X-ray crystal structure and molecular dynamics simulations of silver hake parvalbumin (Isoform B)

ROBERT C. RICHARDSON,¹ NANCY M. KING,¹ DANIEL J. HARRINGTON,² HONGMAO SUN,¹
WILLIAM E. ROYER,² AND DONALD J. NELSON¹

¹Gustaf H. Carlson School of Chemistry, Clark University, Worcester, Massachusetts 01610

²Department of Biochemistry and Molecular Biology, University of Massachusetts Medical School,
Worcester, Massachusetts 01655

(RECEIVED August 30, 1999; FINAL REVISION November 5, 1999; ACCEPTED November 5, 1999)

Abstract

Parvalbumins constitute a class of calcium-binding proteins characterized by the presence of several helix-loop-helix (EF-hand) motifs. In a previous study (Revett SP, King G, Shabanowitz J, Hunt DF, Hartman KL, Laue TM, Nelson DJ, 1997, *Protein Sci* 7:2397–2408), we presented the sequence of the major parvalbumin isoform from the silver hake (*Merluccius bilinearis*) and presented spectroscopic and structural information on the excised “EF-hand” portion of the protein. In this study, the X-ray crystal structure of the silver hake major parvalbumin has been determined to high resolution, in the frozen state, using the molecular replacement method with the carp parvalbumin structure as a starting model. The crystals are orthorhombic, space group C222₁, with $a = 75.7$ Å, $b = 80.7$ Å, and $c = 42.1$ Å. Data were collected from a single crystal grown in 15% glycerol, which served as a cryoprotectant for flash freezing at -188 °C. The structure refined to a conventional R -value of 21% (free R 25%) for observed reflections in the range 8 to 1.65 Å [$I > 2\sigma(I)$]. The refined model includes an acetylated amino terminus, 108 residues (characteristic of a β parvalbumin lineage), 2 calcium ions, and 114 water molecules per protein molecule. The resulting structure was used in molecular dynamics (MD) simulations focused primarily on the dynamics of the ligands coordinating the Ca^{2+} ions in the CD and EF sites. MD simulations were performed on both the fully Ca^{2+} loaded protein and on a Ca^{2+} deficient variant, with Ca^{2+} only in the CD site. There was substantial agreement between the MD and X-ray results in addressing the issue of mobility of key residues in the calcium-binding sites, especially with regard to the side chain of Ser55 in the CD site and Asp92 in the EF site.

Keywords: calcium-binding proteins; EF-hand; molecular dynamics; parvalbumin; silver hake; X-ray structure

Parvalbumins are among one of the 32 subclasses within the EF-hand superfamily of calcium-binding proteins all sharing the common helix-loop-helix (EF-hand) motif (Kretsinger, 1975; Kretsinger & Nelson, 1976). A high degree of homology exists for the primary and secondary structures for members of this subclass. Parvalbumins are small (10–12 kDa), acidic (pI 4.0–5.2), Ca^{2+} - and Mg^{2+} -binding proteins, particularly abundant in the muscle tissue of lower vertebrates. Parvalbumins have been found in nonmuscle tissues (Kagi et al., 1987), in endocrine glands (Endo et al., 1985), in skin cells (MacManus et al., 1985), and in specific neurons (Pfyffer et al., 1987).

Proposed functions for parvalbumins include involvement in the muscle relaxation/contraction cycle (Jiang et al., 1996), calcium buffering (Pauls et al., 1996), and signal transduction (Chard et al., 1993). Other roles proposed for parvalbumins include involvement

in the triggering of gene expression (Chard et al., 1993), in cell division (Rasmussen & Means, 1989), in processes influencing cell shape and motility (Andressen et al., 1995), and in immune system development (Brewer et al., 1989).

From sequence and structural analyses of members of the parvalbumin subfamily, two distinct phylogenetic lineages (α and β) have been identified in lower and higher vertebrates. β -Parvalbumins include a shortened C-terminal end and conserved sequence-specific residues. A cysteine at residue 18 and an aspartic acid at residue 61 are characteristic of β -parvalbumins (Kretsinger, 1980). The two lineages differ in other characteristics, including isoelectric points (β : pI < 5.0; α : pI > 5.0), metal ion binding affinities, crystal structures, cell-type specific expression, chromosomal localization, and physiological functions (Fohr et al., 1993).

The primary cation-binding sites in parvalbumins are localized in the CD and EF loops, with binding constants varying from $1 \times 10^7 \text{ M}^{-1}$ to $3 \times 10^7 \text{ M}^{-1}$ and $2 \times 10^6 \text{ M}^{-1}$ to $5 \times 10^7 \text{ M}^{-1}$, respectively, for the two binding sites (Kretsinger, 1980). Both the CD and EF sites are “mixed” binding sites in the native state with

Reprint requests to: Donald J. Nelson, Department of Chemistry, Clark University, 950 Main Street, Worcester, Massachusetts 01610; e-mail: dnelson@clarku.edu.

a high affinity for Ca^{2+} and a lower affinity for Mg^{2+} ($K_{ca} \sim 10^7 \text{ M}^{-1}$ vs. $K_{mg} \sim 10^4 \text{ M}^{-1}$). Presently, oncomodulin is the only known parvalbumin with a Ca^{2+} -specific binding site and a mixed metal binding site. Spectroscopic and crystallographic work provides evidence for a third binding site in some parvalbumins. Fluorescence studies of terbium-loaded parvalbumin (Nelson et al., 1977) and X-ray studies demonstrate the existence of an additional surface cation-binding site. This site does not contain the consensus residues of a functional calcium-binding loop and displays no selectivity toward Ca^{2+} or Mg^{2+} and has also been shown to bind Mn^{2+} , NH_4^+ , and H_2O (Drakenberg et al., 1985; Declercq et al., 1991; Roquet et al., 1992).

In a previous study, Revett et al. (1997) determined the sequence of the major parvalbumin isoform from the silver hake and placed the protein in the β -lineage category. This paper reports the structure of SHPV-B, determined by X-ray crystallography on a frozen crystal. The SHPV-B structure is compared to those reported for other parvalbumins, with a focus on the previously noted features associated with β -lineage parvalbumins. Molecular dynamics (MD) simulations, on both Ca^{2+} -loaded and Ca^{2+} -deficient species, are also reported, and the results are discussed in light of the crystal structure of SHPV-B. The X-ray results are also correlated with the results of NMR experiments performed on native and apo-SHPV-B and with chemical modification kinetic studies directed at determining the solvent accessibility of Cys18 in the calcium-loaded and calcium-free states.

Results and discussion

X-ray structure determination—General aspects

The SHPV-B structure was refined to a resolution of 1.65 Å with a free R -value of 0.251 and conventional R -value of 0.214 for observed data between 8.0 and 1.65 Å. Table 1 summarizes the results of the refinements. Acceptable deviations from ideality are represented in the final root-mean-square (RMS) values. The overall positional error for the structure, based on Luzzati's method (Luzzati, 1952), is between 0.20 and 0.25 Å while retaining excellent stereochemistry.

A Ramachandran plot for the refined model of the polypeptide chain indicated that all nonglycine residues fall inside the normally allowed regions, with 93.9% of the ϕ, ψ values in most favored regions, and 6.1% in additionally allowed regions. Most values are in the right-handed α -helix and extended chain β -regions. Lys54 adopts ϕ, ψ angles (58.1, 34.3) typical of a left-handed α -helix. Parvalbumins from carp (5cpv; Swain et al., 1989), shark (5pal; Roquet et al., 1992), and rat (1rro; Ahmed et al., 1990) display similar ϕ, ψ angles for residue 54. Shark (5pal) has an additional residue adopting a similar conformation at position 93. PROCHECK (Laskowski et al., 1993) was used to assess the stereochemical quality of the model. All measurements were well within acceptable values.

The main structural features observed in other parvalbumins are also conserved in SHPV-B. The tertiary structure consists of six alpha helices (A–F) adopting a compact globular shape. Helices A and B are folded into an antiparallel arrangement, while helices C/D and E/F pair up and each adopt a perpendicular arrangement. Helix D has a well-defined “kink,” a feature observed in other parvalbumin crystal structures. Loops between the CD and EF helical segments form the two major metal ion binding sites. Other invariant features include the Ile58–Ile97 β -sheet, the Arg75–

Table 1. Crystallographic data and refinement statistics for SHPV-B

Space group	C222 ₁
Unit cell (Å)	$a = 75.7$ $b = 80.7$ $c = 42.1$
Data collection	
Number of crystals	1
Number of reflections	129,847
Number of unique reflections	15,685
Completeness	99.9% (100.0) ^a
R_{merge} of intensities	0.071
Cutoffs	8–1.65 Å, $F > 2\sigma$
R -factor	0.214
R -free	0.251
Model	
Residues in model	108
Number of water molecules	114
Number of calcium atoms	2
$\langle r \rangle$ from Luzzati plot	0.20 Å
Stereochemistry	
RMS nonideal bond lengths	0.012
RMS nonideal bond angles	1.729
RMS improper angles ^b	1.440
Thermal parameters	
$\langle B \rangle$ for protein atoms	21.50 Å ²
$\langle B \rangle$ for water molecules	31.55 Å ²
$\langle B \rangle$ for calcium atoms	18.90 Å ²

^aOver all completeness.

^bPlanarity of peptide bond.

Glu81 salt bridge and the characteristic Phe, Leu, Ile, Val hydrophobic core.

Temperature factors—Main-chain and side-chain flexibility

The temperature factors ranged from 9.72 to 76.95, with an average value of 21.5 for the molecule. The highest average values for backbone atoms exist for residues 92 and 93 and at the two ends of the protein chain. High vibrations at the ends of the chain are common, but the two residues of the EF loop indicate some instability in that region. The B -factors of the EF loop are significantly higher than those for the CD loop (Table 2). [The Ca^{2+} coordinating ligands are discussed later, but it might be pointed out here that an MD simulation on Ca^{2+} -loaded SHPV-B reveals that the Asp92– Ca^{2+} linkage (EF-site) is quite tenuous and tends to rupture during even a modest 200 ps MD run (Fig. 2C).] Although such instability has been seen in the B/C helices (5cpv, 5pal) and the CD loop (5pal) of other parvalbumins, the instability in the EF loop of SHPV-B appears to be unique.

Relative thermal mobility of the main-chain atoms for this model were compared with four other parvalbumin structures (i.e., pike, 1pal; carp, 5cpv; rat, 1rro; shark, 5pal). The average B -value for the main-chain atoms for the model of SHPV-B is 18.74 compared to 11.06 (1pal), 9.02 (1rro), 21.14 (5cpv), and 9.45 (5pal) for the other parvalbumins. Except for carp, whose averages are altered by very high B -values found in a very flexible surface loop (BC loop)

Table 2. Coordination distances^a and temperature factors for the CD and EF sites

Ligand	CD site			Ligand	EF site		
	Distance (Å)		<i>B</i> -factor (Å ²)		Distance (Å)		<i>B</i> -factor (Å ²)
Asp51 OD1 (x)	2.38	(2.21)	16.40	Asp90 OD2 (x)	2.46	(2.24)	22.15
Asp53 OD1 (y)	2.30	(2.34)	17.29	Asp92 OD1 (y)	2.52	(2.37)	29.13
Ser55 OG (z)	2.44	(2.61)	15.58	Asp94 OD1 (z)	2.47	(2.49)	25.61
Phe57 O (-y)	2.33	(2.37)	16.58	Ala96 O (-y)	2.31	(2.42)	21.48
Glu59 OE1 (-x)	2.43	(2.34)	18.14	Wat130 O (-x)	2.25	(2.35)	26.11
Glu62 OE1 (-z)	2.60	(2.61)	17.43	Glu101 OE1 (-z)	2.48	(2.38)	19.78
Glu62 OE2	2.65	(2.40)	16.87	Glu101 OE2	2.68	(2.69)	21.31
Average	2.45	(2.41)	16.90	Average	2.45	(2.42)	23.65
CD Calcium			16.05	EF Calcium			21.75

^aCoordination distances obtained from the carp structure (Kumar et al., 1990) are listed in parentheses. The indications (+x, +y, +z) refer to the traditional description of calcium coordination (Kretsinger, 1973).

and a very unstable terminal region, the silver hake model has the highest overall average values. Although most of the vibrational movement in SHPV-B is seen at the EF-binding site and the carboxyl terminus, SHPV-B and also 5pal differ from other parvalbumins that have extensive vibration of residues at the amino terminus. Comparing the carp model with SHPV-B, it is important to note that although both terminal regions appeared to be attracted to neighboring molecules resulting possibly from crystal packing, *B*-values for this region are significantly different. Apparently regions are packed differently to account for the difference in vibrational stability or possibly in the case for SHPV-B a stable dimer has formed at this regional contact. Evidence of antiparallel-oriented N-terminal subdomains in dimeric forms (such as it appears in this crystal) has recently been shown to occur with the globular protein platelet talin (Isenberg & Goldmann, 1998).

Side-chain atoms for the SHPV-B model naturally show higher thermal factors than main-chain atoms with *B*-values ranging from 9 to 33 Å². Side chains for SHPV-B with lower values are in constrained local environments. For example, the salt bridge residues (Arg75 and Glu81) have *B*-values of 9.2 and 12.7 Å², respectively. Stability of calcium in the binding sites is also observed resulting from local constraints of the coordinating ligands. The highest side-chain values are for EF region residues Lys83 and Lys87, each having values of 40.70 and 43.15 Å², respectively.

There are a few regions of SHPV-B where the main-chain *B*-values are higher than for side-chain atoms. For residues 32–39, located in the BC loop and C helix, lowered values may be attributed to the greater proportion of contacts with waters. Residues 83–98, located in the terminal regions of the E helix and the EF loop, also have lower side-chain values. This variation in the *B*-values might suggest that the calcium is stabilizing the EF loop ligands. Average *B*-values for the ligands of the EF loop binding the calcium average higher values (23.65 Å² for the ligands compared to 21.75 Å² for the calcium). In the CD site, *B*-values for the metal ion in the CD site and side-chain atoms are small (16.05 Å² vs. 16.90 Å², respectively).

Solvent structure and hydrogen bonding

There are 114 ordered water molecules included in the asymmetric part of the unit cell. There are 50 water molecules within hydrogen-

bonding distance (<3.2 Å) of the main-chain carbonyl oxygens, and 17 waters in contact with backbone amides. Another 41 water molecules form at least one hydrogen bond to side-chain nitrogen or oxygens atoms, 58 are involved in forming water–water contacts, and one water is involved in the binding of the EF calcium. The *B*-values of the solvent molecules ranged between 9.72 and 56.9 Å², with an average value of 31.55 Å². The other parvalbumins have comparable average *B*-values (1pal, 28.64; 1rro, 24.43; 5cpv, 33.41; 5pal, 31.48) for solvent.

Numerous waters are conserved in parvalbumin crystal structures, suggesting the importance of the solvent structure in contributing to forming homologous tertiary conformations. One water (water 130) coordinates the calcium in the EF site, and the others help to stabilize helices within the core. Another conserved water (water 111) interconnects the DE looped segment, including the Arg75–Glu81 salt bridge, to the D helix. This water appears to anchor the D helix to the E helix, stabilizing the segment between the CD and EF metal binding loops. This molecule displays the same hydrogen bonding pattern as in other parvalbumins, such as is found in carp 4.25 (Swain et al., 1989) and pike 4.10 (Declercq et al., 1991). This well-ordered water, with a temperature factor of 12.8 Å², is surrounded by two donors (Phe66 N, Leu67 N) and two acceptors (Arg75 O, Lys64 O) and seems to be involved in the “kinking” of the D helix (see Fig. 1C). The position of water 119 is also conserved among these same parvalbumins and connects the C and D helices, which could contribute to individual binding domain stabilization. Conserved waters 171 and 133 may perform the same function in the AB and EF domains, respectively, but from outside the protein structure.

In addition to the conserved waters previously discussed for parvalbumins, 12 waters are unique to SHPV-B and important in stabilizing the structure. For example, water 172 appears to help secure the amino terminus to the loop of the BC segment. Waters 123 and 142 help stabilize the CD Ca²⁺ binding loop and facilitate formation of the third cation binding site, the so-called “satellite” site (Fig. 1A). Other waters (123, 142, 143, 157, 182, 189, 201, 220) link helices to loop segments (D–CD, D–CD, B–BC, C–CD, E–EF, D–EF, B–AB, and E–EF, respectively). These bridged waters may complement the rigidity recently illustrated in EF hand domains using molecular dynamic simulations (Baldellon et al., 1998).

Fig. 1. **A:** The third cation-binding site located on the surface of the SHPV-B molecule. An ordered water molecule (water 221) is coordinated with four other hydrogen bond donors. **B:** Connolly surfaces for residues Cys18, Arg75, and Glu81 each participating in the hydrogen bond network around the salt bridge. The two calcium ions are seen in yellow. The residues are represented in CPK format. **C:** The coordination geometry of the CD site. **D:** The coordination geometry of the EF site.

Examination of the intramolecular polar contacts of SHPV-B reveals numerous side-chain and main-chain hydrogen bonding interactions consistent with other parvalbumins. Overall, the greatest number of interactions occur in the loop regions of the protein. Of particular interest are the side chains of Asp, Glu, Ser, and Thr residues, which are able to loop back to nearby backbone atoms in SHPV-B and other parvalbumins. This pattern, referred to as the "Asx" bond, has been seen in other proteins (Rees et al., 1983; Baker & Hubbard, 1984). Herzberg and James (1985) have proposed that the calcium-binding loops (EF hands) of the type found in parvalbumin contain four Asx hydrogen bonds as part of their consensus structure. In the amino acid sequence of SHPV-B parvalbumin, there are three likely Asx bonds at the CD site (Asp51, Asp53, Ser55) and three at the EF site (Asp90, Asp92, Asp94).

A recent ^{15}N heteronuclear NMR study looking at the dynamics of a homologous calcium loaded parvalbumin (Baldellon et al., 1998) determined that mobility of the DE loop segment that links

the CD and EF domains is strongly dependent on the three-dimensional arrangement around the salt bridge. This is an important finding because interactions occurring between this salt bridge and the AB loop, in addition to hydrophobic contacts involving the A, B, and D helices, are known to play an important role in determining the affinity for Ca^{2+} ions. This additional stability is provided by the presence of a highly structured network between residues in the main-chain and solvent molecules. For SHPV-B, this network features involvement of residues mostly in the AB loop, a few in the DE loop and waters. Amide hydrogens from Asp22, Ser23, Phe24, and Thr78 interact with the carboxylate group of Glu81. Carbonyls of residues Cys18, Glu19, Ala20, and Ala76 are within hydrogen bonding distance of Arg75 N^{η} . Ordered waters are found close to N and N^{η} of Arg75 and the carbonyl group of Arg75 and Glu81. Solvation of the DE loop region is demonstrated by the presence of water molecules in the crystal structure involved in the hydrogen bonds around the Arg75–Glu81

salt bridge. The solvent accessibility of the salt bridge is illustrated by Connolly surface calculations (Connolly, 1983), performed with Insight II, using a probe radius of 1.4 Å (Fig. 1B). Figure 1B also indicates the spatial proximity of Cys18 to the internal salt-bridge residues Arg75 and Glu81. The reactivity toward 5, 5'-dithiobis (2-nitrobenzoic) acid of the just buried Cys18 in SHPV-B, under a variety of experimental conditions, has recently been presented (King et al., 1999).

The CD and EF calcium-binding sites

The coordination at the two cation-binding sites CD and EF for SHPV-B are shown in Figures 1C and 1D. A calcium ion occupies both sites, and each is surrounded by seven oxygen ligands that are common to other high resolution crystal structures of proteins homologous to parvalbumin (Declercq et al., 1988; Strynadka & James, 1989; Swain et al., 1989; Kumar et al., 1990), but can range from five to seven, depending on the calcium-binding protein (Kretsinger, 1987). The observed distances between each calcium and the coordinated oxygen atoms for SHPV-B are illustrated in Table 2. No distance restraints were applied during the refinement resulting in an average value of 2.45 Å for both the CD and EF sites. The Ca–O distances at the CD site in this structure ranges from 2.30 to 2.65 Å, while the EF site distances range from 2.25 to 2.68 Å. All values fall inside the expected range (2.2–2.75 Å) for various types of Ca–O bonding distances as previously described (Herzberg & James, 1985).

The CD calcium in the SHPV-B model is coordinated by the seven oxygen atoms in a distorted octahedral arrangement, which Kretsinger et al. (1989) has described as a pentagonal bipyramid. Glu62 acts as a bidentate ligand, while Asp51 and Asp53 are monodentate ligands.

The geometry of the EF site for SHPV-B differs from the CD site and has been described as a split-vertex octahedron by Kretsinger et al. (1989). Similar to the CD site, coordination involves monodentate ligands (Asp90, Asp92, Asp94) and a bidentate ligand (Glu101). The EF site differs from the CD site in that it allows a water molecule into the coordination sphere at position –9X. The EF site also differs in using a carbonyl or carboxylate ligand at position 5Z instead of a serine hydroxyl group, which is used for the CD site. In the CD site, all ligating atoms originate from the protein.

MD simulations have been performed on calcium-loaded SHPV-B (Fig. 2A,C) and on a calcium-deficient variant with only the CD calcium bound (Fig. 2B). A 200 ps computation on the calcium-loaded form of SHPV-B indicates that particular ligand residues are not as tightly coordinated and may play less of a role in calcium binding. Ligand-to-metal distance trajectories for ligands of residues at positions 3 (Asp53) and position 5 (Ser55) from the CD site (Fig. 2A) and position 3 (Asp92) from the EF site (Fig. 2C) are very mobile, as seen by solvent excursions of their side groups during the MD run. The long excursion of Ser55 is particularly noteworthy. Interestingly, when the EF Ca^{2+} is removed, and the MD computation repeated (Fig. 2B), the dynamic response of the Ser55 side chain is greatly diminished. Calculations of mean-square displacement (MSD) (for the 200 structures generated during the 200 ps MD run) for the ligands of Ca^{2+} ion in the CD and EF sites are shown in Figure 3A. Figure 3B presents a superposition of the CD Ca^{2+} ion ligands from the 20 archived structures generated (one structure every 10 ps). Figure 3C shows the MSD vs. time plots for ligands from both sites from the MD run on the single calcium protein (only the CD Ca^{2+} ion is bound). Figure 3D

shows the superposition of the CD Ca^{2+} ion ligands from the associated 20 archived structures. The dramatic contrast between the MSD plots shown in Figures 3A and 3C indicates an apparent increased structuring (or ordering) of the CD Ca^{2+} ion ligands upon the release of the EF Ca^{2+} ion. This interpretation of the MSD results is supported by the equally dramatic contrast revealed in the superpositions of the 20 ps archived structures (Fig. 3B,D). Clearly, the CD Ca^{2+} ion ligands in the protein containing calcium *only* in the CD site are more structured/organized than these same ligands in the protein fully loaded with calcium. There appears to be an apparent “tension” between the symmetry-related clusters of residues responsible for calcium ion coordination in the CD and EF calcium-binding sites, a tension that can be relieved with the release of the EF calcium ion.

Parvalbumins typically bind calcium very tightly with small K_d s of up to 10^{-9} M (Haiech et al., 1979), yet at high concentrations Mg^{2+} , Na^+ , and K^+ ions can compete with Ca^{2+} for the metal-binding sites of parvalbumins (Pechère et al., 1979) in the native state. (Excess calcium is utilized in native crystal growth studies for this reason.) A two-dimensional ^1H -NMR solution study of two forms of parvalbumin, the fully Ca^{2+} -loaded form and the Mg^{2+} -loaded form established that the sixth ligand, a highly conserved Glu residue at the relative position –12 Z in both cation-binding sites CD and EF, undergoes a conformational rearrangement. The Mg form adopts the less energetically favored *g*-conformation following a $\text{C}_\alpha\text{--C}_\beta$ rotation of 120° (Blancuzzi et al., 1993). Regardless of this change, an analysis of the chemical shifts of the resonances revealed identical symmetrically shaped behavior for both cation-binding sites upon calcium/magnesium exchange. Other cations (e.g., lanthanides, cadmium) have been shown to bind to these sites. For example, confirmation from NMR and crystal studies of carp 4.25 suggests that the addition of one equivalent of Yb^{3+} (1:1 molar ratio of ytterbium chloride to protein) causes an extensive displacement of calcium from the EF metal-binding site (Kumar et al., 1991). A comparison of the ytterbium-substituted model with the native and cadmium-substituted structure showed no significant differences, except around the substituted EF metal-binding region. The displacement of calcium by ytterbium at the EF site caused a movement in the polypeptide backbone of Ser55 and Asp92. The mobile nature of this Asp92 at this position is consistent with MD simulations (Fig. 2C). Although this movement resulted in an increase in the number of oxygen ligands bound to ytterbium in the EF-site from seven to eight, more work, possibly using other multivalent cations, is necessary to determine factors required in changing the number of coordinating ligands.

The third binding site (the CD satellite site)

In addition to the two calcium-binding sites, it has been proposed that an additional cation-binding site exists in parvalbumins. Fluorescence (Nelson et al., 1977) and X-ray crystal studies (Declercq et al., 1991; Roquet et al., 1992) provide evidence for the existence of an additional surface binding site. This site has been shown to bind cations other than calcium and has been described by Declercq et al. (1991) to be a satellite of the CD binding site. Refinement results for SHPV-B support this earlier work. Figure 1A, presented earlier in the context of solvent structure and hydrogen bonding, illustrated an ordered water molecule (water 221) in the third cation-binding site of the SHPV-B model. The oxygen of the water is within hydrogen bonding distance of four neighboring oxygen atoms in the SHPV-B structure. This water

Fig. 2. Ligand-to-metal distance trajectories from a 200 ps MD simulation performed on SHPV-B. **A:** The CD ligand-to-CD Ca^{2+} plots for calcium-loaded SHPV-B. **B:** The CD ligand-to-CD Ca^{2+} plots for SHPV-B containing calcium in the CD site only. **C:** The EF ligand-to-EF Ca^{2+} plots for calcium-loaded SHPV-B. The atoms on the ligands chosen for the distance calculation are indicated on the plots. The Insight II program DISCOVER, with the cff91 force field, was employed for these computations.

was fit into a peak of electron density that lies in the exact position as the ammonium ion found in the third site of pike β 4.10 parvalbumin (Declercq et al., 1991). The water lies 4.5 Å away from the calcium in the primary CD-binding site of SHPV-B compared to 4.4 Å from the ammonium ion and calcium from pike. Coordination involves the oxygen groups for the side chains of residues Asp53, Glu59, and Asp61, as well as two other ordered waters. In this model, the water 221 displays a hydrogen bonding pattern that appears to be tetrahedrally coordinated to four of these, while Asp61 appears to anchor one of the coordinating waters. This third site is now known to be present in both α and β parvalbumins. Three noncontiguous residues and solvent on the surface of these proteins establish the cation-binding pocket. Two of these residues (Asp53 and Glu59) are systemically present in α and β parvalbumins, while Glu appears in position 61 of shark parvalbumin as in

most α -parvalbumins instead of Asp as in most β -parvalbumins. It is not clear what purpose this site serves.

Val106 and MD simulations on native vs. Apo-SHPV-B

Proton NMR data indicate that Val106 of SHPV-B acquires a dramatically different orientation relative to its surrounding residues in the apo-protein compared to that found in the calcium-loaded protein (Revett et al., 1997). In the native protein, the unique upfield-shifted methyl proton resonance associated with Val106 (at about -0.8 ppm relative to tetramethylsilane for SHPV-B) has been attributed to ring current effects from aromatic residues in close proximity to at least one of the γ methyl groups (Padilla et al., 1988). Total correlated NMR spectroscopy (TOCSY) performed on the na-

Fig. 3. **A:** MSD vs. time plots for the CD site ligands (red) and the EF site ligands (blue) from their original position during 200 ps MD simulations performed on native SHPV-B (Ca²⁺ present in CD and EF sites). **B:** Superposition of the CD site ligands in the 20 archived structures (one structure every 20 ps) generated during the 200 ps MD run performed on native SHPV-B (Ca²⁺ present in CD and EF sites). **C:** MSD vs. time plots for the CD site ligands (red) and the EF site ligands (blue) from their original position during 200 ps MD simulations performed on SHPV-B with Ca²⁺ bound only in the CD site. **D:** Superposition of the CD site ligands in the 20 archived structures (one structure every 20 ps) generated during the 200 ps MD run performed on SHPV-B with Ca²⁺ bound only in the CD site.

tive, calcium-loaded SHPV-B provides evidence of the occurrence of this effect on the protons of only one of the γ methyl groups, as two signals are observed in the TOCSY, one for each of the two methyl groups (Laney et al., 1997). When calcium is removed, the unique upfield-shifted signal relocates into the broad band of resonances containing many other methyl proton resonances, indicating that the valine 106 side chain is no longer experiencing the intense electronic shielding. MD computations performed on the apo-protein provide supporting evidence for the effects observed in the NMR experiments. Figures 4A and 4B clearly demonstrate the dramatic alteration in the position and orientation of Val106 (with respect to the aromatic residues Phe30 and Trp102) in the apo-protein (Fig. 4B) relative to that of the calcium-loaded SHPV-B (Fig. 4A). The greatest degree of movement of Val106 in the apo-protein appears to begin ~140 ps into the MD run, and continues throughout the remainder of the simulation as seen in the plot of MSD vs. time (Fig. 4C). Considerably less movement is seen for this residue in the calcium-

loaded system, thereby remaining in close proximity to the nearby aromatic rings of Phe30 and Trp102.

Materials and methods

Materials

Silver hake (*Merluccius bilinearis*) parvalbumin, Isoform B, was purified as previously described by Zhang et al. (1990). The X-ray analysis was performed with the SHPV-B isoform in the presence of a slight excess of Ca²⁺ to assure that both metal ion binding sites on all protein molecules were occupied with calcium, as prior elemental analyses of purified silver hake parvalbumin indicated some loss of calcium during the purification process. The purity of the protein was determined by polyacrylamide gel electrophoresis, carried out under nondenaturing

Fig. 4. **A:** X-ray structure of native, calcium-loaded SHPV-B, showing the close spatial proximity of Val106 (red) to the aromatic side chains of Phe30 (green) and Trp102 (blue). **B:** Structure at the end of a 200 ps MD simulation performed on apo-SHPV-B, showing the excursion of the Val106 γ -methyl groups away from the side chains of Phe30 and Trp102. **C:** MSD of Val106 from its original position during 200 ps MD simulations performed on native SHPV-B (red) and apo-SHPV-B (blue).

conditions. All chemicals used in the purification process were high-grade commercial products.

Methods

Crystallization

For crystallization of the native calcium-bound SHPV-B, the protein was dissolved in 10 mM *N*-2-hydroxyethylpiperazine-*N'*-

2-ethane-sulfonic acid (HEPES) at a concentration of 30 mg/mL, with 1.0 mM dithiothreitol (DTT). Diffraction quality crystals were obtained at 4°C by vapor diffusion using the sitting drop technique in 2.8–2.9 M ammonium sulfate, 100 mM 2-(*N*-morpholino)ethanesulfonic acid (MES), 10 mM CaCl_2 , and 15% glycerol (pH 5.5–6.5).

Data collection and processing

Diffraction data for the calcium-bound crystal were measured on a 30 cm MAR Research image plate detection system (MAR Research, Hamburg, Germany) using a rotating anode generator (RUH3R) equipped with focusing mirrors and employing monochromated CuK-radiation ($\lambda = 1.54178 \text{ \AA}$). For data collection, a single 15-month-old crystal ($0.4 \times 0.2 \times 0.05 \text{ mm}$) was transferred to a cryo-solution (2.96 M ammonium sulfate, 100 mM MES, 10 mM CaCl_2 , 15% glycerol, and 250 mg glucose in each 1 mL), flash frozen, and mounted using a freezing loop in a nitrogen cryo-stream maintained at -180°C . A total of 110 oscillation frames (1°) were collected at a crystal-to-detector distance of 100 mm. Crystal data were auto-indexed, and space groups determined using Denzo (Otwinowski, 1993). The native calcium-bound SHPV-B crystallized in the orthorhombic space group $C22_1$ with unit cell dimensions of: $a = 75.716 \text{ \AA}$, $b = 80.730 \text{ \AA}$, $c = 42.164 \text{ \AA}$.

Data processing involved merging a total of 129,847 measured intensities for the native crystal into 15,685 unique reflections in the resolution range $1.65 < d < 50.0$. The data are 99.9% complete overall (100.0% in the higher-resolution shell, $1.65 < d < 1.70 \text{ \AA}$) with an overall $R_{\text{sym}} = 0.071$ for the merged unique intensities greater than 2σ (I). These data were processed to a resolution of $1.65 \text{ \AA}$ with high redundancy (six- to sevenfold). Diffraction data with intensities greater than 2σ were used for structure determination and refinement.

Refinement

The structural model was determined to $1.65 \text{ \AA}$ resolution by molecular replacement using coordinates of a previously refined carp parvalbumin model (5cpv; Swain et al., 1989) as a search model in the program AMoRe (Navaza, 1994). Information from anomalous scattering data from a heavy atom derivative (Tb^{3+} bound SHPV-B) was employed to verify the molecular replacement (MR) solution and to unambiguously confirm the amino end (i.e., the first five residues). Two heavy atom sites were located from isomorphous and anomalous differences Fourier. The coordinates and occupancies of the heavy atoms were refined with MLPHARE (Otwinowski, 1993). Phases were improved by solvent flattening and histogram matching using DPHASE by G. Van Duyne (Cornell University). Solvent contrast and excellent connectivity were seen in the weighted Fourier map calculated using SIRAS (Single Isomorphous Replacement with Anomalous Scattering) phases. Positional refinement using X-PLOR (Brünger, 1992) was carried out and resolution was extended conservatively in 0.2 increments. Fourier $2|F_o - F_c|$ and $|F_o - F_c|$ maps were computed, and extensive model building was conducted using the interactive molecular graphics program O (Jones et al., 1991). Water molecules were included in the model using the programs WATERGEOM by G. Van Duyne (Cornell University) and PEAKS by Wayne Hendrickson (Columbia University). The final crystallographic R -factor was 0.214 and R_{free} was 0.251. The program PROCHECK (Laskowski et al., 1993)

was used all during the refinement to evaluate the geometry of the structural model.

Insight II MD simulations

MD simulations were performed on an Indigo 2 R10,000 Silicon Graphics Workstation, using the Insight II software package from Molecular Simulation Incorporated (MSI) (San Diego, California). The computer program DISCOVER was used for calculations, employing the cff91 force field. To mimic aqueous solvent conditions, the protein molecule was soaked with two shells of water molecules, having a total thickness of 14 Å. The kinetic energy for MD simulation was provided by a thermal bath at a constant temperature of 300 K for 10,000 steps of 1.0 fs each. Nonbonded interactions were evaluated with a cutoff distance of 12.0 Å and a switch distance of 2.0 Å. Following thermal equilibration, the simulation was continued for 200 ps with the output of data occurring every 1 ps during the trajectory and stored as a history file. The program was also set to collect "snapshots" every 10 ps and store these (20 structures) in an archive file for future analysis.

Acknowledgments

We thank Dr. Dave Lambright of the Program in Molecular Medicine, University of Massachusetts Medical Center (Worcester, Massachusetts) for helpful discussions. This work was supported by instrument grants to D.J.N. from the National Science Foundation (DUE-9650859) and the Ira W. DeCamp Foundation. Work was done during the tenure of an Established Investigatorship from the American Heart Association to W.E.R.

References

- Ahmed FR, Przybylska M, Rose DR, Birnbaum GI, Pippy ME, MacManus JP. 1990. Structure of oncomodulin refined at 1.85 Å resolution. An example of extensive molecular aggregation via Ca^{2+} . *J Mol Biol* 216:127–140.
- Andressen C, Gotzios V, Berchtold MW, Pauls T, Schwaller B, Fellay B, Celio MR. 1995. Changes in shape and motility of cells transfected with parvalbumin cDNA. *Exp Cell Res* 219:420–426.
- Baker EN, Hubbard RE. 1984. Hydrogen bonding in globular proteins. *Prog Biophys Mol Biol* 44:97–179.
- Baldellon C, Alattia J, Strub M, Pauls T, Berchtold W, Cavé A, Padilla A. 1998. ^{15}N NMR relaxation studies of calcium-loaded parvalbumin show tight dynamics to those of other EF-hand proteins. *Biochemistry* 37:9964–9975.
- Blancuzzi Y, Padilla A, Parelo J, Cavé A. 1993. Symmetrical rearrangement of the cation-binding sites of parvalbumin upon $\text{Ca}^{2+}/\text{Mg}^{2+}$ exchange. A study by ^1H 2D NMR. *Biochemistry* 32:1302–1309.
- Brewer JM, Wunderlich JK, Kim DH, Carr MY, Beach GG, Ragland WL. 1989. Avian thymic hormone (ATH) is a parvalbumin. *Biochem Biophys Res Commun* 160:1155–1161.
- Brünger AT. 1992. *X-PLOR version 3.1: A system for X-ray crystallography and NMR*. New Haven, Connecticut: Yale University Press.
- Chard PS, Bleakman D, Christakos S, Fullmer CS, Miller RJ. 1993. Calcium buffering properties of calbindin D28K and parvalbumin in rat sensory neurons. *J Physiol (Lond)* 472:341–357.
- Connolly ML. 1983. Analytical molecular surface calculation. *J Appl Crystallogr* 16:548–558.
- Declercq JP, Tinant B, Parelo 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, Parelo 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.
- Drakenberg T, Sward M, Cave A, Parelo J. 1985. Metal ion binding to parvalbumin. A ^{113}Cd -n.m.r. study of the binding of different lanthanide ions. *Biochem J* 227:711–717.
- Endo T, Takazawa K, Onaya T. 1985. Parvalbumin exists in rat endocrine glands. *Endocrinology* 117:527–531.
- Fohr UG, Weber BR, Muntener M, Standenmann W, Hughes GI, Frutiger S, Banville D, Schafer BW, Heizmann CW. 1993. Human alpha and beta parvalbumins. Structure and tissue-specific expression. *Eur J Biochem* 215:719–727.
- Haiech J, Derancourt J, Pechère JF, Demaille JG. 1979. Magnesium and calcium binding to parvalbumins: Evidence for their differences between parvalbumins and an explanation of their relaxing function. *Biochemistry* 18:2752–2758.
- Herzberg O, James MN. 1985. Common structural framework of the two $\text{Ca}^{2+}/\text{Mg}^{2+}$ binding loops of Troponin C and other Ca^{2+} binding proteins. *Biochemistry* 24:5298–5302.
- Isenberg G, Goldmann WH. 1998. Peptide-specific antibodies localize the major lipid binding sites of talin dimers to oppositely arranged N-terminal 47 kDa subdomains. *FEBS Lett* 426:165–170.
- Jiang Y, Johnson JD, Rall JA. 1996. Parvalbumin relaxes frog skeletal muscle when sarcoplasmic reticulum Ca^{2+} -ATPase is inhibited. *Am J Physiol* 270:C411–C417.
- Jones TA, Zou J-Y, 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 A* 47:10–119.
- Kagi U, Berchtold MW, Heizmann CW. 1987. Ca^{2+} binding parvalbumin in rat testis. *J Biol Chem* 262:7314–7320.
- King NM, Elkins KM, Nelson DJ. 1999. Reactivity of the invariant cysteine of silver hake parvalbumin (Isoform B) with DTNB and the effects of differing buffer species on reactivity. *J Inorg Biochem* 76:175–185.
- Kretsinger RH. 1975. Hypothesis: Calcium modulated proteins contain EF-hands. In: Carafoli E, ed. *Calcium transport in contraction and secretion*. Amsterdam: North Holland Publishing, p 469.
- Kretsinger RH. 1980. Structure and evolution of calcium-modulated proteins. *CRC Crit Rev Biochem* 8:119–174.
- Kretsinger RH. 1987. Calcium coordination and the calmodulin fold: Divergent versus convergent evolution. *Cold Spring Harb Symp Quant Biol* 52:499–510.
- Kretsinger RH, Moncrief D, Persechini A. 1989. The EF-hand family of calcium-modulated proteins. *Trends Neurosci* 12:462–467.
- Kretsinger RH, Nelson DJ. 1976. Calcium in biological systems. *Coord Chem Rev* 18:29–124.
- Kumar VD, Lee L, Edwards BFP. 1990. Refined crystal structure of calcium-liganded carp parvalbumin 4.25 at 1.5 Å resolution. *Biochemistry* 29:1404–1412.
- Kumar VD, Lee L, Edwards BF. 1991. Refined crystal structure of ytterbium-substituted carp parvalbumin at 1.5 Å and its comparison with the native and cadmium-substituted structures. *FEBS Lett* 283:311–316.
- Laney EL, Shabanowitz J, King G, Hunt DF, Nelson DJ. 1997. The isolation of parvalbumin isoforms from the tail muscle of the American alligator (*Alligator mississippiensis*). *J Inorg Biochem* 96:67–76.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291.
- Luzzati PV. 1952. Traitement statique des erreurs dans la détermination des structures cristallines. *Acta Crystallogr* 5:802–810.
- MacManus JP, Watson DC, Yaguchi M. 1985. Rat skin calcium-binding protein is parvalbumin. *Biochem J* 229:39–45.
- Navaza J. 1994. AmoRe: An automated package for molecular replacement. *Acta Crystallogr A* 50:157–163.
- Nelson DJ, Miller TL, Martin RB. 1977. Non-cooperative Ca(II) removal and Terbium(III) substitution in carp muscle calcium binding parvalbumin. *Bioinorg Chem* 7:325–334.
- Otwinowski Z. 1993. Oscillation data reduction program. In: Sawyer L, Isaacs N, Bailey S, eds. *Proceedings of the CCP4 study weekend: Data collection and processing*. Warrington, UK: Daresbury Laboratory.
- Padilla A, Cavé A, Parelo JJ. 1988. Two-dimensional ^1H nuclear magnetic resonance study of pike pI 5.0 parvalbumin (*Esox lucius*). *Mol Biol* 204:995–1017.
- Pauls TL, Cox JA, Berchtold MW. 1996. The Ca^{2+} -binding proteins parvalbumin and oncomodulin and their genes: New structural and functional findings. *Biochim Biophys Acta* 1306:39–54.
- Pechère JF, Haiech J, Derancourt J, Demaille JG. 1979. Magnesium and calcium binding to parvalbumins: Evidence for differences between parvalbumins and an explanation of their relaxing function. *Biochemistry* 18:2752–2758.
- Pfaffy GE, Faivre-Bauman A, Tixier-Vidal A, Norman AW, Heizmann CW. 1987. Developmental and functional studies of parvalbumin and calbindin D 28K in hypothalamus neurons grown in serum-free medium. *J Neurochem* 49:442–451.
- Rasmussen CD, Means AR. 1989. The presence of parvalbumin in a nonmuscle cell line attenuates progression through mitosis. *Mol Endocrinol* 3:588–596.
- Rees DC, Lewis M, Lipscomb WN. 1983. Refined crystal structure of carboxypeptidase A at 1.54 Å resolution. *J Mol Biol* 168:367–387.
- Revett SP, King G, Shabanowitz J, Hunt DF, Hartman KL, Laue TM, Nelson DJ.

1997. Characterization of a helix-loop-helix (EF hand) motif of silver hake parvalbumin Isoform B. *Protein Sci* 7:2397–2408.
- Roquet F, Declercq JP, Tinant B, Rambaud J, Parello J. 1992. Crystal structure of the unique parvalbumin component from muscle of the leopard shark (*Triakis semifasciata*). *J Mol Biol* 223:705–720.
- 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 cadmium-substituted carp parvalbumin using X-ray crystallographic data at 1.6 Å resolution. *J Biol Chem* 264:16620–16628.
- Zhang C, Speno H, Clairmont C, Nelson DJ. 1990. The isolation of an unusual parvalbumin from the white muscle of the silver hake, *Merluccius bilinearis*. *J Inorg Biochem* 40:59–79.