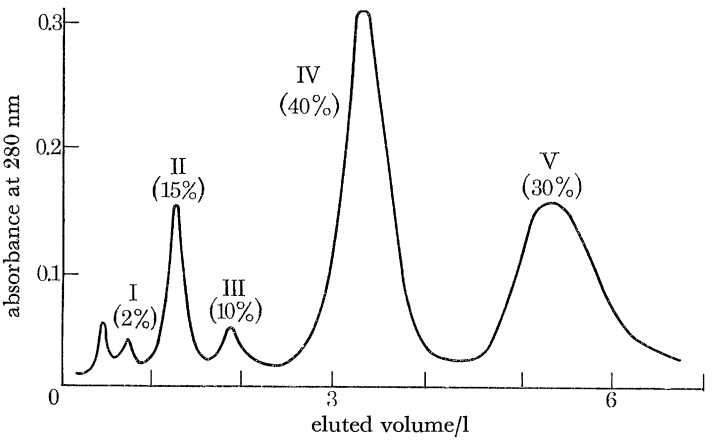

FIGURE 4. Elution pattern of the separation of 800 mg of twice crystallized carboxypeptidase A_{Anson} on a column (2.5 × 90 cm) of DE-52. The column was developed at 93 ml h^{-1}, at ph 7.5 and 4 °C, and the solution used contained (mmol l^{-1}): 50 β-phenylpropionate, 40 LiCl and 50 Tris.

of each protein, isolated after cleavage with cyanogen bromide. The structure of the N-terminal peptide fragments is given in table 2. It is clear that peaks II and III have identical partial sequences at the N-terminus, thus demonstrating that they are derived from carboxypeptidase A_β (cf. table 2). Similar experiments with peaks IV and V showed them to possess a structure

TABLE 2. SEQUENCE OF N-TERMINAL PEPTIDES

CPA_α†	H.Ala-Arg-Ser-Thr-Asn-Thr-Phe-Asn-Tyr-Ala-
peak II	H.Ser-Thr-Asx-Thr-Phe-Asx-Tyr-Ala-
peak III	H.Ser-Thr-Asx-Thr-Phe-Asx-Tyr-Ala-
peak IV	H.Asx-Tyr-Ala-
peak V	H.Asx-Tyr-Ala-

† Sampath Kumar *et al.* (1964*a*).

identical to that found in carboxypeptidase A_γ. The N-terminal sequence of peak I was identical to that shown for carboxypeptidase A_α, establishing the identity of this peak. The investigation of the C-terminal hexapeptide obtained after cyanogen bromide cleavage was also undertaken for peaks II, III, IV and V as is shown in table 3. It is seen that in the case of peaks II and IV,

TABLE 3. AMINO ACID COMPOSITION OF C-TERMINAL PEPTIDES

number of residues per molecule

	Glu	His	Thr	Val	Leu	Asp
CPA pool	1	1	1	**0.6**	**0.4**	2
peak II	1	1	1	1	0	2
peak III	1	1	1	0	1	2
peak IV	1	1	1	1	0	2
peak V	1	1	1	0	1	2

leucine is completely absent from the peptides, whereas the opposite is true for peaks III and V, where valine is totally absent. These data clearly demonstrate an amino acid replacement in the C-terminal region and support the earlier work of Walsh, Ericsson & Neurath (1966) who isolated the valine and leucine forms of carboxypeptidase A from preparations of pancreas glands of single animals homozygous with respect to this trait.

All these species of carboxypeptidase A have the same catalytic properties as shown in table 4. There are no major differences among these fractions in their activity towards the peptide substrate carbobenzoxyglycyl-L-phenylalanine or the ester substrate hippuryl-L-phenyllactic acid (unpublished data by J. Uren). The various enzymes do differ markedly, however, in their heat stability at 50 °C as shown in figure 5. The rate of inactivation appears to depend on two factors. First, the longer the N-terminal region of the enzyme, the slower the rate of inactivation. Carboxypeptidase A_α is more heat stable than carboxypeptidase A_β which in turn is more heat stable than the γ form. Secondly, all the valine allotypes are more heat stable than the leucine allotypes, indicating that the valine allotype possesses a greater conformational stability. Thus the six different forms of carboxypeptidase, differing in both the N-terminal structure and in amino acid sequence, possess different conformational stability, a fact that may reflect altered interaction in different loci of the molecule.

TABLE 4. KINETIC PARAMETERS

J. Uren, unpublished results

	CGP†		HPLA‡	
	K_m/kmol l^{-1}	k/s^{-1}	K_m/kmol l^{-1}	k_{cat}/s^{-1}
control§	2.0	55	12	466
peak I	1.7	38	14	439
peak II	1.9	55	10	500
peak III	1.5	47	23	666
peak IV	1.8	48	14	476
peak V	2.4	62	9	396

† Carbobenzoxyglycyl-L-phenylalanine.
‡ Hippuryl-DL-phenyllactic acid.
§ CPA_γ pool.

Carboxypeptidase A_{Cox} cannot be chromatographed in the same system as just described for the Anson enzyme. Instead a complex nine chamber gradient was developed in which both lithium chloride and β-phenylpropionate are varied simultaneously. The results of the chromatography are shown in figure 6 as a solid line. The four labelled peaks have approximately the same specific activity toward the ester hippuryl-L-phenyllactic acid. The three major peaks were characterized on the basis of their N- and C-terminal fragments as just described. Peak I corresponds to the valine form of the α enzyme, peak III to the leucine form of the β enzyme, whereas peak II was a mixture of the leucine form of the α enzyme with a 20 % contamination of the valine form of the β enzyme. The γ form was completely absent. The nature of this contamination was confirmed by demonstrating that authentic carboxypeptidase A_β^{Val} (isolated as a pure protein by chromatography of an Anson preparation as just described), shown as a broken line, co-chromatographed with carboxypeptidase A_α^{Leu}. Peak I–A has not yet been analysed.

To summarize this aspect of the investigation, chromatographic systems are now available for the separation of all six known forms of carboxypeptidase A present in both Anson and Cox

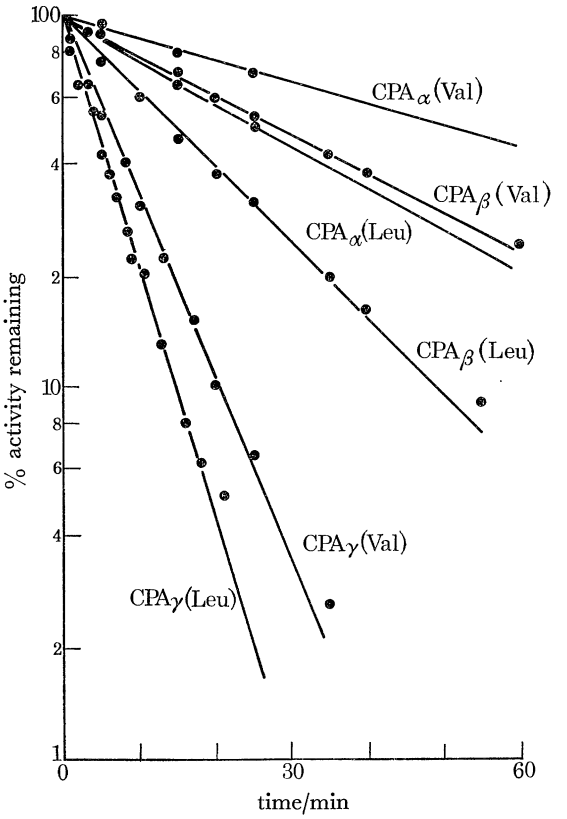

FIGURE 5. Progress curves of heat inactivation of several carboxypeptidases A at 50.0 °C, pH 7.5 in 1 of sodium chloride and 5 mmol l^{-1} of sodium phosphate. Activities were determined with 10 DL-HPLA, pH 7.5.

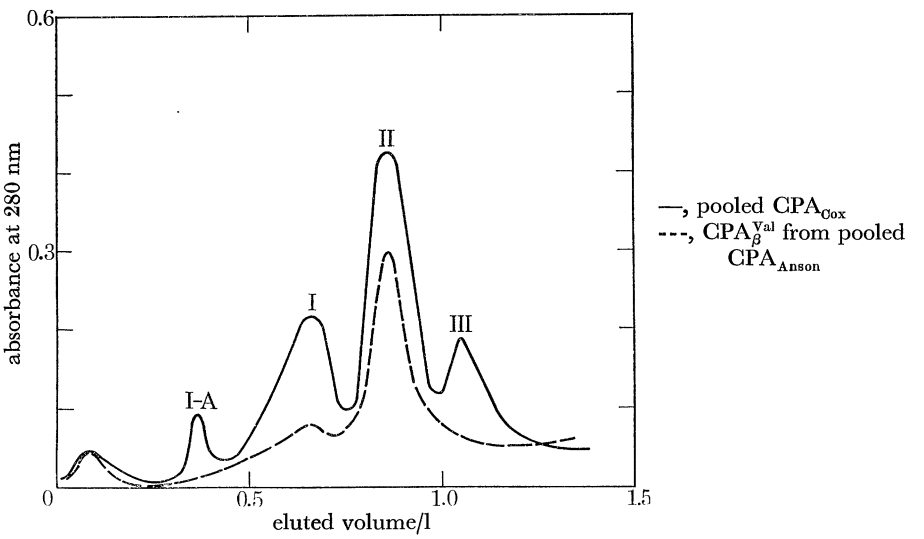

FIGURE 6. Chromatography of bovine carboxypeptidase A_{Cox} obtained from pooled pancreatic glands on a column (1.5 × 90 cm) of DE-52. Elution was carried out at 48 ml h^{-1} (4 °C) with a 9 chamber complex gradient, (chambers 1, 2, 3, and 5 containing 50 mmol l^{-1} of β-phenylpropionic acid, 4 mmol l^{-1} of lithium chloride 50 mmol l^{-1} of Tris pH 7.5, and chambers 4, 6, 7, 8, and 9 containing 122 mmol l^{-1} of β-phenylpropionic acid, 40 mmol l^{-1} of LiCl, 50 mmol l^{-1} of Tris, pH 7.5).

preparations and for the isolation of each in an apparently pure state. Carboxypeptidase A_β, a contaminant of both Anson and Cox preparations, had not been previously obtained as a pure protein.

On the basis of the present identification of the various fractions, preparations of carboxypeptidase A obtained from a single animal should show different elution profiles, depending on the presence or absence of one or the other of the two allotypes. This has in fact been observed. In order to carry out such investigations, pancreatic secretions were collected from live animals in which pancreatic fistulas were introduced by surgical procedures. This approach was first described by Keller *et al.* (1958*a*, *b*) and more recently has been perfected in collaboration with veterinary surgeons at Washington State University in Pullman. The selection of individual animals was somewhat arbitrary, though previous work had already demonstrated that the distribution of the leucine and valine variants among individual pancreas glands appears to follow the predictions of Mendelian genetics (Walsh *et al.* 1966). The chromatographic pattern of crystalline carboxypeptidase A_{Cox} obtained from the pancreatic juice of individual animals is shown in figure 7. The solid line represents the elution pattern of material shown by chemical

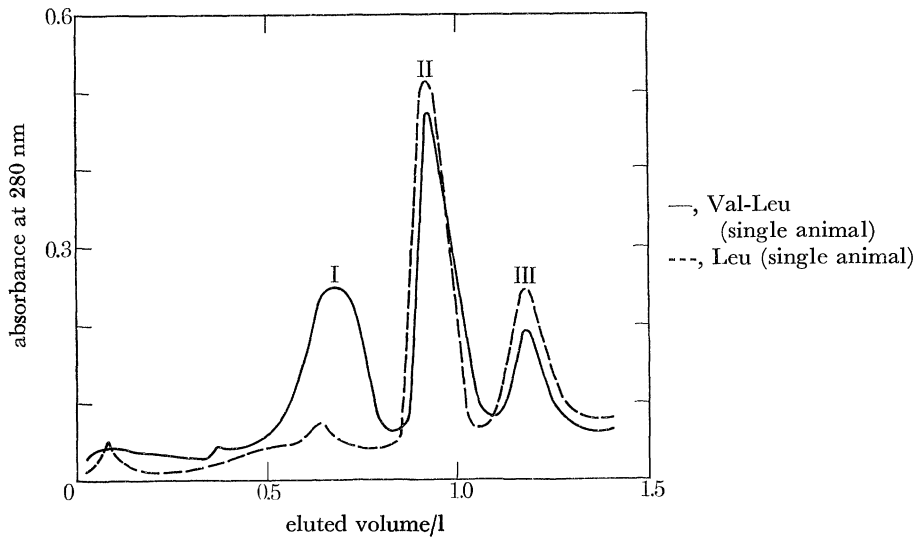

FIGURE 7. Chromatography of bovine carboxypeptidase A_{Cox} obtained from the pancreatic juice of two individual animals. The conditions are identical with those described in figure 6.

analysis to contain both allotypic variants. Peaks I and II corresponding to the valine and leucine forms of the α enzyme were crystallized and were found by activity to occur in a 50:50 ratio, a finding which supports the proposition that both genes, each coding for one allotypic variant, are fully expressed. The broken line represents the elution pattern of material from a different cow shown by chemical analysis to contain only the leucine variant. As expected, peak I was absent and only the α and β forms of the leucine variant were found.

SEQUENCE ANALYSIS

The overall strategy of sequence determination was dictated by the large size of carboxypeptidase A_γ which contains 300 amino acid residues. Since the protein contains three methionyl residues, it was considered advantageous to effect a limited cleavage of the molecule

with cyanogen bromide, followed by sequence analysis of the isolated pieces. This approach has proven successful in that three of the four fragments have been completely structured (Bargetzi, Thompson, Sampath Kumar, Walsh & Neurath 1964; Sampath Kumar *et al.* 1964*a*; Neurath *et al.* 1968), and detailed information about the fourth fragment is now available. Crystallized carboxypeptidase A_{Anson}, purchased from Worthington, was used as a starting material. Despite its heterogeneity, its availability in large quantities was considered sufficiently important to overcome the technical problems introduced by using the mixture. In addition, analysis of this material was ultimately of great value in locating the linked replacements, characterizing the allotypic forms of carboxypeptidase A. Cyanogen bromide cleavage was performed in 70% formic acid, and the fragments produced were recovered after lyophilization. Separation of the fragments was carried out using Sephadex G-75 equilibrated in 0.1 mol l^{-1} of propionic acid as shown in figure 8. The small fragments derived from the N- and C-terminal region of the

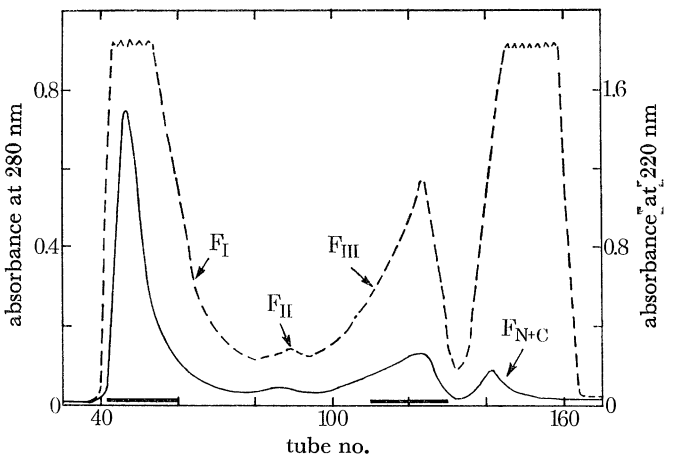

FIGURE 8. Elution pattern of the separation of the cyanogen bromide fragments of carboxypeptidase A_{Anson} on a column (2.5 × 117 cm) of Sephadex G-75 equilibrated with 0.1 mol l^{-1} of propionic acid. The column was developed at 30 ml h^{-1} and monitored at 280 nm (solid line) and 220 nm (dashed line). Fraction volume = 4.0 ml.

polypeptide (F_N and F_C) were eluted as mixture in the final peak. Purification, characterization and complete amino acid sequence analysis for both these fragments have been previously reported and will not be detailed here. In addition to this pool, two major and one minor peak were also found and were designated as F_I, F_{III} and F_{II}, respectively.

Sedimentation equilibrium analysis of the first pool, F_I, revealed that it possessed a molecular weight of approximately 160000. In view of the fact that carboxypeptidase A_γ has a molecular weight of about 34000, it seemed probable that the material in this pool was an aggregate of a lower molecular weight species. Sedimentation analyses in 6 mol l^{-1} of guanidine hydrochloride substantiated this hypothesis. Similar analysis of F_{II} and F_{III} gave molecular weight values of 21000 and 9000 respectively. Rechromatography of pool F_I recovered after lyophilization indicated that there was an equilibrium relation between F_I and F_{II} and that F_{II} represented the monomeric species. The sum of the molecular weight values of F_{II}, F_{III}, F_N and F_C, accounts very satisfactorily for the molecular weight of carboxypeptidase A_γ.

Amino acid analyses were performed on each fragment after acid hydrolysis for 24 h under the usual conditions. Since performic acid oxidation of each fragment established that all of

the residues of half cystine were found in F_I, reduction and alkylation with iodoacetate were carried out on this fragment prior to hydrolysis. The results of these analyses are shown in table 5. The composition of F_N and F_C are those previously reported by Sampath Kumar *et al.* (1964*a*) and Bargetzi *et al.* (1964). In each case, the data obtained have been averaged and rounded off to integral numbers. The four amino acids, glutamic acid, alanine, valine and isoleucine, occur in F_I in fractional quantities due to the presence of allotypic replacements in this fragment. In addition, the values of glutamic acid and isoleucine in F_I are uncertain, in the case of glutamic acid because of the presence of homoserine produced by the cyanogen bromide reaction, while in the case of isoleucine, the presence of several Ile-Ile and Ile-Val sequences make exact extrapolation difficult. As can be seen in table 5, the agreement between the sum of the amino acid content of the four cyanogen bromide pieces and the composition of carboxypeptidase A_γ obtained from acid hydrolysates (Bargetzi, Sampath Kumar, Cox, Walsh & Neurath 1963) is satisfactory.

TABLE 5. AMINO ACID COMPOSITION OF THE CNBr FRAGMENTS OF BOVINE CARBOXYPEPTIDASE A_γ

amino acid	F_N	F_C	F_{III}	F_I	total fragments	CPA_γ
Lys	0	0	4	11	15	15
His	1	1	2	4	8	8
Arg	0	0	4	6	10	10
Asp	3	2	6	17	28	26–27
Thr	2	1	5	16	24	23–24
Ser	0	0	7	24	31	30–31
Glu	1	1	8	15.5–16.5†‡	25.5–26.5	24–25
Pro	0	0	4	6	10	10
Gly	0	0	7	16	23	22.5
Ala	1	0	5	13.5†	19.5	19
Val	0	0.5	4	11.5†	16	16
½Cys	0	0	0	2	2	2
Met	1	0	1	1	3	3
Ile	1	0	7	11.5–12.5†§	19.5–20.5	20
Leu	1	0.5	7	15	23.5	23
Tyr	3	0	3	13	19	19
Phe	1	0	4	10	15	15
Trp	0	0	3	5‖	8	8
total	15	6	81	198–200	300–302	293–298

† The fractional value reported represents an average value for an equal mixture of carboxypeptidase A_γ^{Val} and A_γ^{Leu}.
‡ Contains homoserine.
§ Calculated by extrapolation.
‖ Calculated from alkaline spectra measurements in 6 mol l^{-1} of guanidine hydrochloride.

The correct alinement of the four cyanogen bromide fragments lacks rigorous chemical proof at this time. The assignment of F_C to the C-terminus of the polypeptide can be made on the basis of the lack of homoserine or its lactone in this peptide. Fragment F_N has been judged to be derived from the N-terminus through end group data and the length of the three peptides from A_α, A_β and A_γ. The internal arrangement of F_{III} and F_I is less secure. Limited overlap data suggest that F_N is linked to F_{III}, thus producing the order F_N-F_{III}-F_I-F_C as shown in figure 9. That this is the correct order is given substantial added credence by the X-ray diffraction data of Lipscomb *et al.* (1968) which also indicate this sequence.

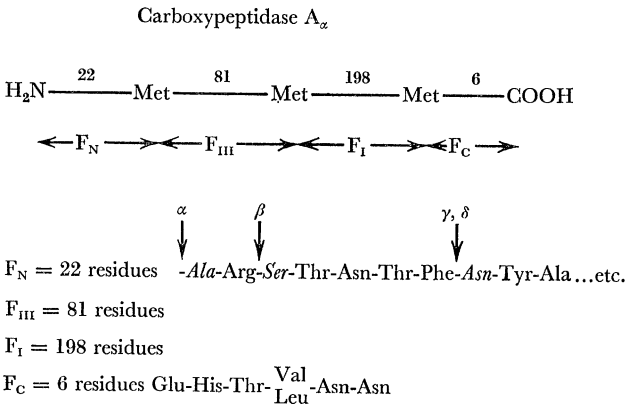

FIGURE 9. Schematic representation of the linear arrangement of the cyanogen bromide fragments of carboxypeptidase A_α.

Detailed sequence analysis of fragment F_{III} was carried out by isolation of tryptic and chymotryptic peptides. A soluble portion from each digest was fractionated on a column of Dowex 50X8. The isolated fractions were subjected to high voltage electrophoresis and those pools found to be impure were further fractionated on Dowex 1X2 and Dowex 50X2. Sequence analysis was carried out by the subtractive Edman procedure and by hydrolysis with leucine-aminopeptidase and carboxypeptidases A and B. Amide assignments were made on the basis of high voltage electrophoresis at pH 6.5, and from enzymic digests with the exopeptidases. On the basis of these data, the sequence of fragment F_{III}, as shown in figure 10 has been deduced. The position of the principle tryptic cleavages are indicated by solid circles. In addition, three minor cleavages marked with broken circles were also observed. Interestingly, each of these

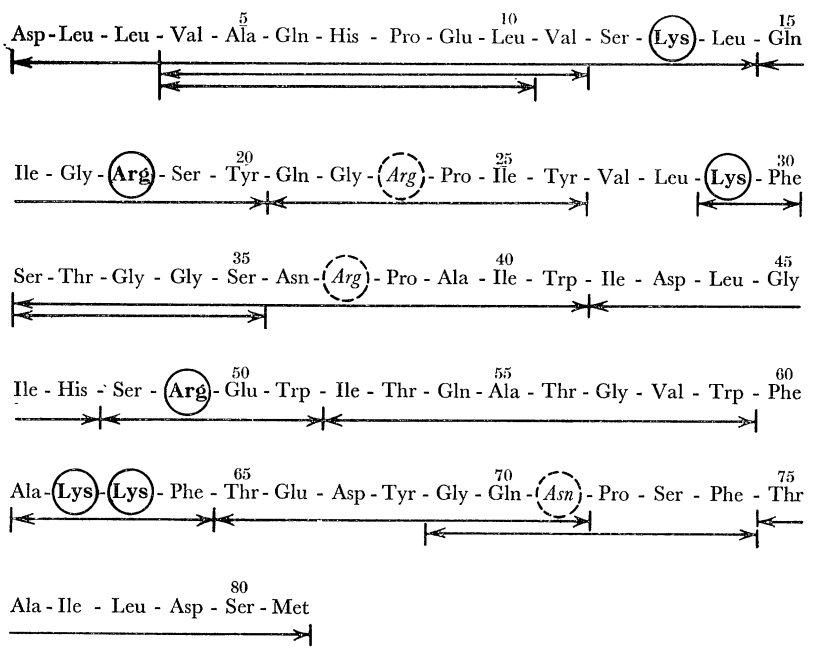

FIGURE 10. Proposed amino acid sequence of cyanogen bromide fragment F_{III} of carboxypeptidase A. The principle sites of cleavage by trypsin are indicated by solid circles and the minor sites by broken circles. The chymotryptic peptides isolated are underlined with double headed arrows.

latter hydrolyses involves a proline bond, a site of hydrolysis traditionally considered inert to attack by trypsin. The peptides obtained from the chymotryptic hydrolysates are shown by the double headed arrows. The specificity exhibited is much along the lines expected of this enzyme. It is noteworthy, however, that chymotrypsin also hydrolyses the asparagine-proline bond at position 71–72 that was cleaved by trypsin. However, chymotryptic hydrolysis of the larger tryptic fragments (residues 63–81) did not occur at this site, suggesting that the sensitivity of this bond to attack is at least in part a structural feature of intact fragment F_{III}. Assignment of the overlapping structure at residues 29–30 on the basis of the tryptic and chymotryptic peptides alone can only be made through a deductive process of elimination. However, recent data derived from thermolysin peptides of this fragment have eliminated this uncertainty by supplying a legitimate overlapping peptide. Thus with the exception of rigorous amide assignments to the glutamic and aspartic acid residues, 66 and 67, the chemical sequence of F_{III} can be taken as established.

Sequence analysis of the larger fragment F_I is currently in progress in our laboratory. Isolation and characterization of the soluble peptides from digests by trypsin, chymotrypsin and thermolysin are now complete. These data have made it possible to structure large fragments of the F_I piece, particularly in the area of peptides already reported. The overlapping peptides necessary to aline these structured fragments are currently being sought.

Linked replacement

Several lines of evidence suggest that besides the leucine-valine replacement in the carboxyl terminal region, additional allotypic substitutions might exist in carboxypeptidase. First, separation effected on the ion exchange medium was difficult to rationalize in terms of a single replacement in the C-terminal region, although the complex nature of the chromatographic system did not preclude this possibility. Secondly, amino acid analysis of the purified enzymes showed systematic variations, particularly with regard to the alanine content. Thirdly, sequence analysis of the largest cyanogen bromide fragment, F_I, indicated the presence of peptides of sufficiently similar structure to suppose that they might be derived from the same portion of the molecule. Therefore, samples of chromatographically purified carboxypeptidase A_γ of the leucine and valine allotypes were subjected to a comparative analysis and the presence of an additional linked amino acid replacement was established. After cleavage by cyanogen bromide, the two major fractions, F_{III} and F_I, were isolated from each enzyme. Amino acid analysis of fragment F_{III} gave identical compositions to that already reported, and indicated that no amino acid replacements resulting from allelomorphism occur in this region of the molecule. After reduction and alkylation with iodoacetate, the S-alkylated fragment F_I from each peak was digested with trypsin and the soluble tryptic peptides of each fractionated on Dowex 50X8. Comparison of the elution profile indicated only one major difference corresponding to tryptic pool 6. After further purification on Dowex 1X2, two fractions, 1 and 2, were isolated. Fraction 2 was a pure peptide containing histidine and was identical in both preparations. However, peptides in fraction 1 were electrophoretically distinct at pH 6.5, and amino acid analysis revealed a single amino acid difference as shown in table 6. Thus whereas the peptide TP6–1 from the valine allotype of carboxypeptidase A_γ contains three residues of alanine and no glutamic acid, the corresponding peptide from the leucine allotype contains two residues of alanine and one residue of glutamic acid. Electrophoresis indicated that the glutamic acid

TABLE 6. COMPOSITION OF TP-6 FROM F_1 OF CARBOXYPEPTIDASE A_γ^{Val} AND A_γ^{Leu}

	pool	
amino acid	A_γ^{Val}-6-I	A_γ^{Leu}-6-I
Ser	0.97	0.96
Glu	—	**1.00**
Ala	**3.04**	2.11
Val	1.01	1.00
Leu	1.00	1.00
Lys	0.98	1.01
total	7	7

TABLE 7. AMINO ACID REPLACEMENTS PROBABLY ARISING FROM ALLELOMORPHISM

Ser Ala Val	ALA	Ala Leu Lys
Ser Ala Val	GLU	Ala Leu Lys
Glu His Thr	VAL	Asn Asn
Glu His Thr	LEU	Asn Asn

residue was in the acid form. Sequence analyses of these peptides gave the results shown in the top half of table 7. Each peptide was identical in structure except at the fourth position, where the peptide from the valine variant contained alanine and that from the leucine variant contained glutamic acid. Absolutely no trace of the glutamic acid peptide was found in the valine enzyme and vice versa. Although rigorous proof that these two peptides are not found in separate portions of the F_I fragment must await the determination of the complete amino acid sequence, sufficient data are now available essentially to rule out this possibility. In addition, the complete absence of the corresponding peptides in the isolated peaks is strong proof that these peptides are derived from the same segment of the molecule and represent an allotypic replacement linked to the valine-leucine replacement in the C-terminus.

The presence of at least one additional replacement has been indicated by amino acid analyses of fragment F_I isolated from carboxypeptidase A_γ^{Val} and carboxypeptidase A_γ^{Leu} and by isolation of the peptide from carboxypeptidase prepared from pooled glands. This replacement involves a valine-isoleucine interchange, but at present it is not known whether it is linked to the other two replacements. It is obvious that even more replacements may be located in those regions of fragment F_I that are difficult to solubilize and will only be detected by improved techniques for the isolation of peptides.

A linked replacement does provide a ready explanation for the chromatographic separations of the 'valine' and the 'leucine' allotypes because the valine type contains alanine in place of glutamic acid in the leucine type. This charge difference accounts for the consistently later elution of the α, β and γ leucine types on the anion exchange columns. Furthermore, the linked replacements provide a more cogent explanation for the differences in heat stability.

At first glance, it seems curious that a linked double replacement should have accumulated in the bovine species to a 60:40 level without the detection of a molecule modified at a single site. In fact, we know of no other example of such linked double replacements within alleles of a single species of animal. Our initial expectation upon finding two loci of mutation was that a statistical distribution of the four possible allotypic proteins would be found. Analogously, if

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

there are three replacement sites, eight possible allotypes should exist. Figure 11 illustrates, however, how each allotypic species could arise by a single mutation from one ancestral protein. Thus if the gene for the ancestral protein was selected against, the presence of only two allotypes would be accounted for. It is not obvious why such selection pressures should exist, since the variants are functionally indistinguishable and it would therefore be surprising if the ancestral protein did not have a similar function.

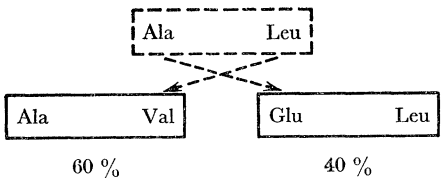

FIGURE 11. Hypothetical scheme for the evolution of the allotypic forms of carboxypeptidase A. Ala, Val, Glu and Leu refer to the alanyl, valyl, glutamyl and leucyl replacements. The percentages refer to the relative distributions of the two allotypes in the population.

HOMOLOGY

Bovine carboxypeptidase A has been the most intensively studied of the carboxypeptidases. This discussion and the following two concern themselves almost exclusively with that enzyme. Yet it may be of interest to examine the similarities and differences that exist between bovine carboxypeptidase A and B and to extend this inquiry to analogous enzymes of different phylogenetic origin. Pancreatic enzymes have been isolated in this and other laboratories from a

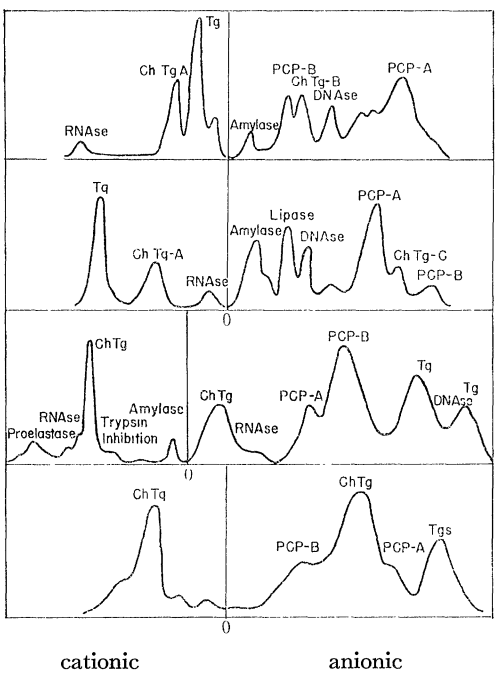

FIGURE 12. Comparative chromatographic profiles of the pancreatic extracts of four animals on DEAE- and CM-cellulose. (*a*) Bovine (Keller *et al.* 1958*a*, *b*); (*b*) porcine (Marchis-Mouren 1965); (*c*) lungfish (Reeck, Winter & Neurath, unpublished experiments); (*d*) dogfish (Prahl & Neurath 1966).

number of species ranging on the evolutionary scale from the dogfish to the cow. A composite cationic and anionic chromatographic elution profile of four such species is summarized in figure 12. The chromatographic elution profiles are quite different in kind: thus, whereas in the bovine and the porcine species, trypsinogen and chymotrysinogen represent the most basic proteins, in the dogfish and in the African lungfish, trypsinogen is among the most acidic proteins. Also, whereas bovine procarboxypeptidase A occurs as an aggregate (Brown *et al.* 1963) together with two other subunits, one of which is a precursor of a chymotrypsin-like enzyme, the dogfish, the porcine, and the lungfish zymogens occur in monomeric form. Yet many of the characteristic chemical and functional features of the carboxypeptidases have been preserved during some 200 millions years of divergent evolution that separates the dogfish from the domestic mammals.

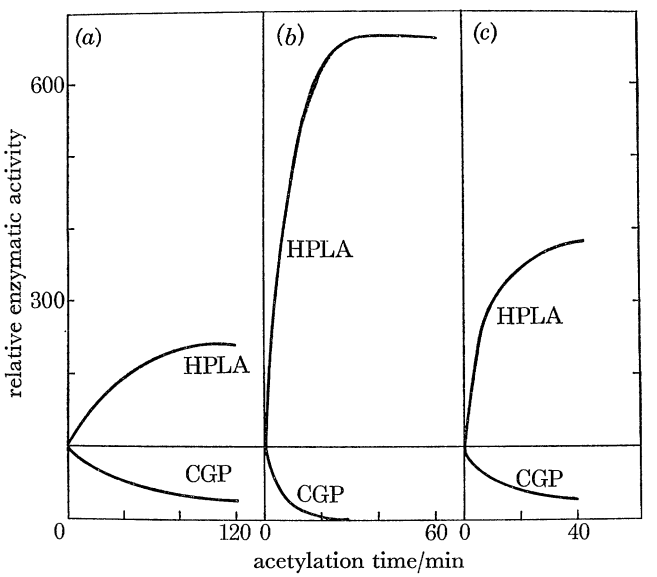

FIGURE 13. Comparison of the effects of acetylation with N-acetylimidazole on the esterase (HPLA) and peptidase (CGP) activities of (*a*) dogfish carboxypeptidase B (Stevens & Neurath 1968); (*b*) bovine carboxypeptidase A (Simpson *et al.* 1963) and (*c*) dogfish carboxypeptidase A (Lacko & Neurath 1967).

A striking similarity in the amino acid composition of carboxypeptidases A and B of bovine, porcine and dogfish origin is illustrated in table 8 and suggests considerable conservation of structure. Furthermore, these four enzymes have common functional features. First, they all appear to be zinc metalloenzymes and the metal is an absolute requirement for function. Secondly, all six enzymes are exopeptidases requiring a free α carboxyl group in the substrate. Thirdly, four of the enzymes that have been tested, namely bovine A and B, and dogfish A and B, exhibit a common sensitivity to reagents which modify tyrosyl residues. As first demonstrated by Simpson, Riordan & Vallee (1963), *N*-acetylimidazole serves to increase the apparent esterase activity and to abolish peptidase activity. As shown in figure 13, the same qualitative effect is produced in dogfish carboxypeptidases A and B, and in each case the effects of acetylation are prevented by prior addition of β-phenylpropionate and reversed by the addition of hydroxylamine. Qualitatively similar effects can be demonstrated by either iodination (Simpson & Vallee 1966) or nitration (Riordan, Sokolovsky & Vallee 1967) of tyrosyl residues. It has been amply demonstrated in the case of bovine carboxypeptidase A that tyrosyl residues are the sites of modification. Current work of Dr R. L. Stevens in our laboratory indicates that a parallel situation holds for dogfish carboxypeptidase B.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

TABLE 8. AMINO ACID COMPOSITION OF CARBOXYPEPTIDASES

enzyme	Lys Arg	Tyr Phe Trp	Ser Thr	Leu Ile Val	Gly Ala	Pro	His	Met Cys
bovine A	26	43	61	59	43	10	8	5
B	30	44	52	50	43	12	7	13
porcine A	24	41	55	52	48	14	9	5
B	28	41	48	51	48	13	6	13
dogfish A	28	43	47	55	50	17	7	13
B	29	39	50	54	44	14	4	13

It is curious in these cases that the esterase activity is increased by tyrosyl modification whereas the peptidase activity is inhibited. The data shown in figure 14 reveal that the effects of either acetylation or nitration of dogfish carboxypeptidase B cannot so easily be categorized as an activation or an inhibition of the esterase activity toward hippuryl-phenyllactate. For example, at a substrate concentration of 30 mmol l^{-1}, acetylation activates and nitration shows no effect.

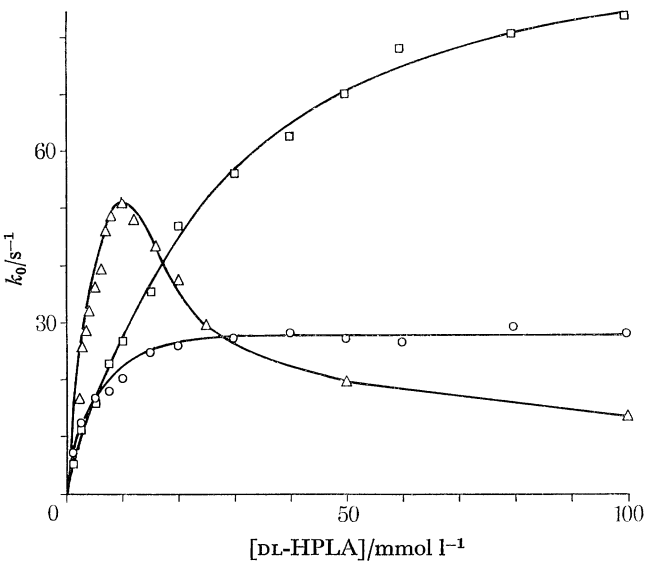

FIGURE 14. Effect of nitration and acetylation of dogfish carboxypeptidase B on the dependence of its esterase activity towards substrate concentration. △-△, native CPB; ○-○, nitrated; □-□, CPB acetylated with *N*-acetylimidazole (Stevens & Neurath 1968).

In contrast, at a concentration of 10 mmol l^{-1}, both treatments seem to inhibit the enzyme. At a substrate concentration of 100 mmol l^{-1} both treatments seem to activate the enzyme. These effects are expressions of the relation of K_m to the range of substrate concentrations in which excess substrate inhibition is manifested (Davies, Riordan, Auld & Vallee 1968). A similar pattern of effects holds for bovine carboxypeptidase A.

These data suggest the hypothesis that both the metal and one or more tyrosyl residues may play a common functional role among carboxypeptidases as a class. The preservation of sensitive tyrosyl residues in two species of enzyme which span 200 million years of evolution suggests a specific and critical role of these residues in the function of these enzymes. Attempts are currently being made to seek evidence of structural homology to assist in pinpointing common functional determinants.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

References (Neurath *et al.*)

Allan, B. J., Keller, P. J. & Neurath, H. 1964 *Biochemistry* **3**, 40.
Anson, M. L. 1937 *J. gen. Physiol.* **20**, 663.
Bargetzi, J. P., Sampath Kumar, K. S. V., Cox, D. J., Walsh, K. A. & Neurath, H. 1963 *Biochemistry* **2**, 1468.
Bargetzi, J. P., Thompson, E. O. P., Sampath Kumar, K. S. V., Walsh, K. A. & Neurath, H. 1964 *J. biol. Chem.* **239**, 3767.
Bergmann, M. & Fruton, J. S. 1941 *Adv. Enzymol.* **1**, 63.
Bethune, J. L. 1965 *Biochemistry* **4**, 2698.
Brown, J. R., Greenshields, R. N., Yamasaki, M. & Neurath, H. 1963 *Biochemistry* **2**, 867.
Coleman, J. E. & Vallee, B. L. 1960 *J. biol. Chem.* **235**, 390.
Cox, D. J., Wintersberger, E. & Neurath, H. 1962 *Biochemistry* **1**, 1078.
Cox, D. J., Bovard, F. C., Bargetzi, J. P., Walsh, K. A. & Neurath, H. 1964 *Biochemistry* **3**, 44.
Davies, R. C., Riordan, J. F., Auld, D. S. & Vallee, B. L. 1968 *Biochemistry* **7**, 1090.
Elkins-Kaufman, E. & Neurath, H. 1948 *J. biol. Chem.* **175**, 893.
Freisheim, J. H., Walsh, K. A. & Neurath, H. 1967 *Biochemistry* **6**, 3010, 3020.
Keller, P. J., Cohen, E. & Neurath, H. 1956 *J. biol. Chem.* **223**, 457.
Keller, P. J., Cohen, E. & Neurath, H. 1958*a* *J. biol. Chem.* **230**, 905.
Keller, P. J., Cohen, E. & Neurath, H. 1958*b* *J. biol. Chem.* **233**, 344.
Lacko, A. G. & Neurath, H. 1967 *Biochem. Biophys. Res. Commun.* **26**, 273.
Lipscomb, W. N., Hartsuck, J. A., Reeke, G. N., Quiocho, F. A., Ludwig, M. L., Steitz, T. A. & Bethge, P. 1968 *Brookhaven Symp. Biol.* **21**, 24.
Ludwig, M. L., Paul, I. C., Pawley, G. S. & Lipscomb, W. N. 1963 *Proc. natn. Acad. Sci. U.S.A.* **50**, 282.
Marchis-Mouren, G. 1965 *Bull. Soc. Chim. Biol.* **47**, 2207.
Neurath, H. & Schwert, G. W. 1950 *Chem. Rev.* **46**, 69.
Neurath, H. 1960 In *The enzymes*, vol. IV (ed. P. D. Boyer, H. Lardy and K. Myrbäck), p. 11. New York: Academic Press.
Neurath, H., Bradshaw, R. A., Ericsson, L. H., Babin, D. R., Pétra, P. H. & Walsh, K. A. 1968 *Brookhaven Symp. Biol.* **21**, 1.
Pétra, P. H. & Neurath, H. 1968 *Fedn Proc.* **27**, 589.
Prahl, J. W. & Neurath, H. 1966 *Biochemistry* **5**, 4137.
Putnam, F. W. & Neurath, H. 1946 *J. biol. Chem.* **166**, 603.
Quiocho, F. A. & Richards, F. M. 1964 *Proc. natn. Acad. Sci. U.S.A.* **52**, 833.
Riordan, J. F., Sokolovsky, M. & Vallee, B. L. 1967 *Biochemistry* **6**, 358.
Roholt, O. A. & Pressman, D. 1967 *Proc. natn. Acad. Sci. U.S.A.* **58**, 280.
Sampath Kumar, K. S. V., Clegg, J. B. & Walsh, K. A. 1964*a* *Biochemistry* **3**, 1728.
Sampath Kumar, K. S. V., Walsh, K. A., Bargetzi, J. P. & Neurath, H. 1964*b* In *Aspects of protein structure* (ed. R. N. Ramachandran), p. 319. New York: Academic Press.
Simpson, R. T., Riordan, J. F. & Vallee, B. L. 1963 *Biochemistry* **2**, 616.
Simpson, R. T. & Vallee, B. L. 1966 *Biochemistry* **5**, 1760.
Smith, E. L., Brown, D. M. & Hanson, T. 1949 *J. biol. Chem.* **180**, 33.
Smith, E. L. 1951 *Adv. Enzymol.* **12**, 191.
Smith, E. L., Lumry, R. & Polglase, W. J. 1951 *J. Phys. Coll. Chem.* **55**, 125.
Smith, E. L. & Stockell, A. 1954 *J. biol. Chem.* **207**, 501.
Snoke, J. E., Schwert, G. W. & Neurath, H. 1948 *J. biol. Chem.* **175**, 7.
Sokolovsky, M. & Vallee, B. L. 1967 *Biochemistry* **6**, 700.
Stevens, R. L. & Neurath, H. 1968 *Abstr. Am. Chem. Soc. Meeting (September), Atlantic City, New Jersey.*
Vallee, B. L. & Neurath, H. 1954 *J. Am. chem. Soc.* **76**, 5006.
Vallee, B. L. 1955 *Adv. Prot. Chem.* **10**, 317.
Vallee, B. L., Rupley, J. A., Coombs, T. L. & Neurath, H. 1960 *J. biol. Chem.* **235**, 64.
Vallee, B. L., Coombs, T. L. & Hoch, F. L. 1960 *J. biol. Chem.* **235**, PC45.
Vallee, B. L. 1964 *Fedn Proc.* **23**, 8.
Vallee, B. L. & Riordan, J. F. 1968 *Brookhaven Symp. Biol.* **21**, 91.
Walsh, K. A., Sampath Kumar, K. S. V., Bargetzi, J. P. & Neurath, H. 1962 *Proc. natn. Acad. Sci. U.S.A.* **48**, 1443.
Walsh, K. A., Ericsson, L. H. & Neurath, H. 1966 *Proc. natn. Acad. Sci. U.S.A.* **56**, 1339.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES