Structure and Dynamics of Calmodulin in Solution

Willy Wriggers[1], Ernest Mehler[2], Felicia Pitici[2],
Harel Weinstein[2,3], and Klaus Schulten[1]

[1] Department of Physics and Beckman Institute
University of Illinois at Urbana-Champaign
405 N. Mathews Ave.
Urbana, IL 61801

[2] Mount Sinai School of Medicine
City University of New York
One Gustave L. Levy Place
New York, NY 10029

[3] Corresponding Author

November 11, 1998

Running Title: Solution Dynamics of Calmodulin

Keywords: target peptide binding, flexible linker, hydrophobic effect, counter ions, diffusion coefficient, molecular dynamics

Abstract

To characterize the dynamic behavior of calmodulin in solution, we have carried out molecular dynamics (MD) simulations of the Ca^{2+}-loaded structure. The crystal structure of calmodulin was placed in a solvent sphere of radius 44 Å, and 6 Cl^- and 22 Na^+ ions were included to neutralize the system and to model a 150 mM salt concentration. The total number of atoms was 32,867. During the 3 ns simulation the structure exhibits large conformational changes on the nanosecond time scale. The central α-helix, which has been shown to unwind locally upon binding of calmodulin to target proteins, bends and unwinds near residue Arg74. We interpret this result as a preparative step in the more extensive structural transition observed in the "flexible linker" region 74-82 of the central helix upon complex formation. The major structural change is a reorientation of the two Ca^{2+}-binding domains with respect to each other and a rearrangement of α-helices in the N-terminus domain which make the hydrophobic target peptide binding site more accessible. This structural rearrangement brings the domains to a more favorable position for target binding, poised to achieve the orientation observed in the complex of calmodulin with myosin-light-chain-kinase. Analysis of solvent structure reveals an inhomogeneity in the mobility of water in the vicinity of the protein which is attributable to the hydrophobic effect exerted by calmodulin's binding sites for target peptides.

Introduction

The calcium-sensing protein calmodulin (CaM) exemplifies the function of protein domain movements in the regulation of cellular processes. CaM is a small protein (Babu *et al.*, 1988; Chattopadhyaya *et al.*, 1992) consisting of 148 amino acid residues that belongs to a class of ubiquitous proteins with similar structure characterized by their distinctive helix-loop-helix Ca^{2+}-binding motif, the so-called EF hand (Kretsinger, 1980). CaM is involved in the regulation of most of the important Ca^{2+}-dependent signaling pathways in the cell (Cohen and Klee, 1988; James *et al.*, 1995). Figure 1 shows that in the crystal CaM adopts a dumbbell conformation in which the long helix connecting the two Ca^{2+}-binding domains is completely extended. In this conformation the two domains are in a *trans*-like orientation. Early Small Angle X-ray Scattering (SAXS) results suggested that the tethering helix in CaM is flexible allowing the two domains to adopt various relative orientations and separations in solution (Seaton *et al.*, 1985; Heidorn and Trewhella, 1988; Matsushima *et al.*, 1989). These results motivated computational studies of both CaM and the related protein troponin C that also supported a flexible interdomain linker (Mehler *et al.*, 1991). The molecular dynamics (MD) simulations suggested that large regions of conformational space were accessible to CaM, allowing it to adopt a continuum of relative separations and orientations, ranging from the fully extended structure observed in the crystal to a very compacted structure suggested by the SAXS studies when CaM was complexed to small peptides like mellitin or mastoparan (Kataoka *et al.*, 1989; Matsushima *et al.*, 1989). This flexibility was also supported by early NMR results (Barbato *et al.*, 1992). The functional significance of this ability of CaM to achieve a large range of relative conformations of its N-terminal and C-terminal domains became clear from the crystal structure of CaM complexed with a peptide analog of smooth muscle myosin light chain kinase (Meador *et al.*, 1992) and by the solution structure of calmodulin with a peptide analog of skeletal muscle myosin light chain kinase (Ikura *et al.*, 1992). In both of these CaM-peptide complexes a short helical segment at the interface of helices D and E (Fig. 1) was found to unwind, leading to a compacted protein with reoriented domains in *cis*-like conformation where the hydrophobic patches face each other and hold the peptide between them (reviewed by Weinstein and Mehler, 1994).

Recently, NMR structures of Ca^{2+}-free (apo) calmodulin were reported (Finn *et al.*, 1995; Zhang *et al.*, 1995; Kuboniwa *et al.*, 1995), which, in comparison with the crystal structures of Ca^{2+}-saturated (holo) calmodulin (Babu *et al.*, 1988; Chattopadhyaya *et al.*, 1992), revealed the details of the conformational transition induced by Ca^{2+}-binding to the protein. In both domains of holo calmodulin, in which all four Ca^{2+}-binding sites are occupied, a large, hydrophobic surface is exposed to the solvent (Tanaka and Hidaka, 1980). These hydrophobic patches, which are nearly completely buried in apo-CaM, form the binding interface with aliphatic helices of CaM's target proteins and contribute most of the binding energy (Ikura *et al.*, 1992; Meador *et al.*, 1992). Each hydrophobic surface is surrounded by a rim of mostly negatively charged residues which may act to orient the target sequences by interactions with positively charged residues of the target proteins, and also contribute to the binding energy in the complex (Zhang *et al.*, 1995).

The early simulations of CaM in solution suggested two types of structural changes relative to the crystal structure: A reorientation of the domains and an overall compaction of the protein (Weinstein and Mehler, 1994). Both changes have been observed in complexes of CaM with targets, as pointed out above, but it was not clear if both are required to prepare isolated CaM for the interaction with a target peptide. The incomplete solvent model used in the earlier simulations (Pascual-Ahuir *et al.*, 1991) may have led to artifactual structural changes not involved in preparing CaM for target interaction. To evaluate these possibilities in the context of a full characterization of the protein-solvent system, we immersed CaM in a large water sphere and added Na^+ and Cl^- ions to provide a neutral system at physiological ionic strength. Simulation of the isolated molecule cannot yield the structure of the complex with target peptides, but, as discussed recently (Weinstein and Mehler, 1996), indicates the intrinsic properties that determine the subtle structural rearrangements necessary for target binding.

Results of the simulation reveal how the presence of the protein alters the physical properties of the surrounding aqueous ionic solution. In ionic solution, calmodulin's negative charge (-15 e) is balanced by counterions. Most of the Na^+ counterions can be expected to form a diffuse atmosphere of ions surrounding the protein (Israelachvili, 1992). We determined the distribution of ions as they explore the accessible conformational space in the MD simulations and compared the results to a continuum model of ionic distribution derived from a numerical solution of the Poisson-Boltzmann equation (Honig and Nicholls, 1995). The dynamics of solvent water emerging from the analysis of the long trajectory identifies a change in the mobility of water molecules near the surface of CaM. Although protein-water interactions have been the subject of intense computational research (Teeter, 1991; Daggett and Levitt, 1993), it is still unclear how exposed hydrophobic sidechains alter the dynamics of the surrounding water. Recently, MD simulations suggested a reduced number of hydrogen bonds in the first solvation shell of hydrophobic compounds (Laidig and Daggett, 1996), confirming Muller's "modified hydration-shell hydrogen-bond" model for the molecular explanation of the hydrophobic effect (Muller, 1988; Muller, 1990). This model is opposed to the classic "clathrate cage" model of water hydrogen bonds (Powell, 1948), which predicts increased order near hydrophobic surfaces relative to bulk solvent. Of special interest, in this respect, is the hydration of calmodulin's hydrophobic target binding patches. In this paper we report deviations of translational water diffusion near calmodulin's hydrophobic patches and compare the results with other theoretical and computational studies of protein-solvent interactions.

Computational Methods

Preparation of Structures, Solvent Model and Molecular Dynamics

The coordinates of Ca^{2+}-loaded calmodulin were obtained from PDB entry 3CLN (Babu *et al.*, 1988). For the simulation, residues 1-4 and 148, which are missing in the crystal structure, were added as described previously (Pascual-Ahuir *et al.*, 1991). The completed protein was then immersed in a sphere of explicit water molecules with a radius of 44 Å. The water shell was constructed by covering the system with a three-dimensional lattice of water cubes which were provided by X-PLOR (Brünger, 1992). To neutralize the system and to model physiological ionic strength, 22 sodium and 6 chloride ions were iteratively placed at the minima and maxima of the electrostatic potential surrounding the protein, following the method of Mark *et al.*, 1991. This results in a system with a nominal ionic strength of 150 mM, comprising 32,867 atoms (10,474 water molecules), shown in Fig. 1.

To prepare the system the solvent was first equilibrated with fixed protein coordinates. After energy minimization, initial velocities were assigned according to a Maxwell distribution and the solvent was heated to 310 K in steps of 30 K in a 5 ps time period, and then equilibrated for 60 ps at 310 K. The total energy of the final system was minimized, followed by heating to 310 K, and equilibration for 40 ps. Free MD simulation was carried out for 3 ns. The parameters of the PAR19 united atom force field of CHARMM (Brooks *et al.*, 1983) were used for the protein with default values of the control parameters, unit dielectric constant and a 1 fs integration step. Ion parameters were obtained from the PAR22 force field for Na^+ and Cl^- and from Hori *et al.*, 1988, for Ca^{2+}. The TIP3(P) water model (Jorgensen *et al.*, 1983) was modified by relaxing internal geometry constraints to provide water flexibility. This modification was motivated by MD studies which demonstrated an improvement of physical properties of flexible water over the rigid TIP3 model (Teeter, 1991; Daggett and Levitt, 1993). Flexible water models such as F3P (Levitt and Sharon, 1988) and F3C (Levitt, 1989) better reproduce dynamical water properties compared to TIP3, which overestimates the translational diffusion coefficient by up to 60% (Daggett and Levitt, 1993; Obst and Bradaczek, 1996). Flexible TIP3P water exhibits values for the water density and heat of vaporization as close to the experimental values as those obtained with the original TIP3P model (Steinbach and Brooks, 1993). Also, the structure of water and of the ionic hydration shells is influenced little by the flexibility of the TIP3 water molecules (Guàrdia and Padró, 1996).

Structural and Statistical Analysis

Fluctuations. Since the two globular regions are joined by a flexible linker that permits a high degree of relative motion of the domains as two rigid bodies, it is necessary to remove this motion in order to compare the calculated B-factors and order parameters with the experimental results. Two least-squares fits (Kabsch, 1976) were carried out for each entry in the trajectory to remove the rigid body motions of the N-terminal (res. 1-77) and C-terminal (res. 78-148) domains, respectively. The atomic fluctuations calculated from the simulation were compared with the B-factors in the crystal from the PDB entry 3CLN (Babu *et al.*, 1988) using the relationship

$$\langle \Delta {r_i}^2 \rangle = 3B_i/(8\pi^2)\,, \tag{1}$$

where Δr_i is the atomic displacement of atom i from its average position and B_i is the corresponding B-factor. The fluctuations reported are averages of the backbone atom fluctuations for each residue, as described below.

Conformational Variability. The conformational changes of the protein during the simulation were described by the root mean square (rms) deviations from chosen reference structures. A linear measure is the time-dependent rms value that gives the departure of the protein from the starting, crystal-like structure. To monitor the time evolution of the simulation structure and its convergence in phase space, mutual rms deviations (rms_{ij}) were calculated for coordinate sets $\vec{r}(t)$ of either the N- or the C-terminal domain, using the relationship

$$\mathrm{rms}_{ij} = \sqrt{\frac{1}{N_D}\sum_{\alpha=1}^{N_D}\left[\vec{r}_\alpha(t_i) - \vec{r}_\alpha(t_j)\right]^2}, \tag{2}$$

where $\vec{r}_\alpha(t_i)$ and $\vec{r}_\alpha(t_j)$ are the least-squares fitted (Kabsch, 1976) coordinate sets of alpha-carbons $\alpha = (1, 2, \cdots N_D)$ at times t_i and t_j, and where N_D is the number of residues in the domain of interest (77 and 71 for the N- and C-terminal domains, respectively).

^{15}N Order Parameters. The orientational procedure described above results in atomic coordinates that are expressed in domain-fixed reference systems. This procedure corresponds to the model assumptions made by Barbato *et al.*, 1992, to determine the order parameters from their NMR measurements. Thus the values calculated here can be directly compared with the experimental results. The ^{15}N order parameters were calculated according to the model free formulation of Lipari and Szabo, 1982.

Identifying Rigid Domains. Hinge-bending movements, where rigid domains are connected by flexible joints, are the dominating type of conformational changes documented in the Brookhaven Protein Data Bank (Gerstein *et al.*, 1994). By means of the new *Hingefind* algorithm (Wriggers and Schulten, 1997) we examined the movements of calmodulin's structural domains that can be considered to undergo rigid body motions. The *Hingefind* algorithm compares two known conformations of a protein and identifies connected regions which exhibit preserved packing within a specified tolerance of positional fluctuations. These relatively rigid regions are found in an iterative procedure, in which poor matching residues are excluded from a selected subset of α-carbons and good matches are included. After dividing the protein into rigid domains, the algorithm determines effective rotation axes (hinges) which characterize the movements of the found domains relative to each other. The identification of hinge axes and of corresponding rotation angles provides a reduced representation of the complex movements exhibited by the protein. We extracted rigid domains by comparing calmodulin's structure after 3 ns simulation to the initial crystal structure. A 2.0 Å tolerance in the position of α-carbons was chosen to filter, as described (Wriggers and Schulten, 1997), the significant movements from local noise and thermal disorder.

Counterion Distribution. Localized clusters in the distribution of sodium counterions were determined from a three-dimensional histogram. Trajectory frames from the production run were least-squares fitted (Kabsch, 1976) to the initial protein structure every 100 fs. The positions of the 22 sodium ions in each of the fitted frames were recorded in bins of $(1.4\,\text{Å})^3$ size. The resulting density map was normalized such that the sum over all histogram densities equals the total number of sodium ions.

The electrostatic potential around CaM was calculated with GRASP (Nicholls *et al.*, 1991). A dielectric constant of 2 was assigned to the protein interior and 80 was used for water. The charges were taken from the PAR19 parameter set of CHARMM (Brooks *et al.*, 1983), resulting in a net charge of -15 e on the protein including the calcium ions. The linear Poisson-Boltzmann equation (McQuarrie, 1976; Honig and Nicholls, 1995) was solved at various ionic strengths and the concentration of sodium counterions was determined from the electrostatic potential calculated from the Boltzmann factor at temperature $T = 310\,\text{K}$.

Diffusion Coefficients. Translational diffusion coefficients can be calculated from the mean-square displacement function

$$6Dt = \langle [\mathbf{r}(\tau + t) - \mathbf{r}(\tau)]^2 \rangle + C, \tag{3}$$

where $\langle \cdots \rangle$ denotes an average over time origins τ and over a species of molecules of interest, $\mathbf{r}(t)$ is the position of a molecule at the offset time t, and C is a constant accounting for short-time effects in the dynamics of investigated molecules (Wong *et al.*, 1989; Sansom *et al.*, 1996; Breed *et al.*, 1996). For short offset times the diffusive model of motion breaks down: at the picosecond time scale the distance traveled by a molecule is only of the order of the inter-atomic distances and atom displacements may originate to a significant degree from local fluctuations in the evolving system. Hence, in simulations with periodic boundary conditions, D is usually estimated from the long time limit of Eq. (3) (Eriksson *et al.*, 1995;

Obst and Bradaczek, 1996). In our case of confined diffusion we evaluated D at short offset times t to prevent finite system size effects. Time offsets $10\,\mathrm{ps} < t < 100\,\mathrm{ps}$ were found to be practical to avoid both finite size effects and short-time correlations. We obtained mean values and standard deviations of the diffusion coefficients by evaluating Eq. (3) for this time window.

Water Mobility. The three-dimensional distribution of water mobility was evaluated as follows. Trajectory frames from the production run were least-squares fitted (Kabsch, 1976) to the initial protein structure every 1 ps. The displacements of water molecules were recorded in histogram bins of $(3.0\,\text{Å})^3$ size. The traveling water molecules force a trade-off between the feasible spatial resolution of the diffusion coefficient and the possible length of the offset time t. Hence, displacements were only recorded for short offset times $t_1 = 1\,\mathrm{ps}$ and $t_2 = 2\,\mathrm{ps}$. The short-time mobility of water surrounding the protein was characterized by the spatial distribution of displacements for $t = t_1$. To compare the short-time mobility with the diffusion coefficient determined by Eq. (3), an effective diffusion coefficient $D_{\mathrm{eff}}^{(1)}$ was computed from

$$6D_{\mathrm{eff}}^{(1)}(i,j,k)\;t_1 = \langle[\mathbf{r}(\tau+t_1) - \mathbf{r}(\tau)]^2\rangle_{ijk}\;, \tag{4}$$

where $\langle\cdots\rangle_{ijk}$ denotes an average over time origins τ and over a histogram bin at a point $\vec{r}_{ijk}$ on the three-dimensional grid of width $3.0\,\text{Å}$.

The diffusive contribution to the local motion of water molecules, $D_{\mathrm{eff}}^{(2)}$, was estimated from the slope of the mean-square displacement function averaged over the recorded displacements in the histogram bins at offset times $t_1 = 1\,\mathrm{ps}$ and $t_2 = 2\,\mathrm{ps}$ (Lounnas *et al.*, 1994),

$$6D^{(2)}_{\text{eff}}(i,j,k)\ (t_2-t_1) = \langle[\mathbf{r}(\tau+t_2)-\mathbf{r}(\tau)]^2\rangle_{ijk} - \langle[\mathbf{r}(\tau+t_1)-\mathbf{r}(\tau)]^2\rangle_{ijk}\ , \tag{5}$$

where $\langle\cdots\rangle_{ijk}$ denotes an average over time origins τ and over a histogram bin at a point $\vec{r}_{ijk}$ on the three-dimensional grid of width 3.0 Å.

Results and Discussion

Dynamics and Stability

In the course of the 3 ns MD simulation CaM undergoes a global conformational change relative to the initial crystal structure, which appears to be completed after about 1.5 ns (Fig. 2). During the second half of the simulation, the system fluctuates around a new average position at 5.2(±0.6) Å rms deviation from the initial structure. While the overall rms deviation steadily increases in the first 1.5 ns, the rms deviations of the individual domains remain much smaller (Fig. 2). The structural fluctuations of the C-terminal domain (res. 78-147) converge rapidly within 100 ps and the structure remains close to the crystal structure [1.3(±0.1) Å rms deviation]. The structural change in the N-terminal domain (residues 5-77) also reaches a plateau after about 100ps [1.5(±0.1) Å rms deviation], but the system then undergoes a further conformational change between 1.3 and 1.5 ns of the simulation, and stabilizes after about 2 ns at 2.3(±0.1) Å rms deviation (Fig. 2). The behavior of the global rms deviation reflects the reorientation of the two domains from the crystal structure to a configuration closer to that observed in CaM-target complexes. This is seen in Fig. 3 where the virtual dihedral angle (VDA) between the four calcium ions is plotted against the simulation time. The VDA was shown earlier to be a reasonable representation of the interdomain orientation (Pascual-Ahuir *et al.*, 1991). The VDA starts at the value measured in the crystal structure ($-134°$) and then steadily decreases until about 1.5 ns. Subsequently the VDA oscillates around a value of $-200°$ ($160°$ in the standard IUPAC convention). This latter value of the VDA is much closer to the value assumed when CaM is complexed to targets ($\approx 110°$ in the IUPAC convention; Ikura *et al.*, 1992). Moreover, comparison of the temporal behavior of the VDA with the global rms deviation plotted in Fig. 2 shows that the VDA stabilizes at about the same time as the rms deviation, which suggests that the relative orientation of the two domains underlies the time-dependence of the global rms deviation.

The time evolution of the structural changes can be characterized further using a two-dimensional representation of the mutual rms deviations of trajectory frames (Eq. 2). Figure 4 reveals distinct features in the sampling of the configurational space associated with each globular domain. The trajectory of the C-terminal domain corresponds to a single family of structures with mutual rms deviations < 1.3 Å. The more flexible N-terminal domain exhibits several stages of conformational changes: During a first stage that lasts until about 1.4 ns, the N-terminal domain remains close to the initial conformation. Subsequently, after increasing fluctuations in the structure starting at 900 ps simulation time, the N-terminal domain undergoes an abrupt conformational change. After 1.4 ns, the system continues to exhibit large fluctuations until about 2 ns simulation time and subsequently appears to settle into the new conformation for the remainder of the 3 ns trajectory. As Fig. 4 shows, the mutual rms deviations of the N-terminal domain structures during the last nanosecond of the simulation do not exceed $1.3(\pm 0.2)$ Å, where the standard deviation is the maximum value computed for each structural region.

The evolution of the conformational changes as shown in Fig. 4 suggests that the trajectory can be divided into a *relaxation* phase (0-2ns) and a *quiescent* phase (2-3ns). In the relaxation phase the protein responds to the absence of the intermolecular interactions present in the crystal lattice (Chattopadhyaya *et al.*, 1992), and then the system fluctuates around a new average structure without undergoing further structural rearrangements during the remainder of the trajectory (Figs. 2 and 3). Since the C-terminal domain is also involved in a large number of intermolecular interactions in the crystal structure (Chattopadhyaya *et al.*, 1992), an eventual change of conformation cannot be ruled out, especially in view of the functional significance of the structural change observed in the N-terminal domain (see below). From Fig. 4 it is seen that the structure of the C-terminal domain late in the trajectory is moving further away from the structure early in the trajectory. The increasing difference among the structures after about 2.8 ns is reminiscent of the similar increase noted in the N-terminal domain after about 900 ps that preceded the structural change.

Chain Flexibility

Atomic fluctuations were sampled over three time intervals: 0-1 ns, 1-2 ns, and 2-3 ns (Fig. 5). The patterns of internal fluctuations observed in the three intervals are fairly similar for most regions of the protein indicating the convergence of this property. One exception was the region comprising residues 39-61 that exhibit increased mobility during the second nanosecond of the simulation. These increased fluctuations probably reflect the conformational change exhibited by the N-terminal domain at 1.5 ns (Fig. 2), and the large fluctuations in the N-terminal domain that preceded and followed the structural change (Fig. 4). Two regions, comprising residues 37-42 and 110-115, exhibit high mobilities in the simulation relative to the rest of the protein. This enhanced local flexibility may be due to missing contacts in the crystal lattice, e.g., of residues 37-42 of protein (000) with residues 114-120 of protein

(011) defined by Chattopadhyaya *et al.*, 1992. Moreover, these two six-residue segments lie in homologous regions in the N-terminal and C-terminal domains (Strynadka and James, 1988), suggesting possible functional significance of the observed increased fluctuations.

The computed fluctuations are compared with those estimated from the crystal B-factors (Babu *et al.*, 1988) in Fig. 5. As was found also for other molecular systems (McCammon and Harvey, 1987), there is a good qualitative agreement between the fluctuations calculated from the last nanosecond of the MD run and the experimental values. On the average, and especially in the C-terminal domain, the fluctuations observed from the crystal structure are larger than the values calculated from the simulation, which is probably due to a certain static disorder in the crystal unrelated to intrinsic movements (Kuriyan and Weis, 1991). Because of the flexible connection of the two globular domains they behave as quasi independent bodies (Barbato *et al.*, 1992) and, therefore, the static disorder in each of them can be quite different.

Order Parameters

Global conformational changes may complicate identification of the intrinsic motional freedom of the protein backbone from the computed fluctuations. Another descriptor of the absolute backbone motion is obtained from the relaxation of bond rotation. Here the autocorrelation functions and generalized order parameters were calculated for each backbone amide bond. The autocorrelation function is given by (Lipari and Szabo, 1982)

$$C_\alpha(t) = \langle P_2(\vec{\nu}_\alpha(\tau) \cdot \vec{\nu}_\alpha(t+\tau)) \rangle, \tag{6}$$

where P_2 is the second order Legendre polynomial, $\vec{\nu}_\alpha(t)$ is the instantaneous N-H unit vector of residue α in the protein fixed-frame reference system, and the angle brackets indicate time averaging. In the "model free" approach proposed by Lipari and Szabo, 1982, the relaxation described by $C_\alpha(t)$ is encapsulated by two parameters: a generalized order parameter and the relaxation time. The order parameter, S^2, is a measure of the restriction of the angular motion of the N-H bond vector. Its value ranges from one, for completely hindered motion, to zero for completely free, isotropic motion. The order parameter of each N-H bond α is calculated from the asymptotic limit of its time autocorrelation function (Lipari and Szabo, 1982), i.e.,

$$S_\alpha^2 = \lim_{t\to\infty} C_\alpha(t) = \frac{4\pi}{5}\sum_{m=-2}^{2} |Y_{2m}(\Omega)|^2, \tag{7}$$

where Y_{2m} are the second order spherical harmonics and Ω is the orientation of the N-H bonds relative to a macromolecular fixed frame.

Autocorrelation functions and order parameters were calculated for all N-H bond vectors from the final nanosecond of the simulation, since during this interval no major changes in the structural properties occur. Similar to results from previous studies of proteins (Chandrasekhar *et al.*, 1992; Eriksson *et al.*, 1993), most residues were found to undergo fast motions that are efficiently sampled in the course of the simulation. The extended length of the present trajectory made it possible to monitor slow relaxation events for 28 residues, as evidenced by the slow convergence of $C_\alpha(t)$. Only four residues were found to undergo motions that are insufficiently sampled on the nanosecond time-scale. The validity of the simulation was further assessed by comparing the computed order parameters with experimental NMR results for CaM (Barbato *et al.*, 1992).

The $C_\alpha(t)$ of most of the N-H bond vectors showed a sharp initial decay, followed by a plateau that extends over the rest of the 460 ps sampling window used. This indicates that these N-H groups undergo small amplitude, fast librational motions along well-defined directions. The 28 residues exhibiting slower relaxation events that converged on the 300 ps time-scale, are mostly located in flexible parts of the protein, and do not participate in intramolecular interactions. Residues 70-82 of the tether belong to this group of slowly convergent $C_\alpha(t)$ and were found to undergo large structural fluctuations that mediate the relative motion of the two globular domains. The corresponding values of the order parameters of residues 77-81 are significantly larger than those observed experimentally. Finally, the correlation functions of residues Gln3, Gly61, Thr26 and Glu45 did not converge.

The N-H order parameters have been averaged over secondary structural elements and are given in Table 1. Overall, the trends observed from the measured values are fairly well reproduced; the α-helical order parameters exhibit the strongest restriction, while the intradomain linkers show the highest degree of rotational freedom. The average errors between measured and calculated order parameters for the N-terminal and C-terminal domains are 0.006 and 0.07, respectively, while the rms differences are 0.02 and 0.07, respectively, so that the calculated order parameters for the N-terminal domain are in slightly better agreement with experiment than for the C-terminal domain. Comparison of the measured order parameters (Barbato *et al.*, 1992) in the two domains shows that, on the average, the C-terminal domain has a slightly greater rotational flexibility than the N-terminal domain, in agreement with the main chain B-factors (Babu *et al.*, 1988). This difference between the two domains was not resolved by the simulation. It is finally noted that the calculation does not reproduce the experimentally observed low order parameters of residues 57 and 130 located at homologous positions of loop 2 and loop 4, respectively. Instead, the calcium-binding loops exhibit uniformly, highly restricted motions.

Domain Movements

The structural origins of the conformational changes predicted by the simulation were investigated by comparing the structure after 3 ns simulation against the initial structure in terms of rigid domains identified with the algorithm *Hingefind*, as described in Computational Methods. To determine the domains, the structure after 3 ns was averaged over the last 20 ps of simulation time and then energy-minimized. Figure 6 presents a backbone trace of this structure compared to the initial structure.

At 2.0 Å resolution, the C-terminal lobe comprises the largest rigid domain found, whereas the N-terminal lobe splits into four rigid subdomains. The two structures were superimposed by their rigid C-terminal domains, and all subdomain translations and rotations were evaluated relative to this reference orientation. The effective rotation axes determined with *Hingefind* permit a coarse representation of the exhibited conformational changes. Helix A (brown in Fig. 6) reorients and moves towards the C-terminal domain. The binding site of calcium 1 with helix B (blue), as well as helix C (yellow), rotates about axes perpendicular to the central helix, thereby opening up and extending the structure. An inspection of the 'animated' protein dynamics revealed that these movements are prominent in the first nanosecond of the simulation.

The major conformational change in the N-terminal domain that occurs around 1.5 ns (Fig. 2) causes the calcium 2 binding site (green) to swing to the side by a 65° rotation about the central helix (Fig. 6). This unwinding of the helix avoids steric clashes with the central helix when other rigid parts (calcium 1 binding site with helix B, and helix C) move to open up the N-terminal domain (Fig. 6). It appears that conformational strain in the structure also produces a kink in the central alpha helix at Met72.

The effect of the conformational changes in the N-terminal domain on the solvent-exposure of the hydrophobic target peptide binding patch is shown in Fig. 7. The re-orientation as described by the change in VDA (Fig. 3) amounts to a movement of the N-terminal domain from a *trans*-like orientation in the crystal structure towards a *cis*-like conformation required for target binding. Although the positioning of the reactive surface of the N-terminal domain of the simulated structure at 3 ns (Fig. 7) suggests that this surface may be more accessible to target peptides, its Solvent Accessible Surface Area (SASA), calculated using a standard probe of radius 1.4 Å, is actually less than the SASA from the crystal structure. This is true of the C-terminal reactive surface as well. However, the radius associated with a target peptide will be considerably larger than for a water molecule and the surface area accessible to a probe representing such a ligand may be considerably different from the SASA. The dependence of the accessible surface area (ASA) on probe size was evaluated, therefore, and is shown in Fig. 8 for both domains. Comparison of the crystal structure shows a decreasing ASA for the domains as the probe radius is increased. This type of behavior of the ASA, which also applies to the structure at 1 ns, i.e., *before* the structural change in the N-terminal domain has occurred, indicates that the reactive surfaces are concave. However, at 2 ns, *after* the structural change, the response of the ASA to probe size is different in the two domains: The ASA of the N-terminal reactive surface (top panel in Fig. 7) increases strongly with increasing probe size, while in the C- terminal (bottom panel in Fig. 7) it continues to show similar behavior as at 1 ns and as the crystal structure. The increase in ASA of the N-terminal domain is even larger at 3 ns, but here again the ASA in the C-terminal surface remains unchanged. The change in response of the

ASA in the N-terminal domain indicates that the reactive surface has changed from concave to convex, with probe radii of 10–15 Å experiencing increases in ASA by a factor of 2 to 3 relative to the crystal structure.

The ability of the N-terminal domain of CaM to relax to conformations better suited for interaction with targets suggests a functional role of its greater flexibility. A more pronounced flexibility of the N-terminal domain relative to the C-treminal domain was also observed, albeit to a lesser degree, in calculations on the isolated domains (Mehler and Weinstein, unpublished results). In contrast, the C-terminal domain remains relatively rigid (Figs. 2, 5) during the 3 ns simulation, but, as mentioned above, the C-domain may eventually undergo conformational changes as well to bring its reactive surface into a more favorable configuration for interaction with target peptides.

Conformational Changes in the Tether

In addition to the observed kink at Met72, the reorientation of the N-terminal domain is associated with an unwinding of the central α-helix at residue Arg74. This is indicated by the replacement of the typical $i, i+4$ alpha helical hydrogen bonding pattern of the crystal structure with an $i, i+5$ hydrogen bond between Arg74 and Leu69. This disruption was first observed after 900 ps simulation time. Figure 9 demonstrates that the unwinding of the α-helix is associated with a perturbation of torsion angles in residues 71-73, in particular the Ψ torsional angles in residues 71 and 72 show large deviations from the accepted range. An unwinding of the helix by one residue yields a relative rotation of 72° of the two separated parts of the alpha helix. This angular change is close to the observed change in the VDA (Fig. 3) and the rotation of the calcium 2 binding site relative to the C-terminal domain

(Fig. 6). The unwinding of the flexible linker at Arg74 observed in the simulation appears to facilitate the reorientation of the two calcium-binding domains of CaM.

In earlier simulations (Mehler *et al.*, 1991; Pascual-Ahuir *et al.*, 1991; Weinstein and Mehler, 1992) three arginine residues (Arg74, Arg86, and Arg90) were found to form and sever H-bonds with successive attachment sites from more distant parts of the structure leading to compaction and interdomain reorientation. Of these three arginines, Arg74 was most clearly involved in the structural reorganization that included both reorientation of the domains and overall compaction of the system. In contrast, the present simulation shows no compaction of the calmodulin structure: Despite the conformational changes, the radius of gyration of the protein remained stable at a value of $22(\pm 1)$ Å. The results indicate a functional role of Arg74 in the bending and unwinding of the flexible linker. In the simple solvent models used in the earlier studies the region between the two domains was free of water and thus exhibited a low dielectric volume. Consequently, it is likely that in the earlier simulations the structure responded in an artifactual way that was similar to the bending response elicited by the presence of target peptide since the latter will remove water from the region between the domains (Elcock and McCammon, 1996). Nevertheless, it is noteworthy that simulations on the isolated tethering helix indicate that it is inherently flexible (Mehler *et al.*, 1991; van der Spoel *et al.*, 1996), and therefore, it remains possible that simulations substantially longer than the current 3 ns may lead to further conformational changes in the flexible linker that would result in structural changes in the protein in the absence of target peptides.

Solvent Properties

Water makes up 86% of the total mass of the simulated system. The solvent was modeled in the absence of boundary conditions (Roberts and Schnitker, 1995), and long range electrostatic forces were truncated at 8.5 Å. These restrictions, necessary to achieve a reasonable speed of the calculation, need to be applied with care (Roberts and Schnitker, 1995; Del Buono *et al.*, 1994). We show below that the dynamic solvent properties agree well with results from generally accepted benchmark simulations and corroborate the validity of the water force field and solvent modeling approach employed in our simulation.

Counterion Distribution. The trajectory permits an improved sampling of the counterion distribution surrounding the protein: Earlier simulations of biopolymer-water systems have suffered from short simulation times that allowed the sampling of only a small fraction of the total accessible ion configuration space (Eriksson *et al.*, 1995). It is of interest to compare the time-averaged distribution of sodium ions calculated from this simulation with the ion concentration obtained from solution of the Poisson-Boltzmann equation (McQuarrie, 1976; Honig and Nicholls, 1995). Figure 10 compares the three-dimensional histogram of sodium ions sampled from the 3 ns trajectory and the local ion concentration ρ_{Na} calculated from the electrostatic potential surrounding the protein, namely

$$\rho_{\mathrm{Na}}(\vec{r}) = \rho_\infty \; exp\left(-\frac{\Xi(\rho_\infty, \vec{r}\,)}{k_B T}\right) , \tag{8}$$

where ρ_∞ is the ionic strength and $\Xi(\rho_\infty, \vec{r}\,)$ is the potential. The Figure demonstrates that the counterions are well distributed and more localized in regions of negative electrostatic potential. We estimated the effective bulk ionic strength ρ_∞ of the system to be 75 mM, which is half of the total ion concentration ([Na]+[Cl]) in our model. An inspection of computed distributions $\rho_{\mathrm{Na}}(\vec{r})$ for values of the ionic strength ranging from 10 mM to 150 mM showed that the local density is insensitive to changes in the bulk ionic strength ρ_∞.

The local densities calculated from the simulation are larger and the distribution is more scattered than predicted by the continuum model for the diffuse counterion cloud (Fig. 10). The truncation of long-range electrostatic interactions may have favored short-range ion-ion and ion-protein interactions as the ions localized in regions of high sodium density near the protein surface, and undersampling may have caused the observed sodium density fluctuations; it is possible, albeit not certain, that longer simulation times than the current 3 ns may produce a better agreement with the continuum model.

Diffusion of Solvent Molecules. The statistical properties of aqueous ionic solutions, such as solvent-solvent and solvent-ion pair distribution functions, are insensitive to changes of boundary conditions (Roberts and Schnitker, 1995), the water model (Guàrdia and Padró, 1996), and the model of Coulomb interactions (Del Buono *et al.*, 1994), in the sense that the calculated properties all are close to the experimental values. In contrast, these and other authors (Sansom *et al.*, 1996; Teeter, 1991; Daggett and Levitt, 1993) have shown that dynamical solvent properties, such as translational diffusion coefficients, *are* model-dependent and may deviate significantly from the experimental values. To test the reliability of our solvent model and of the force field parametrization, we calculated the diffusion coefficients of water molecules, sodium ions, and chloride ions. In Fig. 11 the mean square displacement functions of these solvent molecules are compared. The resulting values of the diffusion coefficients, as determined from Eq. (3), were $D_{\mathrm{water}} = 2.3(\pm 0.1) \cdot 10^{-9} m^2/s$,

$D_{\mathrm{Na}} = 0.55(\pm 0.03) \cdot 10^{-9} m^2/s$, and $D_{\mathrm{Cl}} = 1.1(\pm 0.1) \cdot 10^{-9} m^2/s$. The experimental value of pure water at 300 K is $D_{\mathrm{water}} = 2.3 \cdot 10^{-9} m^2/s$ (Eisenberg and Kauzmann, 1969) and can be extrapolated to $D_{\mathrm{water}} = 2.9 \cdot 10^{-9} m^2/s$ at the simulation temperature of 310 K (Obst and Bradaczek, 1996). The translational mobility of water molecules obtained above for our protein-solvent system is slightly smaller than measured experimentally. The agreement with the experimental value is good compared to simulations of pure TIP3 water ($D_{\mathrm{water}} = 4.0 \cdot 10^{-9} m^2/s$; Jorgensen, 1981). The observed sodium ion diffusion is low possibly due to the preference of sodium ions for positions near the negatively charged protein surface (Fig. 10). The value of D_{Na} obtained here is close to the value $0.5 \cdot 10^{-9} m^2/s$ for simulated Na^+ near DNA (Eriksson *et al.*, 1995) [1]. Other simulations resulted in a value of $D_{\mathrm{Na}} = 1.2$–$1.3 \cdot 10^{-9} m^2/s$ (Norberg and Nilsson, 1994; Obst and Bradaczek, 1996), and the experimental datum is $D_{\mathrm{Na}} = 1.33 \cdot 10^{-9} m^2/s$ (Hertz, 1973). The calculated diffusion of the chloride ions is also low compared to the experimental value of $2.03 \cdot 10^{-9} m^2/s$ (Tyrell and Harris, 1984). Values of $D_{\mathrm{Cl}} = 0.5$–$2.2 \cdot 10^{-9} m^2/s$ (Roberts and Schnitker, 1995; Guàrdia and Padró, 1996) have been reported for simulated Cl^- diffusion. Our estimate of D_{Cl} near the protein, therefore, appears reasonable.

The diffusion coefficient D, as determined by Eq. (3), has a physical meaning only for homogeneous and isotropic systems, in which the solvent mobility is independent of the positions visited by each solvent molecule. Consequently, for an inhomogeneous system such as a solvated protein, Eq. (3) only describes an average property of the system. A physically more meaningful representation of the solvent mobility in the vicinity of a protein should take into account the spatial dependence of the diffusive motion. As described in Computational

[1] Similar to water, one would expect a slight increase in ionic diffusion at 310 K relative to 300 K. Here we only discuss experimental and computational values obtained for the more common temperature 300 K

Methods, it is possible to compute the mobility of solvent molecules in localized bins of a three-dimensional histogram, if the distance travelled within the sampling time t [Eq. (3)] is small compared to the bin size used.

Figure 12 presents contours of constant rms displacements of water molecules for $t = 1.0\,\mathrm{ps}$. The short offset time allows one to determine the distribution of the depicted displacements at a resolution smaller than the 3.0 Å bin size used to calculate the global value of D. Water-protein interactions and the water model used cause a spatial inhomogeneity of observed water mobility. An increase of mobility near the surface of the solvent sphere is artifactual, due to the absence of bulk solvent which would surround the spherical system. A decrease in mobility to values as low as $D_{\mathrm{eff}}^{(1)} = 2.0 \cdot 10^{-9} m^2/s$ can be attributed to the coordination of water molecules by the polar surface of the protein. An inspection of the distribution of $D_{\mathrm{eff}}^{(1)}$ values reveals that the hydrophobic effect exerted on the solvent induces an enhanced mobility to values as high as $D_{\mathrm{eff}}^{(1)} = 10 \cdot 10^{-9} m^2/s$ at calmodulin's hydrophobic target-peptide binding patches.

The observed short-time water mobility corresponds to an effective diffusion coefficient $D_{\mathrm{eff}}^{(1)}$ [Eq. (4)] that encompasses both diffusive and nondiffusive contributions to the short-time water motion. The average of $D_{\mathrm{eff}}^{(1)}$ over all histogram bins (Fig. 12) yields a water diffusion coefficient $D_{\mathrm{water}} = 3.6 \cdot 10^{-9} m^2/s$. The coefficient is overestimated by 36% relative to the value of $2.3 \cdot 10^{-9} m^2/s$ determined above from the averaged mean-square displacement function. The water mobility depicted in Fig. 12 can therefore be attributed to a significant degree to non-diffusive movements from short-time correlations in the solvent dynamics (cf. Computational Methods). Figure 13 shows the contours of effective diffusion coefficient $D_{\mathrm{eff}}^{(2)}$, computed from the slope of the mean-square displacements at times $t = 1\,\mathrm{ps}$ and $2\,\mathrm{ps}$, as described in Computational Methods [Eq. (5)]. A comparison with Fig. 12 reveals two prominent changes in the depicted water mobility: First, there is no noticeable increase of diffusive mobility near calmodulin's hydrophobic patches, i.e., the diffusion coefficient uniformly decreases with decreasing distance to the protein surface. Second, the depicted diffusion coefficient $D_{\text{eff}}^{(2)}$ is lower than the effective coefficient $D_{\text{eff}}^{(1)}$ (Fig. 12), due to the neglect of nondiffusive motion. The average of $D_{\text{eff}}^{(2)}$ over all histogram bins yields a water diffusion coefficient $D_{\text{water}} = 2.9 \cdot 10^{-9} m^2/s$, which is in closer agreement with the value of $2.3 \cdot 10^{-9} m^2/s$ determined above. Therefore, $D_{\text{eff}}^{(2)}$, represents a good approximation of the diffusive contribution to the water motion.

Summary and Conclusion

The present MD simulation of Ca^{2+}-saturated calmodulin in solution has revealed several structural features relevant to calmodulin's ability to form complexes with target proteins. Relaxation from the crystal constraints allows the central helix to bend and unwind at residue Arg74. The simulation yields a structure of calmodulin that has functional significance: Calmodulin's two domains reorient to a more *cis*-like conformation and the hydrophobic binding patches appear more accessible for the binding of a target-peptide helix. Our results imply that calmodulin's structure in the crystal is distorted by packing forces. A marked effect of crystal packing on the structure of a protein is not uncommon. Recently, it was shown that applying osmotic pressure to crystals of profilin:actin complexes brings about a collapse of the crystal unit cell, which corresponds to a significant conformational change in the protein actin (Chik *et al.*, 1996). A slight difference in energy between the open and the collapsed state revealed an extreme sensitivity of the actin conformation to changes in the chemical and thermal environment of the crystal (Chik *et al.*, 1996).

After relaxation from the crystal constraints, the calmodulin structure stabilized during the latter part of the simulation trajectory. In this *quiescent* phase of the simulation, the structure fluctuated about a new average conformation although it is unlikely that the conformational relaxation from the crystal was completed. As discussed recently, sampling errors due to short simulation times preclude the identification of slow modes in the dynamics of proteins (Balsera *et al.*, 1996). Likewise, in the 3 ns simulation, we do not expect to have sampled all possible movements of calmodulin. The two individual domains, however, should be well equilibrated, due to the small system size and the relatively long simulation time.

The relatively long simulation times in the fully solvated system that are required for the appearance of the structural changes associated with CaM function demonstrate that this process cannot be characterized with short simulations of a few hundred picoseconds. Rather, trajectories of at least several nanoseconds are required, and even these cannot provide a full account of the details in the molecular dynamics that control the relationship between structure and function in CaM. These findings are likely to be generalizable to other systems that undergo large structural changes related to their function.

Although the structure-function relationship of calmodulin has been the subject of intense research (Weinstein and Mehler, 1994), one basic question was not answered fully. This question concerns the mechanism whereby calmodulin is able to bind and activate a variety of distinct proteins in the cell. A considerable body of evidence indicates that calmodulin can assume multiple modes of interaction with targets other than the modes realized in the complexes with myosin light chain kinase (Ikura *et al.*, 1992; Meador *et al.*, 1992). For example, in the γ-subunit of phosphorylase kinase two non-contiguous peptides (Phk5 and Phk13) form complexes with calmodulin. Whereas Phk5 has an α-helical conformation in the complex, it appears that Phk13 forms an elongated complex, in which the peptide is bent into a hairpin-shaped structure (Juminaga *et al.*, 1993). The spacing of hydrophobic residues of cyclic nucleotide phosphodiesterase (Charibonneau *et al.*, 1991) and calmodulin kinase II (Bennet and Kennedy, 1987) suggests that their binding to calmodulin is different yet again.

Some of the intrinsic properties of calmodulin that enable it to realize multiple target binding modes have been identified. Experimental observations indicate an asymmetry in target affinity of calmodulin's domains. The tryptic fragment corresponding to the C-terminal domain binds preferentially to target peptides (Newton *et al.*, 1984; Newton *et al.*, 1985), at more than 100 times higher affinities (Sanyal *et al.*, 1988), compared to the fragment corresponding to the N-terminal domain. Our results reveal that calmodulin's C-terminal domain is relatively rigid, whereas the central α-helix and the N-terminal domain are flexible. This relative ordering in the rigidity of the two domains support a binding mechanism where the C-terminal domain binds to its target first, followed by the N-terminal domain (Persechini *et al.*, 1994). The inherent flexibility of both the N-terminal domain and the tethering linker allow the N-terminal domain to assume a suitable binding conformation once the second binding site is recognized. The nature and size of the hydrophobic patches involved in target binding depend on this flexible reorientation and constitute elements of recognition specificity.

A large number of computational and experimental studies focused on the behavior of water molecules near the surface of solvated proteins (Teeter, 1991; Daggett and Levitt, 1993).

Such studies indicate that the dynamic properties of water close to the protein are perturbed, but there is some disagreement on whether water molecules are immobilized (Wong and McCammon, 1986; Brooks and Karplus, 1989; Levitt and Sharon, 1988; Tirado-Rives and Jorgensen, 1990) or hypermobile (Lounnas *et al.*, 1994) relative to bulk water. Our results indicate a uniformly reduced diffusion-coefficient of water near calmodulin. However, in terms of short-time mobility, we need to differentiate between water molecules near the polar surface of the protein and water molecules near the hydrophobic patches. Our results demonstrate an increased short-time mobility of waters near the hydrophobic patches, rooted in nondiffusive motion. Other authors have reported an effect of hydrophobicity on the diffusive motion relative to polar residues, but there is disagreement on whether hydrophobic surfaces increase (Levitt and Sharon, 1988) or decrease (Karplus and Rossky, 1980; Brooks and Karplus, 1989) the diffusion coefficient. Clearly, the disagreement between the computational studies suggests that the simulation of the physical properties of water at a protein surface are model-dependent due to limitations of the various protein and water force fields.

For realistic simulations of solvated proteins molecular dynamics force fields require refinements that will yield better agreement with the experimentally observed behavior of aqueous solutions. Of special interest for such refinements are water molecules near the protein surface, which have been elucidated by recent protein crystal structures. It was hypothesized (Klotz, 1958), that water near hydrophobic residues would form clathrate cages (Powell, 1948), but a study of water distributions in 16 protein structures with resolutions below 1.7 Å did not reveal such patterns (Thanki *et al.*, 1988). A more recent model of the hydrophobic effect suggests a reduced number of hydrogen bonds in the first solvation shell of hydrophobic compounds (Muller, 1988; Muller, 1990). This model has been supported by computer simulations (Laidig and Daggett, 1996). Water molecules with a reduced number of hydrogen bonding partners should be kinetically more labile. In this respect Muller's hypothesis is consistent with the enhanced mobility of water molecules for short times and small displacements observed in this work for calmodulin's hydrophobic patches. On the other hand, the classical picture of clathrate cages seems to pertain to the average structure of the hydrogen-bonding network, which is not inconsistent with the existence of kinetically labile waters at a hydrophobic surface.

Acknowledgements

We would like to thank Paul Adams and Axel Brünger for providing the parallel version of the program X-PLOR. This work was supported by NIH grants PHS 5 P41 RR05969-04 (WW and KS), DA00060 (HW), and GM-41373 (FP, EM and HW), NSF grants BIR-9318159 and BIR-9423827-EQ (WW and KS), Roy J. Carver Charitable Trust (WW and KS) and MCA93S028P computertime grant at Pittsburgh Supercomputing Center (WW and KS).

References

Babu, Y. S., C. E. Bugg, and W. J. Cook. 1988. Structure of calmodulin refined at 2.2 Å resolution. *J. Mol. Biol.* 204:191–204.

Balsera, M. A., W. Wriggers, Y. Oono, and K. Schulten. 1996. Principal component analysis and long time protein dynamics. *J. Phys. Chem.* 100:2567–2572.

Barbato, G., M. Ikura, L. E. Kay, R. W. Pastor, and A. Bax. 1992. Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy: The central helix is flexible. *Biochemistry* 31:5269–5278.

Bennet, M. K., and M. B. Kennedy. 1987. Deduced primary structure of the β subunit of brain type II Ca^{2+}/calmodulin dependent protein kinase determined by molecular cloning. *Proc. Natl. Acad. Sci. USA* 84:1794–1798.

Breed, J., R. Sankararamakrishnan, I. D. Kerr, and M. S. P. Sansom. 1996. Molecular dynamics simulations of water within models of ion channels. *Biophys. J.* 70:1643–1661.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* 4:187–217.

Brooks, C. L., and M. Karplus. 1989. Solvent effects on protein motion and protein effects on solvent motion: Dynamics of the active site region of lysozyme. *J. Mol. Biol.* 208:159–181.

Brünger, A. T. 1992. X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR. The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University.

Chandrasekhar, I., G. Clore, A. Szabo, A. Gronenborn, and B. Brooks. 1992. A 500 ps molecular dynamics simulation study of interleukin-1β in water. *J. Mol. Biol.* 226:239–250.

Charibonneau, H., S. Kumar, J. P. Novack, D. K. Blumenthal, P. R. Griffin, J. Shabanowitz, D. F. Hunt, J. A. Beavo, and K. A. Walsh. 1991. Evidence for domain organization within the 61-kDa calmodulin-dependent cyclic nucleotide phosphodiesterase. *Biochemistry* 30:7931–7940.

Chattopadhyaya, R., W. E. Meador, A. R. Means, and F. A. Quiocho. 1992. Calmodulin structure refined at 1.7 Å resolution. *J. Mol. Biol.* 228:1177–1192.

Chik, J. K., U. Lindberg, and C. E. Schutt. 1996. The structure of an open state of beta actin at 2.65 Å resolution. *J. Mol. Biol.* 263:607–623.

Cohen, P., and C. B. Klee. 1988. Calmodulin. Elsevier Science Publishers, Amsterdam, The Netherlands.

Daggett, V., and M. Levitt. 1993. Realistic simulations of native-protein dynamics in solution and beyond. *Ann. Rev. Biophys. Biomol. Struct.* 22:353–380.

Del Buono, D. S., T. S. Cohen, and P. J. Rossky. 1994. Effects of long-range interactions on the dynamics of ions in aqueous solution. *J. Mol. Liq.* 60:221–236.

Eisenberg, D., and W. Kauzmann. 1969. The Structure and Properties of Water. Oxford University Press, Oxford.

Elcock, A., and J. McCammon. 1996. The low dielectric interior of proteins is sufficient to cause major structural changes in DNA on association. *J. Am. Chem. Soc.* 118:3787–3788.

Eriksson, M. A. L., H. Berglund, T. Härd, and L. Nilsson. 1993. A comparison of ^{15}N NMR relaxation measurements with a molecular dynamics simulation: Backbone dynamics of the glucocorticoid receptor DNA-binding domain. *Proteins* 17:375–390.

Eriksson, M. A. L., T. Härd, and L. Nilsson. 1995. Molecular dynamics simulations of the glucocorticoid receptor DNA-binding domain in complex with DNA and free in solution. *Biophys. J.* 68:402–426.

Finn, B. E., J. Evenas, T. Drakenberg, J. P. Waltho, E. Thulin, and S. Forsen. 1995. Calcium-induced structural changes and domain autonomy in calmodulin. *Nature Struct. Biol.* 2:777–783.

Gerstein, M., A. M. Lesk, and C. Chothia. 1994. Structural mechanisms for domain movements in proteins. *Biochemistry* 33:6739–6749.

Guàrdia, E., and J. A. Padró. 1996. On the structure and dynamic properties of aqueous solutions: Molecular dynamics simulation of Cl^- and Cl^{2-} in water. *Molecular Simulation* 17:83–94.

Heidorn, D. B., and J. Trewhella. 1988. Comparison of the crystal and solution structures of calmodulin and troponin C. *Biochemistry* 27:909–915.

Hertz, H. G. 1973. Nuclear magnetic relaxation spectroscopy. *In* Water: A comprehensive treatise (vol. 3). F. Franks, editor. Plenum Press, New York.

Honig, B., and A. Nicholls. 1995. Classical electrostatics in biology and chemistry. *Science* 268:1144–1149.

Hori, K., J. N. Kushick, and H. Weinstein. 1988. Structural and energetic parameters of Ca^{2+} binding to peptides and proteins. *Biopolymers* 27:1865–1886.

Humphrey, W. F., A. Dalke, and K. Schulten. 1996. VMD – Visual molecular dynamics. *J. Mol. Graphics* 14:33–38.

Ikura, M., G. M. Clore, A. M. Gronenborn, G. Zhu, C. B. Klee, and A. Bax. 1992. Solution structure of a calmodulin-target peptide complex by multidimensional NMR. *Science* 256:632–638.

Israelachvili, J. N. 1992. Intermolecular and Surface Forces. Academic Press, London.

James, P., T. Vorherr, and E. Carafoli. 1995. Calmodulin-binding domains: just two-faced or multi-facetted? *Trends in Biochemical Sciences* 20:38–42.

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

Jorgensen, W. 1981. Transferable intermolecular potential functions for water, alcohol, and esters. application to liquid water. *J. Am. Chem. Soc.* 103:335–340.

Juminaga, D., S. A. Albaugh, and R. F. Steiner. 1993. The interaction of calmodulin with regulatory peptides of phosphorylase kinase. *J. Biol. Chem.* 269:1660–1667.

Kabsch, W. 1976. A solution for the best rotation to relate two sets of vectors. *Acta Cryst. A* 32:922–923.

Karplus, M., and P. Rossky. 1980. Solvation: A molecular dynamics study of a dipeptide in water. *ACS Symposium Series* 127:23–42.

Kataoka, M., J. Head, B. Seaton, and D. Engelman. 1989. Melittin binding causes a large calcium-dependent conformational change in calmodulin. *Proc. Natl. Acad. Sci. USA* 86:6944–6948.

Klotz, I. M. 1958. Protein hydration and behavior. *Science* 128:815–822.

Kretsinger, R. H. 1980. Structure and evolution of calcium-modulated proteins. *CRC. Crit. Rev. Biochem.* 8:119–174.

Kuboniwa, H., N. Tjandra, S. Grzesiek, H. Ren, C. B. Klee, and A. Bax. 1995. Solution structure of calcium-free calmodulin. *Nature Struct. Biol.* 2:768–776.

Kuriyan, J., and W. I. Weis. 1991. Rigid protein motion as a model for crystallographic temperature factors. *Proc. Natl. Acad. Sci. USA* 88:2773–2777.

Laidig, K. E., and V. Daggett. 1996. Testing the modified hydration-shell hydrogen-bond model of hydrophobic effects using molecular dynamics simulation. *J. Phys. Chem.* 100:5616–5619.

Levitt, M. 1989. Molecular dynamics of macromolecules in water. *Chem. Scr.* 29A:197–203.

Levitt, M., and R. Sharon. 1988. Accurate simulation of protein dynamics in solution. *Proc. Natl. Acad. Sci. USA* 85:7557–7561.

Lipari, G., and A. Szabo. 1982. Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. *J. Am. Chem. Soc.* 104:4546–4559.

Lounnas, V., B. M. Pettitt, and G. N. Phillips. 1994. A global model of the protein-solvent interface. *Biophys. J.* 66:601–614.

Mark, A. E., H. J. C. Berendsen, and W. F. van Gunsteren. 1991. Conformational flexibility of aqueous monomeric and dimeric insulin: A molecular dynamics study. *Biochemistry* 30:10866–10872.

Matsushima, N., Y. Izumi, T. Matsuo, H. Yoshino, T. Ueki, and Y. Miyake. 1989. Binding of both Ca^{2+} and mastoparan to calmodulin induces a large change in tertiary structure. *J. Biochem.* 105:883–887.

McCammon, J. A., and S. C. Harvey. 1987. Dynamics of Proteins and Nucleic Acids. Cambridge University Press, Cambridge.

McQuarrie, D. A. 1976. Statistical Mechanics. Harper and Row, New York.

Meador, W. E., A. R. Means, and F. A. Quiocho. 1992. Target enzyme recognition by calmodulin: 2.4 Å structure of a calmodulin-peptide complex. *Science* 257:1251–1255.

Mehler, E. L., J. L. Pascual-Ahuir, and H. Weinstein. 1991. Structural dynamics of calmodulin and troponin C. *Protein Engng.* 4:625–637.

Muller, N. 1988. Is there a region of highly structured water around a nonpolar solute molecule? *J. Sol. Chem.* 17:661–672.

Muller, N. 1990. Search for a realistic view of hydrophobic effects. *Acc. Chem. Res.* 23:23–28.

Newton, D., M. Oldewurtel, M. Krinks, J. Shiloach, and C. Klee. 1984. Agonist and antagonist properties of calmodulin fragments. *J. Biol. Chem.* 259:4419–4426.

Newton, D., C. Klee, J. Woodgett, and P. Cohen. 1985. Selective effects of CAPP1-calmodulin on its target proteins. *Biochim. Biophys. Acta* 845:533–539.

Nicholls, A., K. A. Sharp, and B. Honig. 1991. Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Structure, Function and Genetics* 11:281–296.

Norberg, J., and L. Nilsson. 1994. High pressure molecular dynamics of a nucleic acid fragment. *Chem. Phys. Lett.* 224:219–224.

Obst, S., and H. Bradaczek. 1996. Molecular dynamics study of the structure and dynamics of the hydration shell of alkaline and alkaline-earth metal cations. *J. Phys. Chem.* 100:15677–15687.

Pascual-Ahuir, J.-L., E. Mehler, and H. Weinstein. 1991. Calmodulin structure and function: Implication of arginine residues in the compaction related to ligand binding. *Molec. Engng.* 1:231–247.

Persechini, A., K. McMillan, and P. Leakey. 1994. Activation of myosin light chain kinase and nitric oxide synthase activities by calmodulin fragments. *J. Biol. Chem.* 269:16148–16154.

Powell, H. M. 1948. Clathrate compounds. *J. Chem. Soc.* 1948:61–73.

Ramachandran, G. N., C. Ramakrishnan, and V. Sasisekharan. 1963. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* 7:95–99.

Roberts, J. E., and J. Schnitker. 1995. Boundary conditions in simulations of aqueous ionic solutions: A systematic study. *J. Phys. Chem.* 99:1322–1331.

Sansom, M. S. P., I. D. Kerr, J. Breed, and R. Sankararamakrishnan. 1996. Water in channel-like cavities: Structure and dynamics. *Biophys. J.* 70:693–702.

Sanyal, G., L. Richard, K. Carraway, and D. Puett. 1988. Binding of amphiphilic peptides to carboxy-terminal tryptic fragment of calmodulin. *Biochemistry* 27:6229–6236.

Seaton, B. A., J. F. Head, D. M. Engelman, and F. M. Richards. 1985. Calcium-induced increase in the radius of gyration and maximum dimension of calmodulin measured by small-angle X-ray scattering. *Biochemistry* 24:6740–6743.

Steinbach, P. J., and B. R. Brooks. 1993. Protein hydration elucidated by molecular dynamics simulation. *Proc. Natl. Acad. Sci. USA* 90:9135–9139.

Strynadka, N., and M. James. 1988. Two trifluoperazine-binding sites on calmodulin predicted from comparative molecular modeling with troponin-C. *Proteins: Structure, Function, and Genetics* 3:1–17.

Tanaka, T., and H. Hidaka. 1980. Hydrophobic regions function in calmodulin–enzyme interaction. *J. Biol. Chem.* 255:11078–11080.

Teeter, M. M. 1991. Water-protein interactions: Theory and experiment. *Ann. Rev. Biophys. Biophys. Chem.* 20:577–600.

Thanki, N., J. M. Thornton, and J. M. Goodfellow. 1988. Distributions of water around amino acid residues in proteins. *J. Mol. Biol.* 202:637–657.

Tirado-Rives, J., and W. L. Jorgensen. 1990. Molecular dynamics of proteins with the OPLS potential functions. simulation of the third domain of silver pheasant ovomucoid in water. *J. Am. Chem. Soc.* 112:2773–2781.

Tyrell, H. J. V., and K. R. Harris. 1984. Diffusion in Electrolytes. Butterworths, London.

van der Spoel, D., B. L. de Groot, S. Hayward, H. J. C. Berendsen, and H. J. Vogel. 1996. Bending of the calmodulin central helix: A theoretical study. *Protein Science* 5:2044–2053.

Weinstein, H., and E. L. Mehler. 1992. Structural specificity in the engineering of biological function: Insights from the dynamics of calmodulin. *In* Molecular Aspects of Biotechnology: Computational Models and Theories, . J. Bertran, editor. Kluwer Academic Publishers, Dordrecht, The Netherlands. 153–173.

Weinstein, H., and E. L. Mehler. 1994. Ca^{2+}-Binding and structural dynamics in the functions of calmodulin. *Ann. Rev. Physiol.* 56:213–236.

Weinstein, H., and E. L. Mehler. 1996. Structural change mechanisms in regulatory proteins. *Science* 271:1792–1793.

Wong, C. F., and J. A. McCammon. 1986. Computer simulation and the design of new biological molecules. *Isr J. Chem.* 27:211–215.

Wong, C. F., C. Zheng, and J. A. McCammon. 1989. Glass transition in SPC/E water and in protein solution: A molecular dynamics simulation study. *Chem. Phys. Lett.* 154:151–154.

Wriggers, W., and K. Schulten. 1997. Protein domain movements: Detection of rigid domains and visualization of hinges in comparisons of atomic coordinates. *Proteins: Structure, Function, and Genetics* 29:1–14.

Zhang, M., T. Tanaka, and M. Ikura. 1995. Calcium-induced conformational transition revealed by the solution structure of apo calmodulin. *Nature Struct. Biol.* 2:758–767.

Element	Residues	S^2 Simulation	S^2 NMR
Helix A	5–19	0.84	0.84
Helix B	29–38	0.86	0.83
Helix C	45–55	0.85	0.83
Helix D	65–76	0.84	0.85
Helix E	82–92	0.86	0.78
Helix F	102–111	0.87	0.82
Helix G	118–128	0.85	0.81
Helix H	138–147	0.86	0.80
Loop 1	20–28	0.82	0.83
Loop 2	56–64	0.83	0.80
Loop 3	93–101	0.85	0.79
Loop 4	129–137	0.85	0.77
Linker 1	39–44	0.73	0.75
Linker 2	112–117	0.75	0.64

Table 1: Average order parameters for secondary structural elements of CaM. Comparison of simulation and NMR results (Barbato *et al.*, 1992). The average values do not include those residues for which the correlation functions did not converge during the simulation, and for which there is no experimental data.

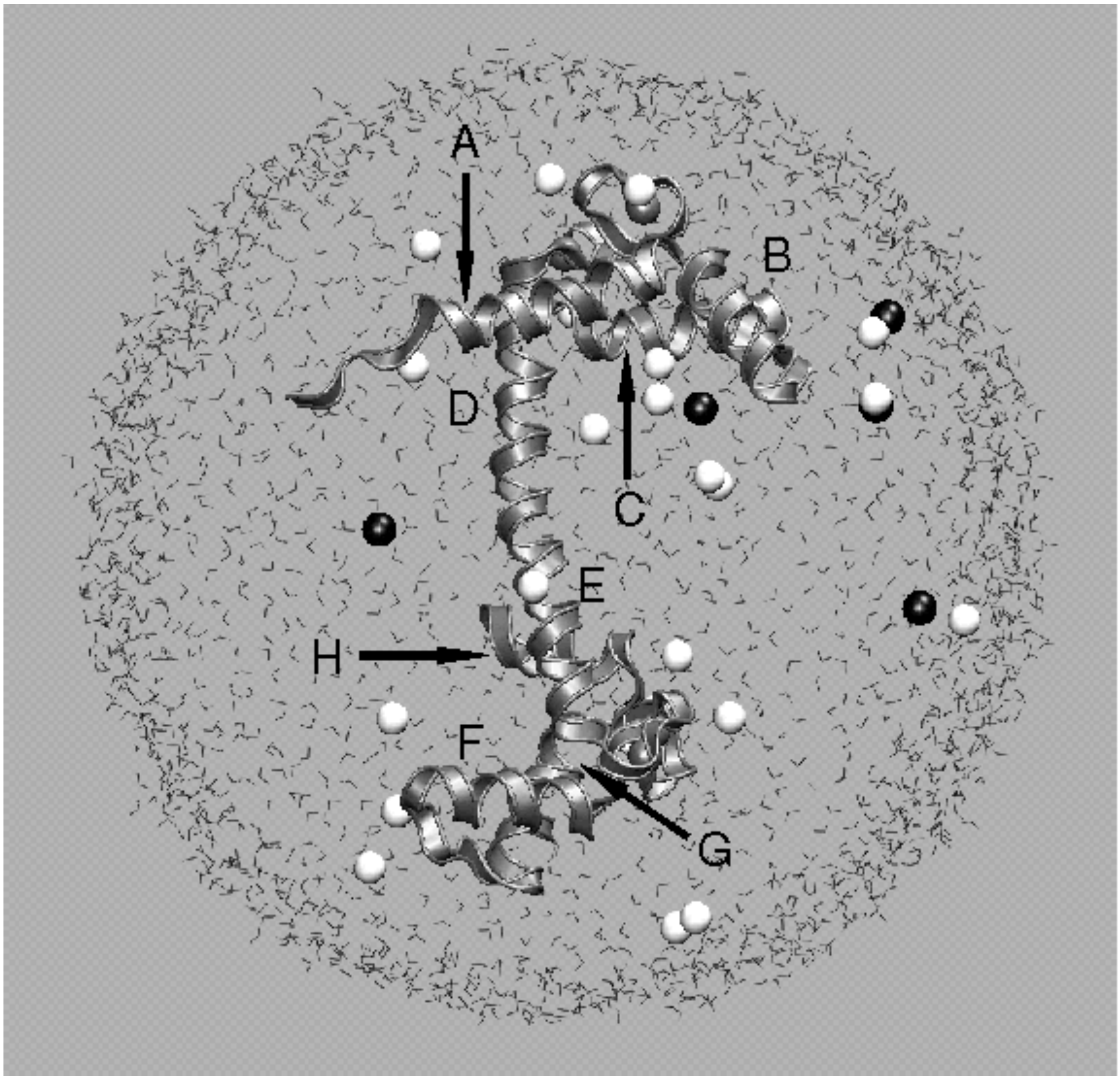

Figure 1: Orthographic projection of calmodulin (ribbon) from the crystal structure (Babu *et al.*, 1988) with bound Ca^{2+} (grey), embedded in a 44 Å radius sphere of 10,474 water molecules (outer 4 Å shell shown). 22 Na^{+} (white) and 6 Cl^{-} ions (black) were placed to neutralize the system at physiological ionic strength. The eight α-helices, comprising residues 5-19 (helix A), 29-38 (helix B), 45-55 (helix C), 65-78 (helix D), 82-92 (helix E), 102-111 (helix F), 118-128 (helix G) and 138-147 (helix H), are indicated. The structure is visualized by the molecular graphics program *VMD* (Humphrey *et al.*, 1996).

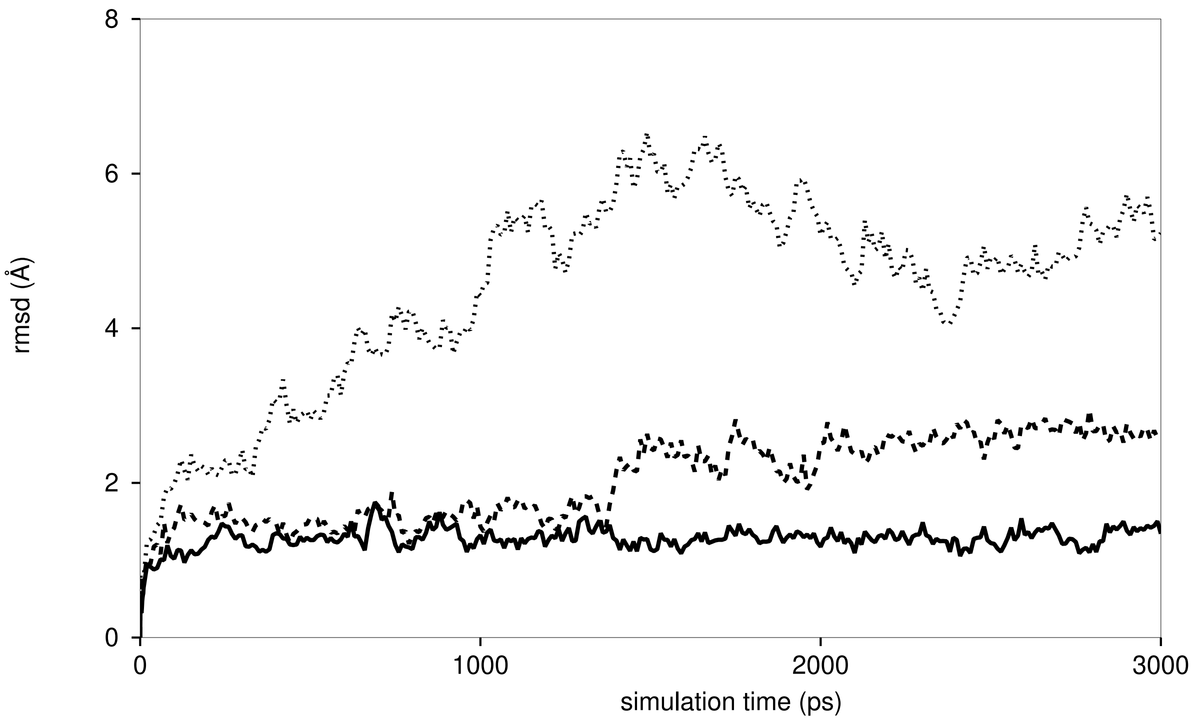

Figure 2: Global rms deviations of calmodulin α-carbons (dotted line) as well as rms deviations of the α-carbons of the N-terminal domain (res. 5-77, dashed line), and of the C-terminal domain (res. 78-147, solid line), as a function of simulation time.

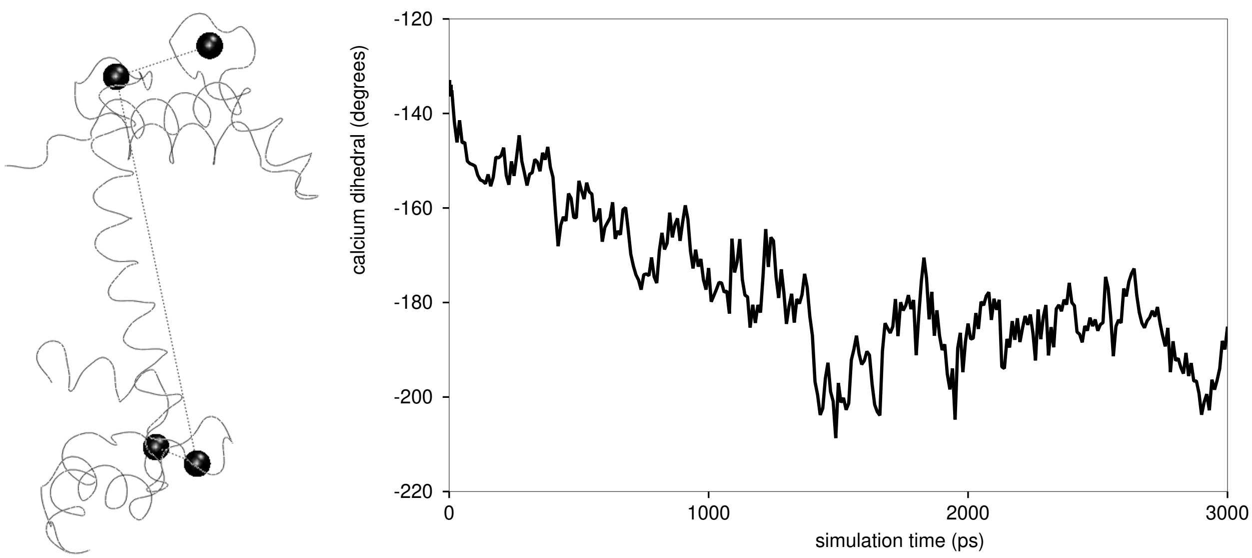

Figure 3: Left: Definition of the calcium dihedral angle. Right: The calcium dihedral angle as a function of simulation time.

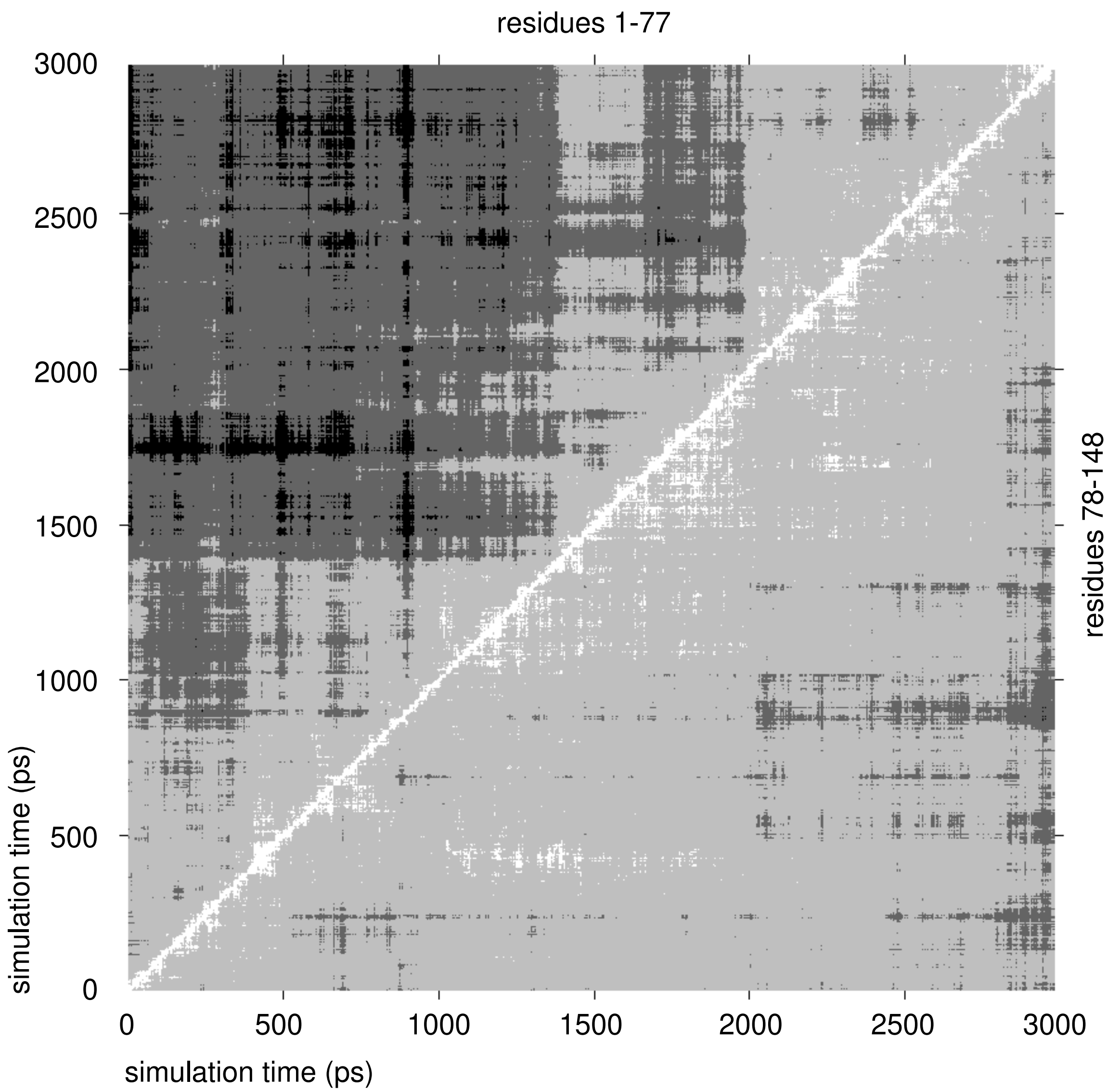

Figure 4: Matrix representation of mutual rms differences rms_{ij} calculated from coordinate sets in the 3 ns trajectory of CaM (Eq. 2). Upper triangle: N-terminal domain; lower triangle: C-terminal domain. $[0.0 \leq \mathrm{rms} < 0.5]$, white; $[0.5 \leq \mathrm{rms} < 1.0]$, light gray; $[1.0 \leq \mathrm{rms} < 1.5]$, dark gray; $[1.5 \leq \mathrm{rms} < 2.0]$, black.

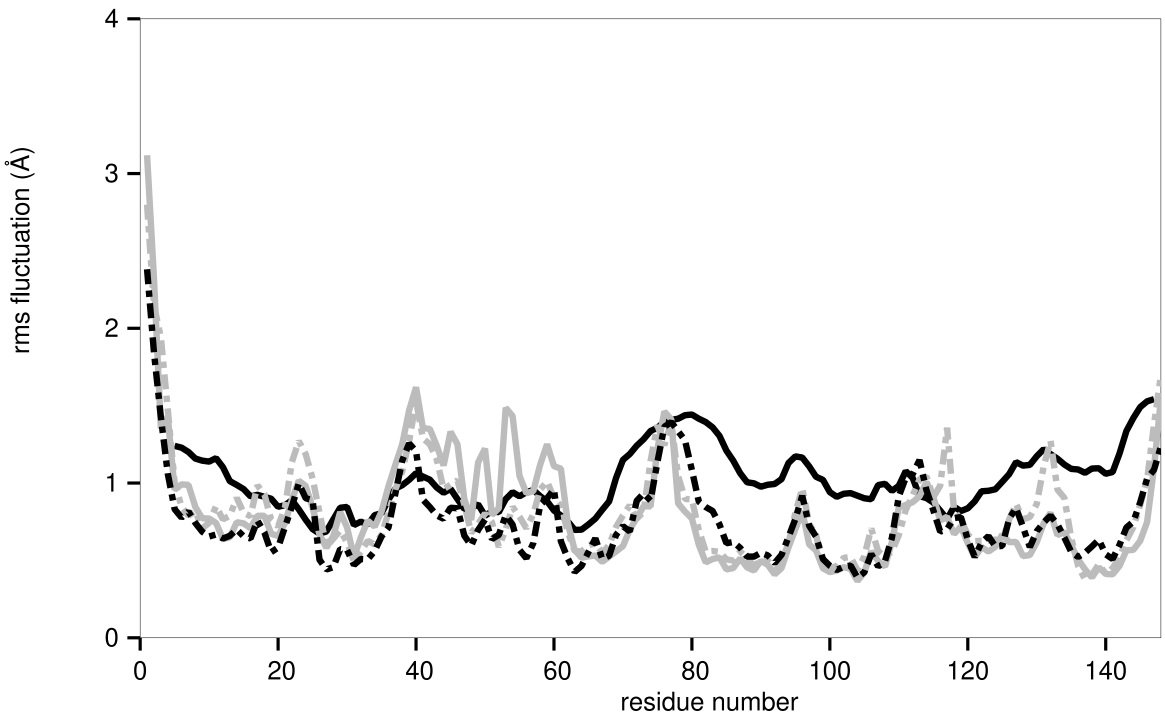

Figure 5: Comparison of rms fluctuations from simulation and from crystallographic B-factors (Babu *et al.*, 1988; solid, black). Fluctuations computed individually for the N-terminal domain (res. 1-77) and C-terminal domain (res. 78-148) sampled from 0-1 ns (dashed, grey), 1-2 ns (solid, grey), and 2-3 ns (dashed, black) are shown.

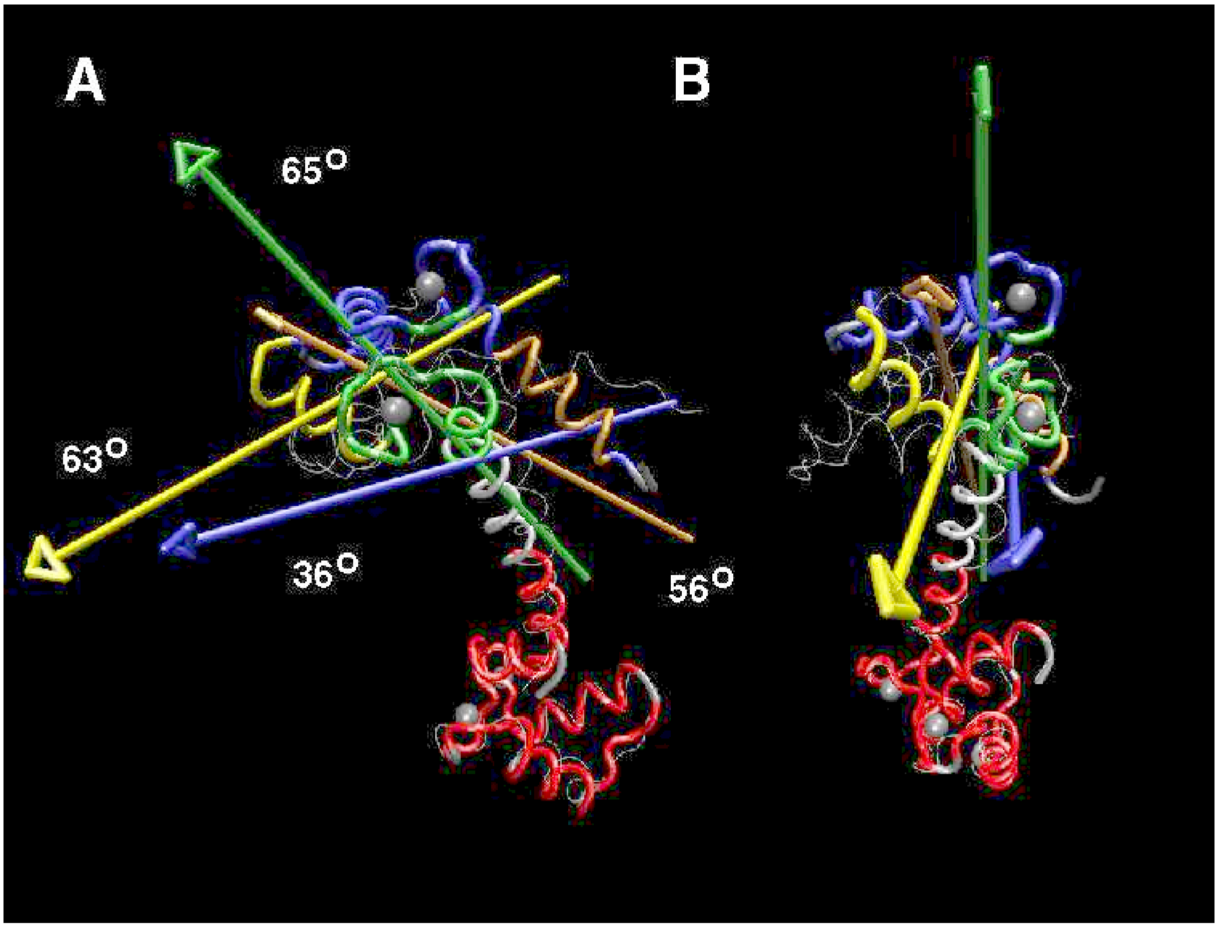

Figure 6: Domain movements of simulated structure (A: side view, B: front view). The colored tube represents the structure after 3 ns (averaged over the last 20 ps) compared to the initial crystal structure (white line). Five rigid domains of more than 10 residues were extracted by *Hingefind* (Wriggers and Schulten, 1997) partitioning at 2.0 Å resolution. Domain 1 (red), the reference domain, comprises 63 residues located in the C-terminal domain. The movements of the four domains in the N-terminal lobe were visualized by the effective rotation axes relative to the C-terminal lobe. The arrows indicate a left-handed rotation from the initial to the final structure. Domain 2 (green, 17 residues, the calcium 2 binding site) rotates by 65°. Domain 3 (orange, 11 residues, helix A) rotates by 56°. Domain 4 (blue, 26 residues, calcium 1 binding loop and helix B) rotates by 36°. Domain 5 (yellow, 11 residues, helix C) rotates by 63°. The movements are visualized by the molecular graphics program *VMD* (Humphrey *et al.*, 1996).

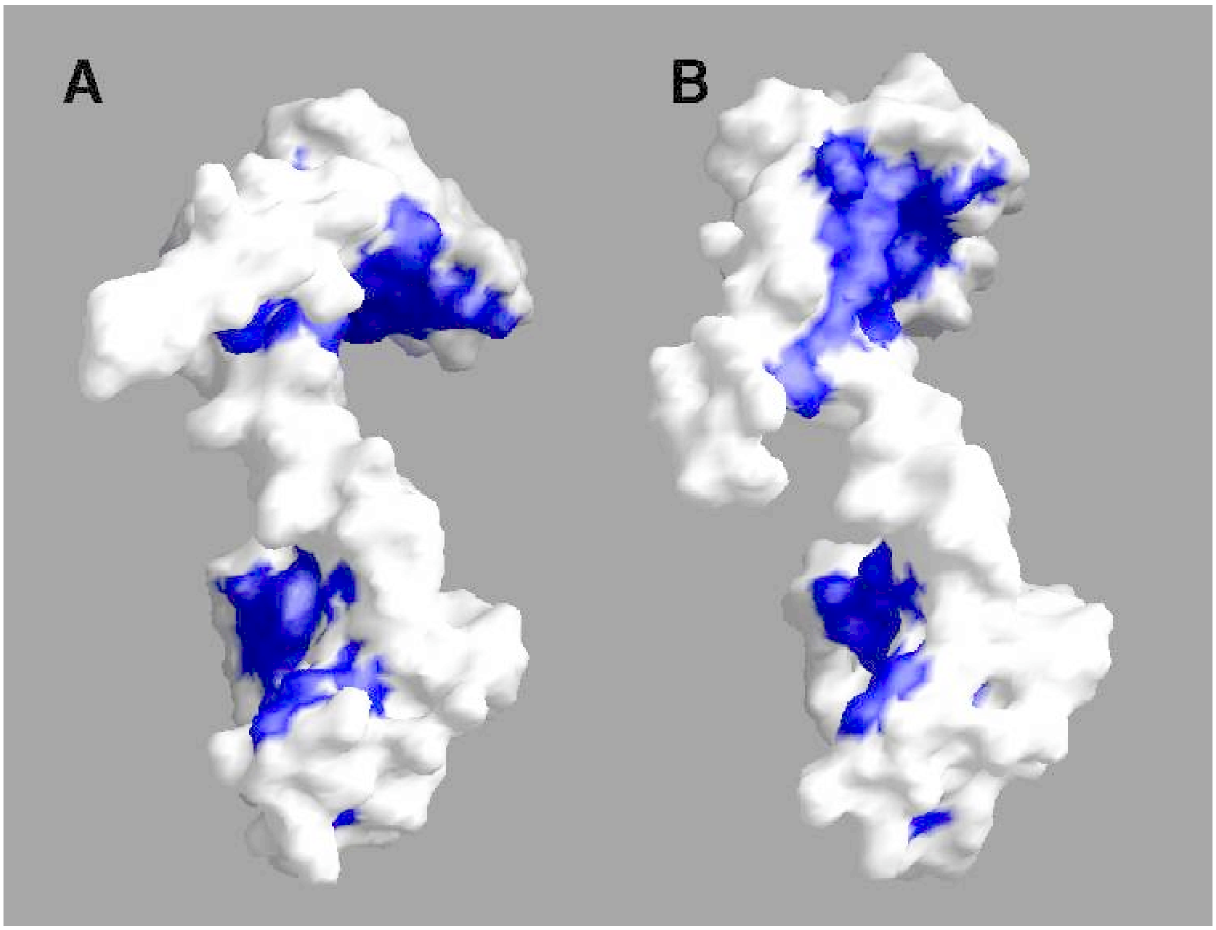

Figure 7: Target peptide binding sites in calmodulin visualized by the program GRASP (Nicholls *et al.*, 1991). A: Initial crystal structure (Babu *et al.*, 1988). B: The structure after 3 ns simulation (averaged over the last 20 ps). Hydrophobic residues for which there exists NMR evidence for binding to myosin light chain kinase (Ikura *et al.*, 1992) are shown in blue.

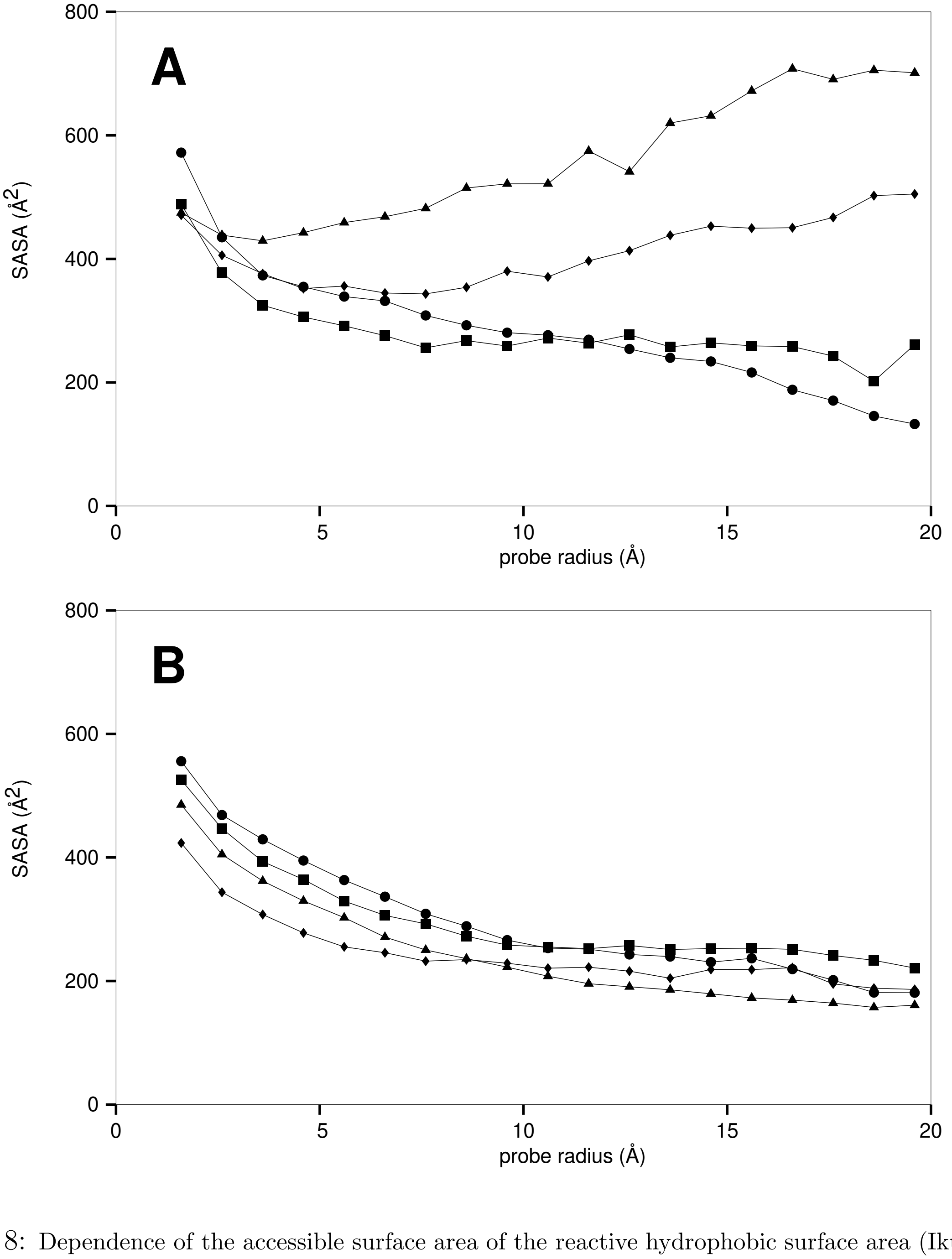

Figure 8: Dependence of the accessible surface area of the reactive hydrophobic surface area (Ikura *et al.*, 1992) on probe size and time. A: N-terminal domain; B: C-Terminal domain. Circles: crystal structure; squares: simulated structure at 1 ns; diamonds: simulated structure at 2 ns; triangles: simulated structure at 3 ns.

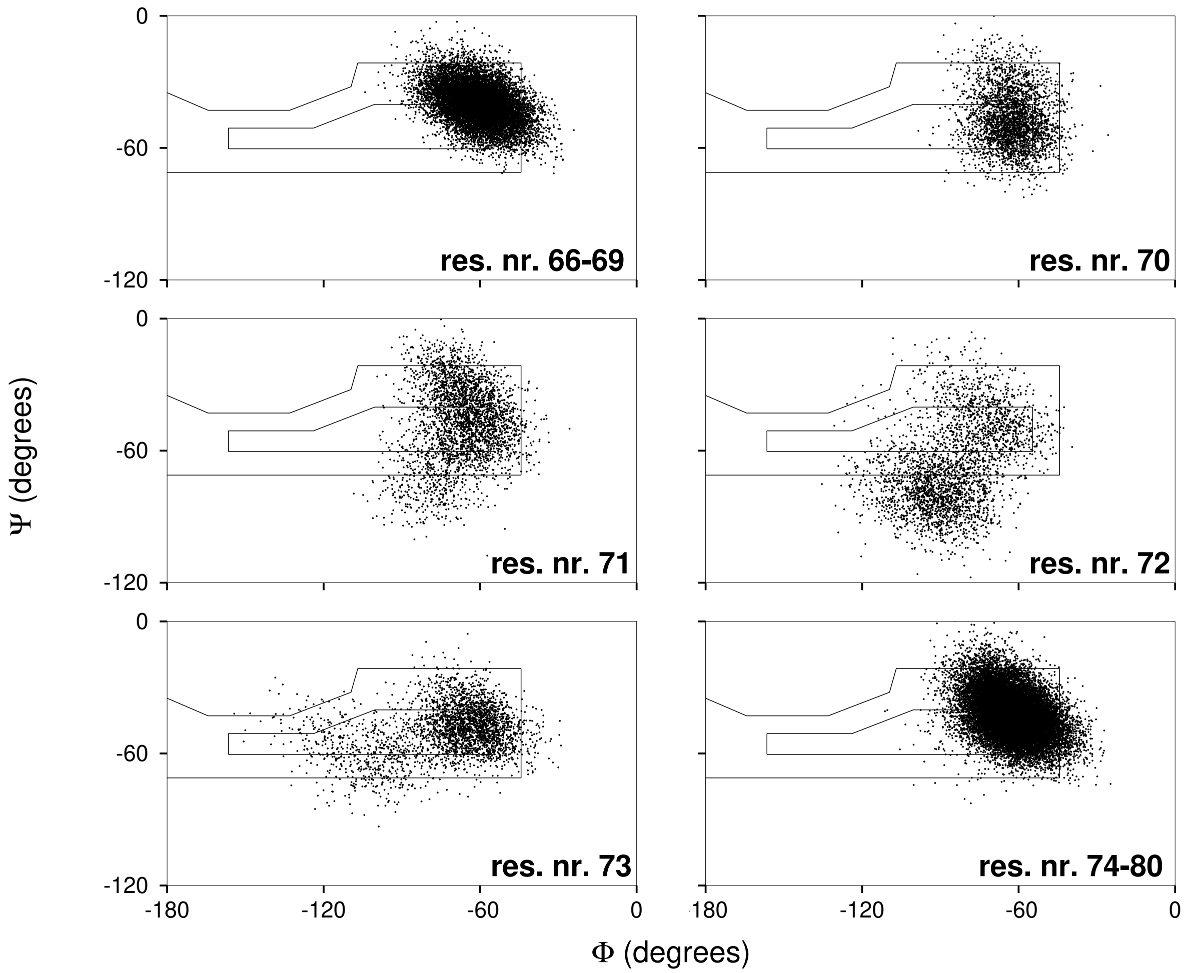

Figure 9: Distribution of torsional Φ and Ψ angles of the polypeptide chain computed from the 3 ns simulation with X-PLOR (Brünger, 1992). Torsional Φ and Ψ backbone angles are shown for selected residues of the central tether. Steric contours for "normal" and "outer limit" angles in right-handed α-helices of L-alanyl (Ramachandran *et al.*, 1963) are given for comparison.

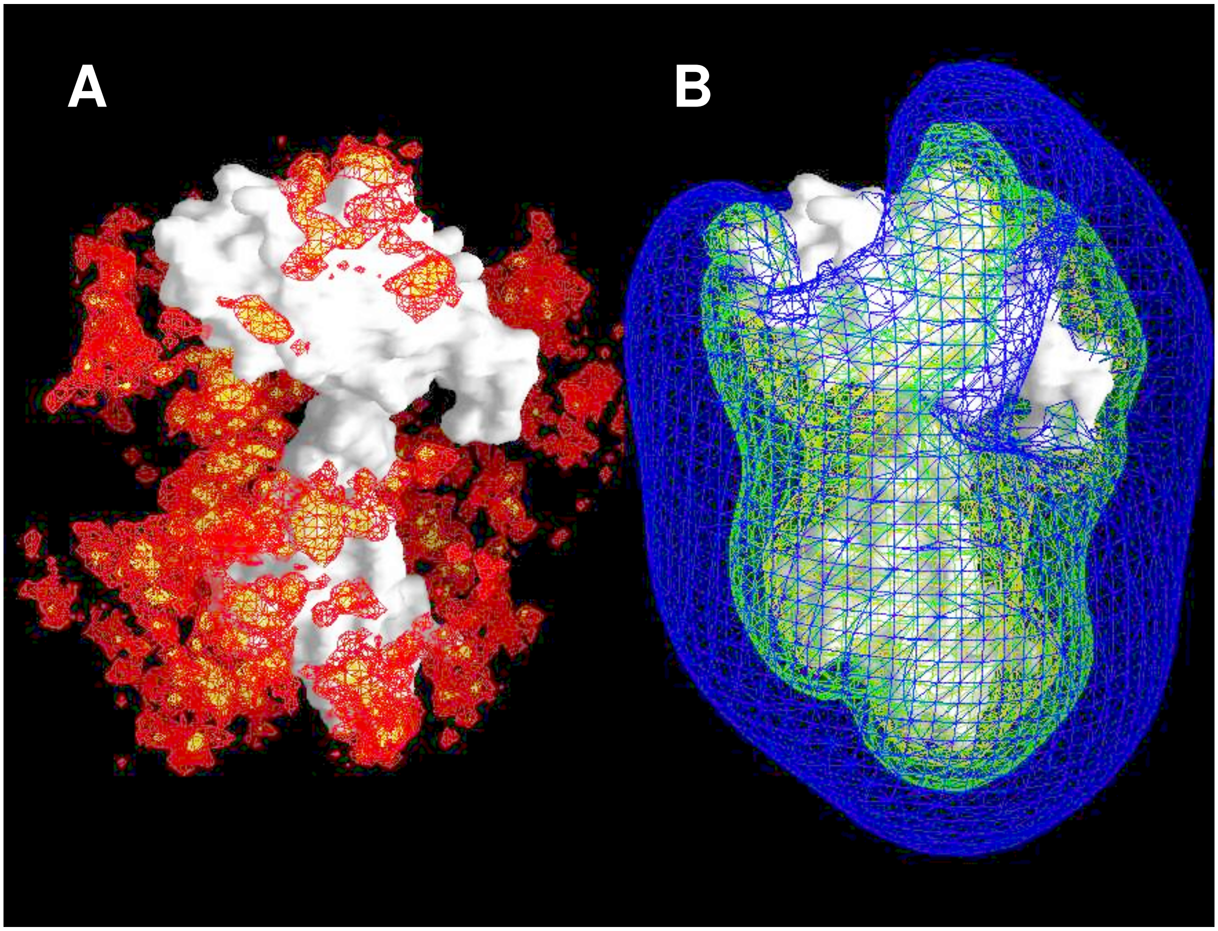

Figure 10: Counterion condensation observed in simulation and predicted by electrostatic theory. Calmodulin is represented by its solvent-accessible surface averaged from 2-3 ns simulation time. A: Three-dimensional histogram of sodium ion distribution in calmodulin's vicinity sampled from the 3 ns trajectory using a $(1.4\,\text{Å})^3$ bin size. The colors code for ion densities corresponding to a concentration $\rho_{\rm Na} = 1.0\,\rm M$ (red), and $\rho_{\rm Na} = 3.0\,\rm M$ (yellow). B: Contours of the sodium concentration $\rho_{\rm Na}(\vec{r})$ and the electrostatic potential $\Xi(\rho_\infty, \vec{r})$ [Eq. (8)] near the surface of the negatively charged (-15 e) protein. The contours were calculated with the program GRASP (Nicholls *et al.*, 1991) at ionic strength $\rho_\infty = 0.075\,M$. The following contours are shown: $\rho_{\rm Na} = 0.1\,\rm M, \Xi = -0.29\,k_BT$ (blue), $\rho_{\rm Na} = 0.2\,\rm M, \Xi = -0.98\,k_BT$ (green), and $\rho_{\rm Na} = 0.5\,\rm M, \Xi = -1.90\,k_BT$ (yellow).

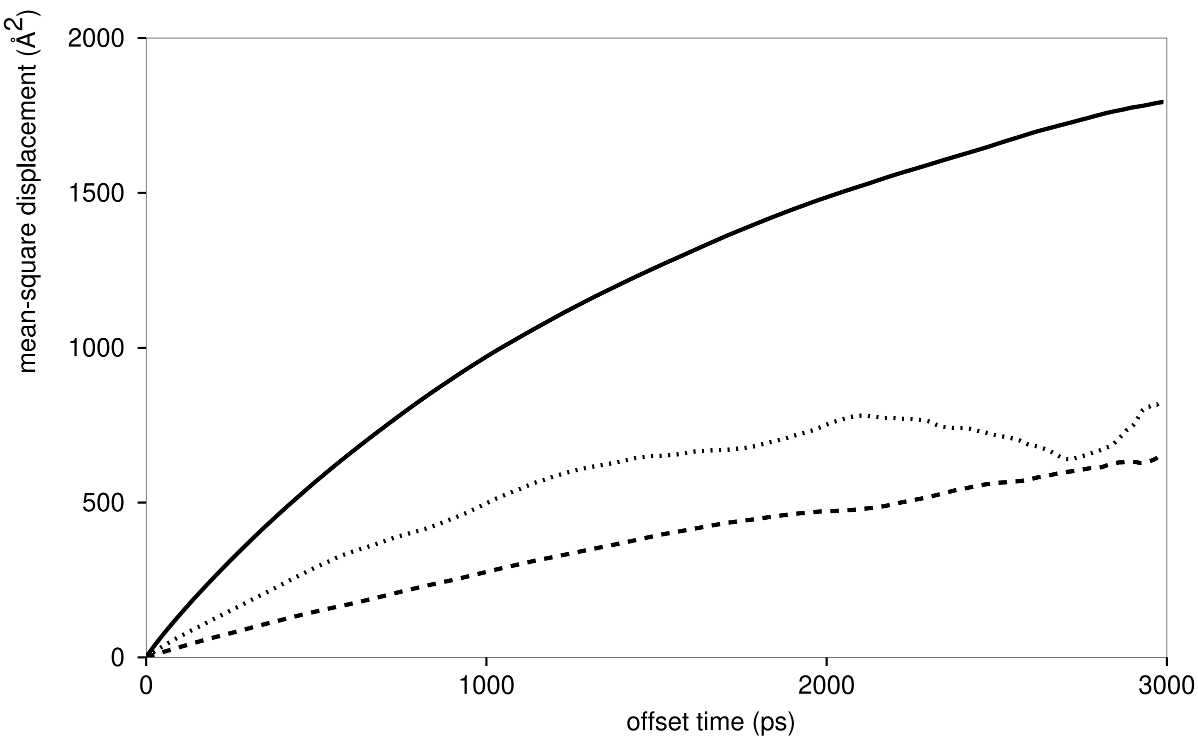

Figure 11: Mean square displacements, $\langle[\vec{r}(t+\tau) - \vec{r}(t)]^2\rangle$, of water molecules (solid line), sodium ions (dashed line) and chloride ions (dotted line) calculated from the 3 ns trajectory. The displacement functions allow one to compute corresponding diffusion coefficients [Eq. (3)].

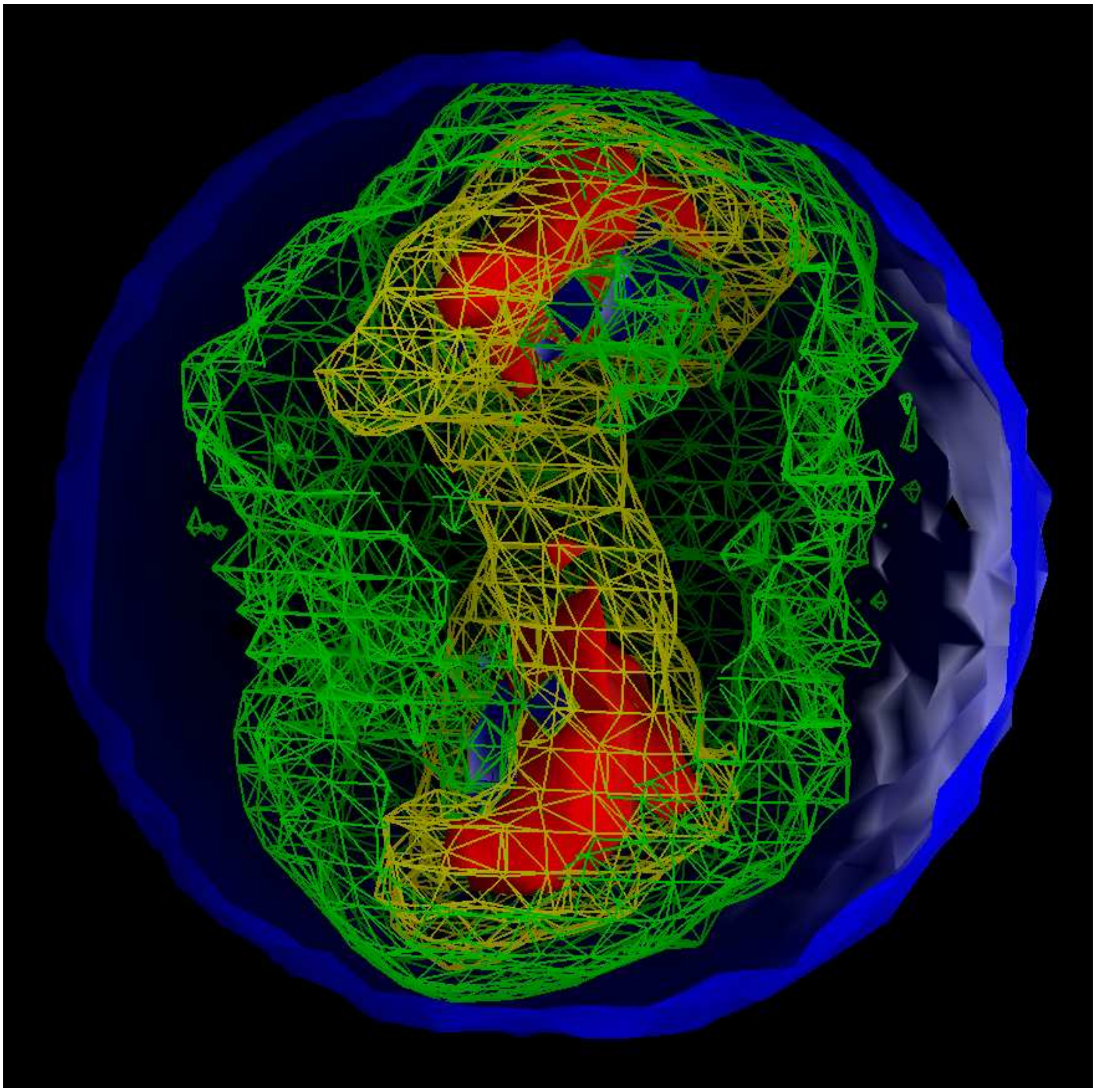

Figure 12: Short-time mobility in the solvent sphere surrounding calmodulin. The contour map (grid width 3.0 Å) was computed from the displacements after 1.0 ps as described in Computational Methods, and visualized by the program GRASP (Nicholls *et al.*, 1991). Contours of constant rms displacement (rmsdp) and corresponding effective diffusion coefficients [$D_{\text{eff}}^{(1)}$, Eq. (4)] are given. The colors code for rmsdp = 0.77 Å, $D_{\text{eff}}^{(1)} = 1.0 \cdot 10^{-9} m^2/s$ (red), rmsdp = 1.1 Å, $D_{\text{eff}}^{(1)} = 2.0 \cdot 10^{-9} m^2/s$ (yellow), rmsdp = 1.3 Å, $D_{\text{eff}}^{(1)} = 3.0 \cdot 10^{-9} m^2/s$ (green), and rmsdp = 1.5 Å, $D_{\text{eff}}^{(1)} = 4.0 \cdot 10^{-9} m^2/s$ (blue). The protein orientation is in accord with Fig. 7. The protein surface (not shown) corresponds approximately to the contour at $D_{\text{eff}}^{(1)} = 2.0 \cdot 10^{-9} m^2/s$.

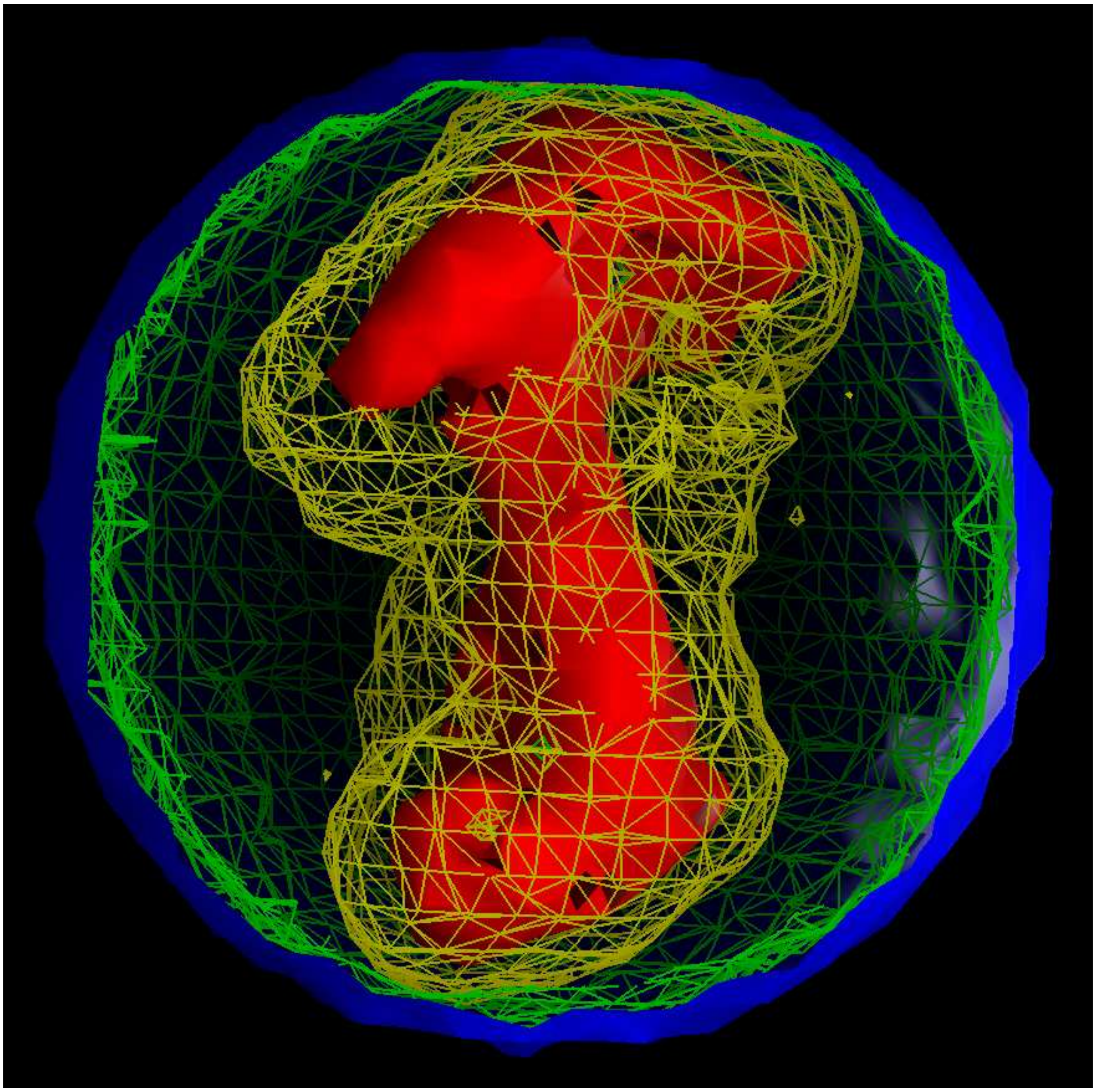

Figure 13: Distribution of the effective water diffusion coefficient $D^{(2)}_{\text{eff}}$ [Eq. (5)] in the solvent sphere surrounding calmodulin. The contour map (grid width 3.0 Å) was computed as described in Computational Methods, and visualized by the program GRASP (Nicholls *et al.*, 1991). Contours of constant coefficient $D^{(2)}_{\text{eff}}$ are given. The colors code for $D^{(2)}_{\text{eff}} = 1.0 \cdot 10^{-9} m^2/s$ (red), $D^{(2)}_{\text{eff}} = 2.0 \cdot 10^{-9} m^2/s$ (yellow), $D^{(2)}_{\text{eff}} = 3.0 \cdot 10^{-9} m^2/s$ (green), and $D^{(2)}_{\text{eff}} = 4.0 \cdot 10^{-9} m^2/s$ (blue). The protein orientation is in accord with Fig. 7. The protein surface (not shown) corresponds approximately to the contour at $D^{(2)}_{\text{eff}} = 1.0 \cdot 10^{-9} m^2/s$.