A model for the calmodulin–peptide complex based on the troponin C crystal packing and its similarity to the NMR structure of the calmodulin–myosin light chain kinase peptide complex

C.Y. SEKHARUDU AND M. SUNDARALINGAM

Laboratory of Biological Macromolecular Structure, Department of Chemistry and Ohio State Biotechnology Center, The Ohio State University, Columbus, Ohio 43210

(RECEIVED August 28, 1992; REVISED MANUSCRIPT RECEIVED December 1, 1992)

Abstract

In the crystal structure of troponin C, the holo C-domain is bound in a head-to-tail fashion to the A-helix of the apo N-domain of a symmetry-related molecule. Using this interaction, we have proposed a model for the calmodulin–peptide complex. We find that the interaction of the C-domain with the A-helix is similar to that observed in the NMR structure of the calmodulin–myosin light chain kinase (MLCK) peptide complex. This similarity in binding has enabled us to make a precise sequence alignment of the target peptides in the calmodulin-binding cleft and to rationalize the amino acid sequence-dependent binding strengths of various peptides. Our model differs from that proposed by Strynadka and James (*Proteins Struct. Funct. Genet.* 7, 234–248, 1990) in that the peptides are rotated by 100° in the calmodulin binding cleft.

Keywords: calcium binding; calmodulin; crystal structure; domains; MLCK peptide complex; troponin C

Calmodulin (CaM) is a calcium-dependent ubiquitous regulatory protein that activates numerous structurally and functionally different enzymes (Cheung, 1980; Klee & Vanaman, 1982; Means et al., 1982). The activation of the target enzymes is triggered by the binding of CaM to a 20-residue peptide segment on the enzymes (Edelman et al., 1985; Lukas et al., 1986; Sikela & Hawn, 1987), and the peptide segments exhibit essentially the same affinity as their parent enzymes. Even though these peptides bind to CaM with high affinity, they do not possess strong sequence similarities (O'Neil & DeGrado, 1990). However, they have a high probability of being basic amphiphilic helices. Synthetic α -helical peptides also bind to CaM with a high affinity (DeGrado et al., 1987; O'Neil et al., 1987). Recently, a structure for the complex of CaM with myosin light chain kinase (MLCK) peptide was presented based on an NMR study of interproton distances between the peptide and the N- and C-domains of CaM (Ikura et al., 1992).

We had noticed that in the crystal structure of chicken skeletal muscle troponin C (TnC), which is homologous

to CaM, the hydrophobic patch of the C-domain interacts with the A-helix of the N-domain of a neighboring molecule (Satyshur et al., 1988; Satyshur et al., unpubl.). The size and shape complementarity of the cylindrical A-helix surface and the crescent-shaped C-domain surface allows them to fit snugly to form a stable “intermolecular complex” (Fig. 1a). The recent report of the NMR structure of MLCK peptide complex of CaM (Ikura et al., 1992) prompted us to compare it with the interactions of the TnC–A-helix in the crystal structure. We find that the mode of binding of the A-helix to the C-domain of TnC is similar to that observed for the MLCK peptide binding to the C-domain of CaM.

Results and discussion

TnC (Herzberg & James, 1988; Satyshur et al., 1988) and CaM (Kretsinger et al., 1986; Babu et al., 1988; Taylor et al., 1991; Rao et al., 1993) belong to the TnC superfamily (Castellani et al., 1980). The X-ray structures show that both proteins are dumbbell shaped with the N- and C-domains spanned by a long central helix. Each domain contains a pair of helix–loop–helix (EF-hand) calcium-binding motifs. The two calcium-binding motifs within each domain are related by a pseudodyad and are held to-

Reprint requests to: M. Sundaralingam, Laboratory of Biological Macromolecular Structure, Department of Chemistry and Ohio State Biotechnology Center, The Ohio State University, Columbus, Ohio 43210.

Fig. 1. **a:** The head-to-tail crystal packing interactions in the chicken skeletal muscle troponin C (TnC). The C-domain of TnC (green) "grabs" the A-helix of a neighboring TnC molecule (yellow). The convex surface of the A-helix snugly fits in the concave surface of the C-domain of TnC. **b:** Ribbon diagram of the A-helix and the C-domain. The A-helix has been "peeled" off from the C-domain and rotated about 180° around the helical axis. The α -carbon atoms of the helix ridges that form the four sites are indicated by different colors: CA1, red; CA2, green; CA3, yellow; and CA4, blue. The complementary sites (CA1', CA2', CA3', and CA4') on the C-domain are indicated by the same colors.

gether by a mini β -sheet. The amino acid sequences of the C-domains of CaM and TnC are highly homologous with no deletions or insertions (Table 1), and the structures are superposable with a root mean square deviation (rmsd) of 1.2 Å. Thus a knowledge of the interactions between TnC and the A-helix from the high-resolution X-ray structure (Satyshur et al., 1988, unpubl.) provides further insights into the interactions between the calcium-dependent regulatory CaM and the target enzymes.

Complementary interaction of the four ridges of the A-helix with the four helices of the C-domain

The A-helix interacts with the C-domain of TnC with its carboxyl-terminus close to the E-helix (long central helix) and the amino-terminus close to the G-helix (Fig. 1b). The A-helix makes an angle of 103° with the long central helix and is nearly perpendicular to it. The binding of the A-helix to the C-domain of TnC is essentially through contacts at four A-helix sites (CA1, CA2, CA3, and CA4), which are complementary to the four C-domain receptor sites CA1', CA2', CA3', and CA4', respectively (Fig. 1b). The residues in the four A-helix sites are: CA1, Glu 16 and Glu 17; CA2, Glu 21; CA3, Ala 20, Lys 23, and Ala 24; CA4, Met 28 and Phe 29. The ridge formed

by the residues of the CA4 site (Met 28 and Phe 29) interacts with the CA4' site of the C-domain comprising two adjacent grooves on the surface of the E-helix, where the first groove is formed by the negatively charged residues Glu 97 and Asp 100 and the second groove by the apolar residues Cys 101, Ile 104, and Phe 105. Although these two grooves are complementary to positively charged residues and apolar residues, respectively, on the CA4 site, the interaction is only with the apolar residues of the A-helix. The CA3 site of the A-helix is formed by the residues Ala 20, Lys 23, and Ala 24. The two small apolar residues Ala 20 and Ala 24 form a ridge and interact with the residues Ile 121, Leu 122, Ala 124, and Thr 125 on the groove of the F-helix (CA3'), while the remaining Lys 23 of the CA3 site is in proximity to the residues Gly 126, Glu 127, and His 128 of the F-G linker (CA3'). Thus the CA3' receptor site on the F-helix is a shallow groove that is highly complementary to small apolar residues on the peptide helix, while the F-G linker can interact with almost any type of residue. The CA2 site residue Glu 21 of the A-helix interacts with the groove of the H-helix (CA2') formed by the residues Phe 154, Met 157, and Met 158. While CA2' is interacting with a negatively charged residue, more favorable interaction will be with an apolar residue. The negatively charged CA1 site, residues Glu 16

Table 1. The C-domain amino acid sequences of calmodulin (top row) and troponin (bottom row)^a

										X	Y				Z		-Y		-X	
E82	E83	E84	I85	R86	E87	A88	F89	R90	V91	F92	D93	K94	D95	G96	N97	G98	T99	I00	S01	
E95	E96	E97	L98	A99	D00	C01	F02	R03	I04	F05	D06	K07	D08	A09	D10	G11	T12	I13	D14	
-Z																				
A02	A03	E04	L05	R06	H07	V08	M09	T10	N11	L12	G13	E14	K15	L15	T17	D18	E19	E20	V21	
I15	E16	E17	L18	G19	E20	I21	L22	R23	A24	T25	G26	E27	H28	V29	T30	E31	E32	D33	I34	
										X	Y		Z		-Y		-X		-Z	
A22	D23	M24	I25	R26	E27	A28	N29	I30	D31	G32	D33	G34	Q35	V36	N37	Y38	E39	E40	F41	
E35	D36	L37	M38	K39	D40	S41	D42	K43	N44	N45	D46	G47	R48	I49	D50	F51	F52	E53	E54	
V42	G43	M44	M45	T46	A47	K48														
L55	K56	M57	M58	E59	G60	V61	N62													

^a The residues that interact with the myosin light chain kinase peptide and the A-helix are boxed. The calcium-binding ligands (X, Y, Z and -Y, -X, -Z) are also indicated.

and Glu 17, are in proximity to the CA1' residues Asp 133, Ile 134, Asp 136, and Leu 137 in the groove of the G-helix. Thus, the ridges on the A-helix are essentially complementary to the grooves on the four C-domain TnC receptor helices (E, F, G, and H). The predominant interactions between the four sites of the A-helix and the C-domain are van der Waals in nature. Because not all the interactions between the A-helix and the C-domain are complementary, the binding to the A-helix is expected to be weak.

Similarity of the A-helix binding to MLCK binding

In the recent NMR structure of the CaM-MLCK peptide complex, the disposition of the MLCK helix in the CaM C-domain is similar to the A-helix in the C-domain of TnC, in that the C-terminus of the MLCK peptide is close to the E-helix and the N-terminus is close to the G-helix. When Trp 4 of the MLCK peptide is placed in register with Glu 17 of the A-helix of TnC, the resulting interactions are very similar to those observed in the NMR study (Fig. 2). In the MLCK peptide, Arg 3 and Trp 4 mimic the site CA1; Asn 7 mimics site CA2; Phe 8, Ala 10, and Val 11 mimic site CA3; and the Lys 14 and Lys 15 mimic the site CA4 of the A-helix. Thus the amphiphilic helical peptides bind in a similar fashion to the C-domains of both TnC and CaM (Fig. 2), and our high-resolution X-ray structure illustrates the interactions between a helical peptide and the calcium-binding EF-hand pair.

Affinities of CaM to various peptide segments

Using our model (Fig. 3), the amino acid sequence-dependent binding strengths of helical peptides to CaM can be rationalized. The amino acid sequences of the two peptides (A-helix and MLCK) were aligned based on the similarity of their interactions with the C-domains of TnC and CaM (Table 2). Except for three locations, the amino acid side chains in all other locations surprisingly differ in their nature (Fig. 2). This variation reveals that the CaM/TnC concave surface can recognize helical peptides with drastically different sequences. This perhaps is important for CaM to bind to a variety of target proteins by recognizing a convex surface characteristic of amphiphilic helical motifs. Even though a helical motif is used as scaffold, it is the individual amino acid side chains emanating from the scaffold that recognize the protein surface. As discussed above, this recognition takes place through the four receptor sites (CA1', CA2', CA3', and CA4') on the surface of the C-domain of TnC/CaM, accommodating the residues on the four sites (CA1, CA2, CA3, and CA4) emanating from the helix scaffold. Each of the receptor sites on the CaM is in register with the pitch of an α -helix, thereby accommodating the side chains that are interspersed 3/4 residues apart along the helix. For a helix to bind to the C-domain with high affinity, the CA2 and CA3 sites in the middle should be apolar in nature, whereas the CA1 and CA4 sites at the ends should be positively charged. Similarly for the N-

Fig. 2. a: Schematic representation of the interactions between the A-helix and the C-domain of TnC. The TnC residues within 5.0 Å of the A-helix residues are listed. Also, the C-domain helices with the highest number of favorable van der Waals contacts to the four A-helix sites are indicated. **b:** Schematic representation of the interactions of the myosin light chain kinase (MLCK) peptide with the C-domain of calmodulin (CaM) constructed using the A-helix-TnC complex as a guide. Open circles denote the residues that are common between the current model and the NMR structure (Ikura et al., 1992). The "x"s denote the common residues between the model and the X-ray structure (Meador et al., 1992). Notice that the elongated Lys 23 of the A-helix (a) has a greater number of neighbors compared to the corresponding but shorter Ala 10 of the MLCK peptide (b). Conversely, Glu 17 of the A-helix has fewer neighbors compared to the bulkier Trp 4 of the MLCK peptide.

domain to bind strongly, the corresponding NA1...NA4 sites should also exhibit two apolar sites in the middle flanked by positively charged sites at the ends. While the first and last receptor sites on CaM can accommodate a variety of amino acid residues with varying size and

charge complementarity, the middle two sites can accommodate only medium-sized apolar residues. Because the amino acid residues of the receptor sites in CaM are remarkably conserved, the affinities of the various helical peptides to CaM arise primarily from their amino acid se-

Fig. 3. a: CaM with the peptide bound to both the N- and C-domains. Notice that the pseudo twofolds within the N- and C-domain are nearly perpendicular. **b:** The structure of CaM after the two bound helices were superposed as described in the Materials and methods. Notice the pseudo 222 symmetry of this molecule. The twofold along the helix axis does not intersect the other two twofold axes: (1) the twofold that relates the EF-hands within the domain and (2) the twofold that relates the N- and C-domains, which passes through the center of the helix and is perpendicular to the helix axis.

Table 2. Primary sequences of some peptide segments that bind to calmodulin (CaM)^a

Peptide	Site														Complementarity to CaM ^b				Remarks	K _d (nM)
	CA1		CA3			CA2		CA3			CA4			CA1'	CA2'	CA3'	CA4'			
A-helix of TnC	E	E	M	I	A	E	F	K	A	A	F	D	M	F	B	F	G	F	Weak	—
M13	R	W	K	K	N	F	I	A	A	V	S	A	A	K	G	F	G	G	Strong	—
Mastoparan	N	L	K	A	L	A	A	L	A	K	K	I	L		G	G	G	F	Strong	0.3
Mastoparan X	N	W	K	G	I	A	A	M	A	K	K	L	L		G	G	G	F	Strong	0.3
Mellitin	A	V	L	K	V	L	T	T	G	L	P	A	L		G	G	F	G	Strong	3.5
P. mastoparan	D	W	K	K	I	G	Q	H	I	L	S	V	L		G	F	G	G	Strong	3.5
VIP	A	V	K	K	Y	L	N	S	I	L	N				G	F	F	B	Weak	50.0
GIP	R	Q	Q	D	F	V	N	W	L	L	A	Q	Q		G	F	B	F	Weak	70.0

^a Boxed residues indicate the peptide residues involved in binding to calmodulin.^b B, bad; G, good; F, favorable.

quence. For the peptides to exhibit tightest binding, all the sites on the helix (CA1...CA4 and NA1...NA4) that interact with CaM should have surface and charge complementarity, and lack of such complementarity will be expected to result in decreased affinity for CaM.

The helical peptides mastoparan and mellitin, which exhibit very high affinity for CaM, have all CA1...CA4 sites highly complementary to the CA1'...CA4' receptor sites (Table 2). On the other hand, VIP (vasoactive intestinal peptide) and GIP (gastric inhibitory peptide) peptides have only one good site (CA1) that can favorably and optimally interact with CaM, thus explaining their lower binding affinity. The sequence of the TnC A-helix belongs to the low affinity class of peptides, whereas that of MLCK belongs to the high affinity class. This is consistent with the observations that the hydrophobic region of the synthetic peptide helices (sites CA2/NA2 and CA3/NA3), interspersed with positively charged residues (sites CA1/NA1 and CA4/NA4), binds strongly to CaM (O'Neil & DeGrado, 1989, 1990).

Note added in revision

After this paper was submitted, an X-ray structure of CaM-sMLCK (smooth muscle) peptide complex at 2.4 Å resolution was reported (Meador et al., 1992). The overall structure of the CaM-peptide complex is similar to our model. The interactions between the peptide and CaM in our model are very similar to those observed in the X-ray structure (Fig. 2b). Trp 4 of MLCK that has the largest interacting surface with CaM in our model interacts with the same residues as the corresponding Trp 5 of sMLCK in the X-ray structure. The sMLCK peptide has two significant sequence changes; the two glycines at position 9 and 16 are replaced by Phe 8 and Asn 15 in MLCK, and these changes are reflected in the interactions. The short Gly residues have fewer neighbors than their bulkier counterparts.

Materials and methods

The coordinates of our 2.0-Å resolution X-ray structure of chicken skeletal muscle troponin C (Satyshur et al., 1988; PDB entry 4TNC) were used to construct the CaM-peptide model. The atomic positions of the N-domain A-helix interacting with the C-domain were generated by applying the $(-y, x - y, \frac{2}{3} + z)$ transformation. The A-helix in the C-domain was replaced by the MLCK peptide by extending the helix in the standard α -helical conformation and appropriately mutating the side chains. Because the MLCK peptide contains some longer and bulkier side chains, their steric contacts were alleviated by rotating the $\chi^1 \dots \chi^n$ dihedral angles of the side chains. The resulting TnC-MLCK peptide complex was compared with the NMR structure of the CaM-MLCK peptide complex (Fig. 2).

The C-domain of TnC is structurally homologous to both the N- and C-domains of CaM. Using the MLCK-TnC structure as a template, a model for CaM with the MLCK peptide bound to both the N-domain as well as the C-domain was constructed (Fig. 3a). The transformations required to superimpose the C-domain of TnC on the N-domain as well as the C-domain of the CaM crystal structure (Babu et al., 1988; PDB entry 3CLN) were determined. By applying these transformations to the TnC-bound MLCK peptide, the coordinates of the MLCK peptide bound both to the N- and the C-domains of CaM were obtained. The model where the N- and C-domain wrap around the same peptide helix (Fig. 3b) was constructed as follows. The transformation relating the α -carbon atoms of residues 1-20 of the MLCK peptide bound to the N-domain and the α -carbon atoms of residues 20-1 of the MLCK bound to the C-domain of CaM was determined and applied to the N-domain residues. This brings the two domains closer together to interact with the same helix. In this model, the N-domain and the C-domain structures are identical to the X-ray structure.

Using the conformation of the linker regions in the N- and C-domains of CaM as a guide, the linker segment bridging the N- and C-domains was modeled. In the X-ray structure (Babu et al., 1988), the pseudo twofold axis of the C-domain is nearly perpendicular to that of the N-domain, but in the folded model they are nearly parallel. Our folded model is very similar to that of the NMR CaM-MLCK peptide complex. Several models with varying amounts of bend in the central helix of CaM have been proposed (Persechini & Kretsinger, 1988a,b; O'Neil & Degrado, 1990; Strynadka & James, 1990; Means et al., 1991) with the bound peptide helix oriented nearly parallel to or perpendicular to the central helix. Our model differs from an earlier model based on the A-helix interactions with the C-domain of turkey TnC (Strynadka & James, 1990) where the tryptophan residue of the target peptide was placed in register with the Met 18 of the A-helix in contrast to our model where it is in register with Glu 17 involving an azimuthal rotation of 100° of the peptide helix bound to the C-domain. Registry with Met 18 partially exposes the tryptophan to solvent and significantly reduces its interactions with the C-domain.

Acknowledgments

We acknowledge the College of Mathematical and Physical Sciences, The Ohio State University, for their continued support. This work was supported by an Ohio Regents endowment to M.S. and in part by NIH grant AR34139.

References

- Babu, Y.S., Bugg, C.E., & Cook, W.J. (1988). Structure of calmodulin refined at 2.2 Å resolution. *J. Mol. Biol.* 204, 191–204.
- Castellani, L., Morris, E.P., & O'Brien, E.J. (1980). Calmodulin as model for troponin C. *Biochem. Biophys. Res. Commun.* 96, 558–565.
- Cheung, W.Y. (1980). Calmodulin plays a pivotal role in cellular regulation. *Science* 207, 19–27.
- DeGrado, W.F., Erickson-Vitanen, S., Wolfe, H.R., Jr., & O'Neil, K.T. (1987). Predicted calmodulin binding sequence in the γ -subunit of phosphorylase b kinase. *Proteins Struct. Funct. Genet.* 2, 20–33.
- Edelman, A.M., Takio, K., Blumenthal, D.K., Hansen, R.S., Walsh, K.A., Titani, K., & Krebs, E.G. (1985). Characterization of calmodulin-binding and catalytic domain in skeletal muscle myosin light chain kinase. *J. Biol. Chem.* 260, 11275–11285.
- Herzberg, O. & James, M.N.G. (1988). Refined crystal structure of troponin C from turkey skeletal muscle at 2.0 Å resolution. *J. Mol. Biol.* 203, 761–779.
- Ikura, M., Clore, G.M., Gronenborn, M., Zhu, G., Klee, C.B., & Bax, A. (1992). Solution structure of a calmodulin-target peptide complex by multidimensional NMR. *Science* 256, 632–638.
- Klee, C.B. & Vanaman, T.C. (1982). Calmodulin. *Adv. Protein Chem.* 35, 213–321.
- Kretsinger, R.H., Rudnick, S.E., & Weissman, L.J. (1986). Crystal structure of calmodulin. *J. Inorg. Biochem.* 28, 289–302.
- Lukas, T.J., Burgess, W.H., Prendergast, F.G., Lau, W., & Watterson, D.M. (1986). Calmodulin binding domains: Characterization of phosphorylation and calmodulin binding sites from myosin light chain kinase. *Biochemistry* 25, 1458–1464.
- Meador, W.E., Means, A.R., & Quirocho, F.A. (1992). Target enzyme regulation by calmodulin: 2.4 Å structure of a calmodulin-peptide complex. *Science* 257, 1251–1255.
- Means, A.R., Tash, J.S., & Chafouleas, J.G. (1982). Physiological implications on the presence, distribution and regulation of calmodulin in eukaryotic cells. *Physiol. Rev.* 62, 1–39.
- Means, A.R., VanBerkum, M.F.A., Bagchi, I., Lu, K.P., & Rasmussen, C.D. (1991). Regulatory functions of calmodulin. *Pharm. Ther.* 50, 255–270.
- O'Neil, K.T. & DeGrado, W.F. (1989). The interaction of calmodulin with fluorescent and photoreactive model peptides: Evidence for a short interdomain separation. *Proteins Struct. Funct. Genet.* 6, 284–293.
- O'Neil, K.T. & DeGrado, W.F. (1990). How calmodulin binds its targets: Sequence independent recognition of amphiphilic α -helices. *Trends Biochem. Sci.* 56, 59–64.
- O'Neil, K.T., Wolfe, H.R., Jr., Erickson-Vitanen, S., & DeGrado, W.F. (1987). Fluorescence properties of calmodulin binding peptides reflect α -helical periodicity. *Science* 236, 1454–1456.
- Persechini, A. & Kretsinger, R.H. (1988a). Toward a model of the calmodulin-myosin light chain complex: Implications for calmodulin function. *J. Cardiovasc. Pharmacol.* 12, 501–512.
- Persechini, A. & Kretsinger, R.H. (1988b). The central helix of calmodulin functions as a flexible tether. *J. Biol. Chem.* 263, 12175–12178.
- Rao, S.T., Wu, S., Satyshur, K., Ling, K.-Y., Kung, C., & Sundaralingam, M. (1993). Structure of *Paramecium tetraurelia* calmodulin at 1.8 Å resolution. *Protein Sci.* 2, 436–447.
- Satyshur, K.A., Rao, S.T., Pyzalska, D., Drendel, W., Greaser, M., & Sundaralingam, M. (1988). Refined structure of chicken skeletal muscle troponin C in the two-calcium state at 2-Å resolution. *J. Biol. Chem.* 263, 1628–1647.
- Sikela, J.M. & Hawn, W.E. (1987). Screening an expression library with a ligand probe: Isolation and sequence of cDNA corresponding to a brain calmodulin-binding protein. *Proc. Natl. Acad. Sci. USA* 84, 3038–3042.
- Strynadka, N.J.C. & James, M.N.G. (1990). Model for the interaction of amphiphilic helices with troponin C and calmodulin. *Proteins Struct. Funct. Genet.* 7, 234–248.
- Taylor, D.A., Sack, J.S., Maune, J.F., Beckingham, K., & Quirocho, F.A. (1991). Structure of a recombinant calmodulin from *Drosophila melanogaster* at 2.2-Å resolution. *J. Biol. Chem.* 266, 21375–21380.