Biochemistry **1992**, *31*, 1011–1020

Three-Dimensional Solution Structure of Ca^{2+}-Loaded Porcine Calbindin D_{9k} Determined by Nuclear Magnetic Resonance Spectroscopy[†]

Mikael Akke,[‡] Torbjörn Drakenberg,[‡] and Walter J. Chazin*

Department of Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037

Received August 8, 1991; Revised Manuscript Received October 17, 1991

ABSTRACT: The three-dimensional solution structure of native, intact porcine calbindin D_{9k} has been determined by distance geometry and restrained molecular dynamics calculations using distance and dihedral angle constraints obtained from ^1H NMR spectroscopy. The protein has a well-defined global fold consisting of four helices oriented in a pairwise antiparallel manner such that two pairs of helix–loop–helix motifs (EF-hands) are joined by a linker segment. The two EF-hands are further coupled through a short β-type interaction between the two Ca^{2+}-binding loops. Overall, the structure is very similar to that of the highly homologous native, minor A form of bovine calbindin D_{9k} determined by X-ray crystallography [Szebenyi, D. M. E., & Moffat, K. (1986) *J. Biol. Chem. 261*, 8761–8776]. A model structure built from the bovine calbindin D_{9k} crystal structure shows several deviations larger than 2 Å from the experimental distance constraints for the porcine protein. These structural differences are efficiently removed by subjecting the model structure to the experimental distance and dihedral angle constraints in a restrained molecular dynamics protocol, thereby generating a model that is very similar to the refined distance geometry derived structures. The N-terminal residues of the intact protein that are absent in the minor A form appear to be highly flexible and do not influence the structure of other regions of the protein. This result is important because it validates the conclusions drawn from the wide range of studies that have been carried out on minor A forms rather than the intact calbindin D_{9k}.

Calbindin D_{9k} is the smallest ($M_r \sim 8500$) member of the calmodulin superfamily of calcium-binding proteins that are characterized by a common helix–loop–helix structural motif [the EF-hand (Kretsinger & Nockolds, 1973)] in their binding sites. The basic functional unit is a pair of EF-hands rather than individual binding sites. In calbindin D_{9k}, both sites bind Ca^{2+} with very high affinity ($K = 10^8\ M^{-1}$). The crystal structure of bovine calbindin D_{9k} in the calcium-loaded state has been determined by X-ray diffraction (Szebenyi & Moffat, 1986). This, along with X-ray crystal structures of parvalbumin (Moews & Kretsinger, 1975; Swain et al., 1989; Kumar et al., 1990), troponin C (Herzberg & James, 1988; Satyshur et al., 1988), and calmodulin (Babu et al., 1988), has provided a wealth of structural information about this family of proteins. However, there remains one primary structural issue that has yet to be resolved: determination of the exact nature and extent of structural differences between ion-free (apo) and Ca^{2+}-bound states of these proteins. These comparisons have been stymied by the inability to crystallize calcium-binding proteins with different levels of Ca^{2+} occupancy, thereby precluding direct comparative analysis by X-ray crystallography. The only currently available model for the conformational changes occurring upon binding of Ca^{2+} is based on a comparison of the apo N-terminal and Ca^{2+}-loaded C-terminal domains of troponin C (Herzberg et al., 1986). Nuclear magnetic resonance (NMR)[1] spectroscopy *in solution* provides an alternative to X-ray diffraction for the determination of the three-dimensional structure of proteins [e.g., see Wüthrich (1989) and Clore and Gronenborn (1989)]. This method is being used in our laboratory to circumvent the crystallization requirement, and to carry out a direct comparative analysis of the structural (and dynamic) consequences of the binding of Ca^{2+} ions by calbindin D_{9k} (Skelton et al., 1990; Akke et al., 1991a). Here we present the three-dimensional solution structure of the intact porcine calbindin D_{9k}, determined from ^1H NMR distance, dihedral angle, and hydrogen-bonding constraints using distance geometry (DG) and restrained molecular dynamics (rMD).

Calbindin D_{9k} (originally known as the intestinal calcium-binding protein) from bovine intestines has been isolated in several different forms (Fullmer & Wasserman, 1973). These species differ from one another by various degrees of degradation at the N-terminus that occurs in the course of purification. The minor A form lacking the two N-terminal residues is the most prevalent and extensively studied, and was therefore utilized in the X-ray diffraction studies. This in turn prompted the selection of the minor A sequence for the recombinant bovine proteins that have been characterized in our laboratories by NMR and various biophysical methods [for a review, see Forsén et al. (1991)]. It is important to verify that the absence of the N-terminal residues does not alter the structure of the native protein, and so we have also undertaken the study of native porcine calbindin D_{9k}, which is readily purified in the intact form (Hitchman et al., 1973). Porcine calbindin D_{9k} has one more residue than the bovine protein, and so the intact porcine protein consists of 78 residues (Hofmann et al., 1979). The porcine protein differs further from the bovine by eight substitutions[2] in the sequence: E4→A; G8→S; L30→Q; L32→I; T34→A; S44→R; E48→D; and E51→Q. Despite

---

[†] Supported by the National Institutes of Health (Grant GM 40120 to W.J.C.), the American Cancer Society (JFRA-294 to W.J.C.), the Swedish Natural Sciences Research Council (operating grants to T.D.), and the Royal Swedish Academy of Sciences (scholarship to M.A.).

* To whom correspondence should be addressed.

[‡] Permanent address: Physical Chemistry 2, Lund University, S-221 00 Lund, Sweden.

[1] Abbreviations: NMR, nuclear magnetic resonance; DG, distance geometry; rMD, restrained molecular dynamics; rEM, restrained energy minimization; NOE, nuclear Overhauser enhancement; NOESY, 2D NOE spectroscopy; COSY, correlated spectroscopy; RMSD, root-mean-square deviation.

[2] The asparagine residue at position 58 in Hofmann et al. (1979) was identified as an aspartic acid subsequent to publication (T. Hofmann, unpublished results).

0006-2960/92/0431-1011$03.00/0

these differences in sequence, the two calbindins are believed to have a very similar structure in solution (Drakenberg et al., 1989; Kördel et al., 1989). For convenience, the numbering scheme of the bovine minor A form (Fullmer & Wasserman, 1981) has been adopted here; thus, the three "extra" N-terminal residues (relative to the minor A bovine sequence) are numbered −2, −1, and 0.

Calbindin D_{9k} exhibits cis–trans isomerism at P43 which is readily detected by the observation of two sets of peaks in a 1:3 ratio in the ^{1}H NMR spectra, and which is specifically identified by the difference in NOEs that characterize the G42–P43 peptide bond (Chazin et al., 1989). The conformational heterogeneity arising from the cis–trans isomerism limited to some extent the amount and quality of the data used to determine input constraints for the structure calculations. The problem is due to the inability to distinguish the relative contribution to the cross-peak intensity from the cis and the trans isoforms in cases where corresponding cross-peaks are overlapped; the degeneracy of chemical shifts for corresponding NOESY cross-peaks does not a priori imply that the distance between the two protons (the conformation) is the same in the two isoforms. Consequently there remains an unknown level of uncertainty in the calculated structures. However, in light of detailed comparative analysis of the cis and trans isoforms showing that the structures are virtually identical except in the region near P43 (Kördel et al., 1990), the reliability of the input data is most likely very high and the structures accurate. By using a conservative approach to assign distance constraints, well-defined structures in good agreement with the experimental data are obtained for the predominant trans P43 isoform of native porcine calbindin D_{9k}. The sequential assignment and identification of secondary structure and global folding pattern have previously been reported (Drakenberg et al., 1989).

MATERIALS AND METHODS

The preparation of the single 8-mg sample used for these studies has been described previously (Drakenberg et al., 1989). The pH meter reading was 6.0 for both the ^1H_2O and ^2H_2O samples. All experiments were performed on Bruker AM500 spectrometers equipped with Aspect 3000 computers and digital phase shifting hardware as described previously (Drakenberg et al., 1989). NOESY spectra were obtained at 300 K with mixing times of 40, 80, and 200 ms in ^1H_2O solution, and 200 ms in ^2H_2O solution. $^3J_{HN\alpha}$ coupling constants were measured from cross-peak multiplets in the COSY spectrum.

Experimental Constraints. The assignment of distance constraints from NOESY cross-peaks was limited to the amide, aromatic, and well-resolved C^α and methyl regions. NOE-derived interproton distance constraints were obtained by empirical distance calibration of the NOE connectivities. Cross-peaks were categorized solely according to the mixing time at which they first appear, irregardless of the actual cross-peak volumes. These three categories were then screened for the following connectivities: d_{NN}, $d_{\alpha N}$, $d_{NN}(i,i+2)$, $d_{\alpha N}(i,i+2)$, $d_{\alpha N}(i,i+3)$, $d_{\alpha\beta}(i,i+3)$, $d_{\alpha N}(i,i+4)$, and long-range $d_{ij}(|i-j|\geq 5)$ (β-sheet-type connectivities only). By comparison of the resulting mixing time/connectivity profile with the corresponding distances in regular secondary structure elements (Wüthrich et al., 1984), an estimate of the upper distance limit was determined for each mixing time (category). These are 3.0 Å for the 40-ms category, 3.8 Å for the 80-ms category, and 5.5 Å for the 200-ms category. The 200-ms constraint was chosen conservatively to account for possible effects of spin diffusion. Corrections for motional averaging (Braun et al., 1981) were then introduced by adding 0.2 Å to the 40-ms constraint, 0.5 Å to the 80-ms constraint, and 1.0 Å to the 200-ms constraint. Distance constraints involving methyl protons were increased further by 0.5 Å in order to account for the higher intensity of cross-peaks involving methyl protons.

The standard pseudoatom representation (Wüthrich et al., 1983) was used for methylene protons and methyl groups. However, a modified pseudoatom correction to the experimental upper bounds was implemented. For cases where NOE connectivities are observed from a single proton to both diastereotopic methylene protons or methyl groups, the corresponding upper bound to the pseudoatom point is calculated by the law of cosines (Akke et al., 1991b), and the resulting value is used in the DG calculations if it is lower than the value obtained from the standard pseudoatom correction as described by Wüthrich et al. (1983). This modification serves to save valuable distance information that would be thrown away with the standard pseudoatom correction, and in our experience, it improves the quality of the resulting DG structures (M. Akke, unpublished results). The lower distance bound was taken to be the sum of the van der Waals radii, except for potential hydrogen-bonding contacts between backbone amide and backbone carbonyls, vide infra. All sequential and intraresidue distance constraints that were redundant with elementary, sterically imposed constraints (Wüthrich et al., 1983) were filtered out prior to the DG calculations.

Hydrogen bonds were assigned between slowly exchanging amide protons [identified in Figure 3 of Drakenberg et al. (1989)] and the acceptor backbone oxygens in α-helical structures if both corresponding $d_{\alpha N}(i,i+3)$ and $d_{\alpha N}(i,i+4)$ connectivities were present. Also, if these NOEs could not be identified due to ambiguity in assigning NOESY cross-peaks, but both sequentially adjacent amide protons showed the characteristic α-helical connectivities, then it was assumed that the circumscribed residue was also α-helical, and the corresponding hydrogen bond was assigned. A hydrogen bond pair between L23 and V61, forming a β-type interaction between the two Ca^{2+}-binding loops, was assigned on the basis of characteristic NOEs between the two strands (Drakenberg et al., 1989). No 3_{10}-helical hydrogen bonds were assigned, due to the ambiguity in establishing whether the $d_{\alpha N}(i,i+2)$ connectivities are direct NOEs or the result of spin diffusion. The explicit hydrogen bond constraints used are the following: NH–O, 1.8–2.0 Å; N–O, 2.7–3.0 Å. The lower bound for all distances between backbone amide protons and backbone carbonyl oxygens was reduced from 2.2 Å (the sum of the van der Waals radii) to 1.8 Å in order to allow all possible backbone hydrogen bond pairs to form during the DG calculations. Without this adjustment, the short- and medium-range interproton constraints typical of secondary structure conformation could become incompatible with the close contact of the backbone amide and carbonyl required for hydrogen bond formation. This adjustment resulted in a substantial improvement in the quality of the helices in the DG structures.

The ϕ dihedral angle constraints were classified on the basis of empirically calibrated (Pardi et al., 1984) Karplus relations (Karplus, 1959) into the ranges $-90° < \phi < -40°$ for $^3J_{HN\alpha} \leq 5.5$ Hz, $-160° < \phi < -80°$ for 8.5 Hz $\leq {}^3J_{HN\alpha} < 10$ Hz, and $-140° < \phi < -100°$ for $^3J_{HN\alpha} \geq 10$ Hz.

The calcium ions were not included in the calculations because experimental constraints between the calcium ions and the protein cannot be obtained by NMR. We have opted not to include any constraints derived from the crystal structure. Furthermore, artificial constraints between the atoms known to be involved in ligation of the calcium ions have not been used to enforce standard Ca^{2+} coordination geometries.

Structure Calculations. The structures were initially generated using the distance geometry program DISGEO (Havel & Wüthrich, 1984). These DG structures were subsequently subjected to restrained energy minimization (rEM) and restrained molecular dynamics (rMD), in order to more completely sample the conformational space consistent with the input constraints, as well as to refine the geometry and energies of these structures.

The DG calculations made full use of the refinement procedures available in DISGEO for the chirality, planarity, and dihedral angle constraints. A total of 80 DISGEO structures were calculated, and the best 40 of these, as determined by low residual violations and good covalent geometry, were further refined with rEM and rMD. The DG calculations were run on a Convex C240 computer, requiring roughly 4 min of CPU time per structure for the embed procedures and a further 26 min for the refinement procedure. The version of DISGEO used here had been encoded in the C programming language and optimized for use on the Convex by Dr. Mike A. Christiansen.

The rEM–rMD refinement was performed with the program AMBER 3.0 (Singh et al., 1986), modified by Dr. David A. Case to include a half-parabola penalty function for the NOE-derived distance constraints, with the force constant (K_d) varied as described below. Dihedral angle constraints are enforced with a penalty function of the form $K_\phi[1 - \cos(\phi - \phi_0)]$, where ϕ_0 is the closest end point for the allowed range and $K_\phi = 32$ kcal/mol (Lee et al., 1989). The ϕ constraints for residues with low $^3J_{HN\alpha}$ values were relaxed to the range $-90° < \phi < 40°$ during the rMD calculations, so as to allow a greater sampling of conformational space that is permissible based on the Karplus-type curve (Moore et al., 1988). No hydrogen bond constraints were included during this phase of the refinement. The calculations were performed in vacuo, employing the AMBER all-atom force field (Weiner et al., 1986), with the charges on the Glu, Asp, Lys, and Arg side chains reduced from ±1 to ±0.2 unit of electronic charge. No calcium ions were included in the calculations. A distance-dependent dielectric function was used in order to partially account for the dielectric screening of water. r^{-6} weighting of the NOE constraints was used for nonstereospecifically assigned protons (Clore et al., 1986). An increased ω torsion barrier of 50 kcal/mol was used during the annealing cycle to prevent the peptide units in regions with few constraints from attaining the cis conformation (Lee et al., 1989). The refinement protocol was originally developed by Dr. David A. Case (Moore et al., 1988; Lee et al., 1989) and consisted of the following. First, an rEM step was carried out, during which the force constant for the NOE penalty function (K_d) was increased by a factor of 2 after each 500 steps, starting from a value of 1 kcal/(mol·Å^2). This was followed by a four-step annealing cycle using a temperature-regulated MD algorithm (Berendsen et al., 1984) as follows: (i) heating toward a target temperature of 1200 K during 1 ps with a temperature relaxation time, τ, of 0.6 ps, and doubling K_d after 0.5 ps, starting from an initial value of 4 kcal/(mol·Å^2); (ii) equilibration at 1200 K during 4 ps with $\tau = 0.2$ ps, doubling K_d after 1.0 and 2.0 ps to a maximum value of 32 kcal/(mol·Å^2); (iii) cooling toward 0 K with $\tau = 2.0$ ps, keeping K_d constant at 32 kcal/(mol·Å^2); (iv) rEM with $K_d = 32$ kcal/(mol·Å^2). Several aspects concerning the choices of parameters for the annealing cycle have been discussed previously (Gippert et al., 1990). The AMBER 3.0 calculations were run on a Cray XMP/SE, requiring roughly 50 min of CPU time per structure.

As a comparative study, a model of porcine calbindin D_{9k} derived from the crystal structure of the bovine protein (Szebenyi & Moffat, 1986; 3ICB of the Protein Data Bank, Brookhaven National Laboratory) was subjected to the same annealing protocol as the DG structures. The model was built using the molecular modeling program UCSD-MMS (Dr. Steve Dempsey, University of California, San Diego), running on a Silicon Graphics Personal IRIS-4D workstation. The side chains of the nonhomologous residues were manually fitted to the crystal structure coordinates as far as possible, and were otherwise oriented so as to avoid van der Waals overlap. The resulting united-atom representation of the porcine model was initially energy-minimized, then hydrogen atoms were added using the program Proton (Dr. Keith Cross, unpublished), and finally the model was energy-minimized in the all-atom force field. The model structure resulting from this treatment was used as the starting point for refinement by an rMD annealing cycle.

RESULTS AND DISCUSSION

Experimental Input Constraints. In total, 371 interresidue and 101 intraresidue NOE constraints, 38 ϕ dihedral angle constraints, and 10 hydrogen bond constraints (DG only) were used in the structure calculations. Figure 1A shows the distribution of NOE constraints along the protein sequence. It is readily apparent that there are more constraints on residues in the helices than in the loops, with the exception of the two dipeptide segments (Q22–L23 in loop I, V61–S62 in loop II) that form the β-type interaction. The side chains of L23 and V61 are packed into the core of the protein and have connectivities to a large number of other residues. As expected, hydrophobic residues typically have a higher number of constraints than hydrophilic residues.

Resulting Structures. The 14 best refined (DG–rMD) solution structures of porcine calbindin D_{9k} are shown in Figure 2. Table I summarizes the residual constraint violations and energies and Table II the structural statistics for these structures (DG–rMD) and the refined model-built structure (model–rMD). The criteria used to select the structures in the DG–rMD set were all trans peptide bonds, low final constraint violations (constraint energies less than 35 kcal/mol; distance violations less than 0.4 Å; ϕ angle violations less than 15°), no spurious occurrences of positive ϕ angles in structurally well-defined regions, and total AMBER energies less than −950 kcal/mol. As a comparison, data are also given for the corresponding unrefined DG structures where applicable. It is readily seen that the rMD annealing cycle in general results in lower violations of the experimental constraints (Table I), as well as improved sampling of the conformational space available within these constraints (Table II). The DG–rMD structures are in good agreement with the NMR data. As stated above, none of these structures have distance constraint violations greater than 0.4 Å. The structures have good nonbonded contacts and deviate only slightly from idealized covalent geometry; the RMSD from ideality of bond lengths and angles are 0.006 Å and 1.6°, respectively. The average value for the ω dihedral angle in the DG–rMD set is 181.2° ± 2.2°.

The superimposed structures shown in Figure 2 clearly reveal that different regions of the protein are defined with different degrees of precision. A more quantitative representation is given in Figure 1B, which shows the average RMSD from the average coordinates as a function of residue number. The precision with which the polypeptide segments

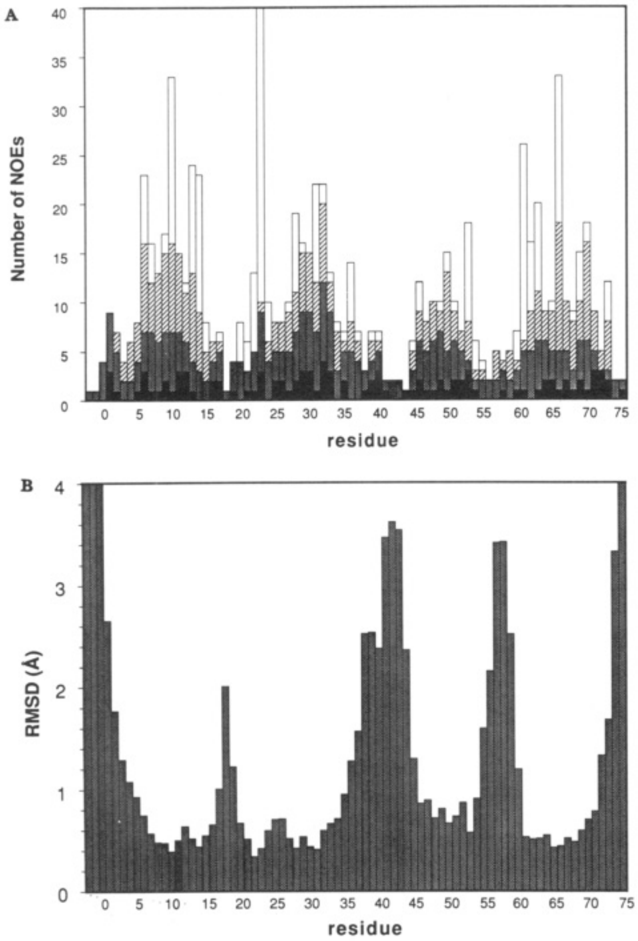

FIGURE 1: Correlation between the number of NOE constraints and the RMSD of the refined structures. (A) Number of NOE constraints per residue. Four different types of constraints are specified: filled, intraresidue; stippled, sequential; hatched, medium range (two to five residues apart along the sequence); open, long range (more than five residues apart). All interresidue constraints are counted twice. (B) Average RMSD from the average structure for the backbone atoms (N, C^α, and C) as a function of residue number for the 14 best porcine calbindin D_{9k} solution structures (DG–rMD set). The structures were superimposed by minimizing the RMSD for the backbone atoms of the helices (see Table II). For clarity, the upper limit of the vertical axis in both (A) and (B) has been set lower than the maximum data point, so that the data values are off-scale for residue L23 (47 NOEs) in (A) and for residues S(−2), A(−1), Q0, and Q75 (6.19, 5.19, 4.00, and 4.35 Å, respectively) in (B).

Table I: Summary of Residual Constraint Violations and Energies[a]

distance violation, d (Å)	av no. of distance constraint violations		
	DG	DG–rMD	model–rMD
$0.1 \leq d < 0.2$	16.9	11.1	15
$0.2 \leq d < 0.3$	13.1	4.6	3
$0.3 \leq d < 0.4$	7.1	2.6	3
$0.4 \leq d < 0.5$	3.9	0	0
$0.5 \leq d < 1.0$	6.4	0	0
$1.0 \leq d$	4.3	0	0
	av max distance violation (Å)		
	0.95	0.33	0.33
ϕ angle violation, ϕ (deg)	av no. of ϕ angle constraint violations		
	DG	DG–rMD	model–rMD
$5 \leq \phi < 10$	5.4	1.1	2
$10 \leq \phi < 15$	4.6	0.6	0
$15 \leq \phi$	0.1	0	0
	av max ϕ angle violation (deg)		
	10.9	9.3	3
	av AMBER energies (kcal/mol)[b]		
		DG–rMD	model–rMD
E_{total}[c]		−987.1 (−1041.0; −954.6)	−910.1
E_{const}[c]		23.2 (17.1; 34.9)	29.6

[a] DG–rMD refers to the best 14 refined DG structures, DG refers to the same set of structures before refinement, and model–rMD refers to the refined model of the porcine protein built from the bovine calbindin D_{9k} crystal coordinates (see the text). [b] DG–rMD and model–rMD structures only. [c] E_{total} = total AMBER energy. E_{const} = constraint energy. The force constants for the distance constraint and the dihedral angle constraint were 32 kcal·mol^{-1}·Å^{-2} and 32 kcal·mol^{-1}, respectively. The range of energies in the set of the best 14 structures is given in parentheses. The AMBER energy entries apply to the DG–rMD and model–rMD structures only.

are defined in the refined structures is clearly inversely correlated with the number of NOE constraints per residue (cf. Figure 1A). The helical regions of the protein and the β-type interaction between the two calcium-binding loops are well-defined, whereas the other parts of the calcium-binding loops are poorly defined. The linker loop between the two EF-hands and the N- and C-termini are also poorly defined.

Figure 3 shows segments of the DG–rMD structures superimposed by minimizing the RMSD for the backbone atom coordinates of the respective segments only. Figure 4 displays the average values and standard deviations for the dihedral angles ϕ and ψ as a function of residue number. The backbone of the helices and many interior side chains are well-defined. Many exterior side chains are totally unconstrained by NOEs, and this is reflected in poorer definition. In total, 23% of the χ_1 dihedral angles are confined to the same rotamer in all DG–rMD structures; these are S8, I9, F10, Q22, E27, K29, L31, I32, L40, D48, L53, V61, F63, E65, F66, and V70. It is interesting to note that the Ca^{2+}-binding loop regions are rather well-defined in Figure 3. This indicates that much of the disorder apparent in Figure 2 is due to variation of only a few dihedral angles in the N-terminal regions of the two loops

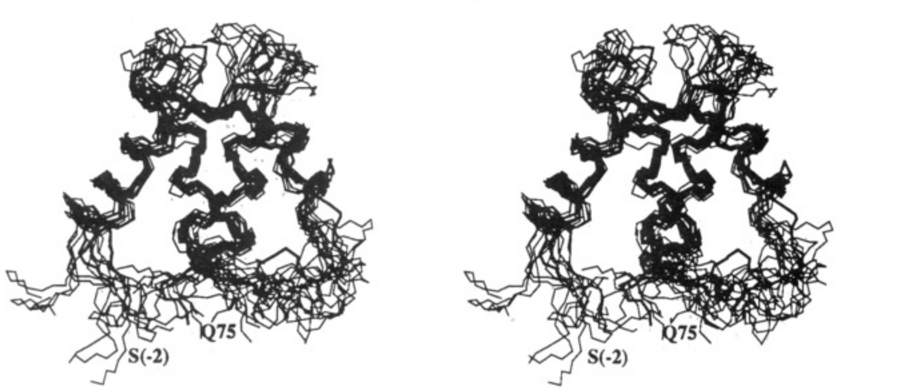

FIGURE 2: Fourteen best porcine calbindin D_{9k} solution structures (DG–rMD set) and the bovine calbindin D_{9k} crystal structure (bold lines) superimposed by minimizing the RMSD for the backbone atoms (N, C^α, and C) of the helical segments (P3–K16, K25–E35, L46–D54, F63–I73).

Table II: Structural Statistics[a]

atoms for RMSD calcn (residue; atoms)[b]	av RMSD vs av coordinates (Å) DG	av RMSD vs av coordinates (Å) DG–rMD	av pairwise RMSD vs DG–rMD set (Å) DG–rMD	av pairwise RMSD vs DG–rMD set (Å) model–rMD
helices; backbone	0.72	0.77	1.13 (0.21)	1.13 (0.24)
helices; all	1.46	1.36	1.99 (0.26)	1.90 (0.24)
helices + loops; backbone	0.85	1.13	1.66 (0.43)	1.54 (0.33)
helices + loops; all	1.56	1.58	2.32 (0.36)	2.30 (0.25)
all; backbone	1.00	1.91	2.80 (0.46)	2.63 (0.22)
all; all	1.72	2.30	3.38 (0.40)	3.27 (0.21)

Comparison to Bovine Crystal Structure[c]

	av pairwise RMSD (Å) DG–rMD	av pairwise RMSD (Å) model–rMD
helices; backbone	1.43 (0.22)	1.86
helices + loops; backbone	1.95 (0.19)	2.61
all; backbone	2.38 (0.25)	2.30

[a] DG–rMD refers to the best 14 refined DG structures, DG refers to the same set of structures before refinement, and model–rMD refers to the refined model of the porcine protein built from the bovine calbindin D_{9k} crystal coordinates (see the text). Standard deviations are given within parentheses. Note that for the DG–rMD set, the average RMSD versus the average coordinates (third column) is related to the average pairwise RMSD within the set (fourth column) through the factor $[(N-1)/2N]^{1/2}$, where N is the number of structures. [b] Heavy atoms only: all, residues (−2)–75; helices + loops, 4–36, 46–73; helices, 4–16, 26–35, 46–54, and 63–73. [c] Coordinates from 3ICB (Szebenyi & Moffat, 1986) of the Protein Data Bank, Brookhaven National Laboratories. For comparison with bovine crystal structure: all, residues 1–75. This comparison is for the DG–rMD and model–rMD sets only.

(see Figure 4). Only glycine residues, G18, G42, G57, and G59 (all situated in loop regions), and the C-terminal residue Q75 have positive average values of ϕ. These residues are all located in poorly defined regions of the molecule and show large standard deviations for the dihedral angles; the positive values for these dihedral angles cannot be taken as structurally significant. However, it is worth noting that residues G18, G57, and G59 have positive ϕ angles in the crystal structure of the bovine calbindin D_{9k} (Szebenyi & Moffat, 1986).

Table III lists the hydrogen bonds that are present in at least 10 of the 14 best final structures (DG–rMD set). These hydrogen bonds were identified using the criteria that the distance between the donor hydrogen and the acceptor oxygen must be less than 2.5 Å and that the angle defined by the donor–hydrogen–acceptor must be larger than 135°. Table III also indicates which amide protons exhibit slow exchange with the solvent (Drakenberg et al., 1989) and the hydrogen bonds that were included in the DG calculations. The following amide protons do not meet the requirement that they are hydrogen bonded to the same acceptor in at least 10 structures, but show slow exchange with the solvent: Q22, S24, E27, F50, D54, K55, D58, G59, E60, and S62. Except for F50, all of these residues are located in the Ca^{2+}-binding loops *and* show high occurrences of hydrogen bonds, but to several different hydrogen bond acceptors. This variability arises from the lack of constraints in combination with the effects of MD calculations in vacuo. For F50, hydrogen bond constraints were included in the calculations. We are unsure as to why this particular hydrogen bond is not consistently formed in the refined structures.

Description of the Structure. Analysis of the average backbone dihedral angles (Figure 4) and the hydrogen-bonding pattern (Table III) together with visual inspection of the structures (Figures 2, 3, and 5) shows that the elements of secondary structure and global folding pattern expected from the qualitative interpretation of NOE data (Drakenberg et al., 1989) are present in the refined structures. The major features of the refined solution structure of porcine calbindin D_{9k} are four helices (defined by the backbone dihedral angles) consisting of residues P3–K16, K25–E35, L46–D54, and F63–I73. The helices are organized pairwise, and in a sequentially antiparallel manner. Helices I and II together with the intervening Ca^{2+}-binding loop (E17–S24) constitute the first EF-hand, while helices III and IV and the Ca^{2+}-binding loop between them (K55–S62) constitute the second EF-hand. A loop segment (P37–T45) links the two EF-hand motifs, which are further coupled through a short antiparallel β-type interaction between two Ca^{2+}-binding loops involving segments Q22–S24 and E60–S62. Overall, the porcine solution structure is highly similar to the bovine crystal structure (Szebenyi & Moffat, 1986), as seen in Figure 2 and described in more detail in the next section.

The five N-terminal residues S(−2) to S2 are largely unconstrained by NOEs and appear to be highly flexible in solution. It is clear that the three "extra" N-terminal residues [S(−2), A(−1), and Q0] of the intact protein do not impact upon the structure of the rest of the protein. The linker region (P37–T45) as well as the C-terminal residues S74–Q75 are also ill-defined and presumably flexible in solution. In each of the Ca^{2+}-binding loops, there is a segment, E17–N21 in loop I and K55–E60 in loop II, that is poorly defined in backbone dihedral angle space. It is a fundamental and as of yet unanswered question whether the poor definition of the structure and the corresponding scarcity of constraints in these parts

Table III: Hydrogen Bonds Identified in at Least 10 of the Best 14 Refined Structures of Porcine Calbindin D_{9k} (DG–rMD Set)[a]

donor[b]	acceptor	occurrence[c]
L6 HN	S2 O	11
K7 HN	P3 O	14
S8 HN	A4 O	14
S8 HO^{γ}	A4 O	14
I9 HN	E5 O	13
F10 HN	L6 O	14*
E11 HN	K7 O	14*
K12 HN	S8 O	10
Y13 HN	I9 O	10*
A14 HN	F10 O	14*
A15 HN	E11 O	14
K16 HN	Y13 O	12
E17 HN	A14 O	10
L23 HN	V61 O	13*
L28 HN	S24 O	14
K29 HN	K25 O	14
Q30 HN	E26 O	13
L31 HN	E27 O	14
I32 HN	L28 O	14
Q33 HN	K29 O	14*
A34 HN	Q30 O	14*
E35 HN	L31 O	14
F36 HN	I32 O	12
Q51 HN	D47 O	11
E52 HN	D48 O	10
L53 HN	L49 O	11
V61 HN	L23 O	12*
F63 HN	N21 O	13
F66 HN	S62 O	14
Q67 HN	F63 O	13*
Q67 $HN^{\epsilon 2}$	E64 O	11
V68 HN	E65 O	13
V70 HN	F66 O	14
K71 HN	Q67 O	14

[a] Hydrogen bonds are listed if they are present in at least 10 of the best 14 refined structures, using the criteria described in the text. [b] Underlined residues exhibit slow solvent exchange (Drakenberg et al., 1989). [c] An asterisk indicates that this hydrogen bond was included in the DG calculations.

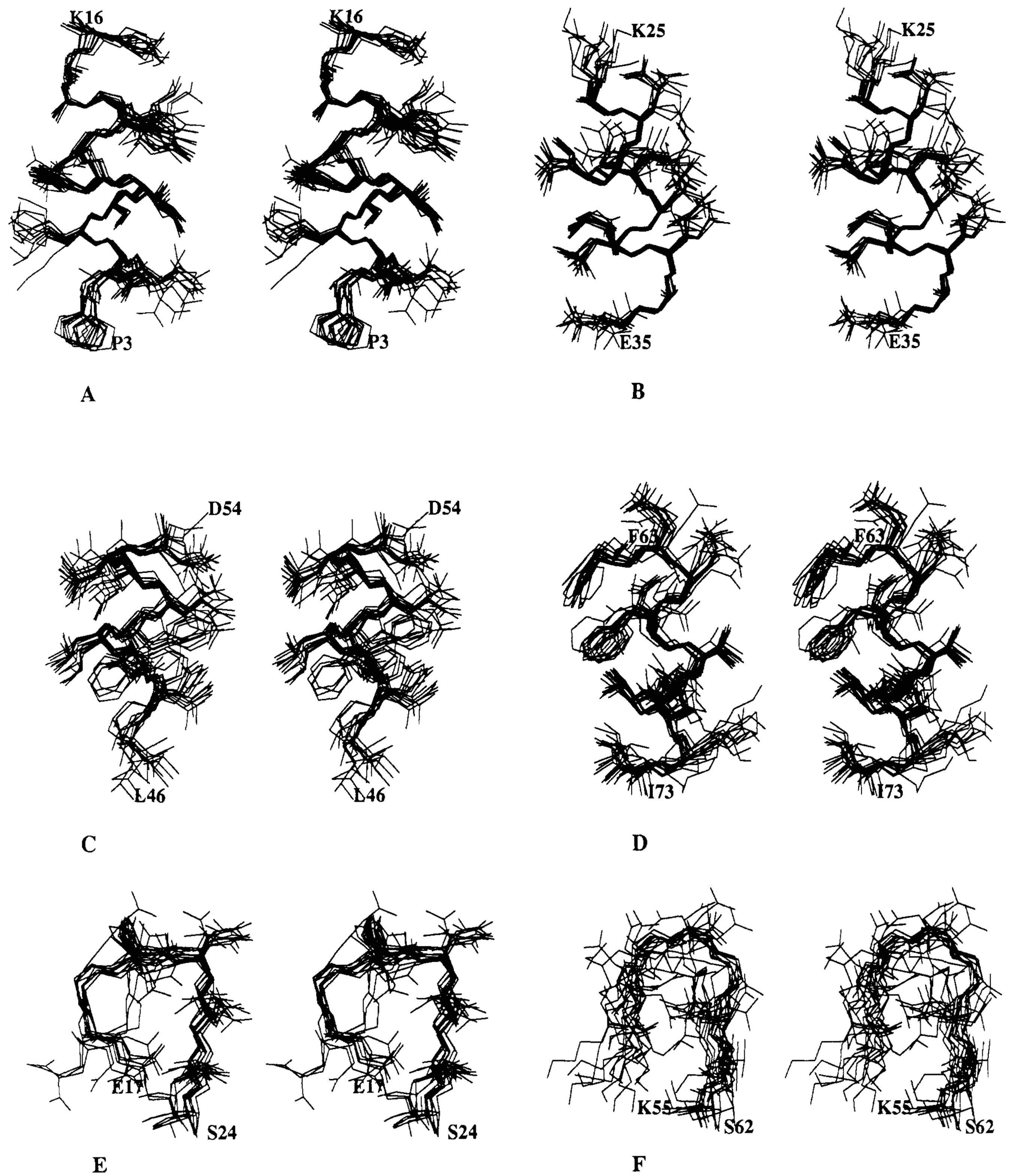

FIGURE 3: Locally superimposed segments of the 14 best porcine calbindin D_{9k} solution structures (DG–rMD set). (A) Helix I (3–16); (B) helix II (25–35); (C) helix III (46–54); (D) helix IV (63–73); (E) Ca^{2+}-binding loop I (17–24); (F) Ca^{2+}-binding loop II (55–62). Each segment was superimposed by minimizing the RMSD for the backbone atoms (N, C^α, and C) of the respective segment only.

of the protein molecule are the result of flexibility or merely because these segments are not tightly packed against the rest of the protein.

Helices I and II of the N-terminal EF-hand are both rather regular α-helices. In helix I, the expected $(i,i-4)$ hydrogen bonds run consecutively from (L6, S2) through to (A15, E11). There is a well-defined kink at Y13, as seen from the nonhelical value for the ϕ dihedral angle of this residue; this is also reflected in the absence of the $(i,i-4)$ hydrogen bond (Y13, I9) in 4 of the 14 structures. There is a type I turn at Y13–K16, as seen from the ϕ, ψ angles of A14 and A15, and the presence of the $(i,i-3)$ hydrogen bond (K16, Y13) in 12 of the 14 structures. This turn is interlocked with another type I turn at A14–E17, thereby extending the helix in a 3_{10}-like fashion into binding loop I. The characteristic $(i,i-3)$ hydrogen bond (E17, A14) is present in 10 of the 14 structures. Although relatively poorly defined, the segment E17–P20 has average dihedral angles indicative of a V′ turn, as is found in the bovine

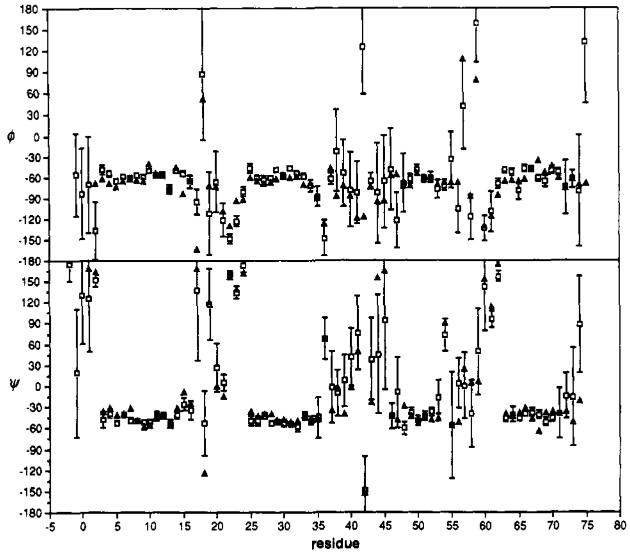

FIGURE 4: Comparison of the ϕ (A) and ψ (B) backbone dihedral angles observed in the porcine calbindin D_{9k} solution structure and the minor A bovine calbindin D_{9k} crystal structure. Average dihedral angles and standard deviations are plotted against the residue number for the 14 best DG–rMD solution structures (open squares), along with the dihedral angles observed in the bovine calbindin D_{9k} crystal structure (filled triangles). The standard deviation is indicated as the height of the vertical bar.

crystal structure. In helix II, (*i,i*–4) hydrogen bonds are present throughout, from (L28, S24) to (E35, L31). F36 at the C-terminus of helix II has backbone dihedral angles indicative of an extended conformation but is involved in a (*i,i*–4) hydrogen bond to I32 in 12 structures.

In the C-terminal EF-hand, helix III is rather poorly defined and irregular, with only three (*i,i*–4) hydrogen bonds: (Q51, D47), (E52, D48), and (L53, L49). Surprisingly, the (F50, L46) hydrogen bond identified on the basis of slow amide proton exchange and the pattern of medium-range NOEs is present in only three of the final structures. Furthermore, the first three residues of helix III, L46–D48, are not well-defined. The average dihedral angles indicate that there are two consecutive type I tight turns in the C-terminus of helix III at F50–L53 and Q51–D54, the latter of which has an (*i,i*–3) hydrogen bond in seven structures. The type I turn from D54 to G57 identified by Drakenberg et al. (1989) is also present in the refined structures, with the requisite (*i,i*–3) hydrogen bond present in 9 of the 14 structures. The putative turn from N56 to G59 (Drakenberg et al., 1989) is not unambiguously identified in the refined structures. Helix IV is also somewhat irregular, with only four (*i,i*–4) hydrogen bonds: (F66, S62), (Q67, F63), (V70, F66), and (K71, Q67). There are two 3_{10}-like turns at E65–V68 and F66–L69, with the (*i,i*–3) hydrogen bond between V68 and E65 present in 13 structures.

The β-type interaction between the two Ca^{2+}-binding loops is fairly well-defined in ϕ, ψ space for the segments N21–S24 and E60–S62. The hydrogen bonds between L23 HN and V61 O and between V61 HN and L23 O are present in 13 and 12 of the 14 structures, respectively. A hydrogen bond between

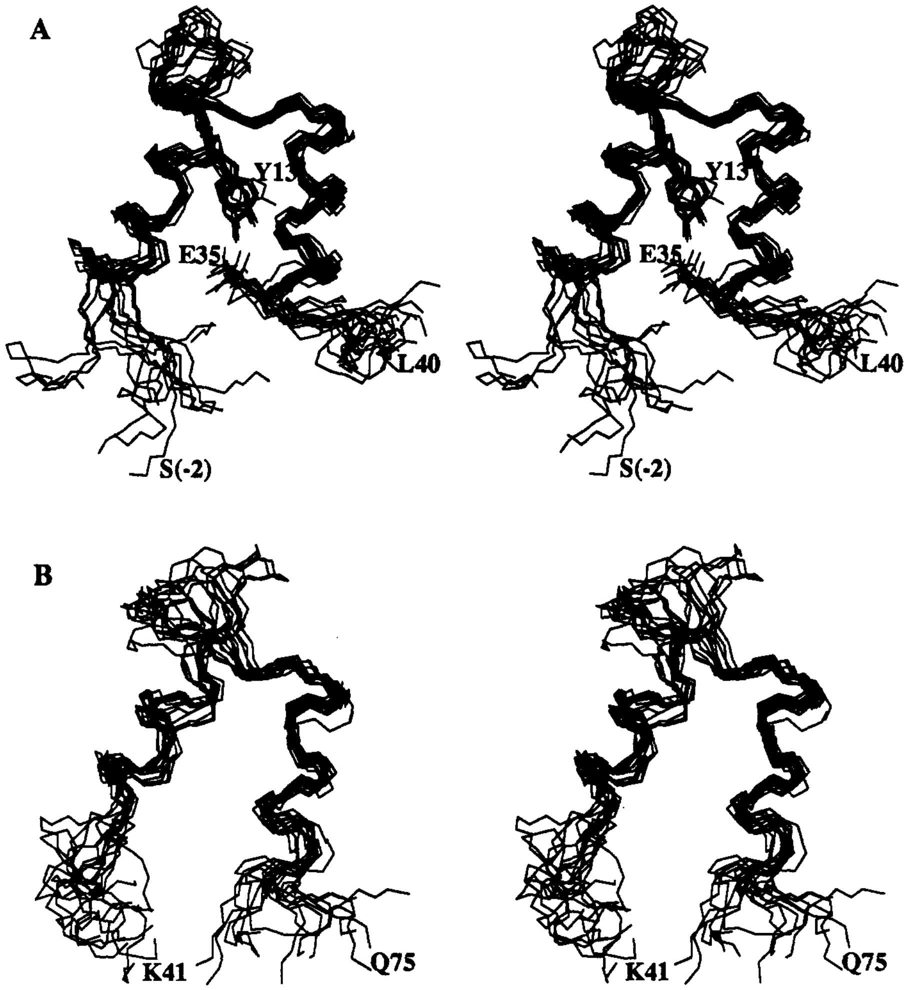

FIGURE 5: Comparison of the structures of the two EF-hands in porcine calbindin D_{9k}. The two halves of the superimposed structures of Figure 2 have been separated and repositioned to facilitate viewing the similarity between the two EF-hands. (A) Residues S(–2)–L40, also showing the interaction between the side chains of Y13 and E35. (B) Residues K41–Q75.

F63 HN and N21 O is present in 13 structures, further stabilizing the β-type interaction between the two Ca^{2+}-binding loops. The formation of this hydrogen bond is clearly indicated by the NOEs involving F63 and nearby residues [cf. Figure 5 in Drakenberg et al. (1989)] and is reflected in the orientation of the corresponding residues in the solution structure. The complementary hydrogen bond between N21 HN and F63 O is clearly not present in solution since F63 has backbone dihedral angles indicative of helical conformation. The presence of the F63 HN–N21 O hydrogen bond is consistent with the low-field chemical shift of the F63 amide proton at 9.61 ppm (Drakenberg et al., 1989). In contrast, the amide proton of F63 does not exhibit slow exchange with the solvent. The increased exchange rate can be attributed to the absence of the adjacent N21 HN–F63 O hydrogen bond and the increased exposure to solvent at the end of the β-type interaction. In the bovine crystal structure, the amide proton of F63 is hydrogen bonded to a water molecule (Szebenyi & Moffat, 1986).

Comparison with the Bovine Calbindin D_{9k} Crystal Structure. The DG–rMD structures have been superimposed onto the bovine calbindin D_{9k} crystal structure (Szebenyi & Moffat, 1986) by minimizing the RMSD for various sets of atoms (Figure 2, Table II). The average RMSD for the backbone atoms of the helices is 1.43 Å, and for all residues except S(−2) to Q0 is 2.38 Å. From Figure 2, it is evident that the orientation of helix III is slightly different from that in the crystal structure, as opposed to the close correspondence of the other helices. The turn of irregular helix observed in the crystal structure (F36–K41) is not present in the porcine solution structure, as previously noted by Drakenberg et al. (1989). In contrast, the distinct kink in helix I at Y13 in the porcine solution structures is also present in the bovine crystal structure (compare the backbone dihedral angles of this residue in Figure 4). Analysis of deviations in the backbone dihedral angles and hydrogen-bonding patterns point at more subtle differences between the bovine crystal and porcine solution structures. There are a limited number of cases where the corresponding backbone dihedral angle in the bovine crystal structure is outside the standard deviation from the average value in the DG–rMD set (Figure 4). In only a subset of these is the difference greater than 15°. This includes 13 ϕ angles (A15, E17, Q22, L23, S38, G42, D47, N56, G57, S62, F63, V68, Q75) and 16 ψ angles (K7, A15, G18, N21, S24, L39, L40, P43, R44, D48, L53, N56, V61, S62, V68, S74). In most cases, these are rather modest differences and do not result in any appreciable local structural deviations between the porcine solution structure and the bovine crystal structure. The largest deviation of a backbone angle from that reported in the X-ray crystal structure is the ϕ angle of G42, in one of the most poorly defined regions of the solution structure. This large difference may be due to the fact that the electron density was not modeled with a distribution of discrete cis and trans isomers of the G42–P43 peptide bond in the crystal structure[3] (Szebenyi & Moffat, 1986). The largest perturbations of well-defined regions involve helix IV, in particular the kinks at E65 and V68. In all cases, the subtle structural deviations are highly localized; thus, the globular fold and side-chain packing of the bovine crystal and the porcine solution structures are very similar. For example, the interaction between the side-chain hydroxyl of Y13 and the side-chain carboxylate of E35 observed in the crystal structure (Szebenyi & Moffat, 1986) is present also in solution. Although the specific hydrogen bond observed in the crystal structure is identified in only five of the porcine solution structures (given the criteria for hydrogen bonding stated above), the side chains of these two residues are oriented toward each other in *all* of the DG–rMD structures (Figure 5). Generally, the occurrence of hydrogen bonds involving side chains is rather sparse in the refined solution structures, and with the exception of (Y13 HO^{η}, E35 O^{ϵ}) and (S74 HO^{γ}, V70 O) which is present in three structures, they do not match those of the bovine crystal structure.

To further compare the structures, the NOE distance constraints of porcine calbindin D_{9k} were checked for violations against the *unrefined* model built from the bovine crystal structure with hydrogen atoms added (i.e., essentially the bovine crystal coordinates, except for nonidentical side-chain atoms of nonhomologous residues). The largest violation (3.9 Å) corresponds to the NOE observed between S74 $C^{\alpha}H$ and Q75 $C^{\gamma}H_2$. Other violations larger than 1.5 Å are D54 $C^{\alpha}H$–E60 $C^{\beta}H_2$ (3.3 Å), T45 $C^{\gamma}H_3$–D47 $C^{\alpha}H$ (3.2 Å), Y13 $C^{\epsilon}H$–L23 $C^{\gamma}H$ (2.5 Å), Q33 $C^{\alpha}H$–L46 $C^{\delta}H_3$ (2.2 Å), Q30 $C^{\alpha}H$–Q33 $C^{\gamma}H_2$ (1.9 Å), Q51 $C^{\alpha}H$–V61 $C^{\gamma}H_3$ (1.7 Å), L40 $C^{\delta}H_3$–T45 $C^{\gamma}H_3$ (1.6 Å), and T45 $C^{\alpha}H$–D48 $C^{\beta}H_2$ (1.6 Å). None of these proton–proton contacts involve atoms in the eight side chains that are different in the two proteins. At first glance, these discrepancies in proton–proton contacts suggest that there may be significant differences between the porcine solution structure and the bovine crystal structure. However, several of the large constraint violations involve residues located in rather flexible segments of the protein, and in fact, 95% of the total number of constraints are satisfied by the model. The main conclusion from this comparison is that the unrefined bovine crystal structure is a good model for the general features of the porcine solution structure.

The distance constraint violations of the unrefined model were efficiently removed by the rMD procedure, resulting in a model–rMD structure that is very similar to the set of refined DG structures. The extent of the shift of the structure during this refinement is reflected in the deviations of the backbone atoms of the model–rMD structure from the bovine crystal structure, given in the lower portion of Table II (column at right). The average pairwise RMSD of the model–rMD structure from the DG–rMD structures is given in Table II, and it is found that the model–rMD structure falls within one standard deviation of the average pairwise RMSD for the DG–rMD set. In addition, the residual constraint violations are very similar to those found for the DG–rMD structures (Table I). On the basis of these criteria, the model–rMD structure is found to be roughly as good a representation of the solution conformation of the porcine protein as the DG–rMD structures. For example, the pattern of violated NOE constraints is highly similar to that of the DG–rMD structures, and there are no "hot-spots", i.e., local regions where the violations are consistently different from those of the DG–rMD set, that could indicate discrepancies between the structures. However, the model–rMD structure has two aberrant cis peptide linkages at N56–G57 and G57–D58 that arose during the first stage of energy minimization (without NOE-derived distance constraints) of the structure built from the bovine crystal coordinates. Other cis peptide linkages were present after the initial energy minimization, but these occurred in regions with higher constraint densities than G57 and D58, and they disappeared during the rMD refinement. This irregularity appears to be highly localized and does not perturb the β-type interaction pair of hydrogen bonds between L23

---

[3] The cis and trans isomers have been observed in the crystalline state of recombinant bovine calbindin D_{9k} in a recent X-ray diffraction analysis (L. A. Svensson, unpublished results).

and V61, nor the hydrogen bond between F63 and N21. The hydrogen-bonding pattern in the model–rMD structure is very similar to that of the DG–rMD set, the only differences being the additional hydrogen bonds between E5 HN and S2 O, F50 HN and D47 O, E65 HN and S62 O, and L69 HN and E65 O, and the absence of a hydrogen bond between Q51 HN and D47 O. The differences in hydrogen bonds to D47 O further emphasize the inability of the experimental NOE constraints to clearly define the N-terminal part of helix III, vide supra.

Back-Calculations of NOE Intensities from the Refined Structure. Calculations of the NOE intensities expected for the lowest energy structure were performed as a final check of the agreement between the refined structures and the input data. NOE intensities were calculated with the program Intense, based on the Remarc module (Yip & Case, 1989) of AMBER 4.0 (Pearlman et al., 1991), provided by Dr. David A. Case. The calculations were carried out assuming isotropic reorientation with a rotational correlation time of 4 ns (fluorescence spectroscopy; Rigler et al., 1990) and experimental mixing times of 40, 80, and 200 ms. Since the original constraint generation did not rely on the quantitative evaluation of NOESY cross-peak volumes, the calculated NOE intensities were calibrated against the experimental *constraints*, rather than the experimental NOESY spectrum. This was performed by determining a lower limit of the calculated intensities at 40-ms mixing time for the large majority of cross-peaks assigned in the corresponding NOESY spectrum. Subsequently, the total intensity matrix was filtered for values larger than this lower limit, and the resulting list was checked for the presence of connectivities that could be identified, but were not present, in the spectrum. No cases were found where an observable connectivity expected from the Intense calculations was absent in all NOESY spectra, whereas 26 instances were identified where a cross-peak was expected in the 40-ms NOESY, but was first observed in the 80-ms spectrum. These minor discrepancies included 15 sequential ($i,i+1$) connectivities (d_{NN}, $d_{\alpha N}$, $d_{\beta N}$), 2 ($i,i+2$) connectivities (d_{NN}, $d_{\alpha N}$), 7 ($i,i+3$) connectivities ($d_{\alpha N}$, $d_{\alpha\beta}$), and 2 $d_{\alpha N}(i,i+4)$ connectivities. All but nine of these discrepancies are located in poorly defined regions of the protein structure. Conversely, the calculated intensities for five cross-peaks observed in the 40-ms NOESY spectrum fell below the lower limit. All of these are connectivities between residues in the poorly defined loop regions. In summary, the results from the Intense calculations indicate a close agreement between the input constraints and the resulting structure, and there is no indication of any large discrepancies.

Prior to the Intense calculations, a refined set of DG structures had been calculated that exhibited several differences from the bovine crystal structure. In particular, these structures exhibited a well-defined kink in helix III at F50–Q51, and the side chain of F50 was well-defined and oriented very differently from the bovine crystal structure, suggesting significant differences between the two proteins. However, the subsequent set of Intense calculations revealed discrepancies between the experimental data and the resulting structures that were not readily apparent as distance constraint violations. Reinspection of the NOESY spectra revealed that these discrepancies corresponded to one misassignment (L23 $C^{\beta}H$–V61 $C^{\gamma}H_3$) and one distance constraint being too restrictive (Q51 $C^{\alpha}H$–V61 $C^{\gamma}H_3$). After the appropriate corrections were made to the distance constraint list, and the structure calculations repeated, the kink in helix III was no longer present, and the side chain of F50 was no longer well-defined nor oriented differently from the bovine crystal structure. It is interesting to note that the average RMSD from the average structure increased slightly. The notion that two incorrect (too restrictive) constraints can result in more well-defined and seemingly better *but inaccurate* structures has also recently been presented for simulated NMR data (Thomas et al., 1991), and has been discussed by Havel (1991a,b). We find the back-calculation procedure very helpful in assessing the accuracy of the calculated structures and suggest that it can be very beneficial even at a rather low level of sophistication, as performed here.

Concluding Remarks. This report represents initial efforts toward generating high-resolution structures of calbindin D_{9k}. It has been demonstrated that well-defined structures in good agreement with the experimental data can be obtained for this helical protein, even in the presence of conformational heterogeneity of the sample. We have established that the additional N-terminal residues of the intact protein do not influence the structure of the rest of the protein but rather appear to be highly flexible and oriented away from the core of the protein. This validates not only the conclusions drawn from the wide range of studies carried out on the minor A form of the protein but also our ongoing high-resolution structural analysis of the apo and Ca^{2+}-loaded states of recombinant calbindin D_{9k}.

The input constraint lists, structural statistics, and coordinates of the 14 best structures of the intact porcine calbindin D_{9k} have been deposited in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY, from which these data can be obtained.

ACKNOWLEDGMENTS

We are greatly indebted to Dr. David A. Case and Garry P. Gippert for help with the AMBER and DISGEO calculations, respectively. We thank Dr. Theo Hofmann for the sample of porcine calbindin D_{9k} and Dr. Paul Weber for pointing out the inconsistency regarding α-helical interproton constraints and nonassigned hydrogen bonds in α-helices.

REFERENCES

Akke, M., Forsén, S., & Chazin, W. J. (1991a) *J. Mol. Biol. 220*, 173–189.

Akke, M., Skelton, N. J., Kördel, J., & Chazin, W. J. (1991b) in *Techniques in Protein Chemistry II* (Villafranca, J. J., Ed.) pp 401–408, Academic Press, San Diego.

Babu, Y. S., Bugg, C. E., & Cook, W. J. (1988) *J. Mol. Biol. 204*, 191–204.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & Haak, J. R. (1984) *J. Chem. Phys. 81*, 3684–3690.

Braun, W., Bosch, C., Brown, L. R., Go, N., & Wüthrich, K. (1981) *Biochim. Biophys. Acta 667*, 377–396.

Chazin, W. J., Kördel, J., Drakenberg, T., Thulin, E., Brodin, P., Grundström, T., & Forsén, S., (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2195–2198.

Clore, G. M., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986) *J. Mol. Biol. 191*, 523–551.

Clore, G. M., & Gronenborn, A. M. (1989) *Crit. Rev. Biochem. Mol. Biol. 24*, 479–564.

Drakenberg, T., Hofmann, T., & Chazin, W. J. (1989) *Biochemistry 28*, 5946–5954.

Forsén, S., Kördel, J., Grundström, T., & Chazin, W. J. (1991) *Acc. Chem. Res.* (submitted for publication).

Fullmer, C. S., & Wasserman, R. H. (1973) *Biochim. Biophys. Acta 317*, 172–186.

Fullmer, C. S., & Wasserman, R. H. (1981) *J. Biol. Chem. 256*, 5669–5671.

Gippert, G. P., Yip, P. F., Wright, P. E., & Case, D. A. (1990) *Biochem. Pharmacol. 40*, 15–22.

Havel, T. (1991a) *Prog. Biophys. Mol. Biol. 56*, 43–78.

Havel, T. F. (1991b) in *Proteins: Structure, Dynamics and Design* (Renugopalakrishnan, V., Carey, P. R., Smith, I. C. P., Huang, S. G., & Storer, A. C., Eds.) pp 110–115, ESCOM Science Publishers, Leiden, The Netherlands.

Havel, T., & Wüthrich, K. (1984) *Bull. Math. Biol. 46*, 673–698.

Herzberg, O., & James, M. N. G. (1988) *J. Mol. Biol. 203*, 761–779.

Herzberg, O., Moult, J., & James, M. N. G. (1986) *J. Biol. Chem. 261*, 2638–2644.

Hitchman, A. J., Kerr, M.-K., & Harrison, J. E. (1973) *Arch. Biochem. Biophys. 155*, 221–222.

Hofmann, T., Kawakami, M., Hitchmann, A. J. W., Harrison, J. E., & Dorrington, K. J. (1979) *Can. J. Biochem. 58*, 737–748.

Karplus, M. (1959) *J. Chem. Phys. 30*, 11–15.

Kretsinger, R. H., & Nockolds, C. E. (1973) *J. Biol. Chem. 248*, 3313–3326.

Kumar, V. D., Lee, L., & Edwards, B. F. P. (1990) *Biochemistry 29*, 1404–1412.

Ko̎rdel, J., Forsên, S., & Chazin, W. J. (1989) *Biochemistry 28*, 7065–7074.

Kördel, J., Forsên, S., Drakenberg, T., & Chazin, W. J. (1990) *Biochemistry 29*, 4400–4409.

Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., & Wright, P. E. (1989) *Science 245*, 565–568.

Moews, P. C., & Kretsinger, R. H. (1975) *J. Mol. Biol. 91*, 201–228.

Moore, J. M., Case, D. A., Chazin, W. J., Gippert, G. P., Havel, T. F., Powls, R., & Wright, P. E. (1988) *Science 240*, 314–317.

Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741–751.

Pearlman, D. A., Case, D. A., Caldwell, J. C., Seibel, G. L., Singh, U. C., Weiner, P., & Kollman, P. A. (1991) AMBER 4.0, University of California, San Francisco.

Rigler, R., Roslund, J., & Forsên, S. (1990) *Eur. J. Biochem. 188*, 541–545.

Satyshur, K. A., Rao, S. T., Pyzalska, D., Drendel, W., Greaser, M., & Sundaralingam, M. (1988) *J. Biol. Chem. 263*, 1628–1647.

Singh, U. C., Caldwell, J., & Kollman, P. A. (1986) AMBER 3.0, University of California, San Francisco.

Skelton, N. J., Kördel, J., Forsên, S., & Chazin, W. J. (1990) *J. Mol. Biol. 213*, 593–598.

Swain, A. L., Kretsinger, R. H., & Amma, E. L. (1989) *J. Biol. Chem. 264*, 16620–16628.

Szebenyi, D. M. E., & Moffat, K. (1986) *J. Biol. Chem. 261*, 8761–8776.

Thomas, P. D., Basus, V. J., & James, T. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1237–1241.

Weiner, S. J., Kollman, P. A., Nguyen, D. T., & Case, D. A. (1986) *J. Comput. Chem. 7*, 230–252.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Wüthrich, K. (1989) *Acc. Chem. Res. 22*, 36–44.

Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949–961.

Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.

Yip, P., & Case, D. A. (1989) *J. Magn. Reson. 83*, 643–648.

Interaction of Tn501 Mercuric Reductase and Dihydroflavin Adenine Dinucleotide Anion with Metal Ions: Implications for the Mechanism of Mercuric Reductase Mediated Hg(II) Reduction[†]

Richard T. Cummings[‡] and Christopher T. Walsh*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

Received August 6, 1991; Revised Manuscript Received October 15, 1991

ABSTRACT: The flavoprotein Tn501 mercuric reductase (MerA) catalyzes the reduction of Hg(II) to Hg(0) through the intermediacy of the tightly bound two-electron-reduced cofactor $FADH^-$. To gain insight into the MerA mechanism, the interaction of the holoenzyme or free $FADH^-$ with various metal ions was investigated. The free two-electron-reduced FAD cofactor, $FADH^-$, readily reduces a variety of metal ions, provided they have suitably high redox potentials. For Hg(II) with various ligands, the rate of reduction is inversely proportional to the stability of the Hg(II)–ligand complex. These results are consistent with the free cofactor reducing metal ions by an outer-sphere electron transfer mechanism. In contrast, MerA can tightly bind several redox labile metal ions, but only Hg(II) is reduced. The inability of MerA to reduce these bound metal ions may suggest that MerA differs from free $FADH^-$ and utilizes an inner-sphere electron transfer mechanism in Hg(II) reduction.

Cellular functions and enzymatic activities are often dependent on critical thiols such as cysteine (Cys) residues in proteins and glutathione. Consequently, organomercurials and mercuric ion are highly toxic to living systems by virtue of tight binding to sulfur [at physiological pH, K_f for $Hg(Cys)_2 > 10^{40.3}$ M^{-2} (Stricks & Koltholff, 1953)]. In contrast, metallic mercury displays no propensity to bind thiols and is much less toxic. Against this background, bacteria have evolved a unique detoxification mechanism for mercury compounds (Moore et al., 1990). Rather than ion removal by sequestration or active

---

[†] This work was supported in part by a grant from the National Institutes of Health (GM21643) and a National Institutes of Health Postdoctoral Fellowship to R.T.C. (GM13357).

* Correspondence should be addressed to this author.

[‡] Present address: Department of Biophysical Chemistry, MSDRL, P.O. Box 2000, Rahway, NJ 07065-0900.

0006-2960/92/0431-1020$03.00/0