Identification of the critical structural determinants of the EF-hand domain arrangements in calcium binding proteins

ARTICLE *in* BIOCHIMICA ET BIOPHYSICA ACTA · JANUARY 2012

Impact Factor: 4.66 · DOI: 10.1016/j.bbapap.2012.01.005 · Source: PubMed

5 AUTHORS, INCLUDING:

Wenchang Zhou
National Heart, Lung, and Blood Institute
8 PUBLICATIONS **49** CITATIONS
SEE PROFILE

Zhiguang Jia
Kansas State University
8 PUBLICATIONS **48** CITATIONS
SEE PROFILE

Qun Wei
Beijing Normal University
106 PUBLICATIONS **1,050** CITATIONS
SEE PROFILE

Biochimica et Biophysica Acta 1824 (2012) 608–619

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Identification of the critical structural determinants of the EF-hand domain arrangements in calcium binding proteins

Ye-dan Feng [1], Jing Li [1], Wen-chang Zhou, Zhi-guang Jia, Qun Wei *

Department of Biochemistry and Molecular Biology, Beijing Normal University, Beijing Key Laboratory, Beijing 100875, PR China

ARTICLE INFO

Article history:
Received 29 July 2011
Received in revised form 2 December 2011
Accepted 4 January 2012
Available online 21 January 2012

Keywords:
EF-hand calcium binding protein
Calmodulin
Calcineurin B
Structural determinant
Molecular dynamics

ABSTRACT

EF-hand calcium binding proteins (CaBPs) share strong sequence homology, but exhibit great diversity in structure and function. Thus although calmodulin (CaM) and calcineurin B (CNB) both consist of four EF hands, their domain arrangements are quite distinct. CaM and the CaM-like proteins are characterized by an extended architecture, whereas CNB and the CNB-like proteins have a more compact form. In this study, we performed structural alignments and molecular dynamics (MD) simulations on 3 CaM-like proteins and 6 CNB-like proteins, and quantified their distinct structural and dynamical features in an effort to establish how their sequences specify their structures and dynamics. Alignments of the EF2–EF3 region of these proteins revealed that several residues (not restricted to the linker between the EF2 and EF3 motifs) differed between the two groups of proteins. A customized inverse folding approach followed by structural assessments and MD simulations established the critical role of these residues in determining the structure of the proteins. Identification of the critical determinants of the two different EF-hand domain arrangements and the distinct dynamical features relevant to their respective functions provides insight into the relationships between sequence, structure, dynamics and function among these EF-hand CaBPs.

1. Introduction

Calcium (Ca^{2+}) signals and calcium binding proteins (CaBPs) of the EF-hand superfamily are involved in the regulation of almost every process in life from birth to death [1,2]. The EF hand is the elementary structural motif in almost all the proteins in this superfamily. Although it is generally assumed that sequence similarity implies shared structure, function, and evolution, the CaBPs are highly homologous and yet are very diverse in their structures and in the targets with which they interact [3]. The CaBPs, such as calmodulin (CaM), calcineurin B (CNB), sarcoplasmic calcium binding protein (SCP) and S100B, which have significant sequence similarity, have different arrangements of their N- and C-terminal domains. Understanding this exception to the sequence-implies-structure paradigm, and addressing the underlying basis of structural diversity, could provide deeper insight into the relationships between sequence, structure, dynamics and function in these CaBPs.

CaM, the best studied and prototypical example of an EF-hand calcium sensor protein, is involved in the regulation of many important Ca^{2+}-dependent signaling pathways by moving between various cellular compartments and interacting with various targets [4]. CaM contains four EF hands: the first two combine to form a globular N-terminal domain, which is separated by a flexible linker from the homologous C-terminal domain containing the other two EF hands. Although CaM varies in the way it interacts with its targets, which is matched by the extensive and diverse list of targets that it regulates [5,6], the most classic conformation of CaM is the dumb-bell shape which exists when CaM is in a Ca^{2+}-saturated state and is not bound to any target [7–10]. A similar extended architecture can also be found in some other EF-hand proteins, such as troponin C (TnC) [11–13] and human calmodulin-like protein (hCLP) [14]. TnC, together with troponin I (TnI) and troponin T (TnT), make up the troponin (Tn) complex, which acts as a trigger in skeletal and cardiac muscle by switching on contraction at a critical Ca^{2+} concentration [15], while hCLP is found in mammary epithelial cells [16], and functions as a myosin-10 binding partner [17].

CNB has a more compact four EF-hand structure: the two globular domains align on the same side and are linked by a U-shaped linker [18,19]. CNB is the regulatory subunit of calcineurin, a Ca^{2+}/CaM-dependent serine/threonine protein phosphatase that plays a critical role in many cellular processes [20]. After the structure of CNB was established, three-dimensional structures of similar architecture were

Abbreviations: Ca^{2+}, calcium; CaBPs, calcium binding proteins; CaM, calmodulin; CNB, calcineurin B; SCP, sarcoplasmic calcium binding protein; TnC, troponin C; TnI, troponin I; TnT, troponin T; Tn, troponin; hCLP, human calmodulin-like protein; CIB1, calcium and integrin binding protein 1; CHPs, calcineurin B homologous proteins; Frq1, yeast frequenin; NCS, neuronal calcium sensor; GCAPs, guanylyl cyclase activating proteins; KChIPs, Kv channel interacting proteins; RECO, recoverin; CNA, calcineurin A; NHE1, Na^+/H^+ exchanger 1; Pik1, phosphatidylinositol 4 kinase; PDB, protein data bank; MD, molecular dynamics; RMSD, root mean square deviation; PME, particle mesh Ewald; ED, essential dynamics; PCA, principal component analysis

* Corresponding author. Tel./fax: +86 10 5880 7365.
E-mail address: weiq@bnu.edu.cn (Q. Wei).

[1] These authors contributed equally to this work.

1570-9639/$ – see front matter.
doi:10.1016/j.bbapap.2012.01.005

found in CIB1 [21,22], CHP1 [23], CHP2 [24], Frq1 (yeast frequenin) [25], NCS1 (neuronal calcium sensor 1, human frequenin) [26], neurocalcin-δ [27], recoverin [28], GCAP1 [29], GCAP2 [30], GCAP3 [31], KChIP1 [32] and KChIP4 [33]. CIB1 (calcium and integrin binding protein 1) was initially identified as a specific binding partner of platelet integrin αIIbβ3; it is also expressed in many other cell types and interacts with a number of other target proteins [34]. CHPs (calcineurin B homologous proteins) are involved in membrane trafficking; they bind multiple effectors and potentially modulate their functions [23], in particular inhibiting NHE1 (Na^+/H^+ exchanger 1) [35]. Also, frequenin, neurocalcin-δ, recoverin, GCAPs (guanylyl cyclase activating proteins) and KChIPs (Kv channel interacting proteins) constitute the 5 subgroups (A to E, respectively) of the NCS protein family; they regulate many cellular events in neurons and retinal photoreceptors, and have specific targets that do not overlap with those of CaM [36,37].

The EF-hand proteins mentioned above can be divided into two groups according to the arrangements of their N- and C-terminal domains: the CaM-like proteins and CNB-like proteins. It is interesting that some members of the opposing groups are more homologous at the sequence level than some proteins in the same group. A typical example is CNB, whose full primary sequence is more homologous to CaM (35% identity) than to recoverin (29% identity). Such proteins provide an opportunity to gain a further understanding of the relationships between sequence, structure, dynamics, and function in these EF-hand proteins. Does any factor other than strong sequence homology have a significant impact on protein folding and the arrangements of the EF-hand domains?

Pawlowski et al. have done comparative modeling of three EF-hand proteins with different arrangements of their N- and C-terminal domains, i.e., CaM, recoverin and SCP; they suggested that correlated mutations on the surface of two units, and the presence of additional fragments, are two factors that might stabilize the desired topologies [38]. Because of the limited numbers of three-dimensional structures available at the time, they could not compare the structures of a large number of different proteins to identify the determinants of the divergence in the EF-hand domain arrangements. Mouawad et al. analyzed the sequences of 17 CaBPs of known structure, and developed a novel procedure to predict the compactness of a CaBP of unknown structure. They suggested that the compactness of a CaBP is determined by the hydrophilicity of a few residues (average of 5) that link the two domains [39].

Here, we report the results of a study of the structural basis of the two EF-hand domain arrangements represented, respectively, by CaM and CNB, and the sequence–structure–dynamics–functional relationships among the CaBPs with four EF hands. We chose nine representative EF-hand proteins from the various subgroups (Table 1), namely 3 CaM-like proteins (CaM, TnC and hCLP) and 6 CNB-like proteins (CNB, recoverin, CIB1, NCS1, GCAP3 and KChIP1), all of whose structures were available in their Ca^{2+}-saturated states. We chose proteins whose functions had been well studied as this would be beneficial for later examination of the relationships between dynamics and function. Using structural alignments and molecular dynamics (MD) simulations on the 9 proteins, we quantified the respective features on structures and dynamics of the two groups of proteins. To address how sequence specifies structure and dynamics, we aligned partial (EF1, EF2–EF3, and EF4) and complete sequences of the 9 proteins, and tentatively identified the critical structural determinants in the EF2–EF3 region. Then we used a customized inverse folding approach by comparative modeling of CaM and CNB, to confirm the importance of these structural determinants. Finally, we surveyed the functions of the two groups of CaBPs as previously reported, especially their manner of interacting with target proteins. The functions of the proteins could be accurately inferred from the identified determinants of the structural and dynamical features of each of the domain arrangements. Our findings provide deeper insight into the sequence–structure–dynamics–functional relationships among the CaBPs.

2. Materials and methods

2.1. Preparation of structures, and structural alignments

Basic information on the 9 chosen proteins is given in Table 1. Their three-dimensional structures obtained from the PDB (protein data bank) [40] rendered with PyMOL [41] are shown in Fig. S1.

For the structural alignments and subsequent MD simulations, residues 133–137 of NCS1, 137–142 of CIB1, and 1–4 and 148 of CaM, which are all not visible in the respective crystal structures, were added with the Insight II Biopolymer Module (Accelrys, Inc., San Diego, CA). In the NMR structure 1JSA (recoverin), the first model in the PDB file was used.

The structural alignment matrix was generated using the Insight II Homology Module with the multiple sequence alignment method. The RMSDs (root mean square deviations) for the backbone between each pair of the 9 proteins were calculated.

2.2. MD simulations

MD simulations were performed using GROMACS [42,43] with the ROMOS96 force field [44]. Each protein was embedded in a box containing an SPC water model [45], with at least 12 Å between the protein and the edge of the box. To neutralize the system and reach physiological ionic strength of 100 mM, appropriate amounts of sodium and chloride ions were added at random positions. After minimization and equilibration for 100 ps under position restraints for the backbone of the protein, each MD production run was performed for 20 ns. In the MD simulations, the LINCS algorithm [46] was used to constrain the lengths of all bonds. The time step for simulations was 2 fs. The simulations were run under NPT conditions, using Berendsen's coupling algorithm to keep temperature and pressure constant ($P = 1$ bar; $T_p = 0.5$ ps; $T = 300$ K; $T_t = 0.1$ ps) [47]. For Van der Waals forces and short-range electrostatic interactions we used a cutoff of 12 Å. Long-range electrostatic forces were obtained by

Table 1
Basic information on the 9 proteins.

Protein	Source	PDB ID	UniProt ID	Length [a]	Ca^{2+}-binding	Reference
CaM	Human	1CLL	P62158	148/149	4	Chattopadhyaya et al. (Ref. [8])
TnC	Rabbit	1TN4	P02586	157/160	4	Houdusse et al. (Ref. [13])
hCLP	Human	1GGZ	P27482	144/149	4	Han et al. (Ref. [14])
CNB	Human	1TCO	P63098	169/170	4	Griffith et al. (Ref. [18])
RECO [b]	Bovine	1JSA	P21457	188/202	2	Ames et al. (Ref. [28])
CIB1	Human	1XO5	Q99828	180/191	2	Gentry et al. (Ref. [22])
NCS1	Human	1G8I	P62166	187/190	3	Bourne et al. (Ref. [26])
GCAP3	Human	2GGZ	O95843	165/209	3	Stephen et al. (Ref. [31])
KChIP1	Human	1S1E	Q9NZI2	181/227	2	Scannevin et al. (Ref. [32])

[a] The first number in each case is the number of residues in the three-dimensional structure in the PDB entry; the second number is the number of residues in the primary amino acid sequence in the UniProt entry.
[b] RECO: recoverin.

the particle mesh Ewald (PME) method [48]. Coordinates were saved every 1 ps.

The RMSDs for the backbone of the whole protein and individual domains were calculated with the g_rms subroutine in GROMACS, in which the first frame of each trajectory is used as reference. The distance between the N- and C-terminal domains was calculated by g_dist in GROMACS, which calculates the distance between the centers of mass of the N- and C-terminal domains as a function of time [49]. The definitions of the N- and C-terminal domains of the 9 proteins used for RMSD and distance analysis are listed in Table 2. In our study, the definitions of the two globular domains refer to the regions of the EF2 and EF3 motifs provided by the UniProt database [50], e.g., the N-terminal domain extends from the first residue of the protein to the end of the EF2 motif, while the C-terminal domain extends from the beginning of the EF3 motif to the last residue of the protein. Thus the few residues corresponding to the linker region are averagely assigned to the two globular domains.

Essential dynamics (ED), also referred to as principal component analysis (PCA) of the MD trajectories [51], was performed essentially as described by Yang et al. [52] and Hu et al. [53]. It was accomplished by g_covar and g_anaeig in GROMACS. The former calculates and diagonalizes a mass-weighted covariance matrix to find the eigenvectors, and the latter generates the projections of trajectories onto the eigenvectors [49].

2.3. Sequence analysis

Sequences of the 9 proteins were obtained from the UniProt entries listed in Table 1. The definitions of each EF motif and the EF2–EF3 region of the 9 proteins are listed in Table 2.

Sequence alignments and phylogenetic tree analysis were performed with ClustalW [54,55]. The protein weight matrix was set to the BLOSUM series; other alignment parameters were set to default.

Structure-based sequence alignments were performed as well with EXPRESSO (3DCoffee) [56] web server to verify whether our pure sequence alignment is accurate or not. The default procedure was used, in which only the sequences were uploaded, and the related structures were automatically incorporated by the server to guide the alignments.

Alignment diagrams were drawn by BioEdit [57], and phylogenetic trees were drawn by Njplot [58].

2.4. Comparative modeling

The use of comparative modeling here is in fact an extension of the inverse folding approach, in which actual models are built and a combination of threading and comparative modeling techniques is applied, as described by Pawlowski et al. [38]. Each protein model consists of the sequence of one protein (wild-type or mutant sequence) and the structure of another.

Before carrying out comparative modeling, 4 hybrid sequences based on CaM and CNB were generated as listed in Table 4. According to Hilbert et al., a difference in sequence length between template and target exceeding 10% fails to produce a reasonable model because of the degree of structural divergence [59]. Meanwhile, to confirm the role of critical structural determinants, the impacts of the N- and C-tails of CNB needed also to be eliminated. For these reasons we added the N- and C-tails of CNB to both the wild-type and mutant sequences of CaM, and in turn deleted the two tails from the CNB sequences before modeling.

Comparative modeling was performed using the MODELLER program [60,61]. Each sequence was used to build a hybrid protein using a structure as template, as shown in Table 4. For example, the sequence of MCaMW was used to build a model employing the structure of CNB as a template. The standard MODELLER routine was used. The best model was selected based on the lowest value of the DOPE assessment score in five models calculated for the same target.

Subsequently, the modeled structures were optimized in GROMACS in order to improve their conformations, as described in Section 2.2.

2.5. Structural assessments and MD simulations for hybrid proteins

The Verify3D program [62] was used to evaluate the conformations and inter-residue contacts of wild-type and hybrid proteins. The Verify3D program provides structural assessments at the residue level, and enables the user to identify regions of proteins that are likely to have the correct conformations, and to look for incorrectly folded regions. The program uses information about local secondary structure, solvent accessibility and fractions of side-chain areas that are covered by polar atoms.

The procedures for the 20 ns MD simulations and analysis of the trajectories of hybrid proteins were similar to those described in Section 2.2.

3. Results

3.1. Distinct domain arrangements and dynamics

3.1.1. Quantification of the differences between the two EF-hand domain arrangements

The structural alignment matrix of the 9 proteins shown in Table S1 indicates that there is a significant structural difference between the CNB-like proteins and CaM-like proteins. Comparison of the 6 CNB-like proteins (CNB, recoverin, CIB1, NCS1, GCAP3 and KChIP1) shows that their structures are very similar. The RMSDs between all pairs of the 6 CNB-like proteins are below 3.5 Å. However the RMSDs between pairs of the 3 CaM-like proteins (CaM, TnC and hCLP) are not as low; the maximum RMSD is about 6.7 Å (between TnC and hCLP). Considering the difference between the mutual orientations of the two globular domains with respect to the central linker, this result is reasonable. Importantly, the RMSD between any CNB-

Table 2
Definitions of the EF hands and globular domains of the 9 proteins [a].

Protein	EF1	EF2	EF3	EF4	EF2–EF3	N-terminal	C-terminal
CaM	8–43	44–79	81–116	117–149	44–116	1–79	81–149
TnC	15–50	51–86	91–126	127–160	51–126	1–86	91–160
hCLP	8–43	44–79	81–116	117–149	44–116	1–79	81–149
CNB	18–51	50–85	87–122	128–163	50–122	1–85	87–170
RECO	24–59	61–96	97–132	147–182	61–132	1–96	97–202
CIB1	30–65	66–101	103–138	148–183	66–138	1–101	103–191
NCS1	24–59	60–95	96–131	144–179	60–131	1–95	96–190
GCAP3	15–50	52–87	88–123	130–165	52–123	1–87	88–209
KChIP1	38–94	97–132	133–168	181–216	97–168	1–132	133–227

[a] The numbers are the residue numbers in the primary amino acid sequences. The definitions of each EF hand were derived from the UniProt Database, except for the EF1 and EF2 motifs of CIB1, which are not available in the UniProt Database, and were estimated by us from the average length of an EF hand and the crystal structure of CIB1.

Table 3
Critical sites in the EF2–EF3 region of the 9 proteins [a].

	CAM	TnC	hCLP	CNB	RECO	CIB1	NCS1	GCAP3	KChIP1
1	E47	E54	E47	L53	Y64	F69	F64	H55	Y90
	−3.5	−3.5	−3.5	3.8	−1.3	2.8	2.8	−3.2	−1.3
2	E54	E61	E54	I60	S71	V76	V71	T62	A97
	−3.5	−3.5	−3.5	4.5	−0.8	4.2	4.2	−0.7	1.8
3	V55	V62	I55	F61	F72	F77	F72	F63	F98
	4.2	4.2	4.5	2.8	2.8	2.8	2.8	2.8	2.8
4	M71	M78	M71	G77	A88	L94	A88	A79	A114
	M72	M79	M72	V78	L89	L95	L89	V80	L115
	1.9	1.9	1.9	1.9	2.8	3.8	2.8	3.0	2.8
5	R74	R81	R74	Q80	M91	V97	V91	L82	I117
	−4.5	−4.5	−4.5	−3.5	1.9	4.2	4.2	3.8	4.5
6	K75	Q82	K75	F81	T92	F98	T92	I83	L118
	−3.9	−3.5	−3.9	2.8	−0.7	2.8	−0.7	4.5	3.8
7	EEE	EEE	EEE	EQK	NQK	DIK	DEK	EQK	HEK
	82–84	92–94	82–84	88–90	98–100	105–107	98–100	89–91	124–126
	−3.5	−3.5	−3.5	−3.6	−3.6	−1.0	−3.6	−3.6	−3.5
8	E87	E97	E87	F93	W103	Y110	W103	W94	W129
	−3.5	−3.5	−3.5	2.8	−0.9	−1.3	−0.9	−0.9	−0.9
9	K94	R104	K94	M100	V110	F117	L110	A101	I136
	−3.9	−4.5	−3.9	1.9	4.2	2.8	3.8	1.8	4.5
10	A102	A112	A102	N108	K118	R125	R118	K109	K144
	1.8	1.8	1.8	−3.5	−3.9	−4.5	−4.5	−3.9	−3.9

[a]Each row represents a conserved site, which consists of 1–3 residues. The numbers on the right of (or below) the one-letter abbreviations of amino acids are the numbers of residues in the three-dimensional structure in the respective PDB file. In each row except rows 4 and 7, the value below the residue is the hydropathy scale for that residue according to Kyte and Doolittle (Ref. [66]). In rows 4 and 7, the value is the average hydropathy scale for 2–3 residues. Hydropathy scales below −2.0 are shown as white (strong hydrophilicity); scales between −2.0 and 0 are highlighted by 20% gray (weak hydrophilicity); those between 0 and 2.0 are highlighted by 40% gray (weak hydrophobicity) and those above 2.0 by 60% gray (strong hydrophobicity).

like protein and any CaM-like protein exceeds 11.5 Å, clearly indicating that the two groups of proteins have distinct EF-hand domain arrangements.

3.1.2. Different dynamical behavior in MD simulations

The CNB-like and CaM-like proteins behaved significantly different in our 20 ns MD simulations. From the RMSDs for the whole protein and individual domains, different ranges of dynamics for their respective trajectories can be identified (Fig. 1, the left Y axis). From Fig. 1D–I, one can see that the structures of the CNB-like proteins were quite stable during the simulations. After equilibration for 2 ns, neither the RMSDs for the whole protein nor the RMSDs for the two globular domains underwent any large changes or significant fluctuations. On the other hand, the RMSDs for the whole protein for the 3 CaM-like proteins all changed significantly: the maximal RMSD for CaM was about 12 Å, and the maximal RMSDs for TnC and hCLP were about 10 Å (Fig. 1A–C). As to the two globular domains of the 3 CaM-like proteins, the fluctuations of the RMSDs were not as obvious as for the whole protein, especially in the case of the two domains of hCLP, which behaved like the CNB-like proteins. This suggests that there are interdomain motions other than internal movements within the domains in the CaM-like proteins; these will be further investigated in the interdomain distance analysis below.

Distance analysis for the N- and C-terminal domains suggests that there were relatively large motions of the two domains in the CaM-like proteins, but not in the CNB-like proteins (Fig. 1, the right Y axis). Changes of distance between the N- and C-terminal domains reveal interactions between the domains of proteins and their global structural stabilities. During the simulations, the interdomain distances for the 3 CaM-like proteins fluctuated by about 15 Å (CaM and TnC) and 8 Å (hCLP). The unstable dynamical behavior and the independence of the two domains of these proteins are consistent with previous conclusions based on MD simulations and NMR experiments on CaM [52,63–65]. In contrast the interdomain distances of the CNB-like proteins CNB, CIB1, NCS1, GCAP3 and KChIP1 only fluctuated within 1 Å after equilibration for 2 ns. The exception was recoverin: its interdomain distance changed from 24 Å to 20 Å, but its fluctuations were all within 4 Å. Note that the interdomain distances of the CNB-like proteins are much smaller than those of the CaM-like proteins: the former are always below 25 Å, while the latter are 25–42 Å. This is consistent with the structural difference between the two groups of proteins as shown in Fig. S1.

To identify the essential modes of motion that are most fundamental to protein activity [51], we employed ED to analyze all of these proteins (Fig. 2). Here fluctuations relative to trajectory average structures were projected on ED modes 1 and 2, which correspond to the first and second largest eigenvectors, respectively. In the trajectory of CaM, mode 1 contributes 65% to the total backbone mean square fluctuations, and the first two modes contribute 85%. Similarly, the first two modes of TnC and hCLP contribute about 80%. In the CNB-

Table 4
Compositions of the 4 sequences prepared for modeling.

Model	Template	Length (residue)	Sequence composition description
MCaMW	CNB	167	Wild-type sequence of CaM with the N- and C-tails of CNB [a]
MCaMM	CNB	167	Mutant sequence of CaM [b] with the N- and C-tails of CNB
MCNBW	CaM	150	Wild-type sequence of CNB without the N- and C-tails of CNB
MCNBM	CaM	150	Mutant sequence of CNB [c] without the N- and C-tails of CNB

[a] The N-tail of CNB refers to the sequence (GNEASYPLE); the C-tail of CNB refers to the sequence (DIHKKMVVDV).
[b] The mutant sequence of CaM refers to the CaM sequence in which the critical sites (Table 3) are substituted by the corresponding sites in CNB.
[c] The mutant sequence of CNB refers to the CNB sequence in which the critical sites (Table 3) are substituted by the corresponding sites in CaM.

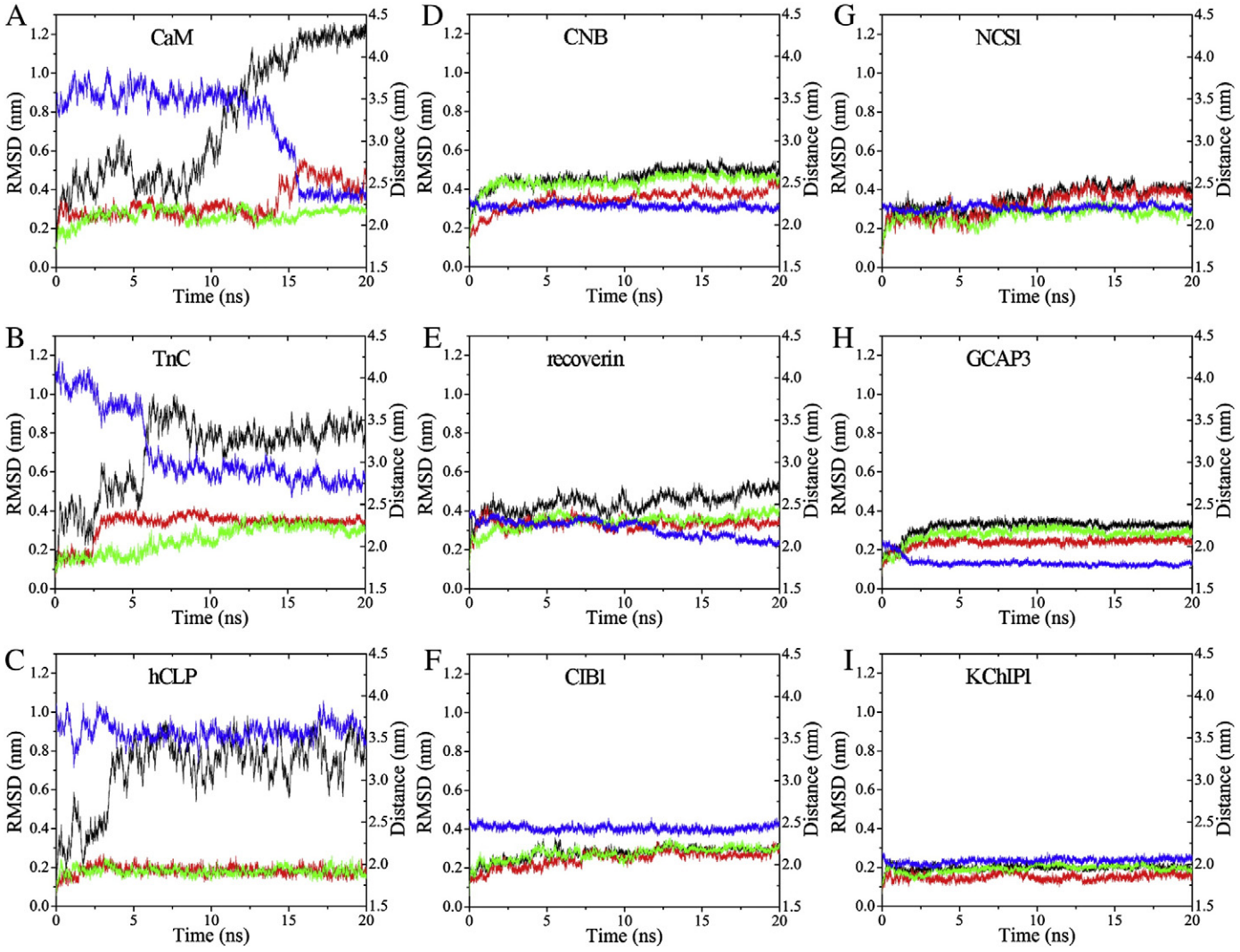

Fig. 1. RMSDs and interdomain distances of the 9 EF-hand CaBPs. The left Y axis shows the RMSDs for the whole protein and individual domains, and the right Y axis shows the distances between the N- and C-terminal domains (interdomain distances). The black, red and green curves show the RMSDs of the whole protein, N-terminal domain and C-terminal domain, respectively, while the blue curve displays the interdomain distances.

like proteins, the contributions of mode 1 to the total backbone mean square fluctuations are all around 35%, except in CNB, in which it contributes 45%. The contributions of the first two modes amount to only about 50%. The major modes of motion of the CaM-like proteins can be categorized as rigid-body domain motions: bending of the central linker around a hinge located in its center, and twisting of the N- and C-terminal domains around the central linker, whereas the modes of the CNB-like proteins can be described as breathing-type motions: elastic deformations of globular proteins rather than extensive mutual motions of domains.

3.2. Structural determinants in the sequences

3.2.1. Identification of the residues differing between the proteins with the two domain arrangements

Since these EF-hand proteins share high sequence homology but clearly differ in their structural and dynamical features, our next concern was how their sequences specify their architectures. We therefore performed sequence alignments and phylogenetic tree analysis on the 9 proteins to address this question.

Surprisingly, alignments of the complete sequences of the 9 proteins (Fig. 3), and the phylogenetic tree in Fig. 4A, indicate that the complete sequences could not distinguish between the CNB-like proteins and CaM-like proteins. CNB and all the CaM-like proteins, i.e., CaM, hCLP and TnC, had high sequence homology, while several other CNB-like proteins, e.g., GCAP3, KChIP1, NCS1 and recoverin were included in the second branch, and CIB1 was categorized as a third branch.

As EF2–EF3 is the critical region linking the N- and C-terminal domains, which includes the linker and interface between the N- and C-terminal domains of both the CaM-like and CNB-like proteins, we thought that sequence alignments of this region might reveal the difference between the two groups more clearly than alignments of the complete sequences. We also performed sequence alignments of other structural elements, i.e., the EF1 and EF4 motifs, to demonstrate the role of the EF2–EF3 region more convincingly.

As shown in Fig. 4B–C, neither alignments of the EF1 motif nor alignments of the EF4 motif could distinguish between the two groups of proteins. The results for alignments of the EF1 motif (Fig. 4B) were very similar to those of the complete sequences. In the alignments of the EF4 motif (Fig. 4C), 2 CNB-like proteins (CIB1 and CNB) and 3 CaM-like proteins were included in one major branch, and the other 4 CNB-like proteins formed the other major branch.

When the EF2–EF3 region was aligned, the difference between the CaM-like and CNB-like proteins was much more significant. In Fig. 4D, the CaM-like proteins all formed one branch, and the CNB-

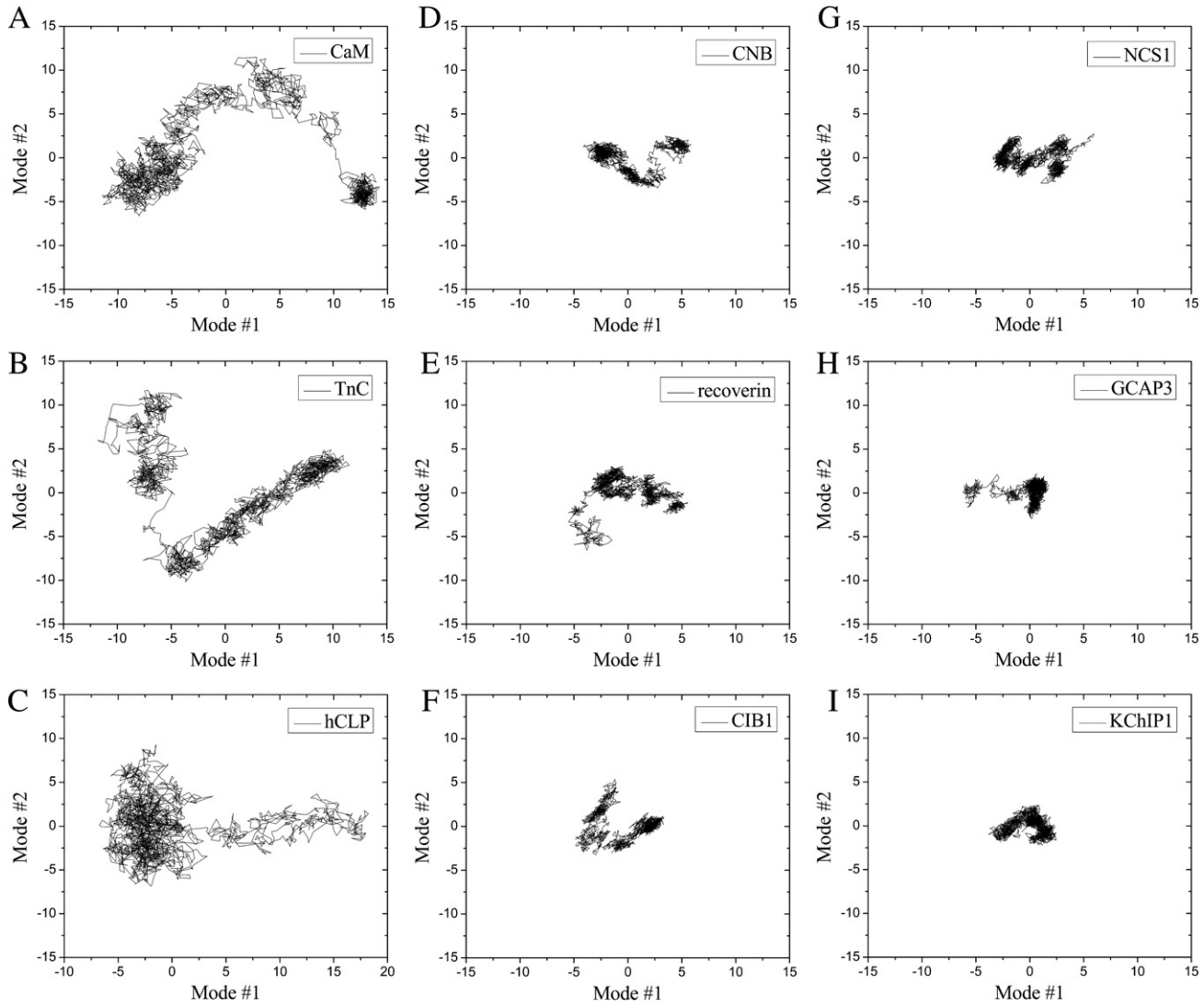

Fig. 2. ED of the 9 proteins. All are shown on the same scale, in which both the X and Y axes extend from −15 to 15 (except the X axis of hCLP, which extends from −10 to 20).

like proteins another. We note that there were two sub-branches in the CNB-like branch. One contained CNB and CIB1, and the other GCAP3, KChIP1, NCS1 and recoverin. These sub-branches agree well with the observation that GCAP3, KChIP1, NCS1 and recoverin are NCS proteins, which mainly function in neurons [36,37], while CIB1 and CNB are widely expressed as multifunctional proteins [20,34]. Thus the EF2–EF3 region is indeed the key region that permits the CaM-like proteins to be distinguished from the CNB-like proteins.

Further detailed comparison of these EF2–EF3 sequences reveals several critical sites which differentiate the two groups of proteins (Table 3). In Table 3, each row represents a conserved site, and the hydropathy scale for each site has been labeled according to Kyte and Doolittle [66]. At sites 1, 2, 5, 6, 8 and 9, the residues in the CaM-like proteins are all strongly hydrophilic, while most of the residues in the CNB-like proteins are hydrophobic or at least relatively weakly hydrophilic. At site 4, the residues in the CaM-like proteins are all weakly hydrophobic methionines, while the residues in the CNB-like protein are almost all strongly hydrophobic. Hence, the overall hydrophobicity of the CNB-like proteins at these sites is higher than that of the CaM-like proteins.

However, the situation with regard to these identified critical sites is more complicated than can be explained by this relative hydrophobicity/hydrophilicity alone. Thus, site 10 is an interesting reverse case, in which the residues in the CaM-like proteins are all weakly hydrophobic, while the residues in the CNB-like proteins are all strongly hydrophilic. At site 3, all the residues in the CaM-like and CNB-like proteins are strongly hydrophobic, but the residues in the CaM-like proteins are valine or isoleucine, while those in the CNB-like protein are highly conserved phenylalanines. Similarly at site 7, although the residues in the CaM-like and CNB-like proteins are almost all strongly hydrophilic, the residues in the CaM-like proteins form the acidic residue-rich motif “EEE”, which may contribute to generating the extended central linker, while the corresponding residues in the CNB-like proteins include adjacent oppositely charged residues. Thus at sites 3, 7 and 10, although the hydropathy scales for the CaM-like and CNB-like proteins are not consistent with the rest of the 10 sites, these sites do differ in the two groups, and probably fulfill some specific structural or functional roles, which will require further experimental or computational studies to identify.

The positions of these important sites are much more widely spread than the central linker and than even the entire long helix region in the CaM-like proteins (including the last helix of the N-terminal domain and the first helix of the C-terminal domain), while they are almost at the interface between the N- and C-terminal domains in the CNB-like proteins (Fig. S2). The only exception is site 10 in the C-terminal loop of the EF3 motif, far from the interface between the two domains.

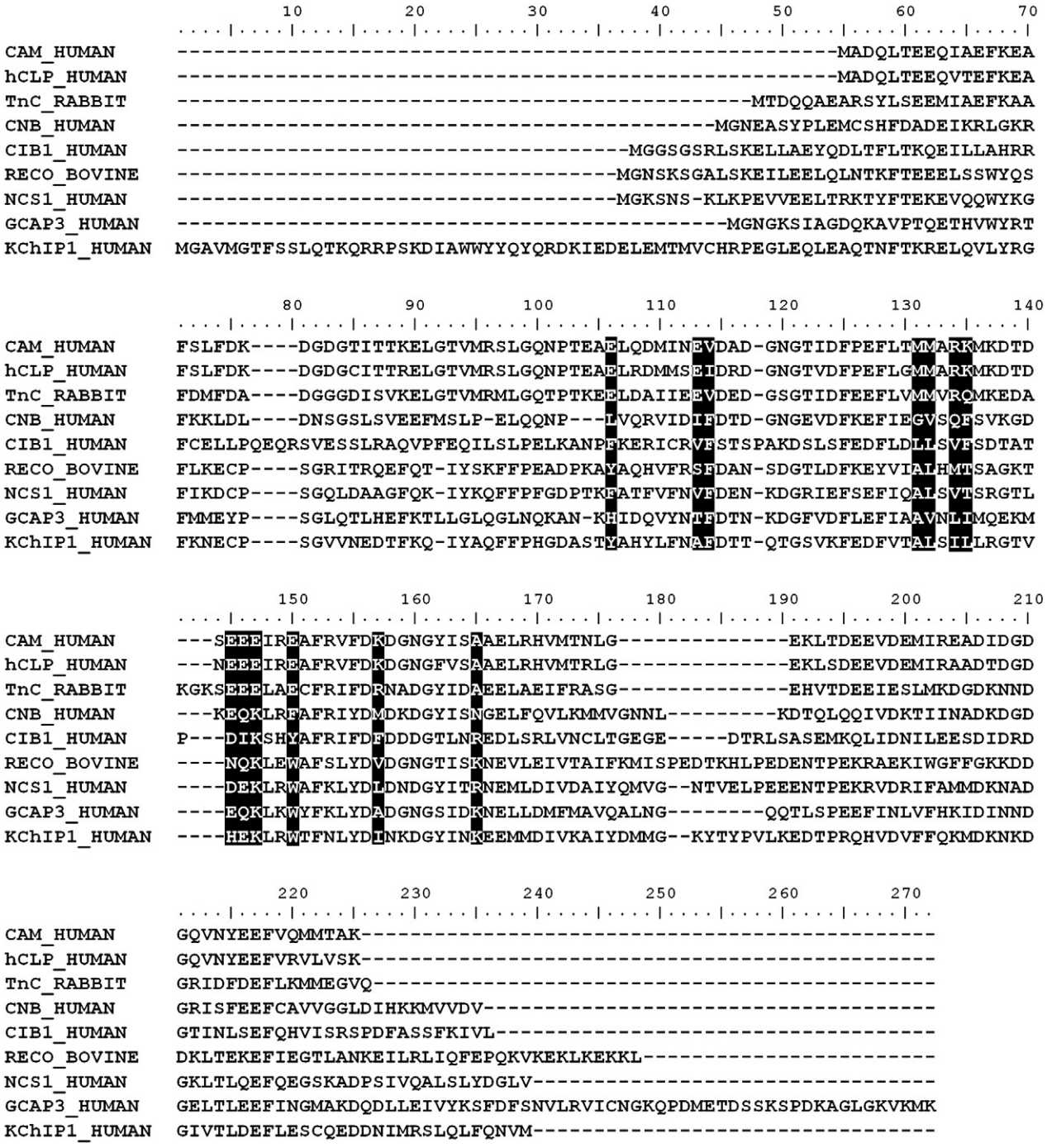

Fig. 3. Sequence alignments of the 9 proteins. The complete sequences were obtained from the UniProt entries listed in Table 1. The residues listed in Table 3 are highlighted by black boxes.

Furthermore, we also performed structure-based sequence alignments on the 9 proteins (Fig. S3), to verify whether our pure sequence alignment (Fig. 3) is accurate or not. It shows that the results of our pure sequence alignments and structure-based sequence alignments are highly consistent, as shown in the two figures (Figs. 3 and S3) the critical structural determinants are exactly at the same position, and the canonical EF-hand sequences are appropriately aligned in both pure sequence alignment and structural-based sequence alignment in our case.

3.2.2. Confirmation of these structural determinants

To verify the importance of the above sites, i.e., to test whether the residues listed in Table 3 determine the domain arrangement, we used a customized inverse folding approach involving comparative modeling of CaM and CNB. If these residues are indeed structural determinants then when the residues in one group are replaced by the corresponding residues in the other group, the folding preference of the mutant sequence should switch. We generated 4 hybrid proteins as described in Section 2.4: we then used both the wild-type, as a control, and mutant CaM sequences to produce new models using the structure of CNB as template; similarly, models of both the wild-type and mutant sequences of CNB were built using the structure of CaM as template. We then carried out Verify3D assessments and MD simulations for the hybrid proteins to estimate the compatibility and stability of each model, as described in Section 2.5.

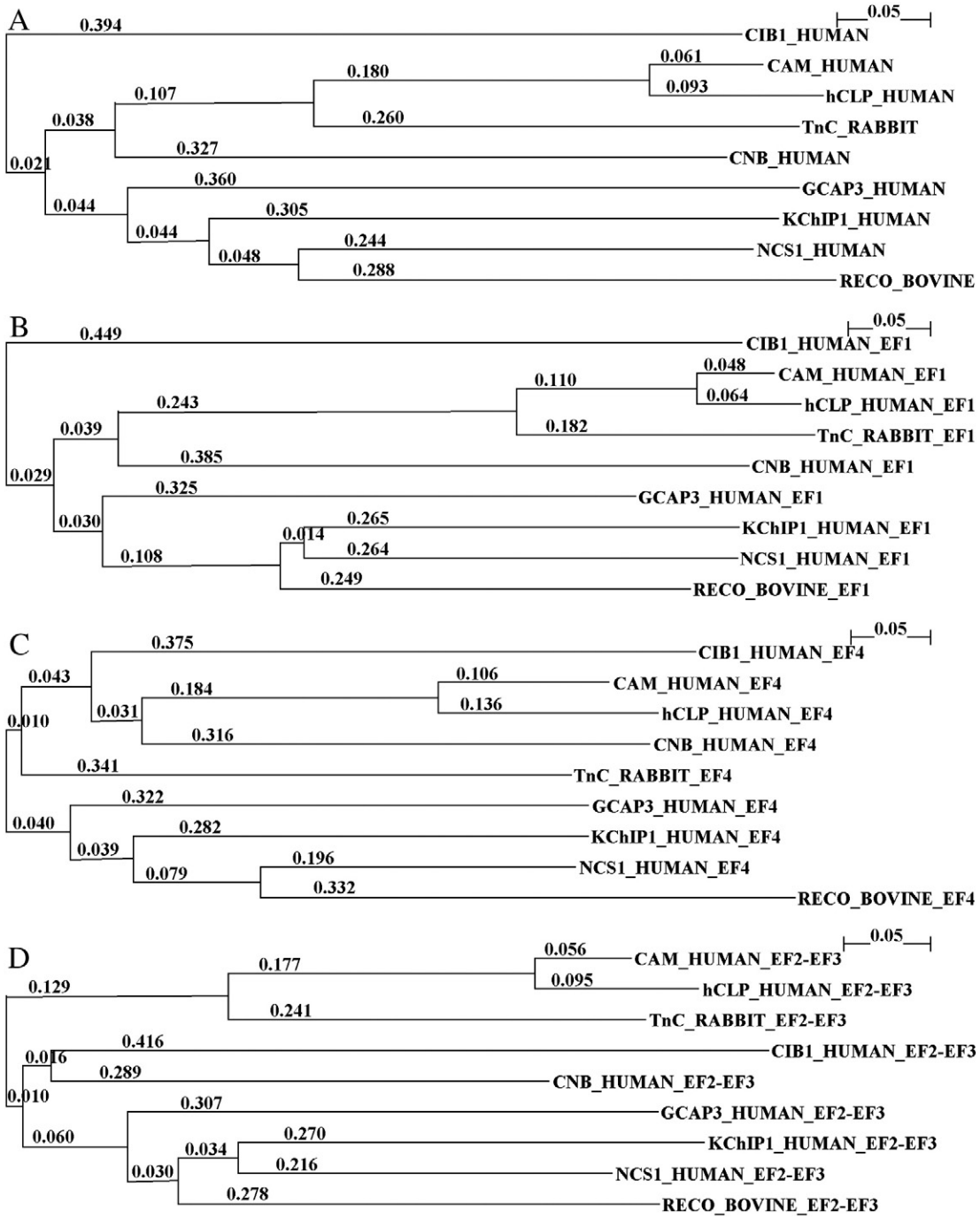

Fig. 4. Phylogenetic tree analysis of the complete and partial sequences of the 9 proteins. Panels (A) to (D) show the phylogenetic tree analysis of the complete sequence, EF1 motif, EF4 motif, and EF2–EF3 region, respectively. The number next to each branch gives the branch length, which represents the number of differences per residue along the branch (Ref. [54]).

3.2.2.1. Verify3D assessments for hybrid proteins. We employed Verify3D to compare and evaluate the compatibilities of the wild-type structure and hybrid proteins as shown in Fig. 5. The average scores of MCaMW, MCaMM and CNB, three structures with the CNB-like architecture, were 0.29, 0.33 and 0.44, respectively (Since Verify3D cannot give precise assessments for the sequences of the N- and C-tails, only residues 13–156 were calculated). The average scores of MCNBW, MCNBM and CaM, three structures with the CaM-like architecture, were 0.23, 0.27 and 0.39, respectively (For the same reason as before, only residues 12–137 were calculated). The average scores were in the order: wild-type structures > hybrid proteins based on the mutant sequence > hybrid proteins based on the wild-type sequence. It can be easily understood that the wild-type structure is the most compatible one, because it has been created by evolution. On the other hand, the average scores of the hybrid proteins based on the mutant sequences were significantly higher than those based on the wild-type sequences for both CaM and CNB (0.23 vs 0.27, 0.29 vs 0.33), suggesting greater compatibility of the mutant sequences than the wild-type sequences.

More importantly, we found that in the central region of the hybrid proteins, which covers all of the mutations, the scores for the mutant sequences were significantly higher than those for the wild-type sequences. For example, the scores for residues 50–95 in MCaMM (between the vertical dashes in Fig. 5A), all with values

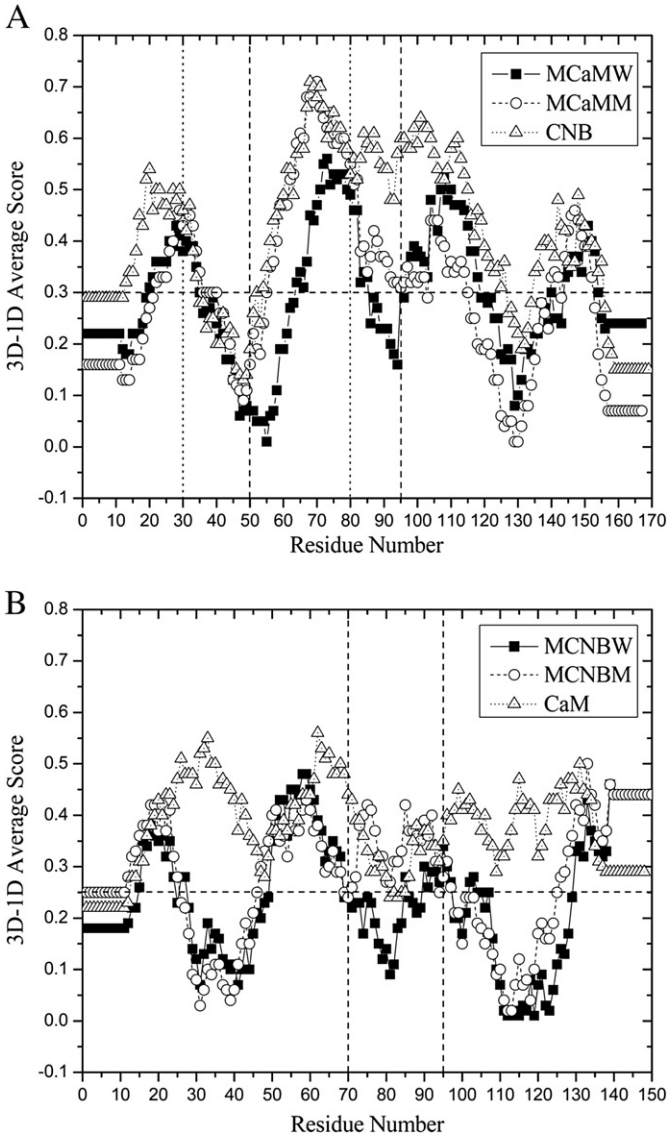

Fig. 5. Verify 3D assessments for the wild-type structures and hybrid proteins. (A) Verify 3D assessments for MCaMW, MCaMM and CNB; (B) Verify 3D assessments for MCNBW, MCNBM and CaM. The 3D–1D average score means the average score for residues in a 21-residue sliding window, the center of which is at the position indicated by the X axis (Ref. [62]). The vertical dashes represent residues 50 to 95 in panel (A) and residues 70 to 95 in panel (B); the vertical dots represent residues 30 to 80 in panel (A); the horizontal dashes represent a score of 0.3 in panel (A) and 0.25 in panel (B).

above 0.3 (the horizontal dashes in Fig. 5A), were significantly higher than those in MCaMW; the scores for residues 70–95 in MCNBM (between the vertical dashes in Fig. 5B), all with values above 0.25 (the horizontal dashes in Fig. 5B) and average value above 0.30, were significantly higher than those in MCNBW. The higher scores in the regions covering all the mutations confirm the importance of the structural determinants listed in Table 3. Furthermore, in the region of residues 30–80 in Fig. 5A (between the vertical dots), corresponding to most of the N-terminal domain of MCaMM, the scores for MCaMM were very close to those of wild-type CNB, further confirming its compatibility.

In many protein modeling procedures, a Verify3D score > 0.3 is considered acceptable [67–69]. Here the average scores for the whole protein, i.e., MCNBW (0.23) and MCNBM (0.27) were all below 0.3. This was caused by the incompatibility of the regions within the individual domains, not the region between the two domains (average above 0.30) which would determine the domain arrangement. The scores for the regions within each domain (residues 25–45 and residues 105–125 in Fig. 5B) of MCNBW and MCNBM were significantly lower than those for wild-type CaM, thus strongly decreasing the average scores. In the light of these considerations we only performed MD simulations on MCaMW and MCaMM, as described in the next section.

3.2.2.2. MD simulations on MCaMW and MCaMM. As shown in the RMSD plot, the structure of MCaMM was quite stable during the simulations (Fig. 6B, the left Y axis). Neither the whole protein nor the two globular domains underwent large changes and significant fluctuations, much like the RMSD plot of CNB shown in Fig. 1D (the left Y axis). The distance between the N- and C-terminal domains of MCaMM was also quite stable, with a value around 23 Å (Fig. 6B, the right Y axis), which is also quite similar to the interdomain distance of CNB as shown in Fig. 1D (the right Y axis).

On the other hand, the structure of MCaMW was unstable during the simulations (Fig. 6A, the left Y axis). The maximum RMSD for the whole protein reached 16 Å, pointing to significant conformational changes. But the RMSDs for the N- and C-terminal domains were comparatively stable. This suggests that the movements of the overall structure of MCaMW do not result from internal movements within each domain, but from relative movements of the two domains. The interdomain distance of MCaMW fluctuated between 24 Å and 40 Å (Fig. 6A, the right Y axis), and it can be seen that the changes of RMSD for the whole protein (black curve in Fig. 6A) were highly correlated with the changes of interdomain distance (blue curve in Fig. 6A), confirming that the independent motion of one domain with respect to the other contributes most of the structural movements of MCaMW.

Comparison of the first and last snapshots of MCaMW and MCaMM during the simulations (Fig. 6C–F) clearly shows that there were global conformational changes of MCaMW at the end of the simulations, with significant rearrangements of the two domains, whereas the structure of MCaMM was much more stable, almost maintaining the arrangements of the modeled domains. After the 20 ns simulations, the structure of MCaMW changed from a compact globular molecule to a scattered one. The structural determinant residues (purple sticks in Fig. 6C) at the interface of the two globular domains were unable to hold the two domains together; instead they tended to be exposed to the solvent (purple sticks in Fig. 6D), and the distance between them increased. On the other hand, during the simulations for MCaMM, the mutated residues (purple sticks in Fig. 6E and F) at the interface of the two globular domains could bind and stabilize the two domains, so that the overall structure was stable.

The stability of MCaMM in the simulations suggests that the CNB-like domain arrangement based on the mutant CaM sequence is much more compatible than that based on the wild-type CaM sequence. It follows that the CNB-like models based on the wild-type sequences of the CaM-like proteins are energetically unfavorable, i.e., the sequences of the CaM-like proteins cannot fold into the CNB-like structures, whereas replacement of the identified determinants in the EF2–EF3 region (Table 3) with the corresponding residues in the CNB-like proteins appears to permit alteration of the domain arrangement. We conclude that the residues identified in our sequence analysis are indeed structural determinants.

4. Discussion

4.1. The critical structural determinants are not restricted to the linker between the EF2 and EF3 motifs

In the critical structural determinant sites, almost all of the residues are located at the interface between the two globular domains of the CNB-like proteins, but are much more widely distributed than the central linker, and even the long helix region of the CaM-like proteins. A previous study compared the hydrophilicity of the linkers

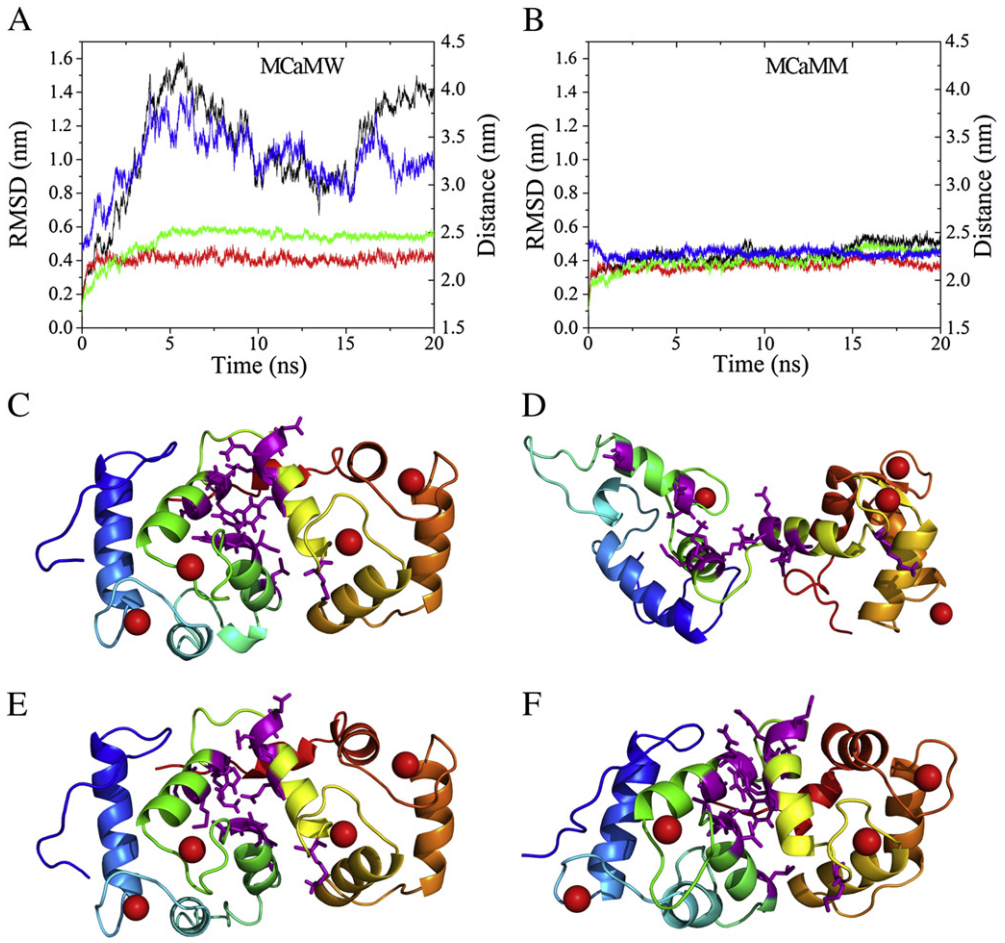

Fig. 6. MD simulations of the hybrid proteins MCaMW and MCaMM. (A) RMSD and interdomain distance analysis of MCaMW; (B) RMSD and interdomain distance analysis of MCaMM; (C and D) the first and last snapshots of MCaMW; (E and F) the first and last snapshots of MCaMM. In panels (A) and (B), the black, red and green curves are the RMSDs of the whole protein, N-terminal domain and C-terminal domain, respectively (the left Y axis); while the blue curve is the interdomain distances (the right Y axis). The definitions of the two domains of MCaMW and MCaMM are the same, i.e., residues 1–88 and 89–167 for the N- and C-terminal domains, respectively. Panels (C) to (F) display the proteins (ribbons), residues listed in Table 3 except for the last row (purple sticks) and calcium ions (red spheres).

between the EF2 and EF3 motifs, and focused on a narrow region consisting on average of 5 residues [39]. In contrast, we analyzed a much more extensive region, i.e., the entire EF2–EF3 region, with an average length of 55 residues. Our data show that the linker between EF2 and EF3 motif is not the only factor determining folding of the proteins into different arrangements.

4.2. The relationships between sequence, structure, dynamics and function

By analyzing the trajectories of the MD simulations, we identified the dynamical features distinguishing the CaM-like proteins from the CNB-like proteins. The CaM-like proteins were highly flexible, whereas the CNB-like proteins consistently exhibited a high degree of stability. Such distinct dynamical can be attributed to their distinct domain arrangements. As shown in Fig. S1, the two globular domains of the CaM-like proteins are relatively independent, and the interactions between them are weak, while the distance between the two domains of the CNB-like proteins is limited, presumably due to the strong interactions between them. Thus in the MD simulations, the CaM-like proteins showed considerable flexibility and extensive rigid-body domain motions, whereas the CNB-like proteins exhibited high stability and breathing-type motions.

On the primary sequence level, the domain arrangements of the CaM-like proteins and CNB-like proteins are determined by several critical residues identified by sequence alignments in the EF2–EF3 region. We confirmed the critical role of these residues by the customized inverse folding approach, together with Verify3D assessments and MD simulations. The structural determinant sites help the CaM-like proteins to form extended helices exposed in solution, while the corresponding sites in the CNB-like proteins provide the structural basis for the more compact domain arrangement.

The differences in structural and dynamical properties resulting from these sequence characteristics have different functional implications for the CaM-like and CNB-like proteins, and these can account very well for the differences in the reported functions of these EF-hand CaBPs, especially their manner of interacting with target proteins.

Current views of how the CNB-like proteins interact with target proteins require their structural stability. A critical structural feature of the CNB-like proteins is a hydrophobic groove at the “bottom” of the protein (opposite all the calcium binding sites), as a result of the compact array of the two globular domains on the same side [18,70,71]. This hydrophobic groove is invariably the site at which the CNB-like proteins bind target proteins and regulate their functions. The binding sites in these target proteins usually form a helix, which can be embedded in the hydrophobic groove, contacting and interacting with the groove. Recoverin forms a hydrophobic groove only at the bottom of the N-terminal domain, where it interacts with rhodopsin kinase [72]. Yamniuk and Vogel have proposed a model for CIBI binding to integrin αIIb, in which only the C-terminal domain

forms a sufficiently hydrophobic groove to bind the helix of αIIb [34], although the structure for this complex has not been resolved. But more importantly, a more general model for the CNB-like proteins specifies that the continuous and elongated hydrophobic groove at the bottom of both the N- and C-terminal domains binds to their target, as in CNB interacting with CNA (calcineurin A) [18,19], CHP1 or CHP2 interacting with NHE1 [24,73], Frq1 interacting with Pik1 (phosphatidylinositol 4 kinase) [74], and KChIP1 interacting with Kv channels 4.2 [75] or 4.3 [76,77] (Fig. S4). In such a mechanism, if the structure of a CNB-like protein was not stable enough to keep the N- and C-terminal domains together, the integrity of the hydrophobic groove would be lost, preventing binding to its target.

In contrast to the CNB-like proteins, the diverse interactions of the CaM-like proteins with their target proteins require the high conformational flexibility of these proteins. For example, the flexibility of CaM enables it to interact with target proteins that are functionally and structurally diverse. The flexible central linker of CaM can bend significantly upon binding to target proteins, allowing independent reorientation of the two globular domains to accommodate the different structures of the target proteins [71]. In addition, CaM can bind target proteins in an extended form or in a dimeric form [5,6], whereas TnC target proteins are limited to TnI and TnT. The former binds actin and inhibits muscle contraction in the absence of Ca^{2+}, and the latter binds tropomyosin and anchors Tn to the thin filament. As the C-terminus of TnC is relatively fixed by binding to TnT and the N-terminus of TnI, the flexible linker between the two globular domains allows the N-terminus of TnC to move between a location "near actin" and a location "far away from actin" [78], and enables TnC to switch between an extended dumb-bell conformation (binding the C-terminus of TnI) and a relaxed conformation (releasing the C-terminus of TnI) [79]. Thus muscle contraction can be regulated.

Thus the distinct domain arrangements of the CaM-like and CNB-like proteins, determined by the several critical sites we have identified in the EF2–EF3 region, provide them with different structural and dynamical properties, and permit them to interact with their respective target proteins.

5. Conclusion

Using structural comparisons and MD simulations of the two EF-hand domain arrangements of CaBPs (3 CaM-like proteins and 6 CNB-like proteins) we have identified the respective structural and dynamical features of the two groups of proteins. During the simulations, the CaM-like proteins displayed high flexibility with large conformational changes that can be mainly described as rigid-body domain motions, while the CNB-like proteins were much more stable dynamically with only local conformational changes resembling breathing-type motions.

Sequence alignments and phylogenetic tree analysis were performed to address how sequence specifies structure and dynamics. Several structural determinant sites were identified that were not restricted to within the linker between the two domains, but were widely distributed throughout the EF2 and EF3 motifs, and these were verified by a customized inverse folding approach.

Finally, we surveyed the previously reported functions of the two groups of CaBPs. The structural and dynamical features of the two groups were clearly shown to accommodate their respective functions, especially the ways they interact with their respective target proteins.

The identification of structural determinants, together with our dynamics simulations connecting structure with function, provides a clearer and deeper understanding of the sequence–structure–dynamics–functional relationships between the EF-hand CaBPs.

Supplementary materials related to this article can be found online at doi:10.1016/j.bbapap.2012.01.005.

Acknowledgements

This work was supported by grants from the National Science Foundation of China, the International Cooperation Project and the National Important Novel Medicine Research Project. The authors appreciate the help received from the High Performance Science Computing Center of Beijing Normal University, and Dr. Hong-wei Tan and Yong Zhang of the Department of Chemistry of Beijing Normal University.

References

[1] E. Carafoli, Calcium signaling: a tale for all seasons, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 1115–1122.
[2] E. Carafoli, The calcium-signalling saga: tap water and protein crystals, Nat. Rev. Mol. Cell Biol. 4 (2003) 326–332.
[3] M.R. Nelson, W.J. Chazin, Structures of EF-hand Ca2+-binding proteins: diversity in the organization, packing and response to Ca2+ binding, Biometals 11 (1998) 297–318.
[4] D. Chin, A.R. Means, Calmodulin: a prototypical calcium sensor, Trends Cell Biol. 10 (2000) 322–328.
[5] S.W. Vetter, E. Leclerc, Novel aspects of calmodulin target recognition and activation, Eur. J. Biochem. 270 (2003) 404–414.
[6] S. Bhattacharya, C.G. Bunick, W.J. Chazin, Target selectivity in EF-hand calcium binding proteins, Biochim. Biophys. Acta Mol. Cell Res. 1742 (2004) 69–79.
[7] Y.S. Babu, C.E. Bugg, W.J. Cook, Structure of calmodulin refined at 2.2 Angstrom resolution, J. Mol. Biol. 204 (1988) 191–204.
[8] R. Chattopadhyaya, W.E. Meador, A.R. Means, F.A. Quiocho, Calmodulin structure refined at 1.7 Angstrom resolution, J. Mol. Biol. 228 (1992) 1177–1192.
[9] C. Ban, B. Ramakrishnan, K.Y. Ling, C. Kung, M. Sundaralingam, Structure of the recombinant *Paramecium tetraurelia* calmodulin at 1.68 Angstrom resolution, Acta Crystallogr. D: Biol. Crystallogr. 50 (1994) 50–63.
[10] M.A. Wilson, A.T. Brunger, The 1.0 angstrom crystal structure of Ca2+-bound calmodulin: an analysis of disorder and implications for functionally relevant plasticity, J. Mol. Biol. 301 (2000) 1237–1256.
[11] K.A. Satyshur, S.T. Rao, D. Pyzalska, W. Drendel, M. Greaser, M. Sundaralingam, Refined structure of chicken skeletal muscle troponin C in the two-calcium state at 2 Angstrom resolution, J. Biol. Chem. 263 (1988) 1628–1647.
[12] K.A. Satyshur, D. Pyzalska, M. Greaser, S.T. Rao, M. Sundaralingam, Structure of chicken skeletal muscle troponin C at 1.78 Angstrom resolution, Acta Crystallogr. D: Biol. Crystallogr. 50 (1994) 40–49.
[13] A. Houdusse, M.L. Love, R. Dominguez, Z. Grabarek, C. Cohen, Structures of four Ca2+-bound troponin C at 2.0 angstrom resolution: further insights into the Ca2+-switch in the calmodulin superfamily, Structure 5 (1997) 1695–1711.
[14] B.G. Han, M.H. Han, H.X. Sui, P. Yaswen, P.J. Walian, B.K. Jap, Crystal structure of human calmodulin-like protein: insights into its functional role, FEBS Lett. 521 (2002) 24–30.
[15] B.D. Sykes, Pulling the calcium trigger, Nat. Struct. Biol. 10 (2003) 588–589.
[16] P. Yaswen, A. Smoll, D.M. Peehl, D.K. Trask, R. Sager, M.R. Stampfer, Down-regulation of a calmodulin-related gene during transformation of human mammary epithelial-cells, Proc. Natl. Acad. Sci. U. S. A. 87 (1990) 7360–7364.
[17] M.S. Rogers, E.E. Strehler, The tumor-sensitive calmodulin-like protein is a specific light chain of human unconventional myosin X, J. Biol. Chem. 276 (2001) 12182–12189.
[18] J.P. Griffith, J.L. Kim, E.E. Kim, M.D. Sintchak, J.A. Thomson, M.J. Fitzgibbon, M.A. Fleming, P.R. Caron, K. Hsiao, M.A. Navia, X-ray structure of calcineurin inhibited by the immunophilin-immunosuppressant FKBP12-FK506 complex, Cell 82 (1995) 507–522.
[19] C.R. Kissinger, H.E. Parge, D.R. Knighton, C.T. Lewis, L.A. Pelletier, A. Tempczyk, V.J. Kalish, K.D. Tucker, R.E. Showalter, E.W. Moomaw, L.N. Gastinel, N. Habuka, X.H. Chen, F. Maldonado, J.E. Barker, R. Bacquet, J.E. Villafranca, Crystal structures of human calcineurin and the human FKBP12-FK506-calcineurin complex, Nature 378 (1995) 641–644.
[20] F. Rusnak, P. Mertz, Calcineurin: form and function, Physiol. Rev. 80 (2000) 1483–1521.
[21] C.J. Blamey, C. Ceccarelli, U.P. Naik, B.J. Bahnson, The crystal structure of calcium- and integrin-binding protein 1: insights into redox regulated functions, Protein Sci. 14 (2005) 1214–1221.
[22] H.R. Gentry, A.U. Singer, L. Betts, C. Yang, J.D. Ferrara, J. Sondek, L.V. Parise, Structural and biochemical characterization of CIB1 delineates a new family of EF-hand-containing proteins, J. Biol. Chem. 280 (2005) 8407–8415.
[23] Y. Naoe, K. Arita, H. Hashimoto, H. Kanazawa, M. Sato, T. Shimizu, Structural characterization of calcineurin B homologous protein 1, J. Biol. Chem. 280 (2005) 32372–32378.
[24] Y.B. Ammar, S. Takeda, T. Hisamitsu, H. Mori, S. Wakabayashi, Crystal structure of CHP2 complexed with NHE1-cytosolic region and an implication for pH regulation, EMBO J. 25 (2006) 2315–2325.
[25] J.B. Ames, K.B. Hendricks, T. Strahl, I.G. Huttner, N. Hamasaki, J. Thorner, Structure and calcium-binding properties of Frq1, a novel calcium sensor in the yeast Saccharomyces cerevisiae, Biochemistry 39 (2000) 12149–12161.
[26] Y. Bourne, J. Dannenberg, V. Pollmann, P. Marchot, O. Pongs, Immunocytochemical localization and crystal structure of human frequenin (neuronal calcium sensor 1), J. Biol. Chem. 276 (2001) 11949–11955.

[27] S. Vijay-Kumar, V.D. Kumar, Crystal structure of recombinant bovine neurocalcin, Nat. Struct. Biol. 6 (1999) 80–88.
[28] J.B. Ames, R. Ishima, T. Tanaka, J.I. Gordon, L. Stryer, M. Ikura, Molecular mechanics of calcium-myristoyl switches, Nature 389 (1997) 198–202.
[29] R. Stephen, G. Bereta, M. Golczak, K. Palczewski, M.C. Sousa, Stabilizing function for myristoyl group revealed by the crystal structure of a neuronal calcium sensor, guanylate cyclase-activating protein 1, Structure 15 (2007) 1392–1402.
[30] J.B. Ames, A.M. Dizhoor, M. Ikura, K. Palczewski, L. Stryer, Three-dimensional structure of guanylyl cyclase activating protein-2, a calcium-sensitive modulator of photoreceptor guanylyl cyclases, J. Biol. Chem. 274 (1999) 19329–19337.
[31] R. Stephen, K. Paiczewski, M.C. Sousa, The crystal structure of GCAP3 suggests molecular mechanism of GCAP-linked cone dystrophies, J. Mol. Biol. 359 (2006) 266–275.
[32] R.H. Scannevin, K.W. Wang, F. Jow, J. Megules, D.C. Kopsco, W. Edris, K.C. Carroll, Q. Lu, W.X. Xu, Z.B. Xu, A.H. Katz, S. Olland, L. Lin, M. Taylor, M. Stahl, K. Malakian, W. Somers, L. Mosyak, M.R. Bowlby, P. Chanda, K.J. Rhodes, Two N-terminal domains of Kv4 K+ channels regulate binding to and modulation by KChIP1, Neuron 41 (2004) 587–598.
[33] P. Liang, H.Y. Wang, H. Chen, Y.Y. Cui, L.C. Gu, J.J. Chai, K.W. Wang, Structural Insights into KChIP4a Modulation of Kv4.3 Inactivation, J. Biol. Chem. 284 (2009) 4960–4967.
[34] A.P. Yamniuk, H.J. Vogel, Insights into the structure and function of calcium- and integrin-binding proteins, Calcium Binding Proteins 1 (2006) 150–155.
[35] X. Lin, D.L. Barber, A calcineurin homologous protein inhibits GTPase-stimulated Na-H exchange, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 12631–12636.
[36] R.D. Burgoyne, Neuronal calcium sensor proteins: generating diversity in neuronal Ca2+ signalling, Nat. Rev. Neurosci. 8 (2007) 182–193.
[37] H.V. McCue, L.P. Haynes, R.D. Burgoyne, The diversity of calcium sensor proteins in the regulation of neuronal function, Cold Spring Harb. Perspect. Biol. 2 (2011) a004085.
[38] K. Pawlowski, A. Bierzynski, A. Godzik, Structural diversity in a family of homologous proteins, J. Mol. Biol. 258 (1996) 349–366.
[39] L. Mouawad, A. Isvoran, E. Quiniou, C.T. Craescu, What determines the degree of compactness of a calcium-binding protein? FEBS J. 276 (2009) 1082–1093.
[40] http://www.rcsb.org/pdb/home/home.do.
[41] http://www.pymol.org/.
[42] H.J.C. Berendsen, D. van der Spoel, R. Van Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, Comput. Phys. Commun. 91 (1995) 43–56.
[43] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (2005) 1701–1718.
[44] W.F. van Gunsteren, S.R. Billeter, A.A. Eising, P.H. Hünenberger, P. Krüger, A.E. Mark, W.R.P. Scott, I.G. Tironi, Biomolecular Simulation: The GROMOS 96 Manual and User Guide, vdf Hochschulverlag AG an der ETH Zürich, Zürich, 1996.
[45] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, Interaction models for water in relation to protein hydration, in: B. Pullman (Ed.), Intermolecular Forces, D. Reidel Publishing Company, Dordrecht, 1981, pp. 331–342.
[46] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[47] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[48] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577–8593.
[49] D. van der Spoel, E. Lindahl, B. Hess, A.R. van Buuren, E. Apol, P.J. Meulenhoff, D.P. Tieleman, A. Sijbers, K.A. Feenstra, R. van Drunen, Gromacs User Manual version 4.0, The GROMACS development team, Groningen, The Netherlands, 2006.
[50] http://www.uniprot.org/.
[51] J. Mongan, Interactive essential dynamics, J. Comput. Aided Mol. Des. 18 (2004) 433–436.
[52] C. Yang, G.S. Jas, K. Kuczera, Structure, dynamics and interaction with kinase targets: computer simulations of calmodulin, Biochim. Biophys. Acta Proteins Proteomics 1697 (2004) 289–300.
[53] C. Hu, J.W. Fang, R.T. Borchardt, R.L. Schowen, K. Kuczera, Molecular dynamics simulations of domain motions of substrate-free S-adenosyl-L-homocysteine hydrolase in solution, Proteins 71 (2008) 131–143.
[54] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.
[55] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Clustal W and clustal X version 2.0, Bioinformatics 23 (2007) 2947–2948.
[56] F. Armougom, S. Moretti, O. Poirot, S. Audic, P. Dumas, B. Schaeli, V. Keduas, C. Notredame, Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee, Nucleic Acids Res. 34 (2006) W604–W608.
[57] http://www.mbio.ncsu.edu/BioEdit/bioedit.html.
[58] G. Perriere, M. Gouy, WWW-query: an on-line retrieval system for biological sequence banks, Biochimie 78 (1996) 364–369.
[59] M. Hilbert, G. Bohm, R. Jaenicke, Structural relationships of homologous proteins as a fundamental principle in homology modeling, Proteins 17 (1993) 138–151.
[60] M.A. Marti-Renom, A.C. Stuart, A. Fiser, R. Sanchez, F. Melo, A. Sali, Comparative protein structure modeling of genes and genomes, Annu. Rev. Biophys. Biomol. Struct. 29 (2000) 291–325.
[61] N. Eswar, B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M.Y. Shen, U. Pieper, A. Sali, Comparative protein structure modeling using MODELLER Chapter 2, Curr. Protoc. Protein Sci. (2007) 9 Unit 2.
[62] R. Luthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, Nature 356 (1992) 83–85.
[63] G. Barbato, M. Ikura, L.E. Kay, R.W. Pastor, A. Bax, Backbone dynamics of calmodulin studied by 15 N relaxation using inverse detected two-dimensional NMR spectroscopy: The central helix is flexible, Biochemistry 31 (1992) 5269–5278.
[64] J.J. Chou, S. Li, C.B. Klee, A. Bax, Solution structure of Ca2+-calmodulin reveals flexible hand-like properties of its domains, Nat. Struct. Biol. 8 (2001) 990–997.
[65] H. Li, K. Yokota, K. Satou, Contractive conformational change of calmodulin caused by the central linker: a molecular dynamics simulation study, in: F. Valafar, H. Valafar (Eds.), METMBS '05: Proceedings of the 2005 International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences, 2005, pp. 225–230.
[66] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, J. Mol. Biol. 157 (1982) 105–132.
[67] A.A. Chmiel, M. Radlinska, S.D. Pawlak, D. Krowarsch, J.M. Bujnicki, K.J. Skowronek, A theoretical model of restriction endonuclease NlaIV in complex with DNA, predicted by fold recognition and validated by site-directed mutagenesis and circular dichroism spectroscopy, Protein Eng. Des. Sel. 18 (2005) 181–189.
[68] A.A. Chmiel, J.M. Bujnicki, K.J. Skowronek, A homology model of restriction endonuclease SfiI in complex with DNA, BMC Struct. Biol. 5 (2005) 2.
[69] K.L. Tkaczuk, A. Obarska, J.M. Bujnicki, Molecular phylogenetics and comparative modeling of HEN1, a methyltransferase involved in plant microRNA biogenesis, BMC Evol. Biol. 6 (2006) 6.
[70] J. Li, Z.G. Jia, W.C. Zhou, Q. Wei, Calcineurin regulatory subunit B is a unique calcium sensor that regulates calcineurin in both calcium-dependent and calcium-independent manner, Proteins 77 (2009) 612–623.
[71] M. Ikura, J.B. Ames, Genetic polymorphism and protein conformational plasticity in the calmodulin superfamily: two ways to promote multifunctionality, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 1159–1164.
[72] J.B. Ames, K. Levay, J.N. Wingard, J.D. Lusin, V.Z. Slepak, Structural basis for calcium-induced inhibition of rhodopsin kinase by recoverin, J. Biol. Chem. 281 (2006) 37237–37245.
[73] M. Mishima, S. Wakabayashi, C. Kojima, Solution structure of the cytoplasmic region of Na+/H+ exchanger 1 complexed with essential cofactor calcineurin B homologous protein 1, J. Biol. Chem. 282 (2007) 2741–2751.
[74] T. Strahl, I.G. Huttner, J.D. Lusin, M. Osawa, D. King, J. Thorner, J.B. Ames, Structural insights into activation of phosphatidylinositol 4-kinase (Pik1) by yeast frequenin (Frq1), J. Biol. Chem. 282 (2007) 30949–30959.
[75] W. Zhou, Y. Qian, K. Kunjilwar, P.J. Pfaffinger, S. Choe, Structural insights into the functional interaction of KChIP1 with Shal-type K+ channels, Neuron 41 (2004) 573–586.
[76] M. Pioletti, F. Findeisen, G.L. Hura, D.L. Minor, Three-dimensional structure of the KChIP1–Kv4.3 T1 complex reveals a cross-shaped octamer, Nat. Struct. Mol. Biol. 13 (2006) 987–995.
[77] H.Y. Wang, Y. Yan, Q. Liu, Y.H. Huang, Y. Shen, L.J. Chen, Y. Chen, Q.Y. Yang, Q. Hao, K.W. Wang, J.J. Chai, Structural basis for modulation of Kv4 K+ channels by auxiliary KChIP subunits, Nat. Neurosci. 10 (2007) 32–39.
[78] S. Takeda, A. Yamashita, K. Maeda, Y. Maeda, Structure of the core domain of human cardiac troponin in the Ca2+-saturated form, Nature 424 (2003) 35–41.
[79] M.V. Vinogradova, D.B. Stone, G.G. Malanina, C. Karatzaferi, R. Cooke, R.A. Mendelson, R.J. Fletterick, Ca2+-regulated structural changes in troponin, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 5038–5043.