Acta Cryst. (1996). D**52**, 165–169

X-ray Structure of a New Crystal Form of Pike 4.10 β Parvalbumin

JEAN-PAUL DECLERCQ,[a] BERNARD TINANT[a] AND JOSEPH PARELLO[b]

[a]*Laboratoire de chimie physique et de cristallographie, Université Catholique de Louvain, 1 place Louis Pasteur, B-1348 Louvain-la-Neuve, Belgium, and* [b]*UA1111 CNRS Laboratoire de chimie organique pharmaceutique, Université de Montpellier I, 15 avenue C. Flahault, F-34060 Montpellier, France*

(*Received 13 March 1995; accepted 21 July 1995*)

Abstract

A new crystal form of pike (p*I* 4.10) parvalbumin has been crystallized in presence of EDTA at pH 8.0. The crystals are orthorhombic, space group $P2_12_12_1$, with $a = 51.84$, $b = 49.95$, $c = 34.96$ Å. Diffractometer data were collected to 1.75 Å. The structure was solved by molecular replacement and refined to $R = 0.168$ for 7774 observed reflections [$I \geq 2\sigma(I)$] in the range 8.0–1.75 Å. In spite of the presence of EDTA, calcium ions are present in both primary binding sites. As compared to the previously reported structures, the main differences concern the conformation of the N-terminal residues and the packing in the unit cell.

1. Introduction

Since the pioneering contribution of Kretsinger & Nockolds (1973), parvalbumins have been intensely investigated by high-resolution X-ray crystallography and several crucial answers have thus been obtained concerning the influence of the binding of different divalent cations on the conformation of the protein, including the two physiologically relevant cations Ca^{2+} and Mg^{2+}, as well the conformational features underlying the classification of parvalbumins into distinct phylogenetic series α and β. The dependence of parvalbumin conformation on the content and the chemical nature of divalent cations bound to the protein is well documented based on a variety of physico-chemical methods (for a review, Ribeiro, Parello & Jardetzky, 1984). In the past years we have determined the crystal structure of a half Ca-/half Mg-loaded form of parvalbumin with the CD site occupied by Ca^{2+} whereas the EF site is occupied by Mg^{2+} (Declercq, Tinant, Parello & Rambaud, 1991). The occurrence of this crystal form suggests that Ca/Mg exchange takes place more freely at the EF site, in agreement with previous observations that the EF site is more susceptible to cation exchange than the CD site (Eberspach, Strassburger, Glatter, Gerday & Wollmer, 1988). So far no structure of a cation-depleted parvalbumin has been determined. Early results by NMR with β-carp p*I* 4.25 parvalbumin suggested that the Ca-depleted form markedly differs from the corresponding Ca-loaded from in the sense that the hydrophobic core of the protein is absent (Parello *et al.*, 1974). However, recent NMR results with α-pike p*I* 5.0 parvalbumin have established that the Ca-free form still retains part if not all of its hydrophobic core organization (Blancuzzi, Padilla, Parello & Cavé, 1993). The crystallization in the presence of EDTA, a Ca-chelating agent, is an unsuccessful attempt to obtain cation-depleted crystals of a parvalbumin.

2. Experimental

2.1. *Purification, crystallization and data collection*

The details concerning the purification of parvalbumin components from pike muscle, the separation of the two components and the titration of the Ca/protein ratio were described by Declercq *et al.* (1991). At the end of this procedure the minor component (p*I* 4.10) was available in the form of a lyophilized powder. For crystallization, this powder was dissolved in deionized water at a concentration of 27 mg ml^{-1}. The crystals were obtained at room temperature by vapor diffusion using the hanging-drop technique. The drops were formed by mixing 4 µl of the protein solution with 4 µl of a solution composed of ammonium sulfate 2 *M*, EDTA 1 m*M*, sodium azide 3 m*M* and pH adjusted to 8.0 with Tris buffer. These drops were equilibrated against a reservoir filled with 5 ml of a solution identical to the latter except that ammonium sulfate was 2.4 *M*. Crystals appeared after a few days and grew to 0.4 × 0.4 × 1 mm. They belong to the orthorhombic space group $P2_12_12_1$, with cell dimensions $a = 51.84$, $b = 49.95$, $c = 34.96$ Å, $V = 90526$ Å^3. The crystals previously obtained in ammonium sulfate but in the absence of EDTA (Declercq *et al.*, 1991) were also orthorhombic, but space group $P2_12_12$ and $a = 59.50$, $b = 59.63$, $c = 26.24$ Å, $V = 93099$ Å^3, $Z = 4$. The comparison of the two volumes indicates that this new crystal form also contains one protein molecule in the asymmetric unit.

Intensity data were collected to a resolution of 1.75 Å from eight crystals on a Huber four-circle diffractometer using Cu $K\alpha$ monochromatized radiation produced by a Rigaku RU200 rotating-anode X-ray generator. For each crystal, the data were empirically corrected for absorp-

tion and crystal decay. A total of 18 280 reflections were collected and merged by the program *PROTEIN* (Steigeman, 1992) into a set of 9484 unique reflections, with $R_{merge} = 5.1\%$. Of these unique reflections, 7774 for which $I \geq 2\sigma(I)$ were considered as observed and retained for structure analysis and refinement. They represent 81% of the total number of reflections measurable to a resolution of 1.75 Å, and 57% of the reflections in the highest resolution shell (1.81–1.75 Å) are observed.

2.2. *Structure determination and refinement*

As the crystal structure of the same molecule crystallized in absence of EDTA was available (Declercq *et al.*, 1991), the corresponding coordinates, with reference 1PAL in the Protein Data Bank (Bernstein *et al.*, 1977), were used as a model for the solution of the structure by molecular replacement. All the water molecules, the two calcium ions and the ammonium ion were discarded and the temperature factors of all the remaining atoms were fixed to 15.0 Å^2. The rotation search was performed in Patterson real space according to the procedure implemented in the program *X-PLOR* (Brünger, 1992). The Patterson of the model was computed in a $P1$ unit cell (60 × 60 × 60 Å) large enough to avoid overlap of intramolecular and intermolecular vectors. The data were restricted between 8.0 and 4.0 Å for both the Patterson of the model and of the unknown structure. The 5000 most intense peaks of the model Patterson, with a vector length between 5 and 30 Å were selected, and a product function was computed with intervals of 5° on an angular grid (θ_1 = 0–360, θ_2 = 0–90, θ_3 = 0–180°) defined in the notation of Rossmann & Blow (1962). According to Rao, Jih & Hartsuck (1980), this grid represents the asymmetric unit of the rotation space. The 100 best solutions were then filtered by the Patterson-correlation (PC) refinement procedure of Brünger (1990) to provide one unique independent solution (θ_1 = 330.48, θ_2 = 75.76, θ_3 = 149.14°) corresponding to the first solution of the rotation search. Using the same resolution range, the solution of the translation problem was also straightforward with the Patterson correlation as target and intervals of about 1 Å on a three-dimensional grid extending to a half unit cell in each direction. The best result occurred for a translation of (0.16 0.08 0.0) in fractional coordinates. This rotated and translated model provided an R value of 0.462 for data between 8 and 3.5 Å. After a rigid-body refinement, this value dropped to 0.377. This model was refined by simulated annealing with the slow-cooling protocol provided in *X-PLOR*, applied to the observed data between 8.0 and 1.75 Å. A starting temperature of 3000 K was gradually decreased to 300 K in stages of 25 K, during which 50 steps of 0.0005 ps of molecular dynamics were run. This process was followed by minimization. For this resolution range, the initial R factor of 44% was reduced to 31%.

At this stage, the examination of the electron-density maps (program *TURBO-FRODO*, Roussel, Fontecilla-Camps & Cambillau, 1990) computed with coefficients $2F_o - F_c$ and $F_o - F_c$ clearly showed that the two calcium ions had to be restored in the CD and EF loops and that it was necessary to completely rebuild the first four residues of the N-terminal side. The refinement was completed by alternating between minimizations and visual examinations. Individual restrained isotropic temperature factors were applied and 72 water molecules were included. Most of them were located automatically by the program *MAIN* (Turk, 1992). Electron density also appeared at the position identified by Declercq *et al.* (1991) as the third cation binding site. According to the arguments developed by these authors this site was fitted with an ammonium ion as the crystallization occurs in the presence of ammonium sulfate.

The final atomic coordinates and the structure factors have been deposited with the Protein Data Bank (Bernstein *et al.*, 1977).*

3. Results

3.1. *Quality of the structure*

The structure was refined to a resolution of 1.75 Å with a final R factor of 0.168 for observed data between 8.0 and 1.75 Å. Table 1 summarizes the results of the

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1PVB). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0438).

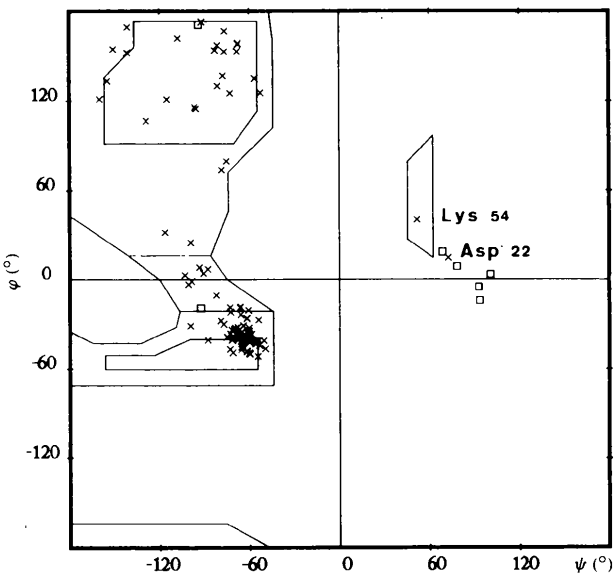

Fig. 1. Ramachandran plot. × Represents the non-glycine residues and □ the glycine residues. All the non-glycine residues show negative φ values except Asp22 and Lys54 (see text).

refinement. Fig. 1 presents a Ramachandran plot showing that all the non-glycine residues have, as expected, negative φ values, except Asp22 and Lys54 Residue 54 appears in the CD ion-binding loop. It is the fourth residue of the loop, and it appears in a region of the map usually occupied by glycine residues, as it occurs for example in this structure: Gly93 is the fourth residue of the EF loop. Roquet, Declercq, Tinant, Rambaud & Parello (1992) concluded from a compilation of the φ values of the fourth residue of the cation binding loops in parvalbumins that a positive φ value was a characteristic of such loops, whatever the identity of the residue occupying this fourth position: a glycine is often

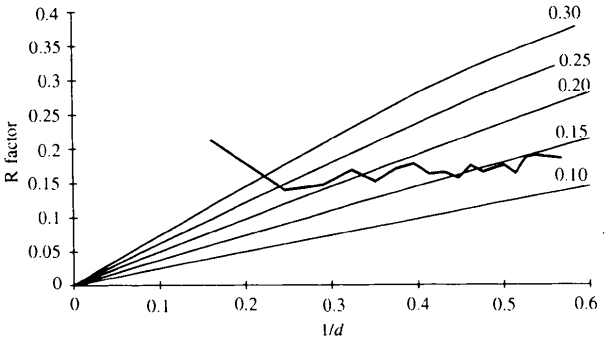

Fig. 2. Luzzati (1952) plot showing that the upper limit of the error on atomic positions is about 0.15 Å.

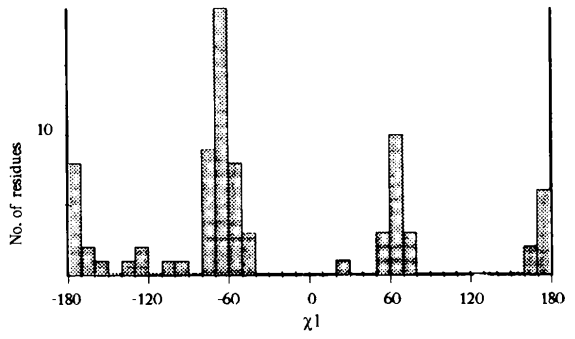

Fig. 3. Histogram presenting the trimodal distribution of the χ_1 angles.

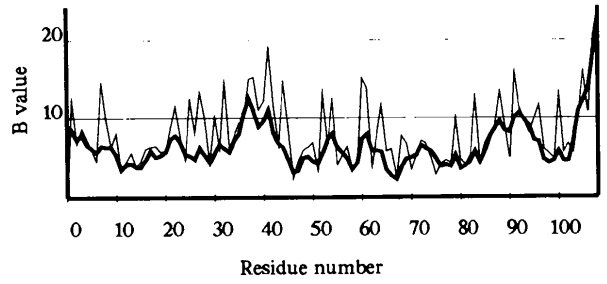

Fig. 4. Average temperature factors (B, Å^2) as a function of residue number: thick line for the main-chain atoms, thin line for the side-chain atoms.

Table 1. *Refinement statistics*

Resolution range (Å)	Reflections with $I \geq 2\sigma(I)$ (%)	R factor
8.0–1.75	81	0.168
8.0–3.42	94	0.166
3.42–2.75	90	0.159
2.75–2.41	85	0.170
2.41–2.20	83	0.163
2.20–2.04	84	0.171
2.04–1.92	79	0.169
1.92–1.83	72	0.187
1.83–1.75	57	0.189

Estimated errors on atomic positions (Luzzati, 1952) (Å)	0.15
No. of solvent molecules	72
Average B values of solvent molecules (Å^2)	21.7
Non-bonded parameters for the refinement of the Ca^{2+} ions	
Charge	+2
ε (kcal mol^{-1})	0.11844
σ (Å)	3.1171
B values of the cations (Å^2)	
CD site	5.3
EF site	4.4
Third site	11.5
R.m.s. deviations from ideality	
Bonds (Å)	0.009
Angles ( )	1.36
Restraints to the B factors (Å^2)	
Target standard deviations	
Bonds backbone	1.5
side chain	2.0
Angles backbone	2.0
side chain	2.5

observed in this position, but even in the other cases, a positive φ value is conserved. Residue Asp22 could represent an extension of this point of view, but applied to the AB loop where no cation binding occurs. The upper limit of the error on the atomic positions is estimated to be about 0.15 Å by means of the Luzzati plot (1952) shown in Fig. 2. A histogram (Fig. 3) of the distribution of the χ_1 torsion angles of the side chains shows the expected trimodal distribution with the largest value around $-60°$. Fig. 4 presents the temperature factors as a function of the residue number for the atoms of the main chain and of the side chains. An overall G factor of 0.46 computed by the program *PROCHECK* (Laskowski, MacArthur, Moss & Thornton, 1993) is also an indication of the excellent quality of the model. It has to be noted that all the individual G factors are larger than 0.2.

3.2. *Description of the structure*

Though the space group and the unit-cell parameters are completely different from those observed by Declercq *et al.* (1991) for the same protein crystallized in absence of EDTA, the two molecular structures are very similar and the principal differences concern the packing in the unit cell and the conformation of the N-terminal residues. Indeed, the r.m.s. deviation between the Cα atoms is only 1.12 Å and Fig. 5 shows a

superposition of these atoms in the two crystal structures. The r.m.s. deviations were also computed as a function of residue number for the atoms of the backbone and of the side chains, as presented on Fig. 6. As concerns the main chain, large differences occur only for the N-terminal part, which had to be completely rebuilt in the electron density. Without these first four residues, the r.m.s. Cα deviation drops to 0.52 Å.

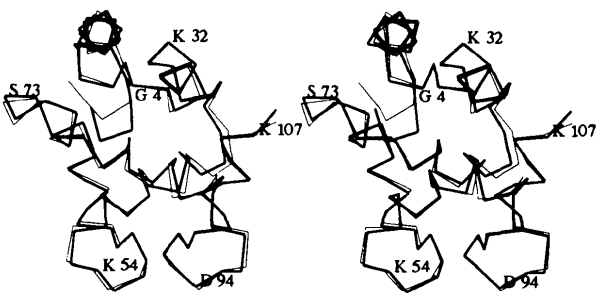

Fig. 5. Superposition of the Cα atoms of pike (p*I* 4.10) parvalbumin crystallized in the presence of EDTA (this study, thick lines) and in the absence of EDTA [1PAL, Declercq *et al.* (1991), thin lines].

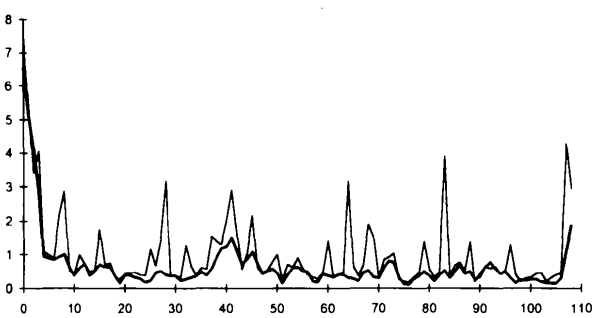

Fig. 6. R.m.s. deviations between the atoms of the backbone (thick line) and of the side chains (thin line) for pike 4.10 parvalbumin crystallized in the presence of EDTA (this study) and in the absence of EDTA (1PAL, Declercq *et al.*, 1991).

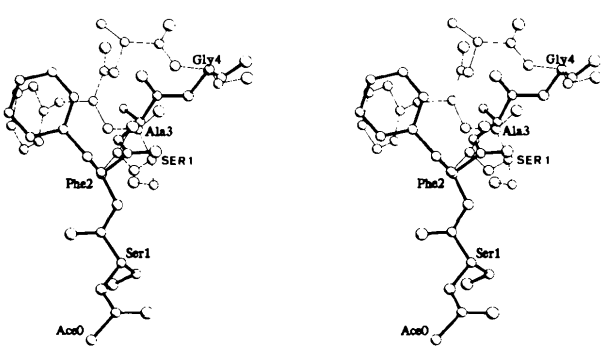

Fig. 7. Superposition of the N-terminal residues in this study (thick drawing) and in 1PAL (Declercq *et al.*, 1991, thin drawing). Program *PLUTO*, (Motherwell & Clegg, 1978).

As a consequence, all the typical features observed in the folding of parvalbumins substituted by Ca^{2+} ions are conserved (Kretsinger & Nockolds, 1973; Declercq, Tinant, Parello, Etienne & Huber, 1988; Swain, Kretsinger & Amma, 1989; Chiadmi, 1990; Kumar, Lee & Edwards, 1990; Roquet *et al.*, 1992) and will not be described here. Special attention should be given to Phe2 which appears in the completely rebuilt N-terminal region. Fig. 7 is a superposition of residues 1–4 in the two crystal structures and it is remarkable to notice that in spite of the very different conformations of these residues, the centres of gravity of the two phenyl rings (Phe2) nearly coincide, with the implication that the position of this aromatic ring with respect to the hydrophobic core is not very affected by the reorganization. For residues 5–102, r.m.s. deviations larger than 2.5 Å occur only for some side chains: Asp8, Glu28, Asp41, Lys64, Lys83, Lys107, Ala108. Except the last one, they are all polar side chains located at the surface of the molecule. Two of these residues (Lys64, Lys83) take part in intramolecular salt bridges (Lys64 N^{ζ}–Asp79 $O^{\delta 1}$ 3.00 Å; Lys83 N^{ζ}–Glu60 $O^{\varepsilon 2}$ 2.64 Å), which were not observed in 1PAL (distance > 3.5 Å). A third new intramolecular salt bridge is observed between Asp42 $O^{\delta 1}$ and Lys45 N^{ζ}. Fig. 6 also shows a relatively large difference for the side chains of these two residues. On the other hand, in 1PAL, an intermolecular salt bridge occurs between Lys107 N^{ζ} and the two O atoms (3.21 and 3.25 Å) of the terminal carboxylate of Ala108 in a symmetry-related molecule.

4. Concluding remarks

The new crystal structure of pike 4.10 parvalbumin establishes that the organization of the N-terminal region is highly dependent on the environment of the protein encountered in the crystal. It is known that the N-termini of parvalbumin can differ during evolution by the number of their amino-acid residues. There are examples of parvalbumins with shortened or elongated N-termini (see Fig. 1 in Roepstorff *et al.*, 1991) thus indicating that the conformational requirements at the level of the N-terminal region are not too stringent in regard to the tertiary folding of the protein.

JPD and BT are indebted to the Fonds National de la Recherche Scientifique (Belgium), and JP to the Centre National de la Recherche Scientifique (France), for financial support.

References

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.

Blancuzzi, Y., Padilla, A., Parello, J. & Cavé, A. (1993). *Biochemistry*, **32**, 1302–1309.

Brünger, A. T. (1990). *Acta Cryst.* **A46**, 46–57.
Brünger, A. T. (1992). *X-PLOR Version* 3.1, *A System for X-ray Crystallography and NMR*. New Haven and London: Yale University Press.
Chiadmi, M. (1990). PhD thèse, Université de Paris-Sud, France.
Declercq, J. P., Tinant, B., Parello, J., Etienne, G. & Huber, R. (1988). *J. Mol. Biol.* **202**, 349–353.
Declercq, J. P., Tinant, B., Parello, J. & Rambaud, J. (1991). *J. Mol. Biol.* **220**, 1017–1039.
Eberspach, I., Strassburger, W., Glatter, U., Gerday, C. & Wollmer, A. (1988). *Biochim. Biophys. Acta*, **952**, 67–76.
Kretsinger, R. H. & Nockolds, C. E. (1973). *J. Biol. Chem.* **248**, 3313–3326.
Kumar, V. D., Lee, L. & Edwards, B. F. P. (1990). *Biochemistry*, **29**, 1404–1412.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.
Motherwell, W. D. S. & Clegg, W. (1978). *PLUTO. Program for Plotting Molecular and Crystal Structures*. University of Cambridge, England.
Parello, J., Cavé, A., Puigdomenech, P., Maury, C., Capony, J. P. & Pechère, J. F. (1974). *Biochimie*, **56**, 61–76.
Rao, S. N., Jih, J.-H. & Hartsuck, J. A. (1980). *Acta Cryst.* **A36** 878–884.
Ribeiro, A., Parello, J. & Jardetzky, O. (1984). *Progr. Biophys. Mol. Biol.* **43**, 95–160.
Roepstorff, P., Klarskov, K., Andersen, J., Mann, M., Worm, O., Etienne, G. & Parello, J. (1991). *Int. J. Mass Spectrom. Ion Phys.* **111**, 151–172.
Roquet, F., Declercq, J. P., Tinant, B., Rambaud, J. & Parello, J. (1992). *J. Mol. Biol.* **223**, 705–720.
Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.
Roussel, A., Fontecilla-Camps, J. C. & Cambillau, C. (1990). *Acta Cryst.* **A46**, C66–C67.
Steigeman, W. (1992). *PROTEIN, Version* 3.1, *A Program System for the Crystal Structure Analysis of Proteins*. Max-Planck-Institute für Biochemie, Martinsried, Germany.
Swain, A. L., Kretsinger, R. H. & Amma, E. L. (1989). *J. Biol. Chem.* **264**, 16620–16628.
Turk, D. (1992). PhD thesis, Technische Universität München, Germany.