THE AMINO ACID SEQUENCE OF THE PIKE (*ESOX LUCIUS*) PARVALBUMIN III

F. FRANKENNE, L. JOASSIN and Ch. GERDAY
Laboratory of General Biology, University of Liège, 22, quai Van Beneden, 4000, Liège, Belgium

Received 6 June 1973

1. Introduction

The primary structure of major parvalbumins of hake (*Merlucius merlucius*) [1, 2] and of carp (*Cyprinus carpio*) [3] have been determined recently. The comparison of the two sequences has confirmed the homology of these proteins typical of aquatic cold blooded vertebrates [4, 5].

We report here the amino acid sequence of the major component of the two parvalbumins extracted from pike muscle. This protein has, by contrast with the rather conservative amino acid compositions of the other parvalbumins, and unusual composition characterized by a high content of 17 lysine residues and by the absence of cysteine, proline, tyrosine and tryptophan. The molecule has an isoelectric point of 5.0 and strongly binds two calcium ions [6, 7].

2. Materials and methods

The protein was prepared and characterized as described previously [6, 7].

Tryptic digestion was performed at 37°C in 0.2 M NH_4HCO_3 brought to pH 7 with formic acid by addition of 2% (w/w) trypsin [8]. The procedure, lasting 8 hr, was repeated twice in order to ensure complete digestion of the native protein.

The chymotryptic digestion was carried out in 0.2 M phosphate buffer, pH 8.5 with 1% (w/w) chymotrypsin for 15 hr at 37°C.

CNBr fragmentation was performed as described earlier [9]. The resulting fragments were isolated by gradient elution chromatography on DEAE Sephadex A-25.

The isolation of the individual peptides produced by enzymatic digestions was according to the following sequence of operations: gel filtration on a Biogel P_4 column (2.5 × 100 cm) equilibrated in 0.05 M NH_4HCO_3, further resolution of the fractions by chromatography on polystyrene sulfonic resins (Dowex 50 × 2, Technicon Chromobeads P) in pyridine–acetic acid buffers and on DEAE cellulose (Whatman DE 32), with final purification by chromatography or electrophoresis on cellulose thin layer (1 mm) and on paper.

The N-terminal acetylated chymotryptic peptide was obtained from a 2 hr, 2% (w/w) chymotryptic digest of the succinylated protein [10] by passing the acidified mixture through a Dowex 50 × 2 column (0.9 × 15 cm), H^+ form. The peptide was recovered by elution with water.

The amino acid analyses of the peptides, after hydrolysis in 6 N HCl for 24 hr, were made with a modified Beckman amino acid analyzer 120B using the two column system [11].

The amino acid sequences of the peptides were elucidated using digestions with carboxypeptidase A (EC 3.4.2.1), carboxypeptidase B (EC 3.4.2.2), leucine amino peptidase (EC 3.4.1.1) and the Edman-dansyl method [12]. Dansyl amino acids were identified on (7.5 × 7.5 cm) polyamide layers (Cheng Chin Trading) [13].

The presence of the N-terminal acetyl-alanine residue and the amino acid sequences of the tryptic octapeptides T_3 and T_{20} have been confirmed by mass spectrometric analysis of the permethylated materials.

The position of the amide groups have been assigned on the basis of the electrophoretic mobilities of the peptides on paper at pH 6.5 [14] and in ambiguous

cases by the same method but on the peptides degraded to suitable extent by the Edman method. The amide group in position 41 has been detected by mass spectrometry.

The tryptic peptides have been aligned mainly on the basis of the CNBr fragmentation and chymotryptic overlaps and partly by comparison with the two other parvalbumin sequences already available [2, 3].

3. Results and discussion

The amino acid sequence of the parvalbumin III of pike, along with the homologous sequences of the hake and carp parvalbumins are shown in fig. 1.

The polypeptide chain is composed of 109 residues in agreement with the amino acid composition previously reported [7] exception being that the later must be corrected for one additional valine and alanine and for one exceeding lysine residue.

When compared to the carp sequence, 39 amino acid replacements and one addition at the C-terminal end have occurred. 13 Substitutions are located within the first 20 residues. A similar figure of 48 substitutions is observed when one compares the pike to the hake sequence, 12 of them occurring within the first 20 residues. An interesting fact is the presence of a valine residue in position 18 found being occupied by the only cysteine residue in the hake as well as in the carp sequence.

Important common features are also striking, notably the N-acetyl terminal group, the doublet –Phe–Phe– in position 29–30, the doublet –Lys–Lys– in position 44–45, the group of four acidic residues in position 59–62 and the presence of Arg in position 75. Moreover, the structure of the two Ca^{2+} binding sites, identified in the carp parvalbumin, using X-ray diffraction data, (R.H. Kretsinger, personal communication) are entirely conserved in the pike and in the hake sequences. The first site implicates residues 51, 53, 55, 57, 59 and 62 and the second the residues 90, 92, 94, 96 and 101. This finding associated with the extreme difficulty of releasing the metal ions from their binding sites emphasize the idea of the role of the two Ca^{2+} ions as structural requirements.

Position 81 is occupied by a Glu residue which was found to form an internal salt bridge with the unique Arg 75 in the carp parvalbumin [3]. These two sites are also conserved in the pike as well as in the hake sequence. Immunochemical studies [15] have provided strong evidence that Arg 75 occupies a key structural position in the parvalbumin molecule.

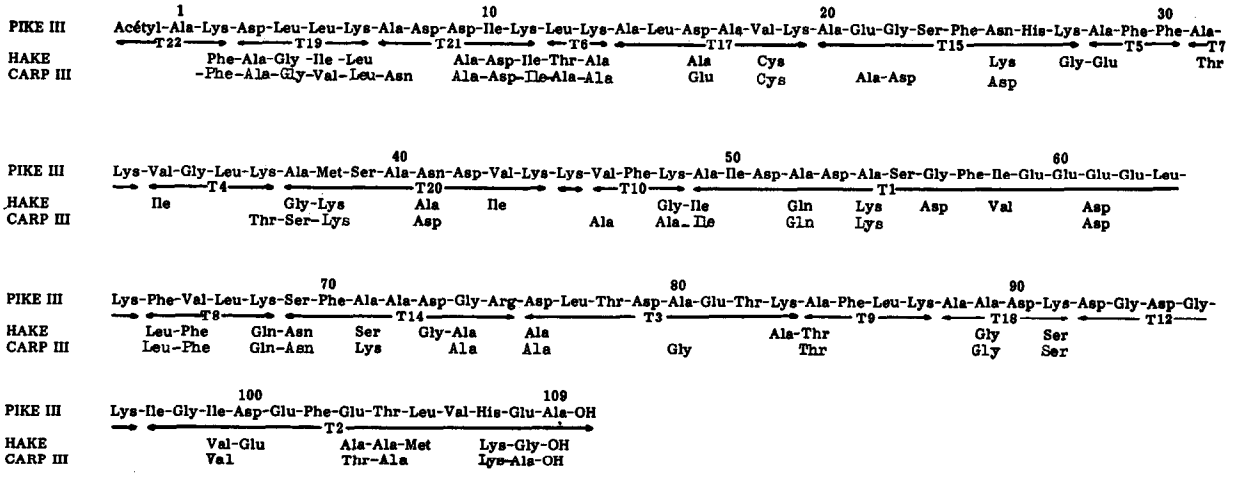

Fig. 1. Proposed alignment of the tryptic peptides of the parvalbumin III of pike compared to the homologous sequences of carp and hake components.

Acknowledgements

F. Frankenne and L. Joassin are holders of post graduate bursaries of the Institute for the development of Industrial and Agricultural Research (I.R.S.I.A.). The authors wish to thank Dr. B.C. Das of the Institut de Chimie des substances naturelles in Gif-sur-Yvette, who has kindly performed the mass spectrometric analyses and Dr. G. Hennen of the University Hospital in Liège who kindly received us in his laboratory for some sequence determination. We are also indebted to Prof. G. Hamoir, head of our laboratory for helpful discussions and his continuing interest in this work.

References

[1] Pechere, J.F., Capony, J.P., Ryden, L. and Demaille, J. (1971) Biochem. Biophys. Res. Commun. 43, 1106.
[2] Capony, J.P., Ryden, L., Demaille, J. and Pechere, J.F. (1973) European J. Biochem. 32, 97.
[3] Nockolds, C.E., Kretsinger, R.H., Coffee, E. and Bradshaw, R.A. (1972) Proc. Natl. Acad. Sci. U.S. 69, 581.
[4] Pechere, J.F., Demaille, J. and Capony, J.P. (1971) Biochim. Biophys. Acta 236, 391.
[5] Gerday, Ch. and Teuwis, J.C. (1972) Biochim. Biophys. Acta 271, 320.
[6] Bushana Rao, K.S.P. and Gerday, Ch. (1973) Comp. Biochem. Physiol. 44B, 931.
[7] Bushana Rao, K.S.P. and Gerday, Ch. (1973) Comp. Biochem. Physiol. 44B, 1113.
[8] Gerday, Ch. and Bushana Rao, K.S.P. (1970) Comp. Biochem. Physiol. 36, 229.
[9] Joassin, L., Frankenne, F. and Gerday, Ch. (1971) Arch. Int. Physiol. Biochim. 79, 834.
[10] Hapner, K.D. and Wilcox, Ph.E. (1970) Biochemistry 9, 4470.
[11] Benson, J.V. and Patterson, J.A. (1965) Anal. Chem. 37, 1108.
[12] Gray, W.R. and Hartley, B.S. (1963) Biochem. J. 89, 379.
[13] Woods, K.R. and Wang, K.T. (1967) Biochim. Biophys. Acta 133, 369.
[14] Offord, R.E. (1966) Nature 211, 591.
[15] Gosselin-Rey, C., Bernard, N. and Gerday, Ch. (1973) Biochim. Biophys. Acta 303, 90.