

The Primary Structure of the Parvalbumin II of Pike (*Esox lucius*)

Charles GERDAY

Laboratoire de Biochimie Musculaire, Université de Liège,
and Laboratory of Molecular Biophysics, Department of Zoology, Oxford

(Received February 23/September 2, 1976)

The amino acid sequence of the parvalbumin II of the pike is reported. The protein has a molecular weight of 11435. It consists of a single polypeptide chain of 107 amino acid residues with an acetyl group blocking the N-terminus and an alanine residue at the C-terminus. The molecule has been enzymically cleaved by trypsin, thermolysin and by the protease of the *Staphylococcus aureus* strain V8. Chemical cleavages make use of the CNBr reaction and of the sulfoethylation method. The comparison of this amino acid sequence with that of the parvalbumin III of the pike indicates that these two homologous proteins belong respectively to two different subgroups derived from an early gene duplication of an ancestral gene at least prior to the differentiation of the Osteichthyes.

The white muscle of the Pike (*Esox lucius*) contains two parvalbumins, the components II and III, which have already been characterized in detail [1,2]. A preliminary report of the amino acid sequence of the parvalbumin III has also been published [3]. The examination of the tryptic peptide maps [2] of the two parvalbumins led to the suggestion that they have rather different primary structure, since only two peptides have the same electrochromatographic comportment.

It had also been suggested [4] that parvalbumins from different species are distributed in at least two different genetic lineages. In order to check the possibility that these different lineages could also cohabit within the same species we have determined the primary structure of the parvalbumin II (pI: 4.2) of the pike. The comparison of the amino acid sequences of the two parvalbumins of the pike will establish that indeed they belong respectively to different evolutionary trends.

MATERIALS AND METHODS

Chemicals were of analytical or sequanal grade. The protein was purified and carboxymethylated as described previously [2,5].

Abbreviation. Dansyl, 5-dimethylaminonaphthalene-1-sulphonyl.

Enzymes. Carboxypeptidase A (EC 3.4.12.2); carboxypeptidase B (EC 3.3.12.3); chymotrypsin (EC 3.4.21.1); thermolysin (EC 3.4.24.4); trypsin (EC 3.4.21.4).

Amino Acid Analyses

Amino acid analyses were made according to Benson and Patterson [6] using a Beckman 120B amino acid analyzer equipped with high-sensitivity cuvettes. Peptides were hydrolyzed by 6 M HCl for 24 h at 107 °C in evacuated sealed tubes.

Prior to the analysis of the peptides resulting from the CNBr fragmentation hydrolysates were incubated for 1 h at 107 °C in sealed tubes containing 1 ml of pH 6.5 buffer (pyridine/acetic acid/water, 250/10/2250, by vol.) in order to convert the homoserine lactone in homoserine.

Chemical Fragmentation

The CNBr cleavage was made by adding to 4.5 µmol of S-carboxymethylated protein dissolved in 10 ml of 70% formic acid two aliquots of 60 mg CNBr (Aldrich, Europe, Belgium) at 6-h intervals.

The reaction was allowed to proceed for 24 h under magnetic stirring and N₂.

The S-cyanocysteinyl derivative was made according to Jacobsen *et al.* [7]: 4 µmol of the native protein dissolved in 3 ml 6 M guanidinium chloride, 0.2 M Tris-HCl, pH 8, containing 3 µmol of dithiothreitol was treated for 15 min by 60 µmol of the half-potassium salt of 2-nitro-5-thiocyanobenzoic acid [8]. The reaction was stopped by adding 3 ml of glacial acetic acid and the derivative was desalted on a Sephadex G-25 column (3.2 × 47 cm) equilibrated in 50% acetic

acid. Cleavage at the *S*-cyanocysteinyl residue was obtained by incubating at 37 °C the lyophilized powder in 5 ml of 6 M guanidinium chloride, 0.1 M sodium borate, pH 9, for 15 h.

Enzymic Hydrolysis

Trypsin digestion was performed at 37 °C during 8 h using 3% (w/w) of enzyme (Worthington, twice crystallized) added in three aliquots at time intervals of 2 h. 13 μ mol of *S*-carboxymethylated protein dissolved in 10 ml of 0.2 M NH_4HCO_3 brought to pH 7 with formic acid was digested.

Thermolysin (Daiwa Kasei, Osaka) digestion was carried out at 37 °C during 3 h on 6.5 μ mol of *S*-carboxymethylated protein dissolved in 7 ml 0.01 M Tris-HCl, 0.01 M CaCl_2 , pH 8, using 0.5% (w/w) of enzyme added in two aliquots at time 0 and 2 h.

The digestion with staphylococcal protease (courtesy of Dr G. Drapeau, Département de Microbiologie, Université de Montréal) was made on 6.5 μ mol of *S*-carboxymethylated protein dissolved in 5 ml 0.05 M KH_2PO_4 , pH 7.8. 3% (w/w) of the enzyme was added at time 0 and another aliquot of 1% after 24 h. The incubation was allowed to proceed for 32 h at 37 °C.

Methods of Isolation of Peptides

The CNBr peptides were isolated by gel filtration on a Sephadex G-50 column (2 $\times$ 135 cm) in 0.005 M NH_4HCO_3 .

The isolation of the fragments resulting from the cleavage of the *S*-cyanoethylated protein was achieved by chromatography on a Sephadex G-50 column (2.5 $\times$ 95 cm) in 0.005 M NH_4HCO_3 .

The trypsin and the staphylococcal protease digests were centrifuged off and the precipitates were washed twice with 1 ml of 0.005 M NH_4HCO_3 . The supernatants of these two digestions along with the thermolysin digest were chromatographed first on Biogel P4 columns in 0.005 M NH_4HCO_3 . All the fractions were analyzed by thin-layer chromatography (butanol/acetic acid/water, 4/1/5) or paper electrophoresis at pH 3.5 (pyridine/acetic acid/water, 10/100/890) and their amino acid composition was determined. The peptides present in the fractions eluted in the void volume were separated at pH 7.3 by ion-exchange chromatography on DEAE-Sephadex A-25 and DEAE-cellulose (Whatman DE-32) using different gradients of NaCl. The other fractions corresponding theoretically to peptides of smaller size were resolved by ion-exchange chromatography on a cation exchanger (Technicon Chromobeads P). Elution was performed by applying gradients of pH and ionic strength generated by mixing different buffers made of pyridine, acetic acid and water. The peptides in

the effluent were detected manually by using the ninhydrin reaction prior to and in some cases after alkaline hydrolysis of aliquots of the collected fractions. All the peptides isolated from the ion-exchange chromatographies using pyridine/acetic acid/water buffers were made clean by chromatography on a Biogel P2 column (1.5 $\times$ 90 cm) equilibrated in 0.005 M NH_4HCO_3 . Ultimate purification by high-voltage electrophoresis on Whatman no. 1 and 3MM paper or on cellulose thin-layer was sometimes necessary.

Dansyl-Edman Sequencing Method

The dansyl-Edman technique described by Gray [9] was followed throughout. The HCl hydrolysis of the dansyl derivatives was usually performed in evacuated sealed ampoules at 110 °C during 4 h, the exception being peptides containing isoleucine for which overnight hydrolysis was used.

The phenylisothiocyanate reagent (Aldrich, Europe, Belgium) was redistilled and kept at -32 °C in nitrogen-sealed ampoules containing 1 ml of the reagent. 1-g samples of trifluoroacetic acid (sequanal grade) were purchased from Pierce Chemical Co. (Rockford, Illinois, U.S.A.). Dansyl chloride was obtained from BDH Chemicals Ltd (Poole, England). All the other reagents were Uvasol grade (Merck, Darmstadt, Germany). The dansylamino acids were identified by chromatography [10] on polyamide sheets (7.5 $\times$ 7.5 cm) (Cheng Chin Trading, Taipei, Taiwan). An additional solvent ethyl/acetate/ethanol/acetic acid (20/1/1 by vol.) was sometimes used in the same direction as solvent II in order to separate clearly the 1-dimethylaminonaphthalene-5-sulfonamide from dansyl-alanine).

Determination of C-Terminal Residues

In some cases the C-terminal amino acid sequences of the peptides were investigated by digestion with carboxypeptidase A (Worthington COA DFP) and carboxypeptidase B (Worthington COB DFP) following the procedure of Dopheide *et al.* [11]. Aliquots (50–100 μ l) were taken at different time intervals. They were inactivated by addition of 0.5 ml citrate-HCl buffer, pH 2.2, and submitted to amino acid analysis.

N-Terminal Acetyl Group

The presence of an acetyl group blocking the N-terminus of the parvalbumin II of the pike has already been identified in our laboratory by gas chromatography analysis of distillates of the protein hydrolysate with *p*-toluenesulfonic acid [2]. The identification of the liberated fatty acid was performed by gas chromatography analysis on Porapak N [12].

Nomenclature of Peptides

The peptides obtained after cyanogen bromide cleavage are designated as CNBr I to CNBr III, in the order of elution on Sephadex G-50. The peptide isolated from the *S*-cyanocysteinyl cleavage is designated as SCN II.

Peptides obtained from enzymic digestion with trypsin, thermolysin and staphylococcal protease are identified respectively by the symbols T, Th, and SP followed by a number corresponding to their order of isolation.

Amide Assignment

The amide groups were assigned on the basis of the relative electrophoretic mobilities of the peptides [13] to aspartic acid at pH 6.5. This was performed on Whatman no. 1 paper or on cellulose thin layer.

RESULTS

Amino Acid Composition

The amino acid composition of the parvalbumin II of the pike has already been published [2]. The results are shown in Table 1.

CNBr Peptides

The separation of the CNBr peptides is shown in Fig. 1. The characteristics of the isolated fractions are depicted in Table 2. Fraction IV corresponds

to by-products of the CNBr reaction. CNBr I is the long N-terminal peptide carrying the acetyl group. This peptide was not characterized further. The amino acid sequences of CNBr II and CNBr III, determined by the dansyl-Edman method, are as follows:

Table 1. Amino acid composition of the parvalbumin II of the pike as determined by Bhushana et al. [2]

Amino acid	Number of residues	
	calculated	deduced from sequence data
Lysine	12.98	12
Histidine	1.02	1
Arginine	0.99	1
Aspartic acid	15.08	15
Asparagine		
Threonine	2.27	2
Serine	8.20	8
Glutamic acid	8.24	8
Glutamine		
Proline	1.03	1
Glycine	7.44	7
Alanine	20.14	20
Cysteine	1.00	1
Valine	4.97	5
Methionine	1.87	2
Isoleucine	3.90	4
Leucine	9.19	9
Tyrosine	0.96	1
Phenylalanine	10.05	10
Tryptophan	0	0
Total		107

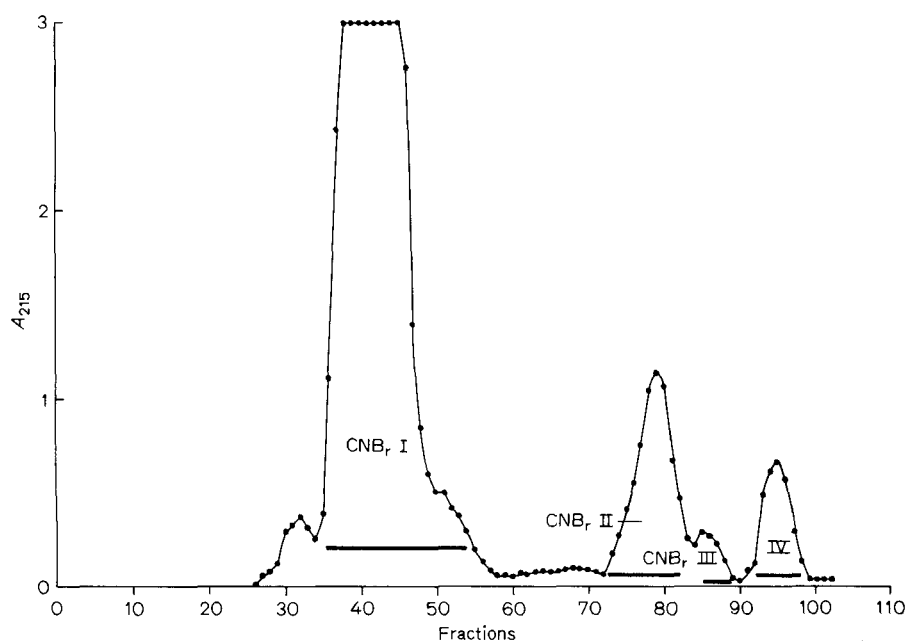


Fig. 1. Chromatography in 0.005 M NH_4HCO_3 on a Sephadex G-50 column (135×2 cm) of the peptide mixture resulting from the CNBr cleavage of the *S*-carboxymethylated parvalbumin II of the pike. Fractions of 4 ml

Table 2. Amino acid composition of the CNBr peptides isolated as in Fig. 1

Results for a 24-h hydrolysis

Amino acid	CNBr I	CNBr II	CNBr III
Lysine	10.8 (11)		0.8 (1)
Histidine	0.9 (1)		
Arginine	1.0 (1)		
Carboxymethylcysteine	0.8 (1)		
Aspartic acid	13.1 (13)	1.0 (1)	
Asparagine			
Threonine ^a	1.9 (2)		
Serine ^a	7.0 (7)		
Homoserine	1.2 (1)	1.1 (1)	
Glutamic acid	6.8 (7)	1.0 (1)	
Glutamine			
Proline	1.0 (1)		
Glycine	7.0 (7)	1.0 (1)	
Alanine	14.0 (14)	1.8 (2)	1.3 (1)
Valine	2.2 (2)	0.8 (1)	
Methionine			
Isoleucine ^a	1.0 (1)	0.8 (1)	0.9 (1)
Leucine	9.3 (9)		
Tyrosine	1.1 (1)		
Phenylalanine	10.4 (10)	1.2 (1)	
Number of residues	89	9	3
N-terminus	blocked	Ile	Ile
Mobility relative to Asp at pH 6.5	—	+ 0.56	— 0.50
Charge at pH 6.5	—	— 2	+ 1

^a Uncorrected data.

CNBr II: Ile-Gly-Val-Asp-Glu-Phe-Ala-Ala-Hse

CNBr III: Ile-Lys-Ala .

The alignment of the three CNBr peptides can therefore be rapidly deduced:

acetyl-CNBr I—CNBr II—CNBr III—OH .

S-CyanocysteinyI Cleavage

The comparison of the titrations by the method of Hooton [14] of the free —SH groups respectively in the native protein and in the derivative indicates that more than 95% of the cysteine residues have been S-cyanoethylated. The profile of the chromatography resulting from the cleavage of the derivative at pH 9 is shown in Fig. 2. Peak I has a composition very similar to that of the native protein, peak III corresponds to the salts, whereas peak II is a peptide of sixteen residues (SCN II) having the following amino acid composition: lysine 1.0, aspartic acid and asparagine 2.0, serine 1.1, glycine 1.1, alanine 6.7, valine 1.0, leucine 1.9, phenylalanine 1.0. It is

obtained with 15% yield and its terminal amino group is blocked. The digestion of 0.5 μmol of this peptide, dissolved in 0.5 ml 0.1 M Na₂HPO₄, pH 7.8, by 2% (w/w) carboxypeptidase A gives the following results (analyses have been made on aliquots of 50 μl of the solution):

Incubation time	Alanine	Leucine	Valine
5 min	0.8		
15 min	1.5	0.4	
30 min	2.4	1.0	
2 h	3.3	1.0	
8 h	4.1	1.0	
23 h	4.5	1.0	0.3

The data are consistent with the following partial amino acid sequence:

SCN II: acetyl-(Lys,Asx 2, Ser, Gly, Ala 2,
Leu,Phe)-Val-Ala-Ala-Ala-Leu-Ala-Ala-OH .

It can be conclude that cysteine occupies position 17 in the protein.

Tryptic Peptides

The insoluble material resulting from the digestion with trypsin was submitted to amino acid analysis; it is the pure peptide T 27 (Table 3). The profile of the chromatography of the supernatant on Biogel P4 is shown in Fig. 3. Six fractions were collected; two of them, namely fractions III and IV, are made of pure peptides T6 and T2 (Table 3). Fractions I, II, and V were further fractionated by ion-exchange chromatography (Fig. 4A,B,C), whereas fraction VI was resolved into three main ninhydrin-positive bands (T3, T5, and T7) by paper electrophoresis at pH 3.5 (Fig. 4D). The characteristics of the isolated peptides are indicated in Table 3. Details concerning the isolation of these peptides can be obtained from the legend to Fig. 4. The peptide T 13 was obtained mainly from Biogel P4 fraction II (18.6%, Fig. 4B) but also from the fraction I (2.1%, Fig. 4A).

Amino acid sequence data relative to the tryptic peptides mentioned in Table 3 are summarized in Fig. 9. The amino acid sequence of the acetylated peptide T23 was carried out as follows: 1.5 μmol dissolved in 0.5 ml Na₂HPO₄ (0.1 M, pH 7.8) were incubated with 2% (w/w) of chymotrypsin (Worthington). The mixture was loaded on a Dowex 50-X₄H⁺ column (0.9 × 5 cm) eluted first with 20 ml of water and then with 20 ml of 4 M NH₄OH. The two collected fractions T23 (H₂O) and T23(NH₄⁺) were lyophilized and submitted to amino acid analysis:

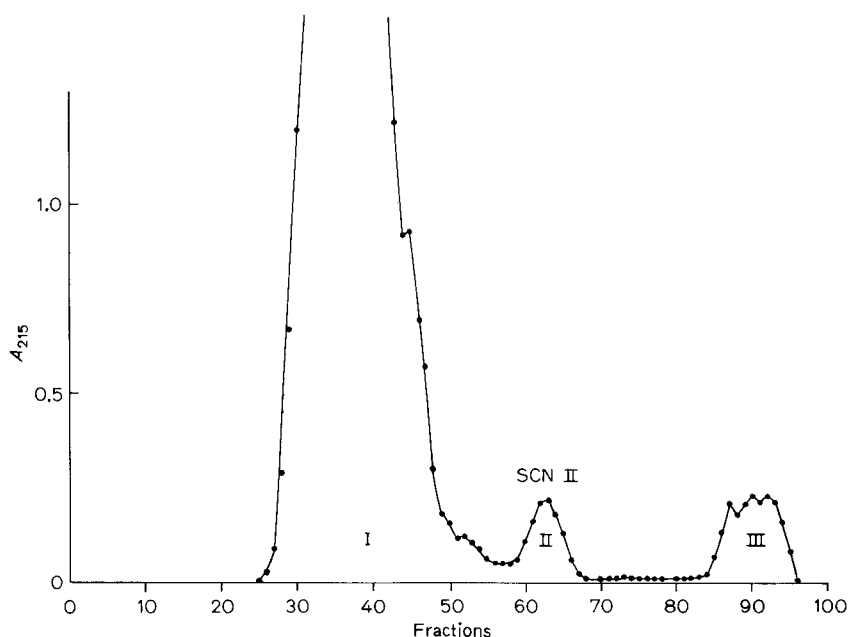


Fig.2. Chromatography in 0.005 M NH_4HCO_3 on a Sephadex G-50 column (95×2.5 cm) of the mixture resulting from the cleavage in borate buffer of the S-cyanoethylated parvalbumin II of the pike. Fractions of 6.2 ml

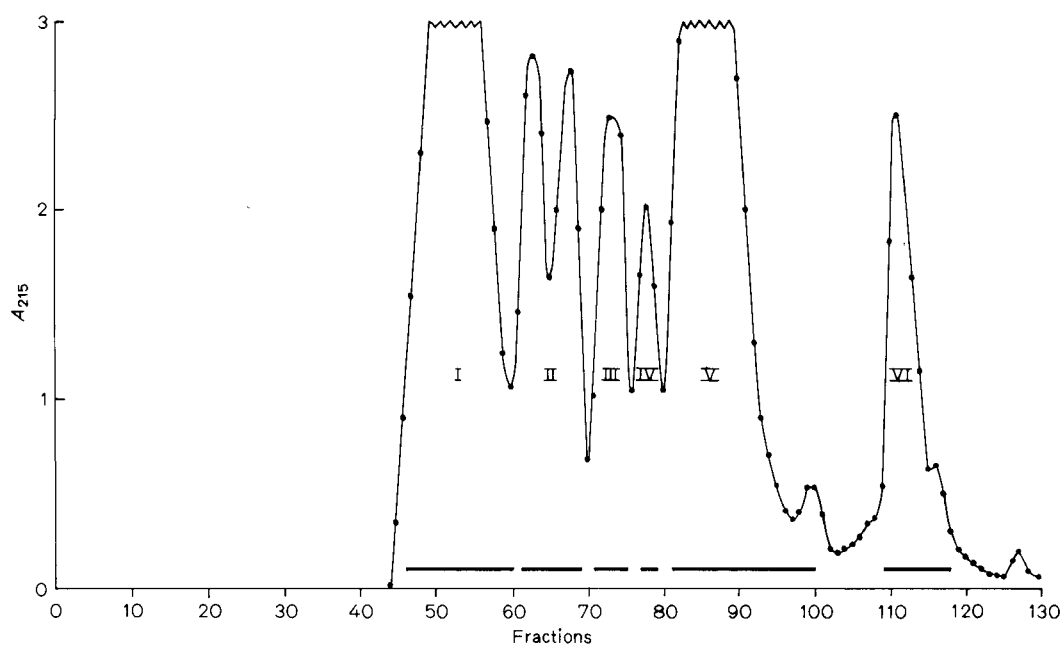


Fig.3. Chromatography in 0.005 M NH_4HCO_3 on a Biogel P4 column (172×2.8 cm) of the tryptic digest of the S-carboxymethylated parvalbumin II of the pike. Fractions of 7.3 ml

T23 (H_2O): serine 1.1; phenylalanine 0.9
 T23 (NH_4^+): lysine 1.0; serine 0.4; glycine 1.0; alanine 1.0; leucine 0.9; phenylalanine 0.4
 T23 (H_2O) has a blocked N-terminus; the digestion with carboxypeptidase A gives the following result at 2 h: serine 0.05; phenylalanine 1.0

T23 (NH_4^+) is obviously a mixture of the uncleaved peptide and of a tetrapeptide liberated by the chymotrypsin hydrolysis.

It has been submitted to the dansyl-Edman degradation (Fig.9). Among the tryptic peptides listed in Table 3 some of them derived undoubtedly from un-

Table 3. Amino acid compositions and other characteristics of the tryptic peptides isolated by different methods from the six fractions obtained by chromatography of the supernatant of the digestion on Biogel P4 (Fig. 3) and from the precipitate (↓) isolated by centrifugation. The peptides are disposed from left to right according to their position from the N-terminus of the polypeptide chain. Yields are expressed in mol/mol protein. p.e., purification by paper electrophoresis at pH 3.5

Amino acid	T23	T13	T12 ^a	T3	T24	T19	T2	T16	T5	T6	T9	T25	T7	T27	T15	T17	T11	T26	T20	T21	T18
Lysine	1.1		1.2	1.0		1.0	1.4	0.9	1.0	1.0	1.0				1.1	0.9		1.8	1.2	1.0	0.9
Histidine				0.9									0.9	1.0							
Arginine																					
Carboxymethylcysteine			1.0																		
Aspartic acid		2.0	3.2					2.1		2.2	1.1	1.0		1.2	1.1	2.1	2.8	4.6			
Asparagine															1.6						
Threonine																					
Serine	0.9		1.9				0.9	0.9			1.0	1.9		1.6							
Glutamic acid																					
Glutamine					0.9	1.1				1.1	2.7	1.0		1.1	0.9	1.0	1.1				
Proline												0.9		1.0							
Glycine	1.0						0.8				1.1					1.0	2.8	3.6			
Alanine	1.0	4.1	7.7		0.9	1.0				1.0			1.0	1.2	2.2	1.9	4.0			1.0	
Valine		0.8	1.1		0.8		0.8	1.1		1.0							0.9	1.0			
Methionine																	0.7	1.5	0.7		
Isoleucine										0.8	0.8						0.9	1.7	1.1	1.0	1.0
Leucine	1.1	1.0	0.9				1.1	1.1			0.9	1.8		1.9	0.9	1.0	0.9				
Tyrosine									0.8												
Phenylalanine	1.0		0.9		2.1	2.1				0.9	1.0	2.0		2.0		1.1	1.0	1.8			
Number of residues	6	8	17	2	3	5	6	6	1	9	10	9	2	11	8	8	11	24	3	2	3
Mobility relative to Asp at pH 6.5	—	0.61	—	—	—	0	—	0.36	—	0.27	0.62	0	—	—	0.33	0.35	0.87	—	—	—	—
Charge at pH 6.5	—	—2	—	—	—	0	—	—1	—	—1	—3	0	—	—	—1	—1	—4	—	—	—	—
Biogel P4 (Fig. 3) fractions	V	I,II	I	VI	V	V	IV	II	VI	III	I	V	VI	↓	II	II	I	I	V	V	V
Purification method (cf. Fig. 4A–D)	C	A,B	A	D	C	C	—	B	D	—	A	C	D	—	B	B	A	A	C	C	C
								(p.e.)								(p.e.)					
Final yield(mol/mol)	0.411	0.207	0.474	0.118	0.054	0.588	0.80	0.288	0.077	0.776	0.667	0.015	0.012	0.67	0.544	0.194	0.431	0.055	0.091	0.215	0.046

^a Results obtained after 72-h hydrolysis.

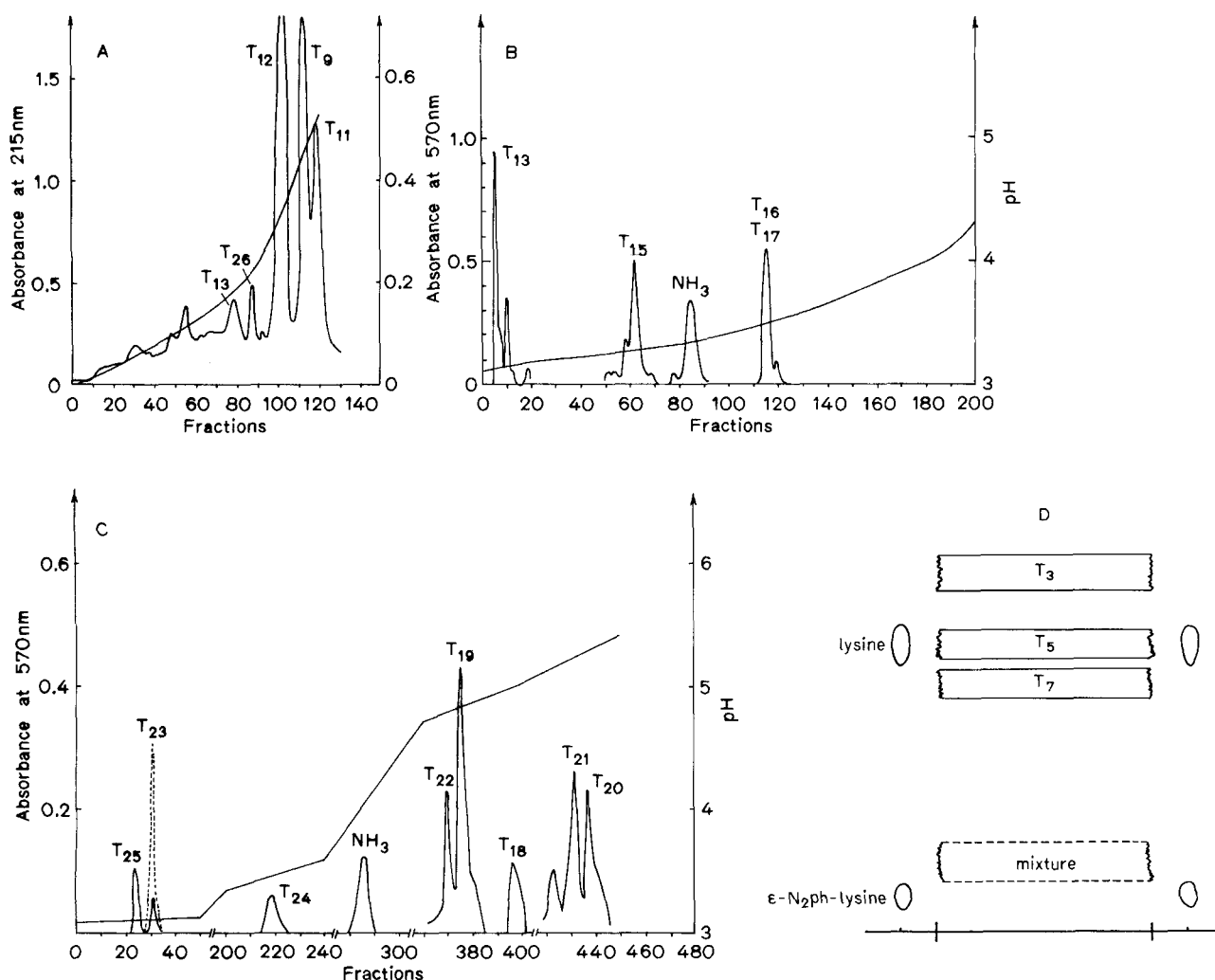


Fig. 4. Purification of tryptic peptides. (A) Chromatography in 0.015 M HCl, piperazine (pH 7.3) on a DEAE-Sephadex A-25 column (17 × 2.1 cm) of the tryptic fraction P4 FI (20 mg). Gradient: 200 ml buffer, 0.15 M NaCl, 200 ml buffer, 0.6 M NaCl. Fractions 4.8 ml. (B) Chromatography on a Technicon Chromobeads P2 column (27 × 2.5 cm) of the tryptic fraction P4 FII (25 mg). Linear gradient of 1 l 0.2 M pyridine/acetic acid (pH 3.1), 1 l 0.5 M pyridine/acetic acid (pH 5). Final washing with 200 ml 2 M pyridine. Fractions of 6 ml. Flow rate 32 ml/h. (C) Chromatography on a Technicon Chromobeads P2 column (80 × 1.5 cm) of the tryptic fraction P4 FV (50 mg). The eluting gradient is made of: 4 × 500 ml 0.2 M pyridine/acetic acid (pH 3.1), 3 × 500 ml 0.5 M pyridine/acetic acid (pH 5), 2 × 500 ml 2 M pyridine. Fractions of 6.5 ml. Flow rate 50 ml/h. (D) Electrophoresis at pH 3.5 (pyridine/acetic acid/water, 10/100/890 by vol.) on Whatman no. 1 paper of 2 mg of the Biogel P4 fraction VI. Conditions of the run were 3200 V, 100 mA. The width of the starting band was 16 cm. The peptides were detected by staining with the ninhydrin reagent two strips of 0.5 cm width cut from each side of the migration band. The ninhydrin-positive bands were then eluted with the above buffer by a centrifugation procedure

specific cleavages; this is the case for T₁₃ (T₁₂), T₂₄ (T₁₉), T₂₅ (T₂₇), T₁₁ (T₂₆). Moreover T₂₀, T₂₁, and T₁₈ seem to be also part of T₂₆. These cleavages occur at the carboxyl end of hydrophobic residues and can probably be attributed to a residual chymotryptic activity of the commercial preparation of trypsin.

T₂₃ is the blocked N-terminal peptide; it therefore carries the acetyl group; it is followed by T₁₂. The sum of these two peptides gives also the peptide SCN II plus the following material (CmCys, Asx, Ser 2, Ala 3)-Phe-Lys. T₂₆; T₁₁, T₁₇, T₁₈, and T₂₀ can be aligned easily using the CNBr fragments to

give the sequence of 25 residues at the C-terminal end of the protein. Indeed T₁₇ is followed by a peptide starting with Asx-Gly (see T₂₆); the only peptide having such characteristic is T₁₁; this latter contains one methionine followed by a sequence -Ile-Gly-Val..., which is the start of CNBr II followed itself by CNBr III (T₁₈).

At this step the molecule is arranged as follows: T₂₃-T₁₂(T₃, T₁₉, T₂, T₁₆, T₅, T₆, T₉, T₂₇, T₁₅)-T₁₇-Asx-Gly-Asx-Gly-Met-CNBr II-T₁₈.

T₂₀ and T₂₁ have not been considered; they most certainly arise from T₂₆; if they are added the total number of lysine residues becomes 15, which is

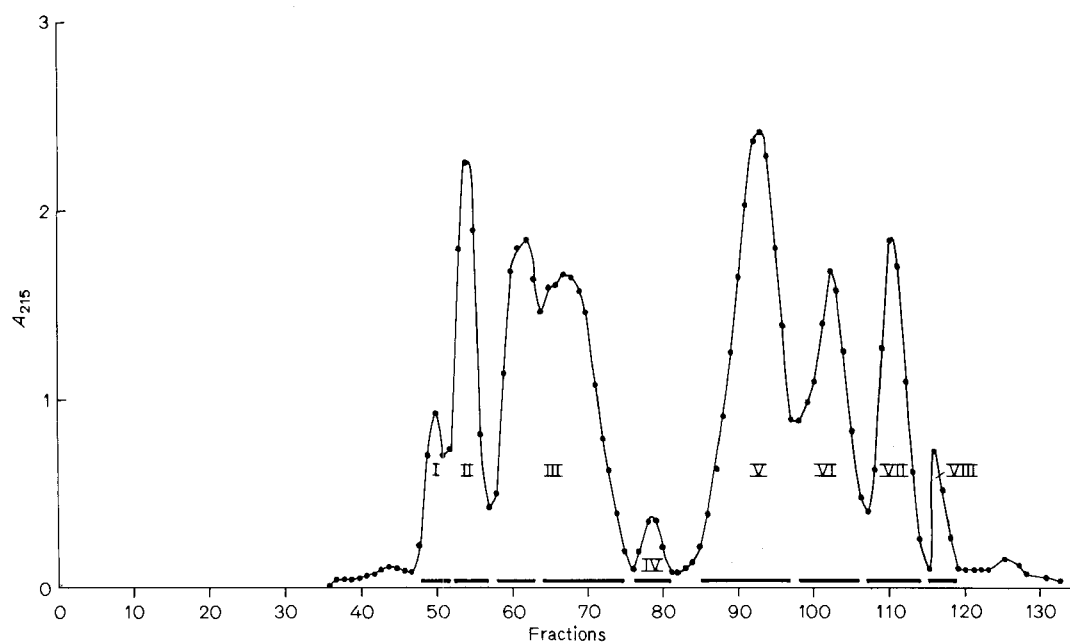


Fig. 5. Chromatography in 0.005 M NH_4HCO_3 on a Biogel P4 column of the thermolytic digest of S-carboxymethylated parvalbumin II of pike

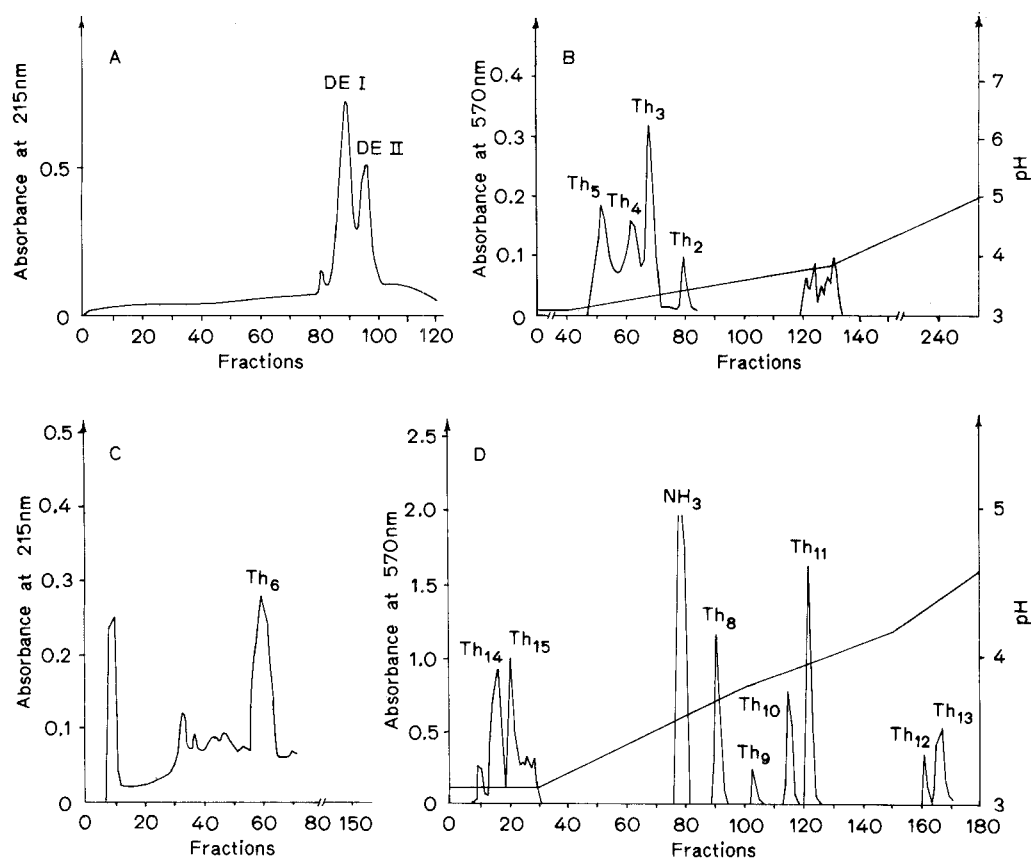


Fig. 6. Chromatography of thermolytic peptides. (A) Chromatography of the thermolytic fractions P4 FII (10 mg) on a DEAE-Sephadex A-25 column (16×2.2 cm) equilibrated in 0.015 M HCl, piperazine, pH 7.3. The eluting gradient is: 2×100 ml, 3×100 ml buffer/0.1 M NaCl, 3×100 ml buffer/0.6 M NaCl. Fractions of 6 ml. Flow rate 30 ml/h. (B) Chromatography on a Technicon Chromobeads P2 column (60×0.9 cm) of the thermolytic fraction DE I (A). The gradient is generated by an Ultrograd apparatus: 4 h 0.2 M pyridine/acetic acid (pH 3.1) followed by a 24-h linear gradient made by mixing the starting buffer with 2 M pyridine/acetic acid (pH 5). Fractions of 6.5 ml. Flow rate 60 ml/h. (C) Chromatography on a DEAE-cellulose column (15×2.2 cm) equilibrated in 0.03 M HCl-piperazine pH 6.5 of the thermolytic fraction DE II (Fig. 6A). The gradient is: 200 ml buffer, 200 ml buffer/0.2 M NaCl, 200 ml buffer/0.6 M NaCl. Fractions of 4.5 ml. (D) Chromatography on Technicon Chromobeads P2 of the thermolytic fraction P4 FIII (Fig. 5). Conditions are as described in (B)

Table 4. Amino acid compositions and other characteristics of the thermolytic peptides isolated from two fractions (II and III) obtained by chromatography of the thermolytic digest on Biogel P4 (Fig. 5)

Peptides are disposed from left to right in the order of position in the polypeptide chain. Yields are expressed in mol/mol protein

Amino acid	Th8	Th14	Th15	Th12	Th13	Th11	Th10	Th9	Th6	Th3	Th5	Th4	Th2
Lysine	1.0			1.8	2.4	2.0	1.9	1.9	1.7	0.9	0.9	1.0	0.9
Histidine				0.9	1.1								
Arginine													
Carboxymethylcysteine		0.7	0.9										
Aspartic acid	2.0		0.9	1.0	0.8	1.8	1.9	1.6	2.3	4.2	4.2	2.9	4.2
Asparagine													
Threonine													
Serine		1.2	1.2	1.5	1.0				0.7				
Glutamic acid				1.0	1.3				4.0				
Glutamine													
Proline													
Glycine									1.0	3.4	3.6	2.2	3.1
Alanine	1.0	4.2	4.1	1.9	0.8		1.0	1.1		1.1	1.0	0.9	1.0
Valine						1.0	0.9	0.9					
Methionine										0.9	1.0		1.1
Isoleucine											0.9		
Leucine	1.0	1.3	1.1			1.0	1.0		1.0	1.1	1.2		
Tyrosine													
Phenylalanine				1.0	2.1				1.3				
Number of residues	5	7	8	10	9	6	7	6	14	11	13	7	10
Relative mobility to ASP at pH 6.5	—	—	—	+ 0.17	—	0	0	0	—	—	—	—	—
Charge at pH 6.5	—	—	—	+ 0.5	—	0	0	0	—	—	—	—	—
Biogel P4 fractions (Fig. 5)	III	III	III	III	III	III	III	III	II	II	II	II	II
Purification method (cf. Fig. 6A–D)	D	D	D	D	D	D	D	D	A,C	A,B	A,B	A,B	A,B
Final yield (mol/mol)	0.307	0.061	0.138	0.054	0.055	0.18	0.10	0.026	0.04	0.08	0.04	0.015	0.03

too far away from the number quoted in the original amino acid composition Table 1.

Thermolytic Peptides

The result of the chromatography of the peptide mixture on Biogel P4 is represented in Fig. 5. Peptide mapping revealed that the eight fractions are each made of rather complex peptide mixtures. As an example fraction V contains 15 ninhydrin-positive spots. This is the reason why only two fractions (II and III) were selected for further examination on the basis of their relatively high content in lysine residues. Both were submitted to ion-exchange chromatography (Fig. 6) to give the pure peptides whose characteristics are given in Table 4. The partial amino acid sequences of the main thermolytic peptides listed in Table 4 are shown in Fig. 9 (Th14, Th13, Th2, and Th4 were not considered, see legend of Fig. 10).

Th8 confirms the alignment T23-T12; Th15 and Th12 allow the alignment of T12-T3-T19; Th10 and

Th9 enable one to pair T16-T5; Th6 aligns T6-T9, Th3 and Th5 are consistent with the alignment of T17-T11. The tryptic peptides can therefore be arranged as follows: T23-T12-T3-T19(T2-T16-T5, T6-T9, T27, T15)-T17-T11-Ala-OH.

Staphylococcal Protease Peptides

From the insoluble material separated by centrifugation a peptide SP1 was isolated by chromatography on Biogel P6 (Fig. 8A) (Table 5). The result of the chromatography of the supernatant on Biogel P4 is shown in Fig. 7. The four fractions obtained were resolved into several peptides by ion-exchange chromatography (Fig. 8). The characteristics of these peptides are shown in Table 5 and their partial amino acid sequences determined by the dansyl-Edman method and by carboxypeptidase A digestion are summarized in Fig. 9 (SP5, SP10, and SP16 were not represented in the figure; see the legend of Fig. 9).

SP13, SP15, SP20, SP9, SP10, SP17, SP16, SP2, and SP4 confirm previous alignments. SP5 and SP6

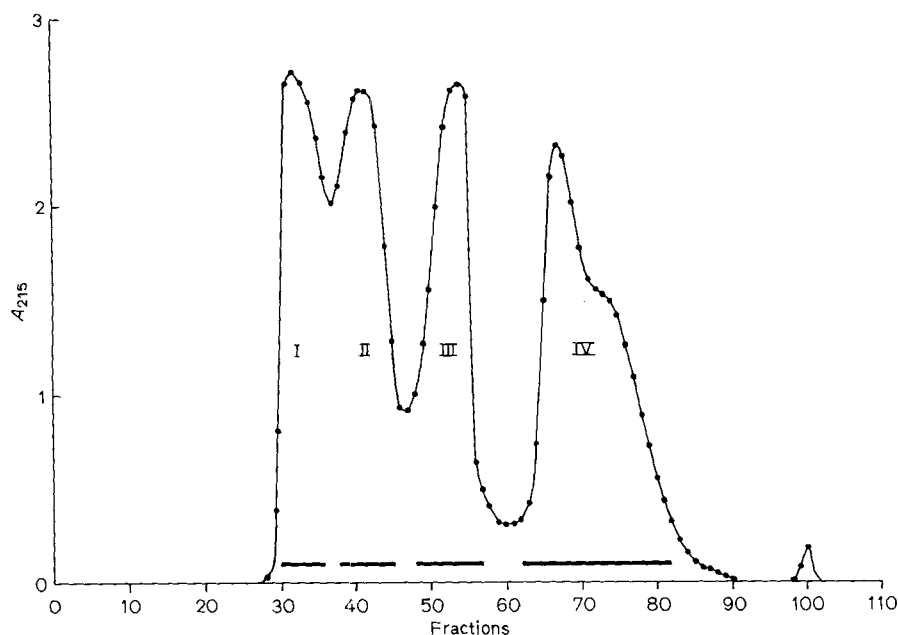


Fig. 7. Chromatography in 0.005 M NH_4HCO_3 on a Biogel P4 column (110×2.5 cm) of the digest of the S-carboxymethylated parvalbumin II of the pike with the protease from *Staphylococcus aureus*

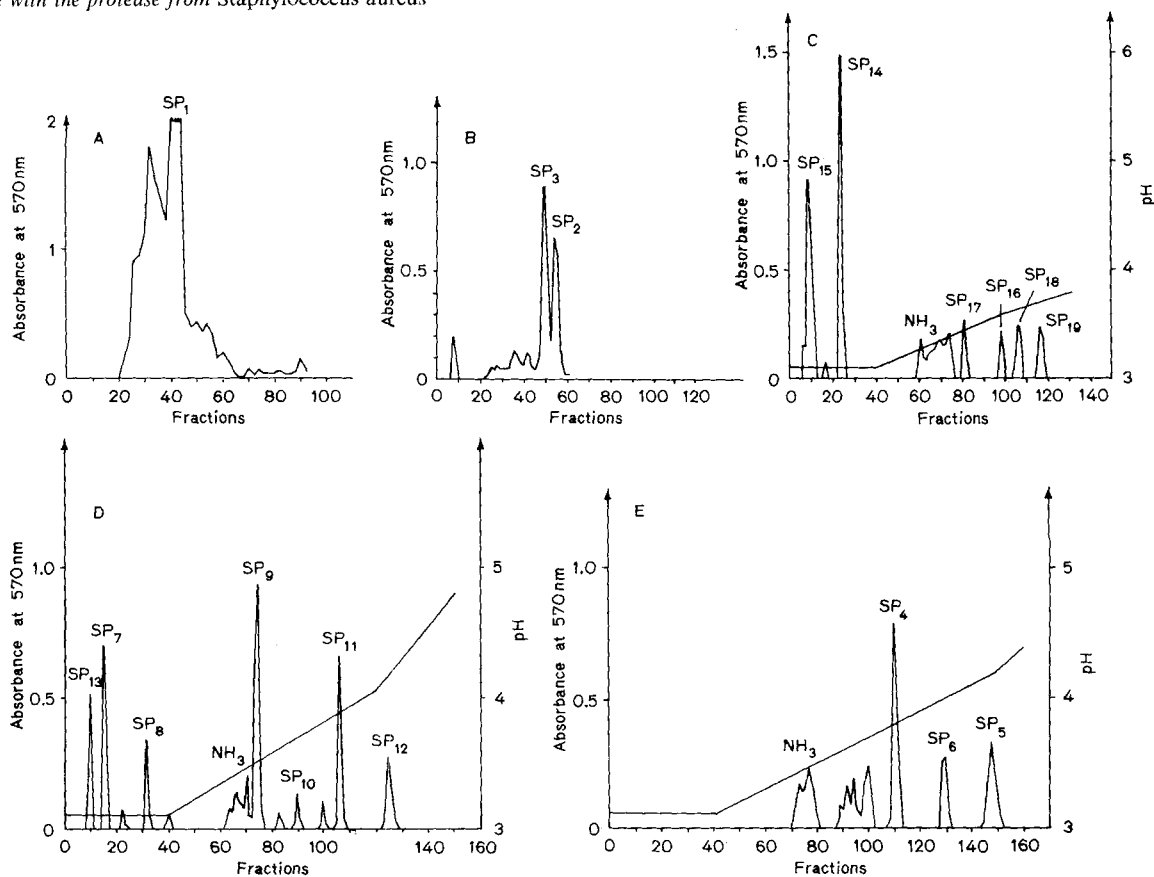


Fig. 8. Chromatography as staphylococcal protease peptides. (A) Chromatography of the precipitate obtained after digestion of the carboxymethylated parvalbumin II of the pike with the protease from *Staphylococcus aureus* on a Biogel P6 column (95×2.5 cm) equilibrated in 30% acetic acid. Fractions of 5.2 ml. (B) Chromatographic profile obtained by passing the fraction P4 FI (20 mg) (Fig. 7) from staphylococcal protease digestion on a DEAE-cellulose column equilibrated in 0.015 M HCl, piperazine (pH 6.5). Gradient is: 200 ml buffer/200 ml buffer, 0.2 M NaCl/200 ml buffer, 0.6 M NaCl. Fractions of 7 ml. Flow rate 32 ml/h. (C) Chromatographic profiles of the fraction P4 FII (Fig. 7) on a Technicon Chromobeads P2 column (60×0.9 cm). The gradients were generated by an Ultrograd apparatus: 4 h 0.2 M pyridine, acetic acid (pH 3.1) followed by 24 h of a linear gradient made by mixing the above buffer (pH 3.1) with 0.5 M pyridine, acetic acid (pH 5). Fractions of 5.5 ml. (D) Chromatographic profile on the fraction P4 FIII (Fig. 7) on a Technicon chromobeads P2 column (60×0.9 cm). Conditions as in (C). (E) Chromatographic separation of the peptides of the fraction P4 FIV (Fig. 7) as in (D) the exception being that the limiting buffer is made of 2 M pyridine, acetic acid (pH 8)

Table 5. Amino acid compositions and other characteristics of the peptides resulting from the staphylococcal protease digestion and isolated from chromatographies on Biogel P4 (Fig. 7) followed by ion-exchange chromatography on DEAE-cellulose (Fig. 8B) and on Technicon chromatobeads P2 (Fig. 8C, D, E)
 SP1 (Fig. 8A) was obtained from the insoluble material (I) generated by the digestion. t. l. means isolation by preparative peptide mapping on cellulose thin layer, (first direction: butanol/acetic acid/water, 4/1/5; second direction: electrophoresis at pH 3.5). We attributed the low yield of SP20 to a strong adsorption effect on the cellulose thin layer

Amino acid	SP13	SP15	SP8	SP7	SP20	SP5	SP6	SP9	SP10	SP17	SP16	SP18	SP12	SP11	SP19	SP1	SP3	SP2	SP14	SP4
Lysine	0.6	0.8			2.0	0.9	0.8	0.9	0.9	1.0	0.9	1.9	2.1	1.0	1.0	1.3	1.9	0.7		1.0
Histidine					0.9															
Arginine																1.0				
Carboxymethyl- cysteine																				
Aspartic acid	1.0	2.0	1.0					1.0	1.0	1.7	2.2	3.0	2.0			2.3	5.3	4.8	1.0	
Asparagine																	0.9	0.9		
Threonine																0.9				
Serine	0.9	1.3			1.7			1.9	1.6	1.6	2.2			0.9	1.1	1.8				
Glutamic acid					1.0							1.0	1.1	1.0	1.9	1.8	1.2	1.1	1.9	
Glutamine																				
Proline																0.8				
Glycine	1.0	1.1				1.0	0.8							1.0	1.1		4.0	3.8		
Alanine	1.0	1.6	1.0	3.7		1.1	1.8	1.0	1.0			1.4	1.0			3.0	2.0	1.0		2.7
Valine				0.9		0.9	1.1					1.8	1.8				1.0	1.1		
Methionine																	0.7	0.8		0.7
Isoleucine												0.7	0.8	0.9	0.8		0.9	0.9		1.0
Leucine	1.2	1.2		1.1			1.2	2.1	1.0	1.6	1.1					3.7				
Tyrosine											0.8	0.9								
Phenylalanine	1.1	0.8			1.0	1.9	1.7				1.1	1.1	1.1	1.0	0.8	1.8	0.9			0.9
Number of residues	7	9	2	6	7	6	8	7	5	8	6	12	11	6	7	19	20	15	3	7
Relative mobility to Asp at pH 6.5	—	—	—	—	—	—	—	0	—	+ 0.36	—	—	—	—	—	—	—	—	+ 0.82	—
Charge at pH 6.5	—	—	—	—	—	—	—	0	—	—	—	—	—	—	—	—	—	—	—	—
Biogel P4 (Fig. 7) fractions	III	II	III	III	IV	IV	IV	III	III	II	II	II	III	III	II	↓	I	I	II	IV
Purification method (cf. Fig. 8A–E)	D	C	D	D	E	E	E	D	D	C	C	C	D	D	C	A	B	B	C	E
Final yield (mol/mol protein)	0.08	0.154	0.064	0.069	0.005	0.084	0.048	0.054	0.051	0.041	0.04	0.034	0.069	0.138	0.031	0.65	0.277	0.256	0.13	0.28

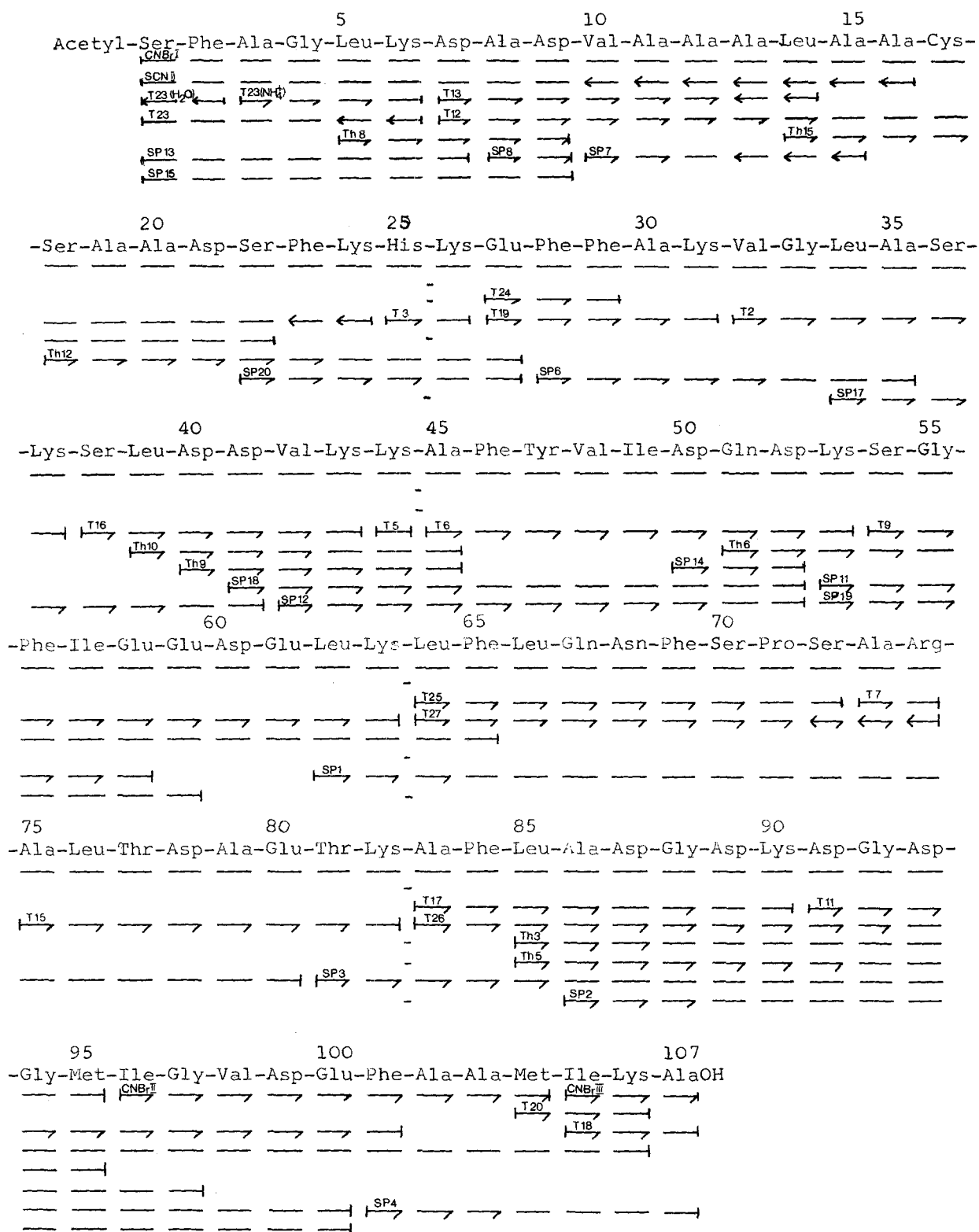


Fig. 9. Summary of the sequence data obtained from the S-carboxymethylated parvalbumin II. The peptides were degraded by the dansyl-Edman method ($\rightarrow$) and by carboxypeptidase digestion ($\leftarrow$). The peptides T21 (Table 3; residues 105–106); Th14, Th13, Th11, Th4, and Th2 (Table 4; residues 14–20, 20–29, 39–44, 86–92, 86–95); SP5, SP10, and SP16 (Table 5; residues 28–33, 36–40, 36–41) are not represented on the figure. Their partial amino acid sequences essentially confirm the results obtained with other peptides isolated from these regions

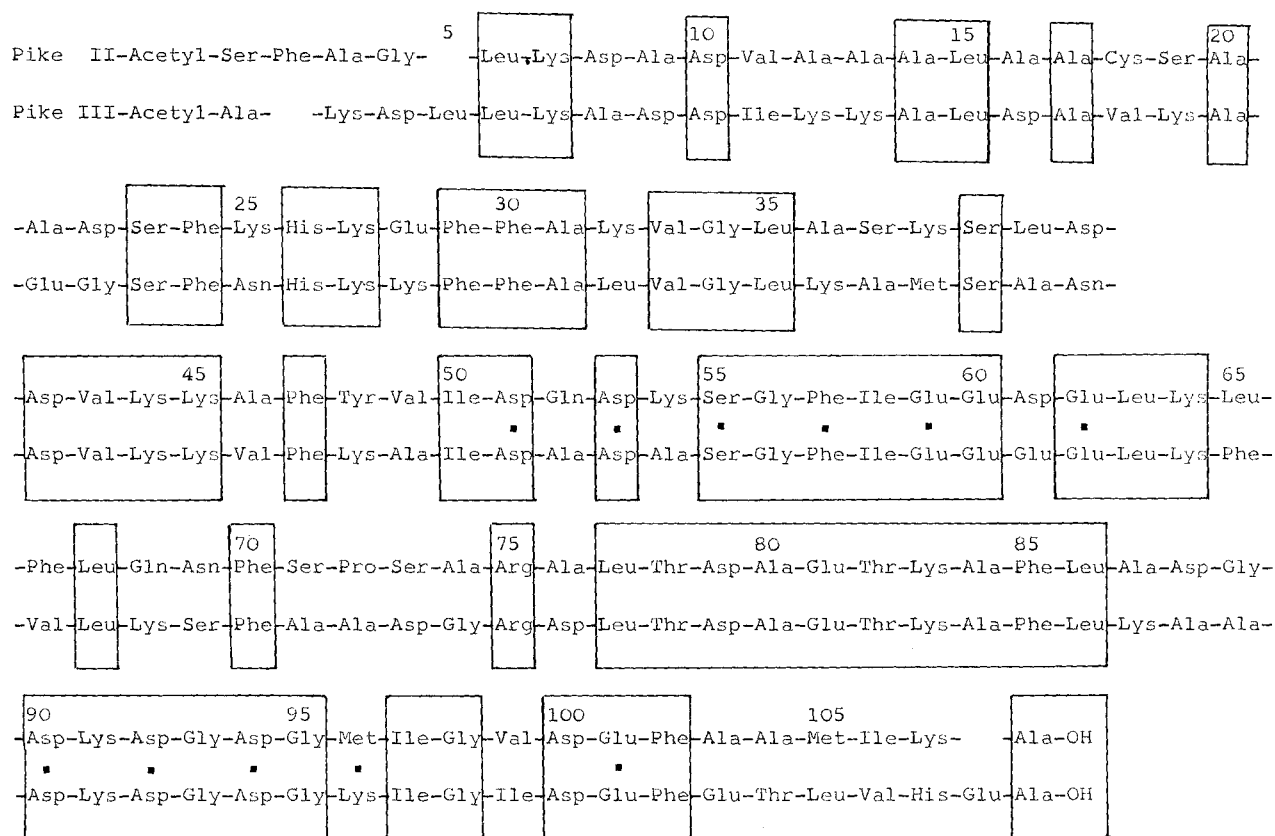


Fig. 10. Alignment of the amino acid sequences of the two parvalbumins of the pike. Residues identified by a black square are implicated in the coordination bonds of the two calcium ions bound to both proteins

allow the alignment T19-T2; SP18 and SP12 are consistent with the alignment T16-T5-T6, whereas SP 1 and SP 3 allow one to align T9-T27-T15-T26. The position of T27 is obvious despite the fact that T23 and T9 are both terminated by -Leu-Lys. Indeed one has already shown without ambiguities that T23 is followed by T12 (SCN II, Th8).

Amide Groups

Among the tryptic peptides mentioned in Table 3, only T6 could carry an amide group (net charge at pH 6.5 = -1). Edman degradation of SP 14 followed by electrophoresis at pH 6.5 of the residual peptides gives the following result:

	Mobilities relative to Asp at pH 6.5	Net charge at pH 6.5
SP14: Asx-Glx-Asx	+ 1.09	- 2
First degradation Glx-Asx	+ 0.77	- 1
Second degradation -Asx	+ 1.00	- 1

These results indicate that the amide group of T6 is located on the Glx residue.

DISCUSSION AND CONCLUSION

The amino acid sequence of the parvalbumin II of the pike is shown in Fig. 10. It is well representative of this protein family showing an acetylated N-terminal group, a cysteine residue in position 18, the doublets of phenylalanine and of lysine residues at positions 30-31 and 44-45 respectively, an arginine residue in position 75 and the quadruplet of acidic residues in the region 59-62.

The primary structure was achieved using the *S*-cyanoethylation method and digestion with the staphylococcal protease [17]. The first method reveals itself to be very specific but gives rather poor yield. This could be due to the experimental conditions: despite the fact that after modification the protein did not react any more with the 5,5'-dithio-bis(2-nitrobenzoic acid) we did not check whether disulfide formation had taken place; this could impair the yield of the derivative. It would be also wise to reduce the amount of dithiothreitol [16].

The protease from *Staphylococcus aureus* [17,18] is very useful in sequencing parvalbumin. Indeed, the relatively high number of acidic residues compared to the lower number of lysine and arginine residues allows one to obtain almost all the overlaps necessary to align the tryptic peptides. The enzyme competes favorably with chymotrypsin and thermolysin, which show a much broader specificity. From the 20 peptides isolated from the digestion with this enzyme 11 are strictly specific; they result from cleavages at the level of aspartic or glutamic residues. The others are produced by one specific cleavage and by rupture at the level of leucine or alanine residues. Bonds involving leucine-34 and alanine-35 are particularly sensitive, since SP 5, SP 6, SP 9, SP 10, SP 16, and SP 17 derived from cleavages at these levels.

The polypeptide chain contains 107 residues and is one residue shorter than most of the other well-characterized parvalbumins. The best alignment of this sequence with that of the parvalbumin III of the pike is shown in Fig. 10; it shows that 49 amino acid replacements, including addition or deletion, can be computed between the two parvalbumins present in the pike white muscles.

This result is very different from what was observed from similar calculation made on the three parvalbumins of the carp, which have rather similar primary structure. We have already suggested [1] that in the case of the carp and other Cyprinidae the extended polymorphism of the parvalbumins could be the result of a rather late polyploidization of the whole genome. This would probably explain their close relationship.

The parvalbumins of the pike arise on the contrary from a far remote gene duplication followed by divergent evolution. In fact, the values of the random evolutionary hits, calculated according to Hölmquist [19] by comparing the primary structure of the parvalbumin II of the pike with the amino acid sequences of ten other parvalbumins, can be divided into two groups: in the first one would class the carp III [4], the carp II [4], the chub V (Gerday, unpublished), the carp IV [4], the whiting [20] and the hake [4] with evolutionary distances ranging from 35 to 51 random evolutionary hits; in the second one would have, the frog [4], the pike III [3], the coelacanth [4] and the rabbit [21] with evolutionary distances of 73, 82, 104, and 104 random evolutionary hits respectively.

This suggests that the parvalbumins of the pike belong respectively to two different subgroups. The pike II is obviously related to the β group defined

by Pechère [4] and the pike III to a non- β group, both originating by gene duplication previous to the divergence of the Osteichthyes. The pike is the first example showing that the two types of parvalbumins coexist within a species. They possess rather different properties and in particular they do not show any immunochemical cross-reactivity (Gosselin-Rey, personal communication). Nevertheless, they keep in common the structural characteristics of the two Ca^{2+} -binding sites (Fig. 10) already recognized in the other parvalbumins.

I would like to thank Prof. D. C. Phillips and Dr R. E. Offord for their hospitality. Dr J. D. Priddle and Dr M. Makinen have generously provided their advice and criticisms but also their warm friendship throughout my stay in Oxford. I thank them very sincerely. I am also indebted to Prof. G. Hamoir for his continued interest in this work. This research has been made possible thanks to the financial support of the NATO Science Fellowship Programme.

REFERENCES

1. Bhushana Rao, K. S. P. & Gerday, Ch. (1973) *Comp. Biochem. Physiol.* **44 B**, 931–937.
2. Bhushana Rao, K. S. P. & Gerday, Ch. (1973) *Comp. Biochem. Physiol.* **44 B**, 1113–1125.
3. Frankenke, F., Joassin, L. & Gerday, Ch. (1973) *FEBS Lett.* **35**, 145–147.
4. Pechère, J. F., Capony, J. P. & Demaille, J. (1973) *Syst. Zool.* **22**, 533–548.
5. Crestfield, A. M., Moore, S. & Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622–627.
6. Benson, J. V. & Patterson, J. A. (1965), *Anal. Chem.* **37**, 1108–1110.
7. Jacobson, G. R., Schaffer, M. H., Stark, G. R. & Vanaman, T. C. (1973) *J. Biol. Chem.* **248**, 6583–6591.
8. Degani, Y. & Patchornik, A. (1971) *J. Org. Chem.* **36**, 2727–2728.
9. Gray, W. R. (1972) *Methods Enzymol.* **25**, 333–344.
10. Woods, K. R. & Wang, K. T. (1967) *Biochim. Biophys. Acta*, **133**, 369–370.
11. Dopheide, T. A. A., Moore, S. & Stein, W. H. (1967) *J. Biol. Chem.* **242**, 1833–1837.
12. Henkel, H. G. (1971) *J. Chromatogr.* **58**, 201–207.
13. Offord, R. E. (1966) *Nature (Lond.)* **211**, 591–593.
14. Hooton, B. T. (1968) *Biochemistry* **7**, 2063–2071.
15. Moore, S. & Stein, W. H. (1954) *J. Biol. Chem.* **211**, 907–913.
16. Degani, Y. & Patchornik, A. (1974) *Biochemistry*, **13**, 1–11.
17. Houmard, J. & Drapeau, G. R. (1972) *Proc. Natl Acad. Sci. U.S.A.* **69**, 3506–3509.
18. Schiltz, E. & Reinbolt, J. (1975) *Eur. J. Biochem.* **56**, 467–481.
19. Hölmquist, R. (1972) *J. Mol. Evolut.* **1**, 211–222.
20. Joassin, L. (1974) *Abstr. Commun. 9th Meet. Fed. Eur. Biochem. Soc.* p. 20.
21. Enfield, D. L., Ericsson, L. D., Blum, H. E., Fischer, E. H. & Neurath, H. (1975) *Proc. Natl Acad. Sci. U.S.A.* **72**, 1309–1313.