Crystal structure of the EF-hand parvalbumin at atomic resolution (0.91 Å) and at low temperature (100 K) . Evidence for conformational multistates within the hydrophobic core

JEAN-PAUL DECLERCQ,¹ CHRISTINE EVRARD,¹ VICTOR LAMZIN,² AND JOSEPH PARELLO^{3,4}

¹Université Catholique de Louvain, Unité CPMC, 1 place Louis Pasteur, B-1348 Louvain-la-Neuve, Belgium ²EMBL c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany

³UPRESA 5074 CNRS, Faculté de Pharmacie, 15 avenue Ch. Flahault, F-34060 Montpellier, France ⁴The Burnham Institute, 10901 North Torrey Pines Rd., La Jolla, California 92037

(RECEIVED May 7, 1999; ACCEPTED July 29, 1999)

Abstract

Several crystal structures of parvalbumin (Parv), a typical EF-hand protein, have been reported so far for different species with the best resolution achieving 1.5 Å. Using a crystal grown under microgravity conditions, cryotechniques (100 K) , and synchrotron radiation, it has now been possible to determine the crystal structure of the fully Ca²⁺-loaded form of pike (component pI 4.10) Parv.Ca₂ at atomic resolution (0.91 Å) . The availability of such a high quality structure offers the opportunity to contribute to the definition of the validation tools useful for the refinement of protein crystal structures determined to lower resolution. Besides a better definition of most of the elements in the protein threedimensional structure than in previous studies, the high accuracy thus achieved allows the detection of well-defined alternate conformations, which are observed for 16 residues out of 107 in total. Among them, six occupy an internal position within the hydrophobic core and converge toward two small buried cavities with a total volume of about 60 \AA ³. There is no indication of any water molecule present in these cavities. It is probable that at temperatures of physiological conditions there is a dynamic interconversion between these alternate conformations in an energy-barrier dependent manner. Such motions for which the amplitudes are provided by the present study will be associated with a timedependent remodeling of the void internal space as part of a slow dynamics regime (millisecond timescales) of the parvalbumin molecule. The relevance of such internal dynamics to function is discussed.

Keywords: conformational substates; cryotechniques (100 K) ; crystal structure at atomic resolution; EF-hand parvalbumin; hydrophobic core and internal cavities; microgravity crystallization; slow dynamics and function; validation tools

The EF-hand proteins are a large family of evolutionary related proteins with Ca^{2+}/Mg^{2+} mixed type binding sites (the EF-hand), including a variety of subfamilies, such as parvalbumin (Parv), troponin C (TnC), calmodulin (CaM), sarcoplasmic calciumbinding protein, the regulatory light chains of myosin, the S100, and VIS subfamilies (for reviews: Kawasaki & Kretsinger, 1994; Schäfer & Heizmann, 1996). In such proteins, Ca^{2+}/Mg^{2+} exchange appears closely related to physiological processes that involve cell excitation and relaxation, as in muscle and in neurons (Berchtold et al., 1985; Celio, 1986; Rüegg, 1989). Upon muscle relaxation, as reviewed by Rüegg (1989), parvalbumin may take up the Ca^{2+} ions bound to TnC and CaM, among other calciproteins, because of its high affinity for Ca^{2+} , while Ca^{2+} may bind first to TnC and CaM, because of a slow off-rate of Mg^{2+} from Parv (White, 1988; Hou et al., 1992). The EF-hand corresponds to a helix-loop-helix (HLH) motif of 30 residues in length. Both helices E and F, with 10 residues each, are flanking a central loop that contains all metal-coordinating residues (at the relative positions 1, 3, 5, 7, and 9) with the exception of the invariant Glu12, which lies in helix F. Both helices E and F adopt a nearly perpendicular orientation within the EF-hand. In all cases, two EF-hands form a pair with both loops connected through hydrogen bonding in a short antiparallel β -strand. This is the case of the EF-hand proteins with two EF-hands (parvalbumin: see references above; calbindin D9k: Andersson et al., 1997) or four EF-hands (TnC, CaM). Most of the hydrophobic residues that belong to helices E and F in each EF-hand have their side chains converging toward a core that contributes to the stability of the pair. Parvalbumin dis-

Reprint requests to: Jean-Paul Declercq, Université Catholique de Louvain, Unité CPMC, 1 place Louis Pasteur, B-1348 Louvain-la-Neuve, Belgium; e-mail: declercq@cpmc.ucl.ac.be.

plays a rather specific structural feature among the EF-hand proteins in the sense that its single pair of EF-hands (labeled EF-3 and $EF-4⁵$) is in close contact with an unpaired HLH motif, i.e., the N-terminal AB domain. The latter displays two antiparallel helices A and B flanking a central loop AB with no capacity of cation binding (also labeled $EF-2^5$ as being the remnant EF-hand of a putative ancestral pair EF-1/EF-2 that has been lost during evolution; Kawasaki & Kretsinger, 1994). The AB domain also contributes to the hydrophobic core of parvalbumin with several invariant nonpolar residues.

The internal dynamics of the parvalbumin hydrophobic core with a high content of Phe residues is well documented (Zanotti et al., 1999 and references therein). Based on a temperaturedependent circular dichroism (CD) study of parvalbumin, Parello and Pechère (1971) initially concluded that its hydrophobic core involves different conformational states for the Phe side chains, possibly interconverting through rotations about their C^{α} -C^{β} and C^{β} -C γ bonds. A direct demonstration of the occurrence of such Phe internal motions was afforded by 13 C NMR relaxation (Opella et al., 1974) and by line-shape analysis of the aromatic resonances in the $1H NMR$ spectra (Cavé et al., 1976). Based on such experimental evidence, it was hypothesized that the conformational mobility of the hydrophobic core of parvalbumin might be related to the ability of this EF-hand protein to adapt its tertiary structure according to the chemical nature of the divalent cations that are bound to the protein (Cavé & Parello, 1981). It was predicted that the substitution of Ca²⁺ (with an ionic radius of 0.99 Å) by Mg²⁺ (with an ionic radius of 0.65 Å) would be accompanied by a "contraction" of the coordination sphere around the central cation and that this local contraction would be followed by variations in the structure and dynamics of the protein. The first determination of the crystal structure of parvalbumin with its EF-4 hand substituted by Mg^{2+} , established the correctness of these predictions (Declercq et al., 1991). All metal-coordinating residues come closer to the central cation in EF-4 upon reduction of the ionic radius with a subsequent reduction of the coordination number from 7 (Ca^{2+}) to 6 $(Mg²⁺)$. The totally invariant Glu residue at the relative position 12 in EF-4 undergoes a characteristic conformational rearrangement upon Ca^{2+}/Mg^{2+} exchange so that the glutamyl side chain switches from the $g(+)$ χ_1 rotamer, in the Ca²⁺-loaded form (bidentate ligand associated with heptacoordination), to the $g(-)$ χ_1 rotamer, in the Mg²⁺-loaded form (monodentate ligand associated with hexacoordination). This selective conformational rearrangement is probably linked to the global mobility of the protein (Allouche et al., 1999). However, an integrated view of the dynamics of the hydrophobic core (with regard to amplitudes and frequencies; see Zanotti et al., 1999) and its role on parvalbumin function is not available so far.

We have recently obtained crystals of pike pI 4.10 Parv.Ca₂ grown under microgravity conditions that diffract X rays to a resolution better than 0.9 Å (Declercq et al., 1999). We report here the crystal structure of this parvalbumin form at an effective atomic resolution of 0.91 Å. To our knowledge, the highest resolution reported so far for an EF-hand protein concerned psoriasin with a resolution of 1.05 Å (Brodersen et al., 1998). By using cryogenic conditions (100 K) , we have been able, at this resolution, to visualize well-defined conformational substates of the protein, in particular within its hydrophobic core. The relevance of these substates to internal dynamics is discussed.

Results and discussion

Structure refinement and quality of the model

Starting with the coordinates at 1.75 Å resolution (Declercq et al., 1996) available with identification 1PVB from the Protein Data Bank (PDB) (Bernstein et al., 1977), the structure was refined with a final *R*-factor of 0.110 for all available data to 0.91 Å (see Table 1; Materials and methods); 95.9% of the nonglycine residues are located in the most favored regions of a Ramachandran plot (not shown) and the remaining 4.1% in the additional allowed regions. The trace of the main chain is very close to that at room temperature (Declercq et al., 1996). The root-mean-square deviation (RMSD) between the C^{α} atoms of the two molecules is only 0.24 Å. A large deviation between the two main chains occurs for the C-terminal residue (Ala108), which was extremely disordered at room temperature and is now very well-defined in the low temperature structure. Coordinates and structure factors have been deposited with the PDB (Bernstein et al., 1977) under accession numbers 2PVB and R2PVBSF, respectively.

The validation tools

It is now well established that a protein crystal structure must be validated (Dodson et al., 1998), and a number of programs are developed for checking the proposed structures. At the same time, it appears that most of these procedures are based either on small peptides (Engh $&$ Huber, 1991) or on well-refined protein structures (program OOPS, Kleywegt & Jones, 1996). As pointed out by the EU three-dimensional (3D) validation network (1998), these restraints are not always appropriate for protein structures. Small molecules are not submitted to the same intramolecular constraints as are proteins and "well-refined protein structures" do not always achieve a sufficient resolution to be considered as references. We describe here a protein structure at 0.91 Å resolution, which probably presents the opportunity to improve some of the quality criteria. The program WHATCHECK (Hooft et al., 1996) has classified as "errors" some contacts corresponding to O_{i-1} ...C_i prohibited distances. This concerns residues with $i = 3, 8, 26, 47, 64$ for which this distance is less than the accepted limit of 2.80 Å. The shortest observed distance is 2.71 Å for $i = 3$. All these contacts correspond to ϕ_i angle values between -52 and -58° , which are perfectly allowed. The atoms are very well-defined in the electron density. Figure 1A presents an example (Lys7–Asp8) of such a short contact within the corresponding electron density. To validate our observations with the short O_{i-1} ...C_{*i*} distances, we searched the Cambridge Structural Database (Allen & Kennard, 1993) for small molecules containing the fragment

in which the distance between the first oxygen atom and the last carbon atom should be less than $2.80~\text{\AA}$ (the shortest allowed

⁵The evolutionary based nomenclature proposed by Kawasaki and Kretsinger (1994) for the EF-hand sites is adopted here: the two-site EFhand parvalbumin is labeled with sites EF-3 (initially site CD between helices C and D) and EF-4 (initially site EF between helices E and F).

No. of observations		481,972			
No. of unique reflections		63,698			
Resolution range (A)	Overall $(25.0 - 0.91)$	Lowest shell $(25.0 - 1.98)$	Highest shell $(0.94 - 0.91)$		
Completeness (%)	99.8	99.2	99.0		
$R_{\text{merge}}(I)$ (%)	3.6	3.5	11.1		
$\langle I/\sigma(I)\rangle$	14.2	37.7	3.6		
R -factor (no cut-off)	0.110	0.133	0.132		
Mean estimated errors (A) on atomic positions (full matrix inversion) All nonhydrogen atoms Protein atoms Main-chain atoms Water molecules Ca^{2+} ions	0.024 0.016 0.011 0.054 0.002				
Mean temperature factors (B, \AA^2)					
All nonhydrogen atoms	11.58				
Protein atoms	8.28				
Main-chain atoms	6.94				
Water molecules	25.36				
Ca^{2+} ions	5.64				

Table 1. *X-ray data processing and refinement statistics*

Diffuse solvent correction: the calculated intensities are multiplied by $(1 - g \cdot \exp(-8\pi^2 U(\sin \theta/\lambda)^2))$ with $g = 0.635$ and $U = 2.66 \text{ Å}^2$

distance in WHATCHECK). Only the structures flagged as error free, without disorder and with $R \le 10\%$, were considered. There were 101 hits, and the shortest O_{i-1} ...C_{*i*} contact was 2.59 Å.

The same WHATCHECK program has flagged as "poor" some main-chain conformational angles. This concerns ϕ, ψ angles of residues Asp22 (75°, 9°), Lys54 (61°, 35°), and the ω angles of residues Arg75 and Ala76 (163° and 164°, respectively). Concerning the ϕ,ψ "outliers," it is now well established (Roquet et al., 1992; Declercq et al., 1996) that such values are characteristic of the *fourth* residues of the cation binding loops (residue 54 in EF-3; residue 93 in EF-4), as well as of residue 22 in the AB loop where no cation binding occurs. Residue 93, which corresponds to a totally conserved Gly in parvalbumin, is also characterized by similar ϕ, ψ values (77°, 12°). Again, the electron density is well defined as shown, for example, in Figure 1B, with residue Asp22 fully located within its electron density. The ω angles with values close to 180° reflect the planarity of the peptide groups. It is now well documented (Dodson et al., 1998; Wilson et al., 1998) that in the past this angle was too tightly restrained and that a value like 163° is not at all prohibited even if it does not occur very frequently. The quality of the electron density corresponding to the segment Arg75–Ala76 (Fig. 1C) confirms the validity of this unusual conformation.

One of the checks performed in OOPS (Kleywegt $&$ Jones, 1996) concerns the orientation of the peptide groups in comparison with a database of high-quality protein structures. The result is a number called the "pep-flip" value. Values larger than 2.5 Å are usually suspect and require inspection. In the present case, a very unusual value occurs for residue 72, i.e., the unique Pro of pike 4.10, with 4.9 Å. Again the electron density $(Fig. 1D)$ confirms that the orientation of this peptide is correct but unusual.

In the future, the availability of more protein structures at atomic resolution will allow the improvement of validation tools used in the determination of protein crystal structures to lower resolution.

Description of the crystal structure at atomic resolution

The helical segments

As expected on the basis of previous crystallographic data with parvalbumin at resolutions between 1.5 and 1.75 Å (Swain et al., 1989; Kumar et al., 1990; Declercq et al., 1991, 1996; Roquet

Fig. 1. Protein model presented with its electron density in some of the regions with unusual features. **A:** Short contact between Lys7 O and Asp8 C. **B:** The main-chain conformation of residue Asp22 is characterized by the unexpected ϕ, ψ angles (75°, 9°, respectively). **C:** The planarity of the peptide bond between Arg75 and Ala76 is not conserved, with an ω value as low as 163°; a strict respect of the planarity would push the atoms outside the electron density. **D:** The residue 72 shows a "pep-flip" value as large as 4.9 Å. The alternate conformations of the side chain of Pro72 (pyrrolidine ring) are also observed. The sigma-A map (Read, 1986) contoured at 2.5 σ (1.75 e \cdot Å⁻³) confirms that the tracing is correct in all these situations showing unusual features (figures produced using the program O, Jones et al., 1991).

et al., 1992), all helices adopt a regular α -fold, with the exception of helix D, which displays a kink in its central part and also displays a 310 helical conformation in its C-terminal region. Most of the amide NH hydrogen atoms that are involved in hydrogen bonding within the parvalbumin helices are clearly identified on the difference electron density map computed without contribution of the hydrogen atoms to the structure factors. Helix A is taken as an example thus showing the location of most of its NH hydrogen atoms (Fig. 2). Subtle distortions from ideality appear at the level of helices A, B, C, and E, i.e., a regular, although slight, bending of the whole helix with its convex face turned toward the exterior of the protein (J.-P. Declercq, C. Evrard, J. Parello, unpubl. results).

The occurrence of a kinked helix D is characteristic of all parvalbumin tertiary structures. The kink angle (about 125°) appears highly conserved in all D helices. It is important to include in this geometrical description of helix D, the water site $W64/67/75$ (the

Fig. 2. A view of the regularly folded α -helix A with the corresponding difference electron density contoured at 3.0σ . The structure factors were computed without the contribution of the H atoms and the difference electron density map clearly shows the location of different hydrogen bonded NH hydrogen atoms. All the heavy atoms and bonds are in black. The NH bonds of the amide groups are in grey (figure produced using the program O, Jones et al., 1991). Distances (dotted) between H and the acceptor atom are indicated.

numbers specify the residues displaying a possible hydrogen bond with the water molecule) that corresponds to an invariant water site in parvalbumin (for a definition see Roquet et al., 1992). This water molecule is very well defined with a small atomic displacement parameter $(B = 5.75 \text{ Å}^2$; see Table 1). The characteristics of such a water-mediated hydrogen bond network in site $W64/67/75$ are given in Table 2, among other parvalbumin invariant water sites. Hydrogen bonding within these invariant water sites displays a narrow range of donor-acceptor lengths (ranging from 2.78 to 2.92 Å; see Table 2), in contrast with the lengths observed within the protein helices (ranging from 2.84 to 3.38 Å). This situation probably translates the fact that water-mediated hydrogen bonding allows a good optimization of the atomic distribution within the hydrogen bond networks due to the fact that water is less constrained than protein elements. The total number of ordered solvent molecules in the low temperature crystal structure of pike 4.10 Parv.Ca2 amounts to 211 instead of 72 in the crystal structure at ambient temperature (Declercq et al., 1996). We provide no detailed description here of the molecular arrangement of the solvation layer (J.-P. Declercq, C. Evrard, J. Parello, unpubl. results).

Table 2. *Hydrogen bonding in the invariant parvalbumin water sites (see Roquet et al., 1992 for definition) as visualized in the present study*^a

Water site	HOH number	Atomic displacement parameters (B, \mathring{A}^2)	X_i	d (Ow – X_i) (\AA)	
W7/33	216	9.00	Lys7 N	2.92	
			Val33 O	2.80	
W50/62	204	6.65	$Ile50$ O	2.86	
			Glu62O	2.82	
W60/95	224	10.31	$Glu60$ N	2.89	
			Gly95O	2.78	
W64/67/75	202	5.75	$Lys64$ O	2.84	
			Leu 67 N	2.80	
			Arg 75 O	2.80	
W89/101	203	7.26	Gly89O	2.79	
			$Glu101$ O	2.78	

^aEach site is defined by the residues at hydrogen bond distance $(X_i =$ donor or acceptor atom) with the water molecule $(0w = w \cdot w \cdot \text{water oxygen}).$

Although most, if not all, of the solvent loci probably include water molecules, it cannot be excluded that some of them are occupied by ammonium ions (see Materials and methods).

The Ca^{2+} *-binding sites*

No restraints on the geometry of the coordination spheres of the $Ca²⁺$ ions were introduced during the refinement. As shown in Table 3, both coordination spheres around the parvalbumin-bound $Ca²⁺$ ions in the primary sites EF-3 and EF-4 correspond to heptacoordination of Ca^{2+} , as initially established (Declercq et al., 1988), with the seven oxygen atoms occupying the apexes of a pentagonal bipyramid (Strynadka & James, 1989; Swain et al., 1989). The tops of the pyramids are occupied by Asp51- O^{δ_1} and Glu59-O^{ϵ 1} in EF-3 and by Asp90-O^{δ 1} and a water molecule in EF-4, along the X and $-X$ directions, if one refers to the traditional description of calcium coordination polyhedra (Kretsinger $& \text{Nockolds}, 1973$. The five remaining oxygen atoms in the coordination spheres constitute the pentagonal basis, which is not a regular pentagon and adopts a skewed conformation. Mean planes were computed through the five atoms forming the pentagonal bases. RMSDs of 0.21 and 0.40 Å are, respectively, observed for the planes through the EF-3 and EF-4 sites, thus showing a larger deformation of the EF-4 site. Both Ca^{2+} ions are very close to these planes $(0.04$ and 0.02 Å, respectively). For a quantitative description of the deformation of the coordinating polyhedra, a regular pentagonal bipyramid was built, based on the mean Ca-O observed distance, 2.385 Å. The best superimposition between the apexes of this regular polyhedron and the seven coordinating oxygen atoms was searched and the distances between each oxygen atom in the coordination sphere and the corresponding apex was computed (see Table 3). Not surprisingly, the largest deviations concern the bidentate ligands (Glu62 in EF-3, Glu101 in EF-4), since the relative positions of their two oxygen atoms are constrained by the geometry of the carboxylate group itself. Again, it appears that the EF-4 site, with a mean deviation of 0.50 Å, is significantly more skewed than the EF-3 site characterized by a value of 0.33 Å. If one compares the individual Ca-O distances, it

^aAll the distances are in Å. Ca-O, distance between the Ca²⁺ ion and the coordinating oxygen atom; d(pla), distance between the oxygen atom and the mean plane through the pentagonal base of the bipyramid (* indicates that the atom was used to define the plane); d(bip), distance between the oxygen atom and the corresponding apex of an idealized pentagonal bipyramid. The indications $(\pm X, \pm Y,$ $\pm Z$) refer to the traditional description of calcium coordination polyhedra (Kretsinger & Nockolds, 1973).

is observed that all the distances concerning a monodentate carboxylate side chain (Asp, Glu) are among the shortest $(2.27–R)$ 2.36 Å) while substantially larger values occur for the bidentate Glu ligands $(2.43-2.54 \text{ Å})$. The significance of the accuracy thus achieved in the description of the Ca^{2+} -binding sites of parvalbumin needs to be assessed by theoretical simulations in comparison to those reported recently using the pI 4.10 Parv.Ca₂ crystal structure at 1.75 Å resolution and at ambient temperature (Allouche et al., 1999).

The conformational substates

Several residues in the tertiary structure of pike 4.10 Parv.Ca₂ at 100 K unambiguously appear as alternate conformational states, essentially at the level of their side chains (see Materials and methods). In all cases these residues can be described by two distinct conformational states. They occupy external as well as internal positions in the tertiary structure. As an example, Figure 3 presents helix B for which three residues, Lys27, Phe30, and Val33, display two alternate conformations each. In most cases, one of the two distinct conformations observed at 100 K nearly coincides with the single conformation observed at room temperature in the same crystal form (Declercq et al., 1996). This is not the case of Phe30. As shown in Figure 4, the single conformation of its side chain, at room temperature, is intermediate between the two substates observed at 100 K. This suggests that in the high-temperature crystal structure the conformation of the Phe30 side chain is timeaveraged between two (at least) conformations. The conformational polymorphism observed at the level of five external lysyl residues (among 12 Lys in total) essentially involves the regions of their side chains close to the ammonium heads (C^{δ} , C^{ϵ} , and N^{ζ} atoms). The rotameric states of externally located residues with two defined substates, including Ser, Glu, and Lys, will not be

Fig. 3. Stereoscopic view of helix B with the sigma-A (Read, 1986) electron density contoured at 2.5 σ (1.75 e \cdot Å⁻³). In this helix, three side chains display alternate conformations $(Lys27, Phe30, Val33)$ (figure produced using the program O, Jones et al., 1991).

Fig. 4. Superimposition of Phe30 at room temperature (grey) and at 100 K (black). The single conformation observed at room temperature is intermediate between the two alternate conformations observed at low temperature (figure produced using the program O, Jones et al., 1991).

discussed further since their exact conformational features could depend on crystal packing effects, which are not taken into consideration here. The rotameric states of internally located residues, with two well-defined substates, are of the highest interest for understanding the packing of the hydrophobic core, as well as its internal dynamics by giving access to the amplitudes of internal motions when these motions are frozen out. As shown in Table 4, among the six internal residues with two substates, i.e., Phe30, Val33, Leu35, Ile50, Ile97, Met105, three of them (Phe30, Ile50, and Ile97) retain the stable χ_1 rotamer $g(+)$ as a common feature in both substates and allow for other torsional angles to be subtly adapted. In contrast, among the remaining three residues, i.e., Val33, Leu35, and Met105, each differs by its χ_1 rotamers in both substates. It is probable that at temperatures of physiological conditions, parvalbumin will undergo dynamic interconversions between the alternate substates of different residues. The interconversion between $g(+)$, $g(-)$ and $t \chi_1$ rotamers will involve rotations about single bonds with energy barriers as those observed in substituted ethane molecules. Such barriers lie in the 6 kcal/mol range (Orville-Thomas, 1974). Since the structural elements (main and side chain) that undergo such dynamic interconversions are embedded in the protein "matrix," many other contacts are certainly at play and will contribute to the energy barriers, which could then be higher than what is known from the internal rotation processes in small molecules. As initially established by Gelin and Karplus (1975), based on a theoretical approach ("flexible geometry barriers"), each rotation barrier in a globular protein will depend on the immediate spatial environment (van der Waals contacts) around the mobile structural elements. If one considers these internal interconversions as unimolecular processes following transition state theory (see Equation 2 in Gelin & Karplus, 1975), expected barriers in the range 7–12 kcal/mol will correspond to reorientation rate constants in the range $10^{7}-10^{4}$ s⁻¹, at ambient temperature. It is probable that the dynamic interconversions in the hydrophobic

	A			$\, {\bf B}$				
	χ_1	χ_2	χ_3	Occ.	χ_1	χ_2	χ_3	Occ.
Phe30	-60 $g(+)$	-28		0.41	-76 $g(+)$	-43		0.59
Val ₃₃	-63 $g(+)$			0.58	93 $g(-)^*$			0.42
Leu35	-82 $g(+)^\ast$	165		0.77	-147 \boldsymbol{t}	40		0.23
Ile50	-53 $g(+)$	-59		0.50	-66 $g(+)$	165		0.50
Ile97	-70 $g(+)$	174		0.38	-57 $g(+)$	-60		0.62
Met105	-84 $g(+)^*$	-174	-67	0.67	7 $g(-)^*$	-137	-13	0.33

Table 4. *Conformational angles (*8*) and occupation factors of the side chains of internal residues having two alternate conformations (A and B)* ^a

^aThe columns labeled "Occ." include the occupation factors of the substates A and B, with a mean estimated error of 0.02. The symbol $*$ denotes a significant departure of χ_1 from the staggered conformation at -60° and 60° for $g(+)$ and $g(-)$, respectively.

core of parvalbumin will correspond to such a slow dynamics range, in agreement with previous NMR results indicating that most of the Phe residues undergo $\pm 180^\circ \chi_2$ flipping motions in the milli- or submillisecond timescales, at ambient temperature (Cavé et al., 1976; Cavé & Parello, 1981). At 100 K in the crystal state, these interconversions are sufficiently slowed down to give rise to discrete substates of the protein contributing to the scattering of X rays.

The observed conformational substates for the internal residues involve not only structural elements from the side chain of a given residue (see Table 4), but also structural elements from the main chain itself. In the case of Phe30, the β -carbons are clearly distinct in both substates (Figs. 3, 4). This situation cannot be accounted for by varying only the χ_1 and χ_2 angles of Phe30 while the adjacent main chain remains strictly unperturbed. However, even at atomic resolution, as achieved in this work, the structural elements of the main chain cannot be described with sufficient accuracy so that different main-chain substates could not be inferred from the examination of the electron density.

The identification of "frozen" conformational substates is of paramount importance for a better understanding of the effect of internal dynamics on the capacity of the protein to select Ca^{2+} vs. Mg^{2+} . The ratio of the binding constants, K_{Ca}/K_{Mg} , is in the range 10^3 – 10^4 and appears to be invariant during parvalbumin evolution (Kawasaki & Kretsinger, 1994). It has been hypothesized that adaptation of parvalbumin conformation to cation size, as well as a concomitant change in its internal dynamics, are crucial to the function of the protein that probably alternates between both Caand Mg-loaded states under physiological conditions (Cavé & Parello, 1981). The adaptative conformational features of parvalbumin upon Ca^{2+}/Mg^{2+} exchange are documented on an experimental (Declercq et al., 1991) and theoretical (Allouche et al., 1999) basis. Our understanding of the dynamic features that underlie such a conformational adaptation is not similarly documented. The present work offers for the first time a rather detailed description of the amplitudes associated with the dynamics in the parvalbumin core of Parv.Ca₂. A parallel study with Mg-loaded parvalbumin appears of the highest interest.

In total a very large number of distinct rotamers, at least 2^{16} $(65,696)$, occurs in parvalbumin, if one takes into account all the residues that are unambiguously defined by two distinct states in the crystal structure of pike 4.10 Parv.Ca₂ at $100 K$ (see above). Each substate for a given residue with two alternate conformations has been arbitrarily labeled A or B in the deposited atomic coordinates 2PVB. As pointed by Frauenfelder (1995), one of the challenges in structure and dynamics of proteins is to establish a connection between the conformational substates of the protein and its function. The conformational substates that are characterized individually on the atomic resolution electron density of pike 4.10 Parv.Ca₂ at 100 K probably correspond to "taxonomic" substates, as defined by Frauenfelder (1995), based on energetics considerations (energy-barrier dependent interconversions). In its turn a "taxonomic" substate will exist in a very large number (probably in the $10⁴$ range and above) of statistical substates, which display only minor differences in their properties and cannot be identified individually and must be described by distributions. The identification of such "taxonomic" substates in parvalbumin by atomic resolution X-ray crystallography at low temperature offers a novel framework for extending our computational analyses of this typical EF-hand protein (Allouche et al., 1999). It is not presently known if all the conformational substates identified so far occur independently or display a given degree of correlation.

The packing of the hydrophobic core

In an attempt to characterize the exact packing of the hydrophobic core, we searched for internal cavities. No internal cavity filled with exogenous components (water molecules) has been found so far in parvalbumin based on high-resolution X-ray crystallography evidence (see Roquet et al., 1992 and references therein). This is also the case for the present crystal structure determined at atomic resolution and at low temperature so that no exogenous molecules are observed within the hydrophobic core. However, using the program VOIDOO (Kleywegt $&$ Jones, 1994), there is evidence that the parvalbumin hydrophobic core contains defined cavities although small in size. Two small cavities were identified in the core using a probe radius of 1.0 Å. For carrying out such an analysis, we used the time-averaged state of $Parv.Ca₂$ with the same crystal form determined at room temperature (Declercq et al., 1996), rather than the 64 possible distinct conformations of the hydrophobic core, which are defined at 100 K, taking into account the six internal residues with two well-defined substates A and B. As shown on Figure 5, both cavities are close to each other and the volume, which can be occupied by the probe, corresponds to 30.7 and 31.5 \AA^3 , respectively. No water molecules are observed inside the cavities though a volume of 30 \AA ³ seems sufficient to accommodate maybe four of them. However, is there still the possibility for a few disordered internal water molecules that would not contribute to X-ray scattering? So far we have no firm answer. As presented in Figure 5, these cavities are delineated by side chains from the following residues: Phe2, Phe30, Leu35, Ala46, Ile50, Ile58, Leu63, Phe66, Phe70, Phe85, Ile97, Phe102, Met105, and Ile106. Five of the six internal residues displaying alternate conformations at low temperature (Table 4) are in this list, and the side chain of Phe30 occupies a central position between both cavities. It is possible that the connection between a buried side chain and an internal empty space is related to the possibility of this side chain alternating between different conformations, at a relatively low energetic cost.

At low temperature $(<150 K)$, internal motions correspond to an harmonic regime, as shown by a study of hydrated Parv.Ca2 powders by inelastic neutron scattering in the 10–320 K range (J. Parello, unpubl. results). All individual substates observed at 100 K in the crystal structure of parvalbumin probably correspond to single local conformations, which are no more time-averaged with the exception of still rapid librational and vibrational motions, in agreement with the hierarchical description of protein dynamics (Nienhaus et al., 1997). The fact that two conformations are observed at 100 K for several side chains of the parvalbumin molecule, in particular within the compact hydrophobic core, indicates that the protein can explore a relatively large conformational space.

Our present crystallographic results with pI 4.10 Parv.Ca₂ at atomic resolution, in combination with other experimental observations with parvalbumin in solution or in hydrated powders (Zanotti et al., 1999 and references therein), suggest that well-defined regions in the protein tertiary structure display selective dynamic properties (wobbling loci). As shown in Table 4, Phe30 is represented nearly equally by its two substates (occupation factors of 0.41 and 0.59), and this is also the case of Ile50, in contrast to other internal residues that display more distinct occupation factors between both substates. A tentative conclusion would be that a direct link could exist between the internal mobility of the protein core and the dynamic properties of the protein as a whole, including the cation-binding loops themselves, so that slow dynamics in the millisecond range could be directly connected to the kinetics of

Fig. 5. Stereoscopic view of the two internal cavities in the hydrophobic core of pike 4.10 Parv.Ca₂ at room temperature (1PVB, Declercq et al., 1996) inferred from the program VOIDOO (Kleywegt & Jones, 1994) showing the side chains of the residues in contact with these cavities.

 Ca^{2+} and Mg²⁺ capture and release, as well as to Ca^{2+}/Mg^{2+} exchange. In agreement with this view, kinetic constants in the range of 500 s^{-1} at ambient temperature have been inferred for $Ca²⁺$ release from parvalbumin, whereas lower constants have been determined for Mg^{2+} (White, 1988).

Materials and methods

Crystallization and data collection

The crystal used for structure determination was grown under microgravity conditions during the flight STS-83 of the NASA space shuttle, by sitting drop vapor diffusion, as previously reported (Declercq et al., 1999). Before data collection, the crystal with dimensions $0.12 \times 0.12 \times 1.0$ mm was transferred for 1 min in the crystallization solution but containing 20% of ethylene glycol as a cryoprotectant and flash-cooled at 100 K. The crystal is orthorhombic, $P2_12_12_1$ with $a = 51.03$, $b = 49.81$, $c = 34.57$ Å, and $Z = 4$. X-ray diffraction data were collected at 100 K on the EMBL X11 beam line at the DORIS storage ring (DESY, Hamburg, Germany) with a 30 cm Mar-Research (Hamburg, Germany) imaging plate scanner and a wavelength of 0.9096 Å. A resolution of 0.91 Å (a geometrical limit of the station at that time) was achieved and, as previously reported (Declercq et al., 1999), the crystal was capable of diffracting to a higher resolution. For measurement of the overloaded low and medium resolution data, three complete sets were collected from the same crystal at resolutions of 0.91, 1.3, and 1.8 Å, yielding a total of 481,972 reflections. All these measurements were indexed and integrated using the program DENZO and then merged with the program SCALEPACK (Otwinowski & Minor, 1997). A summary of data statistics is given in Table 1. Five percent of the reflections were flagged for use in R_{free} calculations (Brünger, 1992b).

Structure refinement

Since the same crystal form had been previously analyzed at room temperature and at a resolution of 1.75 Å (Declercq et al., 1996), the corresponding coordinates with accession number 1PVB in the PDB (Bernstein et al., 1977) were used to initiate the refinement. In this model, all the solvent molecules were removed but the two calcium ions were retained. During the first steps, the low resolution reflections (below 8.0 Å) were not included. A rigid body refinement with X-PLOR (Brünger, 1992a) was followed by a simulated annealing refinement using the slow-cool protocol provided by the same program. A starting temperature of 3,000 K was gradually decreased to 300 K in intervals of 25 K during which 50 steps of 0.5 fs of molecular dynamics were run. This process was followed by minimization. For the rest of the refinement, all the data available between 25.0 and 0.91 Å were included and conjugate gradient least squares with the program SHELXL-97 (Sheldrick $&$ Schneider, 1997) were applied, alternating with visual examinations and rebuilding with O (Jones et al., 1991). Together with the inclusion of low resolution data, a diffuse solvent correction was introduced. Restraints based on the study of Engh and Huber (1991) were applied and the details are given in Table 1. Anisotropic temperature factors were refined for all atoms with the restraints given in Table 1. A rigid bond restraint imposes approximately equal displacement parameters in the direction of the bonds $(1–2$ and $1–3$ distances) while the similar anisotropic displacement parameter restraint imposes approximately equal U_{ii} components for atoms that are spatially close. Solvent molecules were progressively included using first the automated water divining provided in SHELXL-97 (Sheldrick & Schneider, 1997) and later the ARP procedure (Lamzin & Wilson, 1993; Lamzin & Wilson, 1997) providing at the end a total of 211 solvent molecules. Since the parvalbumin was crystallized in the presence of 2.4 M $(NH_4)_2SO_4$ and pure water is ~55 M H₂O, the molar ratio NH_4^+/H_2O is

 \sim 1/12. It can thus be expected that some of these solvent sites are occupied by NH_4^+ instead of H_2O , but except in one case discussed below, we did not find good criteria for identifying the NH₄⁺ ions, if any, and all the solvent molecules were introduced as H_2O . No SO_4^2 ⁻ ions were observed. During the refinement, the water molecules were forced to remain approximately isotropic (approximately isotropic restraint). According to the arguments developed previously (Declercq et al., 1988, 1991, 1996), a third cation binding site should be occupied in the present conditions by an ammonium ion. Suitable electron density appeared at the expected position, in contact with the polar side chains of Asp53, Glu59, and Asp61, and was thus fitted with an ammonium ion. As recently determined by Zanotti et al. (1999), a full coverage of the parvalbumin "surface" by water is achieved with ca. 300 water molecules. The low-temperature crystal structure described here thus accounts for a substantial population (up to two thirds) of the water molecules predicted to occur upon full coverage of the protein "surface." These sites certainly correspond to water molecules firmly bound to the protein in solution besides water molecules interacting more loosely with the rest of the protein. During the manual examinations with O (Jones et al., 1991), it clearly appeared that alternate conformations had to be considered for the side chains of a number of residues: Ser1, Lys27, Phe30, Val33, Leu35, Ser37, Ile50, Glu60, Lys64, Pro72, Lys83, Lys91, Met96, Ile97, Met105, Lys107. These alternate conformations were refined with the constraint that the sum of their occupancies equals 1.0. During examination of difference Fourier syntheses, electron density was observed for nearly all the hydrogen atoms of the amide groups and also for some other hydrogen atoms. All the protein hydrogen atoms were thus introduced on the basis of an idealized stereochemistry and refined using a "riding" model. These hydrogen atoms were kept isotropic with the temperature factor set to 1.2 times (or 1.5 times for OH and methyl) the temperature factor of the parent atom. At the end of the refinement, there remained two large peaks (\sim 1.5 e \cdot Å⁻³), about twice larger than the next ones. Each of them was close to two water oxygen atoms (distance peak-O \sim 1.5 Å and angle O-peak-O \sim 115°) and was not at hydrogen bond distances of any other atom, suggesting something like a formate ion. In the absence of such an anion, but in the presence of EDTA in the crystallization buffer (Declercq et al., 1999), we envisaged the possibility that the "formate-type" densities could be two tails of an EDTA molecule whose remaining atoms are disordered and not observed. Since no interpretation is quite satisfactory, the two entities were kept as formate ions but no constraints were applied to the interatomic distances and angles. At the end of the refinement, the programs PROCHECK (Laskowski et al., 1993) and WHATCHECK (Hooft et al., 1996) were used to identify and correct some minor local problems. The *R*-indices for all data between 25.0 and 0.91 Å is $R = 0.109$ and $R_{\text{free}} = 0.132$, and the electron density is well defined in all parts of the molecule. During the very final cycles of refinement, the R_{free} flagged reflections were included and the final *R*-value for all available reflections is 0.110 (no cut-off). The mean errors on the atomic positions were estimated by full matrix inversion and are reported in Table 1.

Acknowledgments

J.P.D. and C.E. are indebted to the Fonds National de la Recherche Scientifique (Belgium) and to the Fonds de Développement Scientifique (Université Catholique de Louvain) for financial support. J.P. is indebted to the Centre National de la Recherche Scientifique (Paris, France) and to the Burnham Institute (Adjunct Professor 1993–1998, La Jolla, California). The use of the EMBL $X11$ beamline at the DORIS storage ring (DESY, Hamburg, Germany) is gratefully acknowledged. We thank the European Union for support of the work at EMBL Hamburg through the HCMP to Large Installations Project, contract CHGE-CT93-0040. We thank NASA for providing the opportunity of growing crystals under microgravity conditions.

References

- Allen FH, Kennard O. 1993. 3D search and research using the Cambridge Structural Database. *Chem Design Automation News 8*:31–37.
- Allouche D, Parello J, Sanéjouand YH. 1999. Ca^{2+}/Mg^{2+} exchange in parvalbumin and other EF-hand proteins. A theoretical study. *J Mol Biol 285*:857–873.
- Andersson M, Malmendal A, Linse S, Ivarsson I, Forsén S, Svensson LA. 1997. Structural basis for the negative allostery between Ca^{2+} - and Mg^{2+} -binding in the intracellular Ca²⁺-receptor calbindin D9k. *Protein Sci* 6:1139–1147.
- Berchtold MW, Celio MR, Heizmann CW. 1985. Parvalbumin in human brain. *J Neurochem 45*:235–240.
- Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer based archival file for macromolecular structures. *J Mol Biol 112*:535–542.
- Brodersen DE, Etzerodt M, Madsen P, Celis JE, Thogersen HC, Nyborg J, Kjeldgaard M. 1998. EF-hands at atomic resolution: The structure of human psoriasin (S100A7) solved by MAD phasing. *Structure 6:477*-489.
- Brünger AT. 1992a. *X-PLOR version 3.1, A system for X-ray crystallography and NMR*. New Haven, CT: Yale University Press.
- Brünger AT. 1992b. Free *R* value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature 355*:472–475.
- Cavé A, Dobson CM, Parello J, Williams RJP. 1976. Conformation mobility within the structure of muscular parvalbumins. An NMR study of the aromatic resonances of phenylalanine residues. *FEBS Lett 65*:190–194.
- Cavé A, Parello J. 1981. Dynamic aspects of the structure of globular proteins by high resolution NMR spectroscopy. The fluid-like structure of the internal hydrophobic core of muscular parvalbumins. In: Balian R et al., eds. *Les Houches, Session XXXIII 1979. Membranes et Communication intercellulaire*. Amsterdam: North Holland Publishing Company. pp 197–227.
- Celio MR. 1986. Parvalbumin in most γ -aminobutyric acid-containing neurons of the rat cerebral cortex. *Science 231*:995–996.
- Declercq JP, Evrard C, Carter DC, Wright BS, Etienne G, Parello J. 1999. A crystal of a typical EF-hand protein grown under microgravity diffracts X-rays beyond 0.9 Å resolution. *J Cryst Growth 196*:595–601.
- Declercq JP, Tinant B, Parello J. 1996. X-ray structure of a new crystal form of pike 4.10 β parvalbumin. *Acta Crystallogr Sect D 52*:165–169.
- Declercq JP, Tinant B, Parello J, Etienne G, Huber R. 1988. Crystal structure determination and refinement of pike 4.10 parvalbumin (minor component from *Esox lucius*!. *J Mol Biol 202*:349–353.
- Declercq JP, Tinant B, Parello J, Rambaud J. 1991. Ionic interactions with parvalbumins. Crystal structure determination of pike 4.10 parvalbumin in four different ionic environments. *J Mol Biol 220*:1017–1039.
- Dodson EJ, Davies GJ, Lamzin VS, Murshudov GN, Wilson S. 1998. Validation tools: Can they indicate the information content of macromolecular crystal structures? *Structure 6*:685–690.
- Engh R, Huber R. 1991. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr Sect A 47*:392–400.
- Frauenfelder H. 1995. Complexity in proteins. *Nature Struct Biol 338*:623–624.
- Gelin BR, Karplus M. 1975. Sidechain torsional potentials and motion of amino acids in proteins: Bovine pancreatic trypsin inhibitor. *Proc Natl Acad Sci USA 72*:2002–2006.
- Hooft RWW, Vriend G, Sander C, Abola EE. 1996. Errors in protein structures. *Nature 381*:272.
- Hou TT, Johnson JD, Rall JA. 1992. Effect of temperature on relaxation rate and Ca²⁺, Mg²⁺ dissociation rates from parvalbumin of frog muscles. *J Physiol 449*:399–410.
- Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Crystallogr Sect A 47*:110–119.
- Kawasaki H, Kretsinger RH. 1994. Calcium-binding Proteins 1: EF-hands. *Protein Profile 1*:343–517.
- Kleywegt GJ, Jones TA. 1994. Detection, delineation, measurement and display of cavities in macromolecular structures*. Acta Crystallogr Sect D 50*:178–185.
- Kleywegt GJ, Jones TA. 1996. Efficient rebuilding of protein structures*. Acta Crystallogr Sect D 52*:829–832.
- Kretsinger RH, Nockolds CE. 1973. Carp muscle calcium-binding protein. II. Structure determination and general description. *J Biol Chem 248*:3313–3326.
- Kumar VD, Lee L, Edwards BFP. 1990. Refined crystal structure of calciumliganded carp parvalbumin 4.25 at 1.5 Å resolution. *Biochemistry 29*:1404– 1412.
- Lamzin VS, Wilson KS. 1993. Automated refinement of protein models. *Acta Crystallogr Sect D 49*:129–147.
- Lamzin VS, Wilson KS. 1997. Automated refinement for protein crystallography. *Methods Enzymol 277*:269–305.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK. *J Appl Crystallogr 26*:283–291.
- Nienhaus GU, Müller JD, McMahon BH, Frauenfelder H. 1997. Exploring the conformational landscape of proteins. *Physica D 107*:297–311.
- Opella SJ, Nelson DJ, Jardetzky O. 1974. Carbon magnetic resonance study of the conformational changes in carp muscle calcium binding parvalbumin*. J Am Chem Soc 96*:7157–7159.
- Orville-Thomas WJ. 1974. In: Orville-Thomas WJ, ed. *Internal rotation in molecules*. London, New York: John Wiley & Sons. Chap. 1. pp 1–18.
- Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol 276*:307–326.
- Parello J, Pechère JF. 1971. Conformational studies on muscular parvalbumins. I. Optical rotatory dispersion and circular dichroism analysis. *Biochimie 53*:1079–1083.
- Read RJ. 1986. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr Sect A 42*:140–149.
- Roquet F, Declercq JP, Tinant B, Rambaud J, Parello J. 1992. Crystal structure of the unique component from muscle of the leopard shark ~*Triakis semi* $fasciata$). The first X-ray study of an α -parvalbumin. *J Mol Biol 223:705–720*. Rüegg JC. 1989. *Calcium in muscle activation*. Berlin: Springer Verlag.
- Schäfer BW, Heizmann CW. 1996. The S100 family of EF-hand calcium-
- binding proteins: Functions and pathology. *Trends in Biochemical Sciences 21*:134–140.
- Sheldrick GM, Schneider TR. 1997. SHELXL: High-resolution refinement. *Methods Enzymol 277*:319–343.
- Strynadka NCJ, James MNG. 1989. Crystal structures of the helix-loop-helix calcium-binding proteins. *Annu Rev Biochem 58*:951–998.
- Swain AL, Kretsinger RH, Amma EL. 1989. Restrained least squares refinement of native (calcium) and calcium-substituted carp parvalbumin using X-ray crystallographic data at 1.6 Å resolution. *J Biol Chem 264*:16620–16628.
- White HD. 1988. Kinetic mechanism of calcium binding to whiting parvalbumin. *Biochemistry 27*:3357–3365.
- Wilson KS, Butterworth S, Dauter Z, Lamzin VS, Walsh M, Wodak S, Pontius J, Richelle J, Vaguine A, Sander C, et al. (EU 3-D Validation Network). 1998. Who checks the checkers? Four validation tools applied to eight atomic resolution structures. *J Mol Biol 276*:417–436.
- Zanotti JM, Bellissent-Funel MC, Parello J. 1999. Hydration-coupled dynamics in proteins studied by neutron scattering and NMR. The case of the typical EF-hand calcium-binding parvalbumin. *Biophys J 76*:2390–2411.