A Molecular Dynamics Simulation of Bovine Calbindin D_{9k}. Molecular Structure and Dynamics†

Peter Ahlström,* Olle Teleman, Johan Kördel, Sture Forsén, and Bo Jönsson
Physical Chemistry 2, Chemical Centre, POB 124, S-221 00 Lund, Sweden
Received June 7, 1988; Revised Manuscript Received December 13, 1988

ABSTRACT: A molecular dynamics simulation of the Ca^{2+}-binding protein calbindin D_{9k} is reported. The calcium-saturated protein is simulated in an aqueous environment with an X-ray diffraction structure as the starting point. The simulation, which lasted 39 + 124 ps, was performed with the molecular dynamics program MUMOD. Structural and dynamic properties were investigated and compared with experiment. The protein contracts compared to the crystal form during the equilibration. The major contribution to the electrostatic part of the Ca^{2+}-binding energy arises from the amino acids of the Ca^{2+}-binding loops with only marginal contributions from the flanking helices. The estimated global rotational diffusion is faster than in experimental studies. From the simulation a characteristic time for a local reorientation process of the single tyrosine (Tyr 13) is estimated to 0.02–0.06 ns. More long-lived processes are present, but poor sampling precludes comparison with the experimental characteristic time of 0.36 ns obtained from fluorescence depolarization measurements. The simulated relaxation rates for ^{43}Ca in the two binding sites are in qualitative agreement with and corroborate the assignments made in a recent experimental study.

Calcium ions play a central role in a complex intracellular messenger system that mediates a variety of intracellular processes (Rasmussen, 1986a,b). An essential link in the system is a family of regulatory Ca^{2+}–binding proteins that undergo Ca^{2+}-dependent conformational changes and respond to transitory increases in intracellular Ca^{2+} concentrations.

The intracellular regulatory Ca^{2+}-binding proteins show a high degree of amino acid sequence homology, both internally and between different members. Furthermore, X-ray crystallographic studies on four proteins in this family, carp parvalbumin (Kretsinger & Nockolds, 1973; Moews & Kretsinger, 1975), calf calbindin D_{9k} (Szebenyi & Moffat, 1986; Szebenyi et al., 1981), turkey and chicken skeletal muscle troponin C (Herzberg & James, 1985; Sundaralingam et al., 1985), and rat and cow calmodulin (Babu et al., 1985; Kretsinger & Weissman, 1986), have shown that the calcium binding sites are constructed in a similar fashion, colloquially termed "the EF-hand" (Moews & Kretsinger, 1975). This structural unit is made up of two helices separated by a calcium binding loop that is typically 12 amino acids long and wrapped around the ion in such a way that the ion is coordinated to carboxyl and carbonyl oxygen atoms. A second structural feature of this family of proteins is that pairs of EF-hands rather than individual sites seem to be the functional unit.

We have recently initiated a study of the structure–function relations in regulatory Ca^{2+}-binding proteins. The main thrust has been directed toward bovine calbindin D_{9k}, which in many respects constitutes a favorable system for biophysical investigations. It has only two Ca^{2+}-binding sites, and its modest size (75 amino acids, $M_r \approx 8500$), stability, and high solubility in aqueous media make it particularly attractive for NMR studies. The protein has only one tyrosine and no tryptophan, which facilitates the interpretation of time-resolved fluorescence data. The X-ray structure of the minor A form of calbindin, i.e., the native protein with its two N-terminal amino acids proteolytically removed, has recently been refined to a resolution of 2.3 Å. The overall features of the protein are similar to those of the Ca^{2+}-binding globular domains of calmodulin and troponin C. The schematic structure of calbindin D_{9k} is presented in Figure 1. The Ca^{2+}-binding site in the C-terminal half (site II) has the same fold as that of the archetypal EF-hand, while the site in the N-terminal half (site I) constitutes a variant with two extra amino acids in the loop, one of which is a proline (Pro 20).

In order to explore the nature of the Ca^{2+} binding and the interaction between the two Ca^{2+} sites in calbindin, we have recently chosen to combine biophysical methods with the powerful tool of site-directed mutagenesis (Harris, 1982; Smith, 1982). Through studies of mutants in which one of the sites has been perturbed, it has also been possible to establish that the Ca^{2+} binding in the wild-type protein is cooperative with a free energy of interaction between the sites of 5.2 ± 0.4 kJ/mol. Assignment of the complete 500-MHz ^1H NMR spectrum of the Ca^{2+}-loaded form of calbindin is complete (J. Kördel, S. Forsén, and W. J. Chazin, unpublished experiments), and preliminary results are available from time-resolved fluorescence anisotropy measurements on the single tyrosine (Tyr 13) of the protein.

With a view to provide a molecular basis for the interpretation of these biophysical measurements, we have also undertaken molecular dynamics (MD) simulations on the minor A form of bovine calbindin D_{9k}. Since recent MD studies on proteins have indicated the results to be significantly influenced by solvent effects [see, e.g., Ahlström et al. (1987)], in particular the dynamic properties of groups at the protein surface but also of groups in the interior, the simulations have been performed in an aqueous environment. The present simulation also serves as a starting point for further studies of the structural and dynamic changes accompanying the modification and deletion of selected amino acids through site-directed mutagenesis.

Interaction Potential and Simulation Technique

The detailed features of the interaction potential and its parameters have been discussed elsewhere (Ahlström et al., 1987). It is based on inter- and intramolecular atom–atom interactions. All atoms are described explicitly except –CH,

† Financial support from the Swedish Natural Science Research Council is gratefully acknowledged.

0006-2960/89/0428-3205$01.50/0

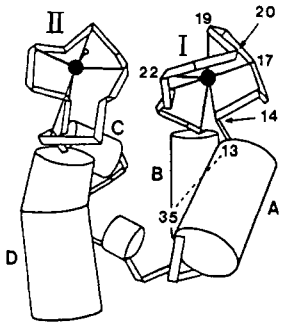

FIGURE 1: Schematic structure of the Ca^{2+}-loaded form of bovine calbindin D_{9k}. The positions of Tyr 13, Ala 14, Glu 17, Asp 19, Pro 20, Gln 22, and Glu 35 are indicated. Helices are named A–D, and the calcium binding sites are numbered I and II according to standard notation [after Szebenyi and Moffat (1986), where also more detailed stereo graphs are given].

$-CH_2$, and $-CH_3$ groups, which are treated as pseudoatoms. The interaction potential between two atoms breaks down into covalent and noncovalent interactions, the latter between atoms in the same molecule more than two bonds apart and between atoms in different molecules. The noncovalent potential, between two atoms i and j, consists of a Lennard–Jones (LJ) plus a pure electrostatic potential

$$U_{ij} = 4\epsilon_{ij}\left\{\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right\} + \frac{1}{4\pi\epsilon_0\epsilon_r}\left(\frac{q_i q_j}{r_{ij}}\right) \quad (1)$$

where ϵ_{ij} and σ_{ij} are the usual LJ parameters and q_i is the charge of atom i. The relative dielectric permittivity ϵ_r was set equal to unity. Parameters were taken from the literature (Hermans et al., 1984; van Gunsteren & Karplus, 1982; Margenau & Kestner, 1969), and the LJ parameters were calculated with the Slater–Kirkwood formula (Margenau & Kestner, 1969) under the assumption that the LJ potential minimum corresponds to the sum of the van der Waals radii of the two atoms. In particular, we used the so-called simple point charge (SPC) model for water (Berendsen et al., 1981) including also intramolecular vibrations.

For the covalent part all internal bond and bond angle vibrations were treated explicitly according to

$$U_{\text{intra}} = \sum_{\text{bonds}} \frac{k_b}{2}(x_i - x_{i,\text{eq}})^2 + \sum_{\text{angles}} \frac{k_a}{2}(\theta_i - \theta_{i,\text{eq}})^2 \quad (2)$$

in which x_i and $x_{i,\text{eq}}$ are the actual and equilibrium bond lengths and similarly for the bond angle with the force constants taken from the literature (Berendsen et al., 1981; Hermans et al., 1984; van Gunsteren & Karplus, 1982; Dolphin & Wick, 1977; Herzberg, 1945).

Internal rotations were handled with periodic dihedral potentials

$$U_{\text{int rot}} = \sum_k \sum_n C_{k,n}\{1 - \cos(n\psi_k)\} \quad (3)$$

where ψ_k is a dihedral angle, $C_{k,n}$ is an interaction parameter (van Gunsteren & Karplus, 1982), k runs over all dihedral angles at each bond, and n goes from 1 to 3.

The protein in its X-ray configuration (Szebenyi & Moffat, 1986), together with 36 crystal waters, was centered in a simulation box with the dimensions 43.4 × 40.3 × 46.5 Å. The box was then filled with 2212 more water molecules placed in a primitive cubic lattice resulting in a protein concentration of 20 mM.

The MD simulation was performed with the program MUMOD (Teleman & Jönsson, 1986), utilizing a double time step

Table I: Simulation Parameters

no. of atoms	7468
box size (Å)	43.4 × 40.3 × 46.5
time steps $\Delta t/\Delta T$ (fs)	0.2/1.2
equilibration time (ps)	39
length of trajectory (ps)	124
velocity scaling interval (fs)	96
av scaling factor	0.99
neighbor list update interval (fs)	4.8
total CPU time/ps (s)	6700
av temp (K)	302
av pressure (MPa)	−36

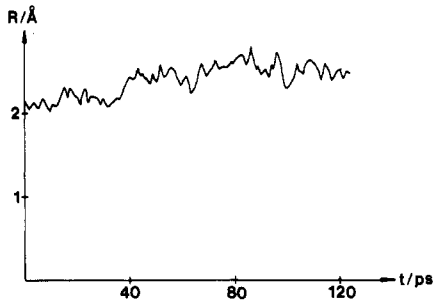

FIGURE 2: Time evolution of the R-factor calculated for all heavy backbone atoms.

Gear algorithm on an IBM 3090-150VF computer. Rapidly varying bond lengths and bond angles were integrated with a time step of Δt = 0.2 fs (fs = 10^{-15} s) whereas for the more slowly varying degrees of feedom a time step ΔT of 1.2 fs was used. Periodic boundary conditions in all directions were used together with a spherical cutoff of 10 Å for nonbonded interactions and a neighbor list technique. The initial X-ray configuration was allowed to equilibrate for 39 ps after which the simulation was continued for another 124 ps. Simulation parameters are gathered in Table I.

Results and Discussion

General Structure of the Protein. In order to compare the simulated protein structure with the initial X-ray configuration, we have superimposed a given configuration on the X-ray structure and minimized the R-factor, defined as

$$R = [\sum_{i=1}^{N} m_i|\mathbf{r}_{i,a} - \mathbf{r}_{i,x}|^2 / \sum_{i=1}^{N} m_i]^{1/2} \quad (4)$$

Here, m_i is the mass of atom i, $\mathbf{r}_{i,a}$ and $\mathbf{r}_{i,x}$ are its coordinates in the actual configuration and the X-ray structure, respectively, and the sum runs over all heavy atoms in the backbone. The R-factor was monitored for the whole trajectory (see Figure 2). Although the simulation was preceded by 39 ps of equilibration, the R-factor continues to drift during the first half of the simulation, but seems to stabilize at a value around 2.5 Å during the latter half. This is a typical behavior seen in many protein simulations. In a 300-ps simulation of myoglobin in vacuo, the R-factor increased from 2.03 Å immediately after the equilibration to 2.82 Å during the last 100 ps of the simulation (Levy et al., 1985).

Another way to study the structural behavior is through changes in the secondary structure. Calbindin has four, pairwise antiparallel, helices. Each helix axis was defined as an average of a set of vectors from backbone atoms at the start of the helix to such at the end of it. The angles between helix A and B and between helix C and D also show large fluctuations on a 10-ps time scale—of the order of ±10°. Their average values during the simulation were found to be 113° and 121°. The same quantities in the X-ray structure are 134° and 125°, respectively.

Table II: Electrostatic Interaction Energy (kJ/mol) between the Calcium Ions and the Remaining System Divided into Sub Categories[a]

	simulation		X-ray structure	
sub category[b]	site I	site II	site I	site II
helix A (2–13)	−140 (−290)	−120 (−200)	−250 (−200)	−110 (−40)
helix B (28–36)	130 (−90)	50 (−90)	70 (40)	50 (20)
helix C (46–53)	−660 (−160)	−830 (−240)	−600 (−160)	−910 (−320)
helix D (66–75)	190 (−100)	260 (−100)	170 (90)	250 (140)
loop I (14–27)	−2020	−480	−2240	−550
loop II (54–65)	−1940	−3480	−1220	−3150
rest (1, 37–45)	270	250	260	230
water	100	140	60[c]	170[c]
total	−4070	−4210	−3750	−4020

[a] Contributions from the peptide bonds are given in parentheses. [b] Amino acid numbers for the categories are given in parentheses. [c] Only crystal waters.

The stability in the helix–loop–helix motif, once the drift is arrested, is reflected by the small fluctuations in the distance between the α-carbons of the amino acid residues beginning and ending the loops. These distances were found to be 4.7 ± 0.2 and 5.7 ± 0.2 Å in loop I and loop II, respectively, which is about 1 Å shorter than in the X-ray structure. The rigidity imposes an obvious restraint on the possible loop conformation, as has recently been pointed out (Hori et al., 1987). Here, an interesting question is to what extent the Ca^{2+} ions contribute to the rigidity. If the calcium-free protein retains its rigidity in the loops, this contributes to the free energy of ion binding by reducing the loss of entropy.

The radius of gyration of the protein r_G is defined as

$$r_G^2 = \sum m_i r_i^2 / \sum m_i \quad (5)$$

in which m_i is the mass of the ith atom, r_i is its distance from the center of mass, and the sum runs over all atoms. For a uniform sphere of radius r

$$r_G = (3/5)^{1/2} r \quad (6)$$

According to the X-ray structure the calbindin molecule has a radius of gyration equal to 11.3 Å, which corresponds to an equivalent sphere radius of 14.6 Å. The protein has contracted slightly in the simulation, $\langle r_G \rangle_t = 10.7 \pm 0.1$ Å, corresponding to a sphere with a radius of 13.8 Å.

Electrostatic Interactions and Calcium Ligation. The charge separation in the amide groups of the protein backbone results in a dipole moment of approximately 3.5 D perpendicular to the C–N bond. In α-helices the curl of the backbone aligns these amide dipoles to make a macro dipole pointing toward the N-terminus. This has been suggested to cause the binding of negative ions to the N-terminus of α-helices (Hol, 1985a,b). It has also been suggested that the presence of charged groups near the N-terminus or C-terminus can stabilize the helix (Shoemaker et al., 1987). These arguments are, however, incomplete since side-chain net charges and dipoles have not been taken into account.

In order to quantify the role played by the α-helices and other parts of the protein in the calcium ion binding, we have evaluated the electrostatic interaction energy between the calcium ions and the remainder of the system

$$U_{el} = \sum_i \frac{q_i q_{Ca}}{4\pi\epsilon_0 r_i} \quad (7)$$

where r_i is the distance between atom i and the calcium ion and q_{Ca} is the charge of the calcium ion. The summation in eq 7 runs over all atoms in the system, and the electrostatic interaction energy was subdivided into different terms according to the atom i. Resulting average energies are shown in Table II together with corresponding energies for the X-ray structure.

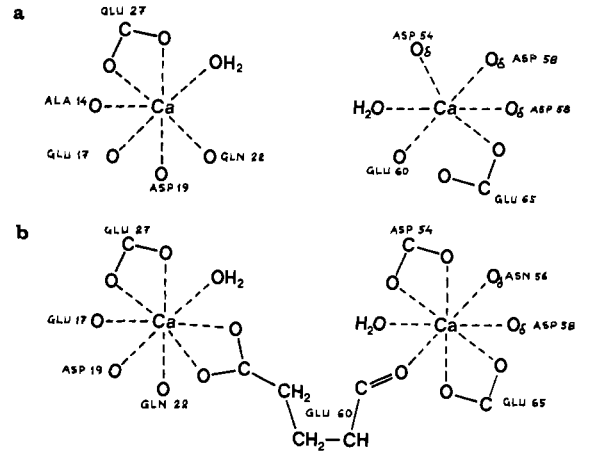

FIGURE 3: Schematic picture of the calcium ligation (a) in the X-ray structure and (b) after the equilibration period. The Ca^{2+} ligands and Glu 60, which binds to both Ca^{2+} ions, are shown.

In both sites, the α-helical peptide bonds contribute only about 15% to U_{el}. The main interaction energy arises from the loops, in particular for site I where even loop II contributes substantially. A surprising result is the very small water contribution, which is even repulsive. The total charge of both sites is negative also when a calcium ion is bound, so that the water dipoles tend to point with their positive end toward the calcium. This accounts for the positive electrostatic interaction energy between calcium and water. However, the water–ligand interaction energy is strongly negative. A calculation of the interaction energy based on the X-ray structure gives qualitatively the same numbers. The largest difference is the increased contribution from loop II to the binding of calcium in site I. This is mainly due to Glu 60 from loop II, the ionized carboxylic end group of which approaches site I and actually binds to the calcium ion in that site (see Figure 3). It is difficult to explain why this happens, but possibly Glu 60 is protonated in the crystal and if so less disposed to bind calcium. The high negative net charge of the region supports this explanation while the high pH of the solution, from which the protein was crystallized (Szebenyi et al., 1981), does not.

It has been suggested that the helix macro dipoles may be important for the stability of the helix–loop–helix super secondary structure (Hol et al., 1981; Sheridan et al., 1982). To investigate this, we have evaluated the helix–helix interactions explicitly

$$U_{el} = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (8)$$

where atoms i and j belong to helices 1 and 2, respectively. Table III shows that helices C and D have an extremely strong

Table III: Electrostatic Interaction Energies (kJ/mol) between the α-Helices[a]

helix	A	B	C	D
	(a) In the X-ray Structure			
A				
B	−10 (90)			
C	220 (70)	−160 (40)		
D	0 (80)	40 (70)	−680 (50)	
	(b) Averaged over the Simulation			
A				
B	−110 (110)			
C	220 (80)	−200 (40)		
D	−60 (90)	0 (70)	−900 (70)	

[a] The contributions from the peptide bonds are shown in parentheses.

interaction of −900 kJ/mol. The other pair, A and B, has a much weaker interaction of −110 kJ/mol, although it is still much larger than kT. These results are qualitatively the same as for the X-ray structure. It has to be remembered that these interactions will be modified by the molecular environment, notably the water. The contribution from the peptide linkages is not negligible, except for the very strong interaction between helices C and D. In general, however, we see no clear evidence for the stabilization of the helix–loop–helix structure by the helix macro dipoles.

The average electrostatic calcium-binding energy during the simulation is close to that found for the X-ray structure. Nonetheless, calcium ligands were exchanged in both sites. Site I loses one carbonyl ligand and gains two new carboxyl ligands, resulting in an average calcium coordination number of 8. Site II gains one carboxyl ligand and the original crystal water is exchanged for another water molecule; thus the average coordination number in this site is also 8.

Present data agree with simulation results for the calcium binding in the EF-hands of parvalbumin (Ahlström et al., 1987). For calcium ions in water hydration numbers of approximately 9 (Probst et al., 1985) and 7 (Teleman & Ahlström, 1986) have been reported. The differences are probably due to slightly different interaction parameters, and it must be borne in mind that the interaction parameters for Ca^{2+}, being a divalent ion, are to be regarded with a bit of scepticism. An X-ray diffraction study (Probst et al., 1985) states a value of 7 for the calcium ion coordination number in water.

Overall Rotation and Diffusion. For a vector, a reorientational autocorrelation function, $C_j(t)$, can be defined as

$$C_j(t) = \langle P_j(\cos \theta(\vartheta,t)) \rangle_\vartheta \quad (9)$$

where $\theta(\vartheta,t)$ is the angle between the vector at time ϑ and itself a time t later, P_j is the jth-order Legendre polynomial, and the brackets indicate a time average. The jth correlation time τ_j is defined by

$$\tau_j = \int_0^\infty C_j(t)\, dt \quad (10)$$

Since most correlation functions do not decay to zero during the simulation period, we have found it convenient to introduce a characteristic time, τ'_j, defined as

$$\tau'_j = \frac{t_b - t_a}{\ln [C_j(t_a)] - \ln [C_j(t_b)]} \quad (11)$$

where t_a and t_b are two different times. If $C_j(t)$ is a single exponential, the characteristic time will be identical with the correlation time. If t_b and t_a are small, the characteristic time for a vector between two nearby atoms will mainly contain information about local dynamics. In this paper we use t_a = 1 ps and t_b = 10 ps unless otherwise stated.

Table IV: Reorientational Characteristic Times (ns) for the Helix Axes

	τ'_1[a]	τ'_2[b]
helix A	2.9	0.9
helix B	2.3	0.7
helix C	1.1	0.4
helix D	1.0	0.3
mean	1.8	0.6

[a] Characteristic time for the time correlation function for the first-order Legendre polynomial. [b] Characteristic time for the time correlation function for the second-order Legendre polynomial.

The overall rotation of a flexible molecule is best evaluated by means of the angular momentum around the center of mass. The angular momentum is a function of positions and velocities—only the former were recorded in the present simulation. As the velocity autocorrelation function decays quickly compared to the sampling interval (0.096 ps), velocities obtained by numerical derivation may be inaccurate. Instead, we have characterized overall rotation in terms of time correlation functions, as in eq 9. The vectors used have to be as tightly coupled to the overall rotation as possible. In a previous study of parvalbumin we chose the helix axes as probes for global reorientation. Since these structural elements appear stable also in the present simulation, we have used the same method here.

The characteristic times τ'_1 for the four helices range from 1.0 to 2.9 ns, in fair agreement with what was found for the parvalbumin molecule in aqueous solution (0.9–2.5 ns). The average characteristic time for the second-order Legendre polynomial was found to be 0.6 ns (Table IV). This is considerably shorter than experimentally determined rotational correlation times; τ_2 = 4.6 ns from fluorescence depolarization measurements (Rigler et al., unpublished experiments) and τ_2 = 2 ns as determined by ^{43}Ca NMR relaxation experiments (Vogel et al., 1985). An estimate of τ_2 can also be obtained from the Stokes–Einstein relation, which states that for a spherical protein τ_2 = 1 ns/2400 dalton (Cantor & Schimmel, 1980). For calbindin this produces a value of 3.5 ns, in good agreement with the experimental values but substantially longer than the simulated one.

The translational diffusion coefficient D can be calculated from

$$D = \lim_{\Delta t \to \infty} \frac{\langle |\mathbf{r}(t+\Delta t) - \mathbf{r}(t)|^2 \rangle}{6\Delta t} \quad (12)$$

where $\mathbf{r}$ is the protein center of mass. The average translational diffusion coefficient was found to be 1.6×10^{-10} m^2/s. This number is subject to large uncertainties, and it is only over the first few picoseconds that the mean square displacement curve is linear. The value is slightly larger than experimental diffusion coefficients obtained for other proteins of similar size [cf. Kumosinski and Pessen (1985)]. The protein rotation and diffusion are mainly determined by the interaction with solvent molecules. Simulations of pure SPC water (Postma, 1985; Teleman et al., 1987) have shown its dynamics to be 2.5–4 times faster than that of real water. This explains the better part of the too fast overall motion. Moreover, the determination from a 124-ps trajectory of characteristic times in the nanosecond range is obviously uncertain, and an unidentified fluctuation of moderate amplitude on an intermediate time scale would produce the observed result.

Fluorescence Depolarization of Tyr 13. The fluorescence depolarization anisotropy of a chromophore in a protein is affected not only by global reorientation of the protein but also

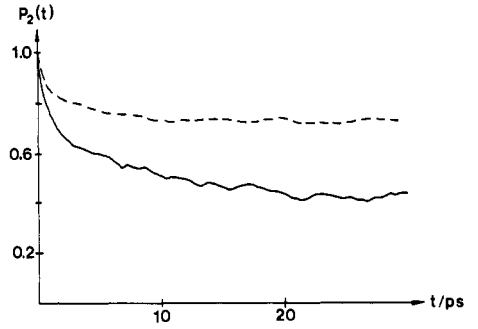

FIGURE 4: Correlation functions for the second Legendre polynomial for vectors in the plane of the tyrosine ring (Tyr 13) (—) for the $C_{\delta 1}$–$C_{\delta 2}$ ($C_{\epsilon 1}$–$C_{\epsilon 2}$) vector and (- - -) for the C_γ–C_ζ vector.

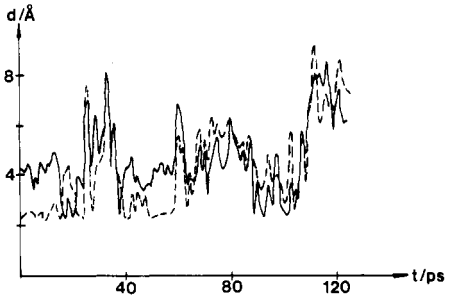

FIGURE 5: Distances between the hydrogen atom in the hydroxyl group of Tyr 13 and the two carboxylic oxygen atoms in Glu 35.

by local dynamics of the chromophore. If the two processes are uncorrelated, we can approximate the emission anisotropy $r(t)$ with

$$r(t) = Ae^{-t/\tau_c}\langle P_2(\cos\theta(\vartheta,t))\rangle_\vartheta \qquad (13)$$

where θ is the angle between the absorption transition dipole, μ_a, at time ϑ and the emission transition dipole, μ_e, a time t later. In eq 13 θ does not contain the global reorientation of the protein, but all other reorientation processes. The exponential term describes the global reorientation of the protein. The inital anisotropy $A\langle P_2(\cos\theta(\vartheta,0))\rangle_\vartheta$ is 0.4 if the absorption and emission dipoles coincide and if the excited chromophore relaxes via a unique path. In practice, this number is always smaller. The comparison between simulated ring dynamics and experimental correlation times is complicated by the fact that the direction of the transition dipole is unknown. In order to perform a qualitative comparison, we have calculated the characteristic times τ'_2, referring to the full correlation function $r(t)$, for two perpendicular vectors in the tyrosine ring plane, namely, the vectors from $C_{\delta 1}$ to $C_{\delta 2}$ (and $C_{\epsilon 1}$ to $C_{\epsilon 2}$) and from C_γ to C_ζ. The correlation functions are shown in Figure 4. With reasonable accuracy it is only possible to estimate characteristic times in the 10-ps regime, giving 0.06 and 0.02 ns, respectively. More long-lived processes are present in $r(t)$ and appear to be in the correct order of magnitude. A far longer trajectory is necessary, however, to assess them quantitatively. The correlation time τ_2 has been determined experimentally and found to be 0.36 ns (Rigler et al., unpublished experiments).

From Figure 4 it is clear that the local fast dynamics is constrained. This may be partly due to a hydrogen bond between the tyrosine hydroxyl group and the carboxyl group of Glu 35. Figure 5 shows the distance between the hydrogen atom in the OH group and the two equivalent carboxyl oxygens. The hydrogen bond is broken and formed several times during the simulation, and both carboxyl oxygen atoms of Glu 35 act as acceptors. The hydrogen bond between Tyr 13 and Glu 35 is also present in the X-ray structure.

Table V: Reorientational Characteristic Times (ps) for Vectors in the Plane of the Phenylalanine Rings[a]

residue	$\tau'_{2,\parallel}$[b]	$\tau'_{2,\perp}$[c]
Phe 10	40	30
Phe 36	90	50
Phe 50	160	120
Phe 63	50	30
Phe 66	120	100

[a] See eq 11. [b] For the vector between C_γ and C_ζ. [c] For the vectors from $C_{\delta 1}$ to $C_{\delta 2}$ (and $C_{\epsilon 1}$ to $C_{\epsilon 2}$).

We have evaluated the corresponding correlation functions for the phenylalanine residues. Their dynamic behavior is similar to that of the tyrosine ring with respect to both the magnitude and characteristic time of the local dynamics (see Table V). The magnitude of the fast motion is slightly larger than for tyrosine, but even the nonspecific interactions acting on the phenyl ring impose considerable constraints on the ring motion.

The dihedral angles formed by C_α, C_β, C_γ, and $C_{\delta 1}$ in the phenylalanines show fast rotational motions with the substantial average amplitude of 30°, which also indicates that the internal mobility of the protein is not largely affected by the protein overall contraction.

^{43}Ca NMR Relaxation. The ^{43}Ca NMR spectrum of ^{43}Ca-loaded calbindin contains one signal at 7 ppm with a line width of 350 Hz, typical of EF-hands, and one signal at 9 ppm with a line width $\Delta\nu_{1/2} \approx 1000$ Hz (Vogel et al., 1985). On the basis of ^{43}Ca NMR data of homologous calcium-binding proteins, Vogel et al. assigned the first signal to the calcium ion in the proper EF-hand site, site II, and, in consequence, the second signal to the calcium ion in site I.

For a Lorentzian NMR signal the line width $\Delta\nu_{1/2}$ is related to the effective transverse relaxation rate R_2 by

$$R_2 = \pi\Delta\nu_{1/2} \qquad (14)$$

For a quadrupolar nucleus such as calcium, with $I = {}^7/_2$, the relaxation is mainly induced by the interaction of the nuclear electric quadrupole moment Q with the fluctuating electric field gradient at the nucleus. In the limit of so-called extreme narrowing the only relaxation rate, R, can be written as

$$R = \frac{3}{8}\frac{2I+3}{I^2(2I-1)}\left(\frac{eQ}{\hbar}\right)^2(1+\gamma)^2\int\langle V_{zz}(0)V_{zz}(t)\rangle\,dt \qquad (15)$$

(Abragam, 1961; Engström & Jönsson, 1981), where V_{zz} is a diagonal element of the symmetric and traceless field gradient tensor, here evaluated in a laboratory coordinate system, and $1 + \gamma$ is the Sternheimer (anti)shielding factor. For the purpose of comparing the line width of two signals it is sufficient to write

$$\Delta\nu_{1/2} \sim \langle V_{zz}(0)V_{zz}(t_f)\rangle \qquad (16)$$

where t_f is of the order of a few picoseconds, i.e., so that the fast components of the correlation function have decayed (see Figure 6).

On the NMR time scale (and for a sufficiently long simulation) the surroundings of the two calcium ions are isotropic, which allows the calculation of $\Delta\nu_{1/2}$ as an average over the three coordinate axes; that is, V_{zz} in eq 16 can be replaced by V_{xx} and V_{yy}, and the final line width is obtained as the arithmetic average of three integrals. The correlation functions are shown in Figure 6. From eq 15 we find

$$\Delta\nu_{1/2,\text{site I}}/\Delta\nu_{1/2,\text{site II}} = 6$$

which agrees reasonably well with the experimental value of

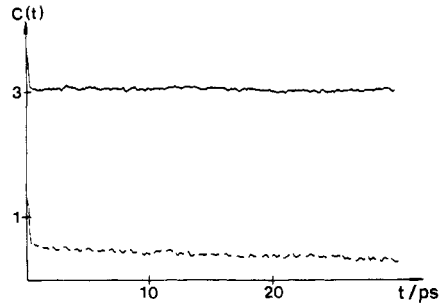

FIGURE 6: Average field gradient correlation function for the diagonal elements of the field gradient tensor $C(t) = 10^{-39} \langle V_{zz}(0)V_{zz}(t)\rangle$ (V^2/m^4) at calcium site I (—) and site II (- - -).

3. That the ^{43}Ca quadrupolar relaxation is faster in site I supports the assignments of Vogel et al. (1985).

We have also evaluated the line widths explicitly by replacing the integral of the time correlation function with $\langle V_{zz}(0)V_{zz}(t_f)\rangle\tau_c$, where τ_c refers to the global reorientation process for which we have used the above estimated value of 0.6 ns and $\langle V_{zz}(0)V_{zz}(t_f)\rangle_{I,II}$ were evaluated from Figure 6. The time t_f was chosen so that the initial fast decay was neglected, since it does not contribute significantly to the relaxation. With a Sternheimer factor of 13.12 (Lucken, 1969) and a quadrupole moment of -4.92×10^{-30} m^2 (Olsson & Salomonsson, 1982), the line widths were found to be 280 and 50 Hz in sites I and II, respectively. These numbers are a factor of 4–5 too small, the most important reason for which is the too fast overall rotation of the protein. Uncertainties in the Sternheimer factor and in the calculated field gradients may also explain part of the discrepancy. Note that none of these error sources applies to the ratio of line widths.

Conclusion

The general behavior of calbindin D_{9k} in terms of deviation from crystal conformation, calcium binding site structure, overall dynamics, and hydrogen-bond duration is similar to that found for parvalbumin in an earlier simulation. Parvalbumin and calbindin are homologous, and it is not expected that the molecular organization of calbindin will make its properties very different from those of parvalbumin.

The various structural elements of calbindin mainly interact via side chains, both internally and with the calcium ions. The role of properties such as the α-helix macro dipole is minor. One residue in the loop of site II actually binds to the calcium ion in site I. Both calcium ions possess eight ligands as in the parvalbumin simulation.

In spite of the short duration of the simulation the characteristic time for the overall rotation is of the correct order of magnitude. It is shorter than experimental values, which can be partly explained by the properties of the water model. Evaluation of motion in the 10-ns regime from a 124-ps simulation is fraught with uncertainty, purely statistical on one hand and on the other because of neglecting motion on intermediate time scales.

Dynamic properties are correctly reproduced in relative terms under certain circumstances. This is the case for the ratio of quadrupolar relaxation rates for the calcium ions, which agree with ^{43}Ca NMR data and corroborate previous spectral assignments. The relaxation rates themselves do not agree with experiment but suffer from several sources of error. A considerable part of the discrepancy can, however, be accounted for by the deficient properties of the water model.

The theoretical and methodological basis of MD simulations does not allow reliable extraction of data from first principles. Rather, validation requires experimental data. Here, such have been used to assess the quality of simulation findings and as a basis for calibration. They thus provide a foundation for the interpretation of the vast amount of information in the trajectory. An important conclusion is that MD simulations of proteins, in order to be really useful, have to be much longer than in the present work so that equilibrium is achieved and proper ensemble averages can be established. In view of this, we will attempt to extend the present simulation into the nanosecond regime, which will also allow accurate evaluation of a large number of NMR relaxation times. The results of that study will be the subject of future communication.

Registry No. Ca, 7440-70-2.

References

Abragam, A. (1961) *The Principles of Nuclear Magnetism*, Oxford University Press, Oxford.

Ahlström, P., Teleman, O., Jönsson, B., & Forsén, S. (1987) *J. Am. Chem. Soc. 109*, 1541–1551.

Babu, Y. S., Sack, J. S., Greenhough, T. G., Bugg, C. E., Means, A. R., & Cook, W. J. (1985) *Nature (London) 315*, 37–40.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., & Hermans, J. (1981) in *Intermolecular forces* (Pullman, B., Ed.) pp 331–342, Reidel, Dordrecht.

Cantor, R. C., & Schimmel, P. R. (1980) *Biophysical Chemistry*, Part II, p 461, W. H. Freeman, San Francisco.

Dolphin, D., & Wick, A. E. (1977) *Tabulation of Infrared Spectral Data*, Wiley-Interscience, New York.

Engström, S., & Jönsson, B. (1981) *Mol. Phys. 80*, 1235–1253.

Harris, T. (1982) *Nature (London) 299*, 298–299.

Hermans, J., Berendsen, H. J. C., van Gunsteren, W. F., & Postma, J. P. M. (1984) *Biopolymers 23*, 1513–1518.

Herzberg, G. (1945) *Molecular Spectra and Molecular Structure: Infrared and Raman Spectra of Polyatomic Molecules*, van Nostrand, Princeton, NJ.

Herzberg, O., & James, M. N. G. (1985) *Nature (London) 313*, 653–659.

Hol, W. G. J. (1985a) *Prog. Biophys. Mol. Biol. 45*, 149–195.

Hol, W. G. J. (1985b) *Adv. Biophys. 19*, 133–165.

Hol, W. G. J., Halie, L. M., & Sander, C. (1981) *Nature (London) 294*, 532–536.

Hori, K., Kushick, J. N., Factor, A., & Weinstein, H. (1987) *Int. J. Quantum Chem., Quantum Biol. Symp. 14*, 341–345.

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

Kretsinger, R. H., & Weissman, L. J. (1986) *J. Inorg. Biochem. 28*, 298–302.

Levy, R. M., Sheridan, R. P., Keepers, J. W., Dubey, G. S., Swaminathan, S., & Karplus, M. (1985) *Biophys. J. 48*, 509–518.

Kumosinski, T. F., & Pessen, H. (1985) *Methods Enzymol. 117*, 154–182.

Lucken, E. A. C. (1969) in *Advances in Nuclear Quadrupole Resonance* (Sen, K. D., & Narasimhan, P. T., Eds.) Academic Press, London.

Margenau, M., & Kestner, N. R. (1969) *Theory of Intermolecular Forces*, Pergamon, New York.

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

Olsson, G., & Salomonsson, S. (1982) *Z. Phys. A307*, 99–107.

Postma, J. P. M. (1985) Thesis, Chapter V:8, Rijksuniversiteit te Groningen, The Netherlands.

Probst, M. M., Radnai, T., Heinzinger, K., Bopp, P., & Rode, B. M. (1985) *J. Phys. Chem. 89*, 753–759.

Rasmussen, H. (1986a) *N. Engl. J. Med. 314*, 1094–1101.

Rasmussen, H. (1986b) *N. Engl. J. Med. 314*, 1164-1170.
Sheridan, R. P., Levy, R. M., & Salemme, F. R. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4545-4549.
Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M., & Baldwin, R. L. (1987) *Nature (London) 326*, 563-567.
Smith, M. (1982) *Trends Biochem. Sci. (Pers. Ed.) 7*, 440-442.
Sundaralingam, M., Bergström, R., Strasburg, G., Rao, S. T., Roychowdhury, P., Greaser, M., & Wang, B. C. (1985) *Science 227*, 945-948.
Szebenyi, D. M. E., & Moffat, K. (1986) *J. Biol. Chem. 261*, 8761-8777.
Szebenyi, D. M. E., Obendorf, S. K., & Moffat, K. (1981) *Nature (London) 294*, 327-332.
Teleman, O., & Ahlström, P. (1986) *J. Am. Chem. Soc. 108*, 4333-4341.
Teleman, O., & Jönsson, B. (1986) *J. Comput. Chem. 7*, 58-66.
Teleman, O., Jönsson, B., & Engström, S. (1987) *Mol. Phys. 60*, 193-203.
van Gunsteren, W. F., & Karplus, M. (1982) *Macromolecules 15*, 1528-1544.
Vogel, H. J., Drakenberg, T., Forsén, S., O'Neil, J. D. J., & Hofmann, T. (1985) *Biochemistry 24*, 3870-3876.

Escherichia coli Thioredoxin Folds into Two Compact Forms of Different Stability to Urea Denaturation†

Knut Langsetmo, James Fuchs, and Clare Woodward*
Department of Biochemistry, University of Minnesota, St. Paul, Minnesota 55108
Received July 14, 1988; Revised Manuscript Received November 9, 1988

ABSTRACT: The urea-induced denaturation of *Escherichia coli* thioredoxin and thioredoxin variants has been examined by electrophoresis on urea gradient slab gels by the method of Creighton [Creighton, T. (1986) *Methods Enzymol. 131*, 156-172]. Thioredoxin has only two cysteine residues, and these form a redox-active disulfide at the active site. Oxidized thioredoxin-S_2 and reduced thioredoxin-$(SH)_2$ each show two folded isomers with a large difference in stability to urea denaturation. The difference in stability is greater for the isomers of oxidized than for the isomers of reduced thioredoxin. At 2 °C, the urea concentrations at the denaturation midpoint are ~8 and 4.3 M for the oxidized isomers and 4.8 and 3.7 M for the reduced isomers. The difference between the gel patterns of samples applied in native versus denaturing buffer, and at 2 and 25 °C, is characteristic for the involvement of a *cis*-proline–*trans*-proline isomerization. The data very strongly suggest that the two folded forms of different stabilities correspond to the cis and trans isomers of the highly conserved Pro 76 peptide bond, which is cis in the crystal structure of oxidized thioredoxin. Urea gel experiments with the mutant thioredoxin P76A, with alanine substituted for proline at position 76, corroborate this interpretation. The electrophoretic banding pattern diagnostic for an involvement of proline isomerization in urea denaturation is not observed for oxidized P76A. In broad estimates of $\Delta G°$ for the native–denatured transition, the difference in $\Delta G°$ (no urea) between the putative cis and trans isomers of the Ile 75–Pro 76 peptide bond is ~3 kcal/mol for oxidized thioredoxin and ~1.5 kcal/mol for reduced thioredoxin. Since cis oxidized thioredoxin is much more stable than trans, folded oxidized thioredoxin is essentially all cis. In folded reduced thioredoxin, cis and trans interconvert slowly, on the minute time scale at 2 and 25 °C. In the absence of urea, the folded reduced thioredoxin is less than a few percent trans. Three additional mutants with additions or substitutions at the active site also show electrophoresis banding patterns consistent with a difference in stability between cis and trans isomers.

Escherichia coli thioredoxin is a small, soluble protein involved in diverse redox and regulatory reactions. It participates in oxidation/reduction reactions with several types of reductases (Holmgren, 1985; Tsang & Schiff, 1976; Fuchs & Carlson, 1983) and can serve as a protein disulfide reductase (Holmgren, 1979; Scheibe, 1984). It has high sequence homology to protein disulfide isomerase (Edman et al., 1985), and folding studies suggest that thioredoxin facilitates protein folding by disulfide interchange (Pigiet & Schuster, 1986). Thioredoxin is also required in the replication cycle of some bacteriophages. DNA polymerase activity of bacteriophage T7 gene 5 protein requires prior complex formation with host thioredoxin (Modrich & Richardson, 1975; Nordström et al., 1981; Huber et al., 1986). Coat assembly in the filamentous phages M-13 and f1 also requires the presence of host thioredoxin (Russel & Model, 1985, 1986; Lim et al., 1985). In both phage systems it appears that reduced thioredoxin is the active species of the protein.

E coli thioredoxin contains 108 amino acids and has a molecular weight of 11 700. It has only two cysteine residues, at positions 32 and 35, which form the redox active S–S at the active site. *E. coli* thioredoxin has high sequence homology to other prokaryotic thioredoxins (Eklund et al., 1984; Gleason, 1986; Clement-Metral et al., 1988). Completely conserved residues include the active site sequence Trp-Cys-Gly-Pro-Cys, as well as Pro 76, which is in a cis imide bond in the crystal structure of the oxidized protein. The disulfide bond in oxi-

† This work is supported by grants from the National Science Foundation–Industry–University Cooperation Research Center for Biocatalytic Processing and the Graduate School of the University of Minnesota.
* Author to whom correspondence should be addressed.

0006-2960/89/0428-3211$01.50/0