Eur Biophys J (2007) 36:225–237
DOI 10.1007/s00249-006-0123-1

ORIGINAL PAPER

Electrostatic control of the overall shape of calmodulin: numerical calculations

A. Isvoran · C. T. Craescu · E. Alexov

Received: 14 July 2006 / Revised: 20 November 2006 / Accepted: 28 November 2006 / Published online: 7 February 2007

Abstract The paper reports the results of numerical calculations of the pKa's of the ionizable groups and the electrostatic interactions between calmodulin lobes in three different states of calmodulin: calcium-free, peptide-free; calcium-loaded, peptide-free; and calcium-loaded, peptide-bound. NMR and X-ray studies revealed that in these states the overall structure of calmodulin adopts various conformations referred as: disordered semi-compact, extended and compact conformations, respectively. In addition, a new X-ray structure was recently reported (Structure, 2003, 11, 1303) showing that calcium-loaded, peptide-free calmodulin can also adopt a compact conformation in addition to the well known extended conformation. The calculated energy changes of calcium-loaded, peptide-free calmodulin along the pathway connecting these two conformations provide a possible explanation for this structural plasticity. The effect of pH and organic compounds in the solution phase on the preference of calmodulin to adopt compact or extended conformations may be thus rationalized. Analysis of the contribution of the ionization changes to the energy of association of calmodulin lobes suggested that the formation of the compact forms requires protonation of several acidic residues. However, two different protonation scenarios are revealed: a protonation due to internal lobe organization and thus independent of the lobes association, and a protonation induced by the lobes association resulting to a proton uptake. In addition, the role of the individual residues on the energy of association of calmodulin lobes is calculated in two compact conformations (peptide-free and peptide-bound) and is shown that a set of residues always plays a dominant role in inter-domain interactions.

A. Isvoran
Department of Chemistry, West University of Timisoara, Pestallozi 16, 300115 Timisoara, Romania

C. T. Craescu
INSERM/Institut Curie, Centre Universitaire Paris-Sud, Bâtiment 112, 91405 Orsay, France

E. Alexov (✉)
Computational Biophysics and Bioinformatics, Department of Physics, Clemson University, Clemson, SC 29634, USA
e-mail: ealexov@clemson.edu

Abbreviations

CaM	Calmodulin
CFPF	Calcium-free, peptide-free
CLPF	Calcium-loaded, peptide-free
CLPB	Calcium-loaded, peptide-bound

Introduction

Calmodulin $(CaM)^{++}$, an EF-hand calcium binding protein, functions as a calcium sensor and becomes activated upon Ca^{2+} binding (Ikura 1996). The overall structure of calmodulin is made up of two lobes (the N-terminal lobe CaM-N and the C-terminal lobe CaM-C) each comprising two EF-hand motifs (Strynadka and James 1989). The lobes are linked by an α-helix with significant mobility around residues 78–81, forming a flexible tether that allows lobes to adjust their relative position and orientation

(Persechini and Kretsinger 1988). During its function, calmodulin undergoes significant conformation changes concerning both the EF-hand motifs themselves and the mutual orientation and distance between the lobes. Three distinct conformations of the overall structure are known—(a) a compact conformation in which the residues make many contacts across the lobes interface (Meador et al. 1992); (b) quite disordered and semi-compact conformation of the average structure which reveals several transitory contacts across the lobes interface (Zhang et al. 1995); and (c) extended conformation in which lobes are separated from each other and have no mutual contacts (Chattopadhyaya et al. 1992). These structural changes are usually attributed to the functional states of calmodulin which are: (a) calcium-free, peptide-free (CFPF); (b) calcium-loaded, peptide-free (CLPF); and (c) calcium-loaded, peptide-bound (CLPB) calmodulin.

Calcium-free, peptide-free calmodulin is quite flexible as it can be seen from the existing NMR structures (Zhang et al. 1995). Some of the structures in the NMR ensemble show a semi-compact shape, while others may be classified as extended conformations. In all NMR structures the lobes adopt practically identical internal conformation, which indicates that variability of the lobes distance and orientation is not caused by the change of the internal energy of the lobes.

Upon calcium binding calmodulin EF-hand domains undergo a conformational change resulting in a reorganization of existing secondary structure elements. The two α-helices in each EF-hand motif move from an antiparallel position to an almost perpendicular position, while the short antiparallel β-sheet linking the two motifs remains roughly unchanged. Consequently, the domains adopt open conformation exposing to the solution a considerable hydrophobic surface area (Ikura 1996). Thus, the free energy change due to the ion binding is large enough to compensate the unfavorable contribution of the hydrophobic patch exposure. Then, the internal reorganization of the lobes induces change of their orientation and lobes move away from each other resulting to extended conformation of calmodulin (dumbbell shape) as suggested by X-ray experiment (Chattopadhyaya et al. 1992). In solution, however, both CFPF and CLPF calmodulin probably sample multitude of conformations and thus the transition from CFPF to CLPF calmodulin may not result to such dramatic conformational change. The picture is even more complicated, since recently CLPF calmodulin was found to adopt a compact conformation (Fallon and Quiocho 2003), although the experimental crystallization conditions were quite similar to the conditions of the X-ray experiment that revealed the extended conformation. The overall structure of the compact CLPF calmodulin is drastically different from those of the extended conformation. To the best of our knowledge no explanation of the existence of two forms of CLPF calmodulin is provided in the literature, and the energetic basis of this conformational diversity has not been analyzed.

The binding of the target peptide to calcium-loaded calmodulin affects mainly lobes distance and orientation, while the internal structure of the lobes is not changed significantly (Meador et al. 1992). Peptide binding causes calmodulin to adopt the compact conformation, as the two domains wrap around the target peptide, with each lobe interacting with a different half of the peptide sequence (Hoeflich and Ikura 2002).

The structures of compact and extended forms of calmodulin were experimentally determined by either X-ray or NMR techniques. The availability of these structures allows us to study the conformation and ionization changes associated with calmodulin function. Most of the existing studies were focused on the conformational changes of EF-hands motifs and the plasticity of the linker helix and to the least extend to the changes of the distance and orientation of the lobes. Specifically, what causes lobes separation was not addressed so far, although an attempt was made to explain the extended conformation as a result of repulsive coulombic interactions between negatively charged lobes (Uchikoga et al. 2005). However, the triggering event that turns calmodulin from compact to extended structure is the binding of four calcium ions. This dramatically reduces the net charge of the lobes making it close to zero and thus weakens the repulsive coulombic interactions between lobes. Thus, from electrostatic point of view, the binding of calcium ions should favor the compact structure of calmodulin. In this work we study the energetics of the conformational charges of the position and orientation of calmodulin lobes and what are the contributions of the individual residues to these changes. Specific attention is paid on the protonation states of the ionizable groups. In addition, we study what is the effect of the parameters of the environment to the preference of calmodulin for different conformations.

Methods

Structures used in the work

Four structures (Berman et al. 2000) were used in this work, representing the three functional states of calmodulin. These structures have identical amino acid

sequences and thus the changes of the lobes distance and orientation are not caused by sequence differences. The apo form, the CFPF calmodulin in semi-compact form is modeled on 1cfd coordinate file (Kuboniwa et al. 1995) that is the average NMR structure and a comparison to the NMR ensemble seen in 1dmo file (Zhang et al. 1995) reveals that the apo form adopts very different overall structures. Thus, we will keep in mind that 1cfd.pdb is simply one of the many overall conformations that CFPF calmodulin adopts in solution. The CLPF calmodulin was crystallized in two forms—extended conformation as shown in 1cll file (Chattopadhyaya et al. 1992) and compact conformation as shown in 1prw file (Fallon and Quiocho 2003). Both structures have four calcium ions. The CLPB calmodulin is studied using 1cdl structure (Meador et al. 1992), which also has four calcium ions. 1cdl structure was chosen to represent CLPB state, because all other available X-ray structures have missing residues in the middle of the sequence.

The X-ray structures of CLPF calmodulin were determined at low pH in the range of 5–6. The crystals were grown in mixture of water and organic compounds. The compact from of CLPF calmodulin was crystallized in 10% grycerol (Fallon and Quiocho 2003) while the extended form was grown in 35% MPD (2-methyl-2,4-pentanediol) and 15% ethanol (Chattopadhyaya et al. 1992). The peptide-bound form was obtained at pH = 4.6 (Meador et al. 1992). The NMR structures (Zhang et al. 1995) of the apo form were determined at neutral pH of about 7.

Calculations of the ionization states of titrable groups and the ionization free energy of folding

Multi-conformation continuum electrostatics (MCCE) (Alexov 2003; Alexov and Gunner 1997; Georgescu et al. 2002) method was used to calculate the charges of the ionizable groups. The default parameters were applied as internal dielectric constant of 4 and ionic strength of 0.15 M. Calcium ions were also included in the calculations and were fixed in their crystallographic positions. To be consistent with the experimental data, the occupancy at the calcium binding sites was kept 100%. X-ray/NMR structures were used in MCCE calculations without minimization.

The ionization free energy of folding was calculated using a protocol implemented in MCCE. The details are described elsewhere (Alexov 2004b). The protocol estimates the electrostatic component of the partition function of folded and unfolded states and then calculates the energy difference as a function of pH.

Lobes boundaries

The boundaries of the lobes were determined by visual inspection of the extended conformation of calmodulin, file 1cll. Thus, the first lobe was assumed to consist of residues 1–74, while the second lobe of residues 89–148. It should be mentioned that the domain boundaries are not crucial point for the aims of the present study and similar boundaries were obtained using a variety of automatic domain parsers (Alexandrov and Shindyalov 2003). In addition, residues 1–3 are disordered in 1cll and are not reported in the X-ray structure as well as residue 4 does not have side chain atoms. Thus, to reduce the complexity and the artifacts arising from modeling, the first four residues were deleted from each of the structures. The last two residues were deleted too, since they are not seen in 1cdl structure. Thus, the first lobe is considered to be made of residues 5–74 and the second of residues 89–146.

Generation of intermediate states

The transition from extended (1cll) to compact (1prw) structures of CLPF calmodulin was studied using the protocol developed by Gerstein and co-workers (Alexandrov et al. 2005; Krebs and Gerstein 2000). The structures of extended and compact CLPF calmodulin were inputted to the MORPH server at Yale University (http://www.molmovdb.mbb.yale.edu/molmovdb/) and ten intermediate states were generated (see Fig. 1 for a schematic presentation of the transformation). Thus, the relative coordinate runs from zero (corresponding to the structure of 1prw) and then goes through 10 intermediate states to end at position 11 corresponding to the extended structure of 1cll. It should be mentioned, however, that these intermediate

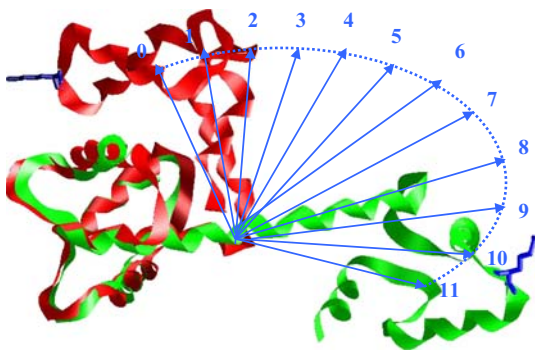

Fig. 1 Superimposition of the first lobes of 1prw (*red*) and 1cll (*green*). The *blue arrows* show schematically the translocation of the center of the mass of the second lobe from extended to compact structures and vice versa. Lys115, which is chemically modified in 1prw is shown in *sticks*

states do not necessarily represent ''real'' physical states and may not follow the most favorable trajectory from the compact to the extended conformation. The positions of calcium ions were assigned by superimposing the oxygen atoms of the coordinating residues taken from 1cll structure onto the same atoms of each MORPH generated structure.

Energy minimization

Each structure (entire structure including the helix and the peptide, if applicable) was minimized with Tinker package (Ponder 1999). The implicit solvent was modeled using the AGB model of Ron Levy (Felts et al. 2004) that was modified as described in Zhu et al. (2005). Charmm27 (Brooks et al. 1983) force field was used. The internal dielectric constant was set to 1 to be consistent with the force field parameters (MacKerell et al. 1998). All residues were kept in the ionized form, except His107 that was deprotonated. A charge of +2 was attributed to Ca^{2+} ions. ''minimize.x'' module of Tinker was used in the minimization. We attempted to keep the structures close to their original X-ray/NMR conformation and thus a very weak convergence criteria was applied, rms gradient per atom = 1.0. The backbone rmsd of the minimized structures are 0.81 Å for 1cfd, 0.53 Å for 1cll and 0.54 Å for 1cdl.

However, in case of 1prw, we noticed that the standard protocol led to significant deformation of the initial structure. Visual inspection of the minimized structure showed that many side chains at the interface of the lobes undergo complete reorientation as compared to the X-ray structure. The distortion of the X-ray structure made us to realize that the protonation states of some of the interfacial residues may not be standard. Thus, using the predictions made by MCCE, we turned off the net charge of Glu7, Glu11, Glu114, Glu120 and Glu127 by reducing the partial charge of the carboxyl oxygens. Their pKa's were calculated by MCCE to be higher than 6.0 and thus they were predicted to be protonated at the pH of the X-ray experiment. Under such conditions, the minimization protocol kept the structure very close to the initial conformation, resulting to backbone rmsd = 0.43 Å. Visual inspection revealed that the conformation of the side chains was also preserved as they were found in X-ray structure.

It should be mentioned that Glu67 and Glu82 were calculated to be protonated in 1cfd structure (Table 1). However, Glu67 is away from the interface and because of that was kept in the standard ionization state. Glu82 belongs to the linker connecting D and E helices, and since the paper focuses primarily of the lobes interactions, this residue was also kept ionized in the minimization protocol of 1cfd.

The MORPH structures were also minimized with charged states of Glu7, 11, 114, 120 and 127 properly assigned according to MCCE predictions. The same protocol as above was used.

Energy calculations

The energy of association of the two lobes is calculated as:

$$\Delta G = G(\text{two lobes}) - G(\text{lobe1}) - G(\text{lobe2}) \quad (1)$$

where G indicates the energy of the corresponding structure or structural domain. The central helix was deleted from the corresponding structures.

A rigid body approximation was adopted such that the structures of the lobes were assumed to be the same in the protein and in the isolated state. The advantage of adopting the rigid body approach is that the internal energy of the entire system is equal to the sum of the internal energies of the parts. Thus, the change of the internal energy is zero in Eq. (1) and only non-bonded energies, namely electrostatic, vdW, the energy associated with surface tension and the ionization energy were calculated. The extra proton of neutral Glu residues was added with MCCE (Forrest and Honig 2005).

Electrostatic energy

The electrostatic energy was calculated as described in Rocchia et al. (2001, 2002). Delphi was used to calculate the coulombic and reaction field energies. For sake of simplicity the salt concentration was set to zero. Scale of 2 grids/Å and filling of 80% was used. Convergence criterion was set to 0.0001. Water probe of radius 1.4 Å was used to build the molecular surface (Alexov 2003). The total energy of the system was calculated as a sum of the coulombic and reaction field energies. The energy contribution of a given residue to the stability was calculated by: (a) charging the residue, collecting the potential at all other residues and multiplying this potential by the partial charges of the corresponding atoms and (b) calculating the reaction field energy. The calculations were done on free lobes and on lobes at particular conformation of calmodulin, and then the results subtracted to obtain the contribution of a given residue to the stability.

Internal dielectric constant [eps(in)] was kept equal to 1 in order to match the conditions of minimization

Table 1 Calculated pKa's of selected acidic residues

Amino acid	1cfd	1cll	1prw	1cdl	1clb (calbindin)
Glu7	4.5	4.7	**10.5**	3.3	–
Glu11	4.6	5.1	**>14.0**	4.7	Glu5 = 3.6
Asp20	5.0	<0.0	1.5	<0.0	–
Glu31	6.5	5.0	<0.0	1.5	Glu27 = 6.5
Asp56	5.0	<0.0	0.7	4.2	–
Glu67	**7.9**	2.5	5.4	<0.0	Glu65 = 5.6
Asp80	3.9	4.7	4.8	2.4	–
Glu82	**8.3**	3.4	5.0	4.1	–
Glu84	5.0	5.2	4.3	4.5	–
Glu114	6.9	4.9	**10.4**	1.0	–
Glu120	5.0	4.7	**8.1**	1.7	–
Asp122	5.8	4.8	4.4	2.6	–
Glu127	4.2	5.1	**9.5**	<0.0	–

The last column provides the experimental pKa's of selected residues in apo calbindin. For the residue correspondence between calmodulin and calbindin see Fig. 2

protocol and to be consistent with MD force field (MacKerell et al. 1998). Keeping the parameters consistent is crucial for the accuracy of Molecular Mechanics Poisson–Boltzmann (MM-PBSA) method. However, to check the sensitivity of the results in respect to the choice of the value of eps(in), the energies were recalculated with eps(in) = 2. The external dielectric constant was 80 in most of cases, however, in several cases we used an external dielectric constant of 40 and 20 to check the effect of the parameters of the solution to the stability of the lobes position and orientation. An external dielectric constant lower than 80 mimics the presence of organic compounds in the solution. As it was reported above, the compact and extended CLPF calmodulins were crystallized in the presence of different organic compounds and this may result to different effective dielectric constants of the solution. The extended form can be considered as grown at conditions corresponding to lower effective dielectric constant of the solution, since the dielectric constant is 24.3, for ethanol and 42.5 for glycerol (see for example Seedher and Bhatia 2003).

vdW energy

vdW energy of interaction between lobes was calculated with ''analyze.x'' subroutine of Tinker package using the minimized structures. Charmm27 force filed was used. The vdW energy of interaction was calculated as a difference of the vdW energy of the MORPH generated structure of the lobes (without the helix) and the vdW energy of separated domains. As the distance between lobes increases the magnitude of the vdW energy of interactions decreases.

Surface tension energy

The accessible surface of each residue was calculated with the program ''surfv'' (Nicholls et al. 1991). The accessible surface area was calculated on the entire structure as well as on the isolated lobes, and then the results were subtracted. Thus, the residues that change their accessible surface area from isolated lobes to the entire molecule are at the interface of the lobes. The energy contribution to the association energy was calculated with a surface torsion coefficient of 0.005 kcal/Å^2 (Sitkoff et al. 1994).

Ionization energy

The ionization energy resulting from the protonation of the amino acids upon lobes association is calculated as $1.36(pH–pK_{standard})$ (Alexov 2004a), where $pK_{standard}$ are the standard pKa's of ionizable groups. Thus, a residue that is unprotonated in the free lobes but becomes protonated upon lobes association contributes unfavorably to the energy of the association. In our case, only Glu residues change their ionization states and the standard pKa of Glu residue is 4.4.

Results

Calculation of the protonation states of the ionizable groups

The results of MCCE calculations are summarized in Table 1. Reported are only the pKa of those groups that were calculated to be neutral in some of the conformational states. The most right column shows experimentally determined pKa's of several acidic groups in homologous protein calbindin (Kesvatera et al. 2001). These acids were selected using structural superimposition of the N-lobes of 1cbl and the peptide-free structures studied in this work (Fig. 2). To reduce the complexity of the graphical presentation, 1cll is deleted form the figure. The bottom of the figure shows the resulting sequence alignment. Five acidic groups superimpose well for all N-lobes—three of them are within Table 1 and other two are not shown. Except for the alignment of Glu27(1cbl) onto Glu31(1cfd,1prw,1cll), all other superimposed residues have completely different side chain orientation among the calbindin and calmodulins. Such structural differences should affect the pKa's of these groups. However, comparing our calculated pKa's to the experimental data reported in Kesvatera et al. (2001) we found that the results are quite similar. The

residues that are not shown in Table 1 are Glu48(1cbl) having pKa(exp) = 4.76 which superimposes on Glu47(1cfd,1prw) with calculated pKa(1cfd) = 3.65, pKa(1prw) = 3.0 and pKa(1cll) = 4.07. Residue Glu51 on calbindin superimposes onto Asp50 of calmodulins. The experimental pKa(exp) is 5.15 and the calculated pKa's are 4.56, 4.93 and 4.73 for 1cfd, 1prw and 1cll, respectively. Table 1 shows the experimental pKa of Glu7 in calbindin. This residue superimposes onto Glu11 of calmodulins. The calculated and experimental pKa's are within 1 pH unit, except for pKa of Glu11 in 1prw. The reason is the compactness of 1prw, which buries Glu11 between the lobes. The best superimposition was found in case of Glu27(1cbl) onto Glu31(1cfd, 1prw, 1cll). The calculated and experimental pKa's are in very good agreement except for Glu31 in 1prw, which is ligand to the calcium ion. The average value of the calculated pKa's for Glu67 over all calmodulins is very close to the experimental pKa of Glu65 in calbindin. Thus, despite of different side chain orientation, the experimental and calculated pKa's are in good agreement even taken from different proteins.

Except for the CLPB, all other compact conformations have protonation states that are different from protonation states of the extended form. All acids are predicted to be fully ionized in the extended CLPF (1cll) calmodulin since the X-ray experiment was done at pH about 6 while the highest calculated pKa is 5.2. Thus, the acidic residues within the extended CLPF calmodulin are not involved in strong unfavorable interactions and their pKa's are not perturbed relative to the standard values. However, the compact structures have several amino acids with elevated pKa's.

The simplest case is the semi-compact structure of CFPF calmodulin (1cfd). Only Glu67 and Glu82 were calculated to have pKa's larger than of the pH of the experiment (pH = 7.0). However, we carried MCCE calculations on the isolated lobes and the central helix and the results showed that the pKa's of Glu67 and Glu82 are the same as calculated with the entire protein. Thus, the abnormal pKa's are not caused by the inter-lobes unfavorable interactions but by desolvation and electrostatic interactions of these residues within the corresponding lobe/helix. Thus, a hypothetical conformational change that brings lobes from extended to semi-compact structure will not induce proton uptake/release.

The most interesting case of compact structure is 1prw. Five amino acids are calculated to be neutral at pH = 5.4, the pH of the X-ray experiment. In contrast to the previous case, these elevated pKa's are induced by inter-lobe interactions. The acids shown in bold in Table 1 form a complicated network across the interface of the lobes in which Glu7(lobe#1) interacts with Glu127(lobe#2), Glu11(lobe#1) interacts with Glu127 (lobe#2), Glu114(lobe#2) pairs with Glu14(lobe#1) and Glu120(lobe#2) interacts with Glu14(lobe#1). In many of these pairs, the inter-atomic distance between carbonyl oxygens of two acids is less than 3 Å and thus suggesting that they should be bridged by a hydrogen bond. This observation suggests the structural origin of the elevated pKa's. Thus, a hypothetical process that will bring the extended conformation (1cll) to the compact conformation (1prw) should induce significant proton uptake of five protons. The protonation of the above mentioned amino acids not only removes the unfavorable interactions across the lobes interface, but

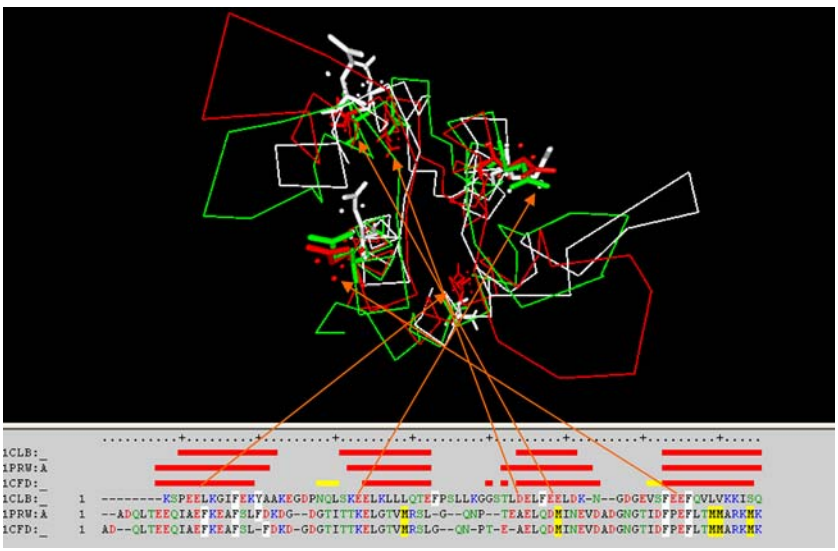

Fig. 2 Grasp 2 (Petrey and Honig 2003) structural alignment and the resulting sequence alignment of apo calbindin and 1cfd and 1prw structures. Acidic residues that are structurally aligned in all three structures are pointed with *orange arrows*. From left to right these residues are: (1) Glu5(1clb) = Glu11(1cfd); (2) Glu27(1clb) = Glu(31)(1cfd); (3) Glu48(1clb) = Glu47 (1cfd); (4) Glu51(1clb) = Asp50(1cfd) and (5) Glu65(1clb) = Glu67(1cfd)

also reduces the negative net charge of the lobes and favors the compact form formation.

The last case is the compact CLPB calmodulin (1cdl). All acidic residues are calculated to be ionized at the pH of the experiment. Majority of the amino acids have pKa values lower than those of the standard values. This is because of the favorable interactions with basic residues of the peptide that provide specific interactions with the negatively charged acids. It should be pointed out that the residues that were calculated to be protonated in CLPF compact form (1prw) are now predicted to have very low pKa's. Comparing 1prw and 1cdl structures, it can be seen that the binding of the target peptide increases the distance between lobes in 1cdl as compared to 1prw structures. In some cases, the peptide inserts a basic side chain between acidic chains that belongs to opposite lobes as in the case of Glu14(lobe#1) and Glu114(lobe#2) whose interactions are now mediated by Lys802 (peptide). In other cases, the peptide binding results in separation of the acidic pairs and each acidic side chain is bridged with a basic side chain of the peptide. Typical cases are Glu7 and Glu14 pairing with Arg798 (peptide), and Glu122 making a bridge with Lys799 (peptide).

Internal conformational changes of the lobes in different states

Available structural data for calmodulin suggests that there are four distinctive conformational states associated with distance and the orientation of the lobes: (a) CFPF adopting semi-compact structure but some of the structures in the NMR ensemble are quite extended too, CLPF compact, CLPF extended and CLPB compact forms. Only CFPF form is calcium-free while all other forms are calcium-loaded. Before analyzing the interactions between lobes, we should investigate how different are the internal conformation of the lobes in these states. To address this question we will take the conformation of the lobes in the extended form (file 1cll) as a reference. Comparison of the CFPF lobes (1cdf) to lobes in 1cll file results to backbone rmsd of 4.38 Å and of 4.20 Å for the N- and C-terminal lobes, respectively. This large difference was expected since the calcium binding induces significant internal rearrangement of EF-hand motifs (Ikura 1996). However, structural superimposition of CLPF lobes, files 1prw and 1cll, also shows noticeable differences. The backbone rmsd is 2.38 and 1.00 Å for the N- and C-terminal lobes, respectively, indicating that the internal organization of the domains is not identical in extended and in compact CLPF forms. This difference may arise either from different experimental conditions or from the interactions between lobes. The internal conformation of the lobes in CLPB form (1cdl) compared to the lobes in 1cll file results in rmsd of 0.62 and 0.72 Å for the N- and C-terminal lobes, respectively. This difference is even of smaller magnitude as for the compact CLPF form of calmodulin and indicates that the internal organization of the lobes is not only determined by the calcium binding. Comparison between CLPF compact (1prw) and CLPB compact (1cdl) conformations resulted to backbone rmsd of 2.25 and 1.00 Å for the N- and C-terminal lobes, respectively. In contrast, the backbone difference between CFPF (1cfd) and the compact CLPF (1prw) structures is 2.83 and 4.48 for N- and C-terminal lobes, respectively. At the end, the backbone rmsd difference between CFPL (1cfd) and CLPB (1cdl) resulted to 4.20 and 3.98 for N- and C-terminus lobes, respectively.

Thus, the internal structure of the calmodulin lobes is not identical among these states, but clearly the tertiary structures of the two domains fall roughly into two categories, corresponding to the calcium-loaded and calcium-free states. Since the internal organization of the lobes is quite similar in CLPF states, which were found to adopt either compact (1prw) or extended (1cll) conformations, we will first study the energetic of CLPF state.

The protonation states and energies of CLPF state: comparison of compact and extended structures

In order to estimate the role of the non-bonded energies to the extended and compact forms of CLPF calmodulin we used the intermediates generated by the MORPH server, as described in the Methods section (Fig. 1). The relative coordinate of this transformation begins with zero, corresponding to the X-ray structure of 1prw, runs over ten intermediates, and ends with 1cll structure.

Calculation of the protonation states

Each of the MORPH structures was subjected to MCCE calculations to reveal the protonation states. The results are shown in Fig. 3. It can be seen that all five residues mentioned above are protonated in the compact structure and remain protonated at the next transformation step. The first residue that becomes ionized is Glu7, which is fully ionized at the second transformation and remains so up to the end of the trajectory. Glu11 is also quickly ionized, since it breaks the unfavorable contacts with the amino acids in the lobe#2 as the transformation progresses from the

compact to the extended structure. Glu114 and Glu120 are the residues that remain protonated up to the sixth step. Thus, the protonation state of the acidic residues at the interface of the lobes is an important factor affecting the change of the association energy of the lobes along the transformation trajectory.

Energy calculations

All structures were minimized prior the energy calculations with appropriate ionization states of Glu7, 11, 114, 120 and 127. To reduce the bias toward particular crystal or MORPH generated structure, we adopted the methodology described in the Methods section, which assumes rigid body calculations. Since we are focusing on lobe interactions and the structure of the central helix changes dramatically as the compact form approaches the extended form, the helix was excluded from the calculations. Thus, in Eq. (1) the term named ''two lobes'' is simply the non-bonded energy of the lobes in a particular conformation without the central helix, while other energy terms were calculated on the isolated domains keeping their conformation the same as they have in the entire protein. The results are shown in Fig. 4. It should be noted, however, that each point of the graph was calculated relative to a different reference state, since the conformation of the lobes changes along the MORPH trajectory. Thus, the energy profiles shown in Fig. 4 should be considered only as an estimation of the ''real'' energy profile of a transition from extended to the compact CLPF calmodulin. In addition, the transformations generated by the MORPH server do not necessary reflect the actual conformational changes. Despite these approximations, Fig. 4 provides interesting insights into the possible role of different energy components on the energy of association of calmodulin lobes.

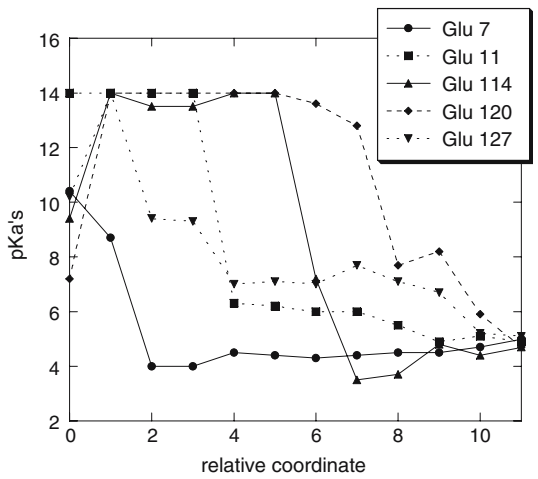

Fig. 3 pKa's of the acidic residues that were calculated to change significantly along the transformation trajectory of 1prw (→ 1cll

Two different scenarios were calculated, and labeled in Fig. 4a as *ionized* and *protonated*. In the first case all residues were kept in their default ionization states, e.g. all residues were ionized except for His107, which was kept neutral in both separated and assembled domains. Thus, the lobe association was considered not to induce ionization changes. In the second protocol the acidic residues that were predicted by MCCE to be protonated in the protein were kept neutral, while they were ionized in the isolated lobes. Thus, in the second case, the association induces proton uptake. As it can be seen, at large distance between lobes both protocols give similar results, with non-bonded energy slightly negative or close to zero. At medium distances, the association energy is positive and the general trend is to increase as the distance between lobes decreases (the bumpy energy landscape is a result of non physical nature of MORPH generated structures (1–10) while the structures at coordinates 0 and 11 are native structures). The major difference between both protocols is seen at very short distances between lobes where the *ionized* protocol calculates large unfavorable energy while the *protonated* protocol predicts favorable energy of association. The association energy of the compact form (coordinate 0) is even more negative than of the association energy of extended form (coordinate 11). The energies were recalculated with eps(in) = 2 (Fig. 4a) and similar behavior was obtained. The major difference in respect to the results with eps(in) = 1 is that the amplitude of the changes is reduced. What are the contributions of different energy terms to the total energy of association?

Figure 4b shows that the electrostatic and vdW energies dominate. The electrostatic energy calculated with the *ionized* protocol strongly opposes the formation of the compact structure and rapidly decreases as the lobes move away from each other adopting the extended conformation. In compact conformation (1prw), there are several acidic residues that form pairs as Glu7(lobe#1)–Glu127(lobe#2), Glu11(lobe#1)–Glu127(lobe#2), Glu14(lobe#1)–Glu114(lobe#2) and Glu14(lobe#1)–Glu120(lobe#2). Since they are at the interface, they are highly desolvated and interact unfavorably to each other. This gives rise to highly positive energy of association between lobes. The magnitude of the energy decreases as the relative coordinate increases and this results to a gradient (a repulsive force) that intends to separate lobes. Turning off the charges of these groups (excluding Glu14) as

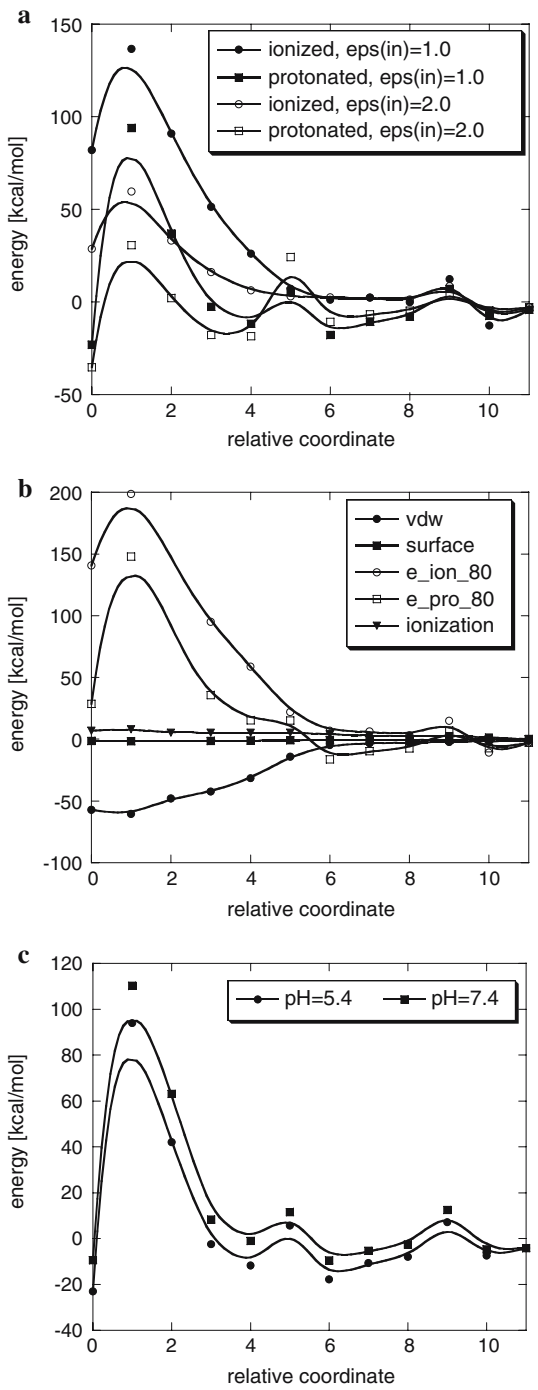

Fig. 4 **a** The total non-bonded energy of association of calmodulin lobes along the MORPH trajectory. **b** The energy components along the MORPH trajectory. ''e_ion_80'' is the electrostatic component of the association energy calculated with *ionized* protocol using dielectric constant of 80 for the solution, while ''e_pro_80'' is the association energy calculated with *protonated* protocol with the same dielectric external dielectric constant of 80. **c** The total association energy calculated at pH = 5.4 and 7.4

done in the *protonated* protocol makes the electrostatic component of the association energy less unfavorable even at close distances between lobes. The electrostatic energy difference between *ionized* and *protonated* protocols at coordinate 0 (1prw compact form) is more than 100 kcal/mol. The unfavorable association energy is reduced with the *protonated* protocol because the desolvation penalty of burial of a neutral group is much less than that of an ionized group. In addition, the *protonated* protocol removes the unfavorable pairwise interactions between closely situated negative charges. The second major energy component is vdW energy. Since the structures were minimized, this energy is favorable and slowly decreases along the trajectory. At coordinate 6 and further the vdW energy is almost zero because there are no contacts between lobes. The other two energy components, namely the energy associated with the surface tension and the ionization energy are relatively small, however, the ionization energy is pH dependent and its magnitude at pH = 5.4 at coordinate 0 is not negligible (+6.8 kcal/mol).

How the parameters of the environment affect this precise balance? To address this question we carried energy calculations assuming that the dielectric constant of the solution is smaller than of water. (In this case, only the *ionized* protocol was used since MCCE was calibrated for external dielectric of 80 and therefore the calculations of pKa's with external dielectric of 40 and 20 may not be accurate). This was done as an attempt to mimic the effect of organic solution on the energy of the association. A better approach would require explicit modeling of the binding of organic molecules onto calmodulin. However, the binding sites are unknown which led us to adopt the above simple approach and to mimic the effect of organic compounds by lowering the dielectric constant of the solution. The calculations indicate that lower dielectric constant of the solution makes the association energy less favorable (results now shown). The energy is extremely positive for compact structures and its magnitude decreases with the coordinate. Such energy profile would result to an energy gradient inducing a force that tries to separate the lobes. Thus, at low dielectric constant of the solution the lobes will prefer to be further away from each other and only the linker keeps them together.

Another important parameter of the environment is pH. Figure 4c shows the total non-bonded association energy calculated at pH = 5.4 and 7.4. The ionization states were kept the same in both calculations and they were identical to those reported above. Thus, the only energy component that is pH dependent under such an assumption is the ionization energy. As it can be seen, the increase of the pH to 7.4 makes the energy of the compact structure less favorable by +13.6 kcal/mol. This is a significant number that could turn the precise

balance of forces and clearly disfavors the compact form. Thus, with increase of the pH, the extended form is more energetically favorable.

pH-dependence of the stability of compact and extended CLPF calmodulin

Alternative way of addressing the effect of pH on the preferentiality of calmodulin to adopt compact or extended conformations is to calculate the ionization free energy of folding as a function of pH for both forms. The calculations were done allowing side chains and polar protons to sample different orientations as well as allowing ionization changes of titratable groups (see Alexov 2004b for details). The results are shown in Fig. 5. Both energy curves were normalized to zero at pH of maximal stability (called optimum pH). It can be seen that compact structure has optimum pH of 5.0 (solid vertical arrow), while extended has optimum pH of 8.0 (empty vertical arrow). Thus, compact structure is mostly stable at low pH and quite rapidly the folding energy increases with increase of pH. The ionization free energy of folding becomes less favorable by about 25 kcal/mol as pH changes from 5.0 to 8.0. The energy profile for extended conformation is flatter, which can be expected since the interactions between lobes are weaker as compared to the compact structure. The ionization free energy of folding becomes less favorable as pH changes from pH 8.0 to 5.0 for the extended conformation. Thus, these calculations confirm the findings of the previous section that pH is a strong regulator of the over shape of calmodulin. The compact form is most stable at low pHs while the extended at high pHs.

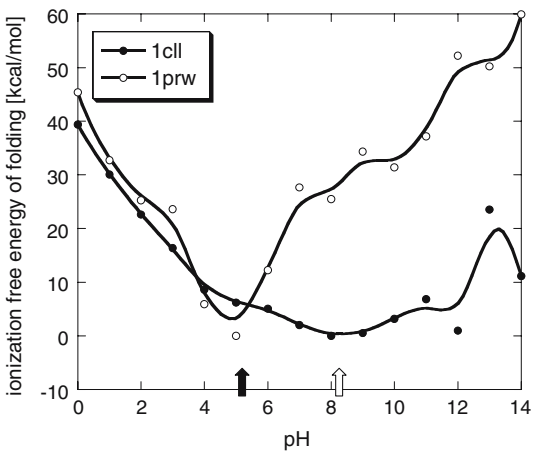

Fig. 5 The ionization free energy of folding calculated for compact and extended structures of CLPF calmodulin

Residues with strong contribution to the lobe association

The first question that we attempted to address is how similar are the interfacial residues in the two compact forms considered in the paper—CLPF compact form and CLPB compact form. The change of the accessible surface area for each residue was calculated for 1prw and 1cdl and the results are shown in Fig. 6a. The residues within the linker helix were deleted form the graph since the paper focuses on lobes interactions. The patterns of 1prw and 1cdl curves are very similar. The same residues that form the interface in the compact CLPF calmodulin are buried in the CLPB form too. In the CLPB form some of the changes in the surface accessibility are caused by the presence of the target peptide. This similarity indicates that interfacial residues of both compact forms are the same and suggests the formation of compact structures requires similar association modes.

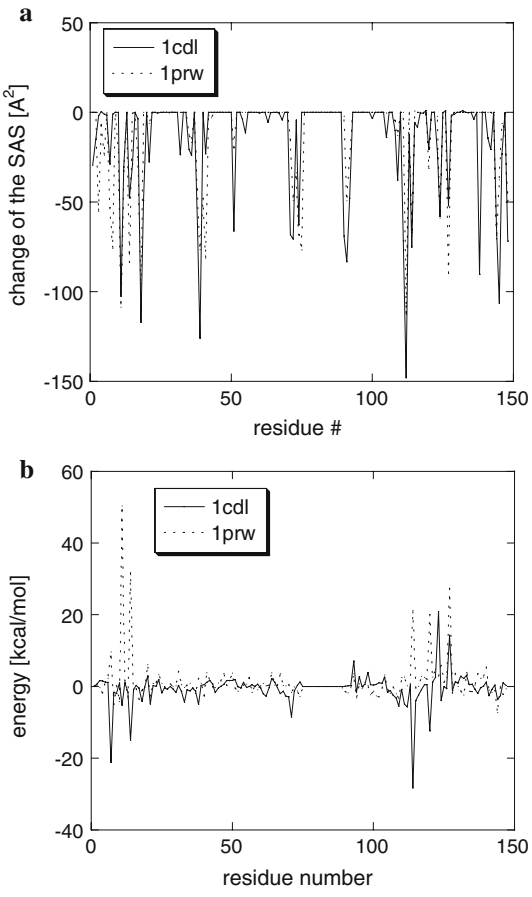

Fig. 6 **a** Change of the accessible surface area of residues caused by the association of the lobes. **b** Change of the electrostatic energy per residue upon lobes association

The energy contribution of the individual residues to the stability of the compact forms is shown in Fig. 6b. The residues within the linker helix are deleted from the figure. The figure shows that the residues having large energy contribution are preserved in 1prw and 1cdl structures. However, in CLPB form (1cdl) most of these acidic residues interact with the bases of the target peptide and thus have favorable energy. In contrast, the same residues in CLPF structure (1prw) are calculated to oppose the formation of the compact structure. However, despite of this fact, they are the residues providing the largest (favorable or unfavorable) energy to the stability.

Discussion

The analysis of the pKa of the ionizable groups in different conformational states of calmodulin revealed three major cases. In the simplest case of extended calmodulin, all amino acids are fully ionized and their pKa are close to the standard values. All amino acids are fully ionized in peptide-bound (1cdl) structure too, but some of them have pKa values much lower than the standard values, indicating that these residues strongly contribute to the peptide binding. Thus, acidic side chains with apparent pKa's smaller than the standard values represent residues that contribute to the stability of protein. The second case is manifested by the apo structure (1cfd). Several acidic residues are calculated to be neutral, but their charges are turned off by the internal structure of the lobes. Thus, these structures may have the tendency to form a compact conformation, because the negative net charge of the lobes is a priori reduced in separated domains. The third case is demonstrated in the compact structure of CLPF calmodulin (1prw). The formation of the compact form is calculated to require protonation of five amino acids and this should be observed experimentally as a large proton uptake.

Protonation of a residue at pH away from its pKa is an unfavorable process (see for example Alexov 2004a). However, the magnitude of this unfavorable energy decreases as the pH of the solution approaches the standard pKa of this type of residue. The compact CLPF calmodulin (1prw) was crystallized at pH = 5.4, which is only a 1 pH unit above the standard pKa of Glu residues. Thus, the magnitude of the unfavorable ionization energy is relatively small. However, at higher pH the magnitude of the unfavorable ionization energy will increase and the protein may not be able to pay the price. Thus, our prediction is that CLPF calmodulin in water should prefer mostly the compact structure at low pH, but an increase of pH should reverse this tendency and at higher pHs calmodulin should exists only in extended conformations. At pH values where the domains are neutral, the intermolecular interactions may be preponderant, resulting in low solubility and aggregation.

The theoretical prediction is supported by recent single-pair fluorescence resonance energy transfer experiments (Slaughter et al. 2005b) showing that the populations of compact and extended calmodulin structures are controlled by pH and salt concentration. It was shown (Slaughter et al. 2005b) that at low pH (pH = 5.0) calmodulin adopts predominantly compact conformations. The same effect was demonstrated experimentally by single-molecule tracking (Slaughter et al. 2005a) showing that at pH = 5 calmodulin adopts the compact conformation while at pH = 7.4 it samples both the extended and compact conformations.

The present pKa calculations support the notion that acidic residues play a dual role. In the peptide-free protein their ionization states determine the overall conformation of calmodulin. At higher pH, these amino acids are ionized and they repel each other resulting to large lobe separation and thus to an extended shape. As pH lowers, some of the amino acids become protonated and since the repulsion between the lobes is reduced, the calmodulin may have the preference to bind smaller targets and to form compact complexes. Thus, pH could serve as an important factor that modulates calmodulin affinity to target with different sizes. The binding of a basic peptide results in a favorable environment for acidic residues, which become fully charged even at low pH, thus providing specific interactions (Bhattacharya et al. 2004; Hultschiga et al. 2004) with the peptide residues.

The parameters of the environment play a significant role on the preference of calmodulin for compact or extended conformations. The salt concentration is an additional regulator that reduces the repulsive interactions between negatively charged lobes. As it was shown experimentally (Slaughter et al. 2005a, b), an increase of the salt concentration increases the probability of the compact form of calmodulin. Thus, both the pH and the salt concentration modulate the inter-lobe distance of calmodulin. In addition, the presence of alcohol in the water phase also affects the distribution of compact and extended conformations. As it was shown in our energy calculations, the lowering of the dielectric constant of the solution favors the extended form of calmodulin by increasing the strength of the repulsive interactions between negatively charged lobes. This finding is in accordance with experimental data (Slaughter et al. 2005b) that

show that an addition of 50% ethanol broadens the distribution of the conformational states of calmodulin as compared to the distribution in water. The increase of the repulsion forces between the lobes as the alcohol concentration increases could be the reason for the decreased affinity for the target (Ohashi et al. 2004).

In this study, the role of the central helix on the conformational states of calmodulin was ignored. This is an obvious simplification and many of the above discussed effects could be affected by the helix plasticity at different pHs, ionic strength and organic compounds. For example, a recent molecular dynamics study showed that the presence of trifluoroethanol makes the central linker more rigid (Broxk et al. 2004). Thus, a combination of lobe interactions and helix plasticity should be taken into account simultaneously in order to uncover the detailed energetic of the calmodulin conformational changes (Yang et al. 2004).

Our energy calculation (data not reported) showed that the long-range interactions between calmodulin lobes are not affected by internal lobe organization. This is in agreement with the recent X-ray structure (Grabarek 2005) that shows that in case of CLPF calmodulin a change of the internal conformation of the lobes do not change the dumbbell shape of extended calmodulin. The comparison of the Ca^{2+} loaded calmodulin structures showed that the variation of the internal lobe conformations is within 3 Å. Thus, the transition from extended CLPF calmodulin (1cll) to either the compact form (1prw) or to CLPB calmodulin (1cdl) could be approximated as a rigid body motion. Modeling of the electrostatic interactions along a putative trajectory that converts the extended to compact conformations of calmodulin and vice versa revealed that the electrostatic energy does not depend on the lobes orientation too (data not shown). Thus, neither the internal conformation nor the orientation of the lobes plays a role at long distances. However, at close distances specific electrostatic and vdW interactions determine the exact 3D structure of the compact forms of calmodulin. It appears that these specific interactions are preserved in the compact forms since the interface between lobes formed in the compact CLPF and CLPB forms are quite similar (see Fig. 6). The fact that the binding mode of peptide-free and peptide-bound calmodulin are similar indicates that either the binding mode is constrained by the central linker or that the lobes interfaces are complementary.

The results obtained in the present work demonstrate that the electrostatic energy calculations may provide a rationale insight into the structural plasticity of a two-domain protein, calmodulin. More than two-third of the eukaryotic proteins are predicted to have also a modular structure, with compact domains linked by flexible linker (Broxk et al. 2004). In most cases this intrinsic flexibility has consequences on the solubility, aggregation or crystallography, and the structure is only known for isolated domains. Analysis of the electrostatic properties of these domains and the simulation of inter-domain interactions by a rationale similar to that proposed in this paper would help in improving the physicochemical conditions of experimental studies.

References

Alexandrov N, Shindyalov I (2003) PDP: protein domain parser. Bioinformatics 19:429–430

Alexandrov V, Lehnert U, Echols N, Milburn D, Engelman D, Gerstein M (2005) Normal modes for predicting protein motions: a comprehensive database assessment and associated Web tool. Protein Sci 14:633–643

Alexov E (2003) The role of the protein side chain fluctuations on the strength of pair wise electrostatic interactions. Comparing experimental with computed pKa's. Proteins 50:94–103

Alexov E (2004a) Calculating proteon uptake/release and the binding free energy taking into account ionization and conformation changes induced by protein-inhibitor association. Application to plasmepsin, cathepsin D and endothiapepsin-pepstatin complexes. Proteins 56:572–584

Alexov E (2004b) Numerical calculations of the pH of maximal protein stability. The effect of the sequence composition and 3D structure. Eur J Biochem 271:173–185

Alexov E, Gunner M (1997) Incorporating protein conformation flexibility into the calculation of pH-dependent protein properties. Biophys J 74:2075–2093

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235–242

Bhattacharya S, Bunick C, Chazin W (2004) Target selectivity in EF-hand calcium binding proteins. Biochim Biophys Acta (BBA)—Mol Cell Res 1742:69–79

Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4:187–217

Broxk R, Scheek R, Weljie A, Vogel H (2004) Backbone dynamic properties of the central linker region of calcium-calmodulin in 35% trifluroethanol. J Struct Biol 146:272–280

Chattopadhyaya R, Meador WE, Means AR, Quiocho FA (1992) Calmodulin structure refined at 1.7 A resolution. J Mol Biol 228:1177–1192

Fallon J, Quiocho F (2003) A closed compact structure of native Ca^{2+}-Calmodulin. Structure 11:1303–1307

Felts AK, Harano Y, Gallicchio E, Levy RM (2004) Free energy surfaces of beta-hairpin and alpha-helical peptides generated by replica exchange molecular dynamics with the AGBNP implicit solvent model. Proteins 56:310–321

Forrest L, Honig B (2005) An assessment of the accuracy of methods for predicting hydrogen positions in protein structures. Proteins 61:296–309

Georgescu R, Alexov E, Gunner M (2002) Combining conformational flexibility and continuum electrostatics for calculating residue pKa's in proteins. Biophys J 83:1731–1748

Grabarek Z (2005) Structure of a trapped intermediate of calmodulin: calcium regulation of EF-hand proteins from a new perspective. J Mol Biol 346:1351–1366

Hoeflich K, Ikura M (2002) Calmodulinin action: diversity in target recognition and activation mechanisms. Cell 108:739–742

Hultschiga C, Hecht H, Frank R (2004) Systematic delineation of a calmodulin peptide interaction. J Mol Biol 343:559–568

Ikura M (1996) Calcium binding and conformational response in EF-hand proteins. TIBS 21:14–17

Kesvatera T, Jonsson B, Thulin E, Linse S (2001) Focusing of the electrostatic potential at EF-hands of calbindin D(9k): titration of acidic residues. Proteins 45:129–135

Krebs W, Gerstein M (2000) The morph server: a standartized system for analyzing and visualizing macromolecular motionsin a database framework. Nucleic Acids Res 28:1665–1675

Kuboniwa H, Tjandra N, Grzesiek S, Ren H, Klee CB, Bax A (1995) Solution structure of calcium-free calmodulin. Nat Struct Biol 2:768–779

MacKerell AD Jr, Bashford D, Bellot M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem 102:3586–3616

Meador WE, Means AR, Quiocho FA (1992) Target enzyme recognition by calmodulin: 2.4 A structure of a calmodulin–peptide complex. Science 257:1251–1254

Nicholls A, Sharp KA, Honig H (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 11:281–296

Ohashi I, Pohoreki R, Morita K, Stemmer P (2004) Alcohols increase calmodulin affinity for Ca^{2+} and decrease target affinity for calmodulin. Biochem Biophys Acta 1691:161–167

Persechini A, Kretsinger R (1988) The central helix of calmodulin functions as a flexible tether. J Biol Chem B 101:12175–12178

Petrey D, Honig B (2003) ''GRASP2: visualization, surface properties and electrostatic of macromolecular structures. Methods Enzymol 374:492–509

Ponder JW (1999) TINKER-software tools for molecular design. Washington University, St. Louis

Rocchia W, Alexov E, Honig B (2001) Extending the applicability of the nonlinear Poisson–Boltzmann equation: multiple dielectric constants and multivalent ions. J Phys Chem 105:6507–6514

Rocchia W, Sridharan S, Nicholls A, Alexov E, Chiabrera A, Honig B (2002) Rapid grid-based construction of the molecular surface and the use of induced surface charges to calculate reaction field energies: applications to the molecular systems and geometrical objects. J Comput Chem 23:128–137

Seedher N, Bhatia S (2003) Solubility enhancement of Cox-2 inhibitors using various solvent systems. AAPS PharmSciTech 4:1–9

Sitkoff D, Sharp KA, Honig B (1994) Accurate calculation of hydration free energies using macroscopic solvent models. J Phys Chem 98:1978–1988

Slaughter B, Bieber-Urbauer R, Johnson C (2005a) Single-molecule tracking of sub-millisecond domain motion in calmodulin. J Phys Chem B 109(26):12658–12662

Slaughter B, Unruh J, Allen M, Bieber R, Johnson C (2005b) Conformational substates of calmodulin revealed by single-pair fluorescence resonance energy transfer: influence of solution conditions and oxidative modification. Biochemistry 44:3694–3707

Strynadka N, James M (1989) Crystal structures of the helix-loop-helix calcium-binding proteins. Annu Rev Biochem 58:951–998

Uchikoga N, Takahashi SY, Ke R, Sonoyama M, Mitaku S (2005) Electric charge balance mechanism of extended soluble proteins. Protein Sci 14:74–80

Yang C, Jas GS, Kuczera K (2004) Structure, dynamics and interaction with kinase targets: computer simulations of calmodulin. Biochim Biophys Acta 1697:289–300

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

Zhu J, Alexov E, Honig B (2005) Comparative study of generalized born models: born radii and peptide folding. J Phys Chem B 109:3008–3022