# The evolving model of calmodulin structure, function and activation

Recent high-resolution crystal and solution structures have answered many long-standing questions about calmodulin and its various conformational states. However, there is still much to learn.

**Structure** 15 January 1995, **3**:7–11

The ubiquitous eukaryotic protein calmodulin (CaM) plays a major role in the $Ca^{2+}$-dependent regulation of wide variety of cellular events. The ~150 amino acid long sequence of CaM from different species is highly conserved and the CaMs constitute their own subfamily that belongs to the family containing the EF-hand structural motif [1]. The regulatory role of CaM is intimately connected with the prevailing intracellular concentration of free $Ca^{2+}$ ions. In a resting cell, the intracellular $Ca^{2+}$ concentration is of the order of 50–100 nM but as a result of an external stimulus the level may transiently rise to 1–10 μM. The $Ca^{2+}$ affinity of CaM is such that the $Ca^{2+}$ binding sites (of which there are usually four, although in a few cases only three are functional) will be largely unoccupied at the resting cell $Ca^{2+}$ concentrations, but highly occupied at the higher levels after stimulus. As a consequence of $Ca^{2+}$ binding, CaM undergoes a conformational change which renders it active and ready to bind to and influence the activity of more than 20 enzymes or structural proteins.

The first amino acid sequences of vertebrate CaMs were presented around 1978-1980. On the basis of the sequence and the only crystal structure of an EF-hand $Ca^{2+}$-binding protein available at that time, that of parvalbumin, several attempts were made to predict the tertiary structure of $Ca^{2+}$–CaM. Many of these early efforts have recently been reviewed by Kretsinger [2] and will not be reiterated here. It suffices to say that most proposals fell short of the unique structure that was revealed by the first successful X-ray diffraction study of rat calmodulin published in 1985 [3], and later refined to 2.2 Å resolution [4]. In the crystal, the $Ca^{2+}$-loaded form of CaM was found to be shaped like a dumbbell, the 'end-weights' of which were the amino-terminal and carboxy-terminal domains, each of which bound two $Ca^{2+}$ ions. The intervening 'handle' is a 28 amino acid long α-helix extending from residues 65 to 92. Shortly before the structure of $Ca^{2+}$–CaM was announced, the crystal structure of the homologous EF-hand protein, troponin C, was reported [5,6] and was also found to have a dumbbell-like structure.

Since the pioneering structural studies on $Ca^{2+}$–CaM in 1985, three additional crystal structures have become available, those of human [7], *Drosophila* [8], and *Paramecium* [9] $Ca^{2+}$–CaM. The structure of recombinant human $Ca^{2+}$–CaM has been refined to a very high resolution (1.7 Å) and is shown in Fig. 1.

The dumbbell shapes of troponin C and $Ca^{2+}$–CaM came as a surprise [10]. The structure of $Ca^{2+}$–CaM seemed difficult to reconcile with biochemical data on the interaction of $Ca^{2+}$–CaM with its many target proteins. The stability of the long central helix was called into question from the very start [10], as was that in troponin C [5]. Temperature (B-) factors, which are

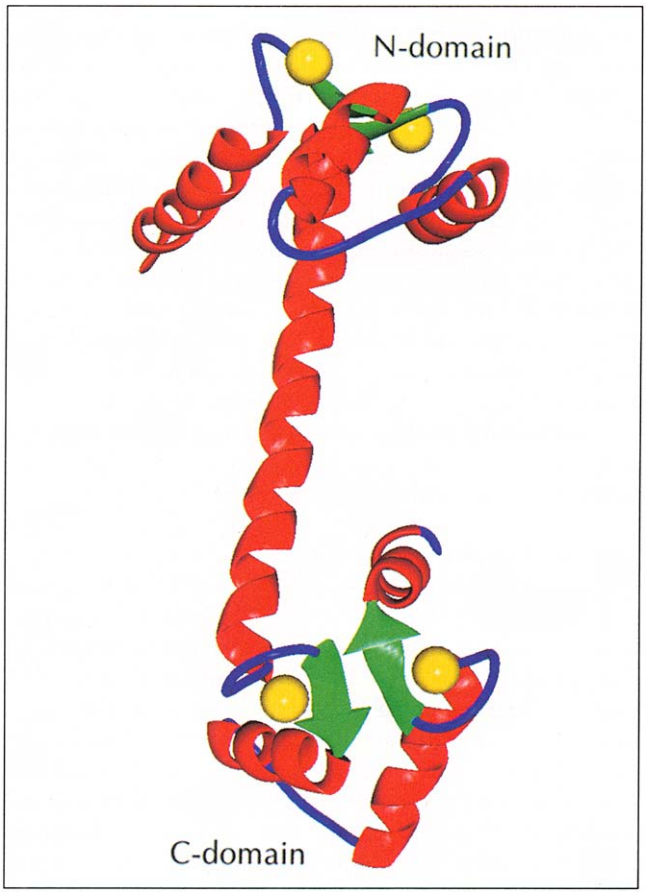


**Fig. 1.** A ribbon representation of the 1.7 Å X-ray crystal structure of human $Ca^{2+}$–CaM [7]. The color scheme is as follows: α-helical regions are red, β-strands are green, loops are blue and the calcium atoms are yellow.

 ISSN 0969-2126

proportioned to the mobility of the main-chain and side-chains, are high not only for the last few amino-terminal and carboxy-terminal residues but also in the central helix, with a maximum around positions 78–82 [7–9]. While the central helix is straight in the structures reported by Chattopadayay *et al.* and Rao *et al.* [7,9], there is a kink at Asp80 in the refined structure of Babu *et al.* [4] and in the *Drosophila* structure there are two small kinks at residues Lys75 and Ile85 [8]. It should be noted that the B-factors in the *Drosophila* $Ca^{2+}$–CaM are lower in the central helix than in the other structures.

Shortly after the presentation of the $Ca^{2+}$–CaM crystal structure, small-angle X-ray scattering and neutron scattering were employed in order to obtain information about the solution structure of the molecule [11,12]. The solution form of $Ca^{2+}$–CaM that emerged from these experiments is more compact than the crystal structure. Heidorn and Trewhella [12] proposed that the central helix could easily be distorted and that the observed time-averaged solution structure was bent in the middle so as to bring the two globular domains some 10 Å closer than in the crystal structure. Other biophysical and biochemical studies gave parallel results (for a review, see [13]).

High resolution information on the solution structure of $Ca^{2+}$–CaM was not obtained until the ground-breaking multidimensional NMR studies performed by Bax, Ikura and co-workers [14]. The secondary structure was found to be in general accordance with the crystal structure but the central linker between the amino- and carboxy-terminal domains was found to have a non-helical region at residues 78–81. Furthermore, $^{15}N$ relaxation studies indicated that this region had considerable flexibility [15]. The NMR relaxation experiments also gave rotational correlation times for the amino- and carboxy-terminal domains (7.1 ns and 6.3 ns, respectively) that were too small to correspond to a 16.7 kDa stiff dumbbell-shaped molecule. Based on this evidence, it seems clear that the solution structure of $Ca^{2+}$–CaM differs from the dumbbell shape found in the crystal form and is most likely to be highly flexible in the central helix which might more accurately be termed the central tether.

### Target binding and the tether effect

On moving from $Ca^{2+}$–CaM to $Ca^{2+}$–CaM–peptide complexes, there is greater agreement between the crystal and solution conformations. Indeed, it was recognized early on that a $Ca^{2+}$–CaM with a dumbbell shape would place the two halves of the binding site away from one another by a larger distance than the length of any of the target peptides. This fact led to other models for the $Ca^{2+}$–CaM–target complex in which the central helix was bent in order to allow the two sites to come together to interact with a peptide [16]. Evidence from small angle X-ray scattering studies [17] supported such a model of a more compact complex. Mutational studies, in which $Ca^{2+}$–CaMs with 1–4 amino acids deleted from the central helix retained binding activity [18], showed that not only did the length of the central helix have little importance, but also that the central helix did not control the orientation of the two domains relative to one another. A distinct demonstration of the flexibility of the central linker came with the three-dimensional NMR structure of $Ca^{2+}$-loaded CaM complexed with a 26 residue peptide comprising the CaM-binding domain of skeletal muscle myosin light chain kinase [19]. This structure showed that residues 74–81 are disordered in the complex (Fig. 2).

What role does this central tether play in the formation of the $Ca^{2+}$–CaM–peptide complexes? Many of the biophysical and biochemical properties of $Ca^{2+}$–CaM can be explained as a superposition of the corresponding properties of the amino- and carboxy-terminal domains.

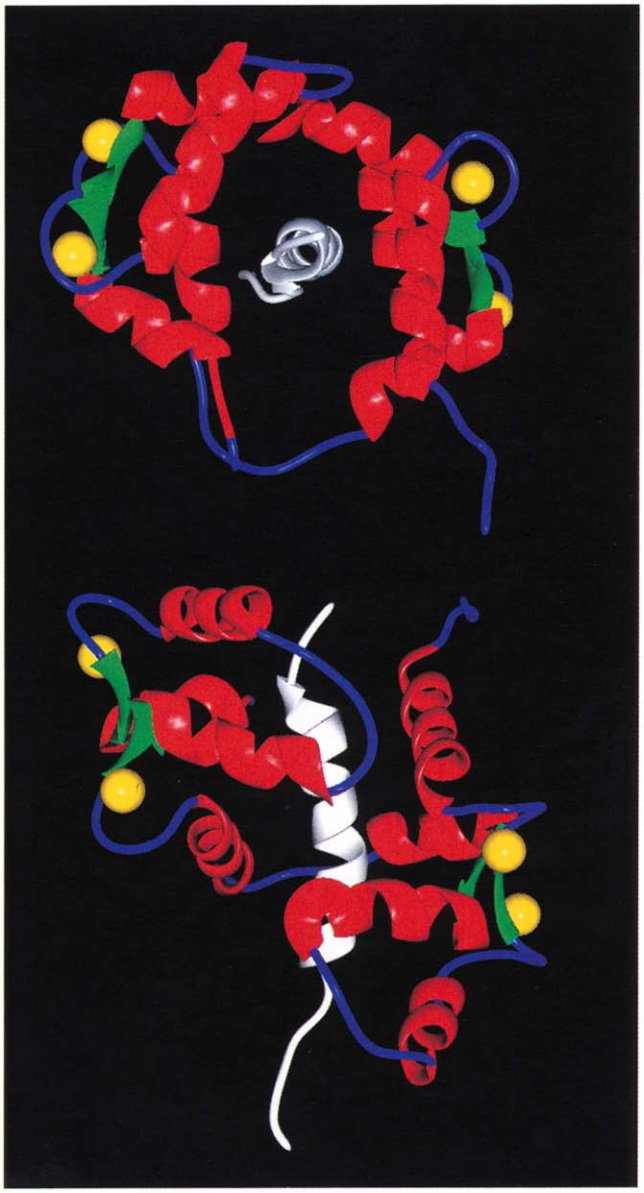

**Fig. 2.** Two views of a ribbon drawing of *Drosophila* $Ca^{2+}$–CaM in complex with rabbit skMLCK as determined by NMR [19]. The coloring is as in Fig. 1 with the exception that the bound skMLCK peptide is colored white.

However, target activation cannot. While high concentrations of the tryptic fragments $TR_1C$ (residues 1–77) and $TR_2C$ (residues 78–148) are able to bind to some target proteins, they by and large fail to activate most of them, unless present in several hundred-fold excess, and even then very poorly ([20] and references cited therein). Thus the presence of a central tether seems a prerequisite for activation.

One benefit of a flexible tether linking two hydrophobic binding sites may well be understood in terms of the 'chelate effect'. This term is usually applied to small molecule complexes to account for the fact that a bidentate ligand — that is, two ligand molecules linked together — tends to form more stable complexes than two monodentate ligands. This is primarily due to the fact that two monodentate ligand molecules have more translational and rotational entropy than one bidentate molecule. An alternative way of looking at this effect (which can give an estimate of the contribution of the tether to the binding) is by considering two globular domains connected with a completely flexible tether of length $a$. From the center of one of the domains, the other domain will be found within a volume, $V=4\pi a^3/3$. This means that the effective concentration, $C_{eff}$ of the second domain near the first will be $C_{eff}=(V\cdot N_{av})^{-1}$ , where $N_{av}$ is Avogadro's number. For example, if $a$=30 Å for $Ca^{2+}$–CaM, then $C_{eff}$=15 mM. At a typical intracellular CaM concentration of 1 μM (or concentration within the individual domains) we see that the intact tethered $Ca^{2+}$–CaM molecule has the upper hand, in that the formation of a complex with a target molecule is favored by a factor of ~$10^4$ over the formation of a complex of two separate domains with the target. The correspondingly higher free energy of binding of the intact $Ca^{2+}$–CaM molecule may be necessary to promote the activation of the target protein. The beneficial effect of the flexible tether will be even higher at lower concentrations while a 'break-even' situation is only reached at 10–20 mM concentrations — far above the physiological range. We should obviously not take this flexible tether model of $Ca^{2+}$–CaM too far and attribute all advantages to its mere presence as a linker as it also contributes interactions between side chains and target peptides. However, it indicates one important role that the tether plays.

### Target binding and molecular recognition

The binding of $Ca^{2+}$–CaM to its many target enzymes is a unique case in molecular recognition because it binds so many different target sequences and yet binds with such high affinity that dissociation constants lie in the nanomolar range. The target sequences are usually short (14–26 amino acids long), mostly α-helical, and made up of a high proportion of hydrophobic and basic amino acids, hence their nickname 'Baa helices' (basic, amphiphilic alpha-helices) (for review, see [21]).

The structures of three complexes have been determined to atomic resolution. One is a structure of a complex of $Ca^{2+}$–CaM with a peptide from skeletal myosin light-chain kinase (skMLCK) determined by NMR [19]. The other two structures, which have been solved by crystallography, are a complex with a smooth muscle myosin light-chain kinase (smMLCK) peptide [22] and a complex with a CaM-dependent protein kinase II (CaMKII) peptide [23]. While each of these peptides has a different sequence, all show the expected predominance of hydrophobic and basic residues. As shown in Fig. 3, if we align the segments of the peptides shown to interact with the same residues from $Ca^{2+}$–CaM, some similarities are apparent, notably that each has patches of basic residues near the beginning and end, as well as patches of hydrophobic residues distributed throughout its sequence. Complementary patches of hydrophobic and acidic residues on $Ca^{2+}$–CaM interact with these to stabilize the complex. The complexes are not, of course, identical with one another. In order to accommodate different peptides, not only do individual side chains adjust to conform to a particular peptide but also the entire orientation of the two domains is altered, thanks to the flexibility of the central tether.

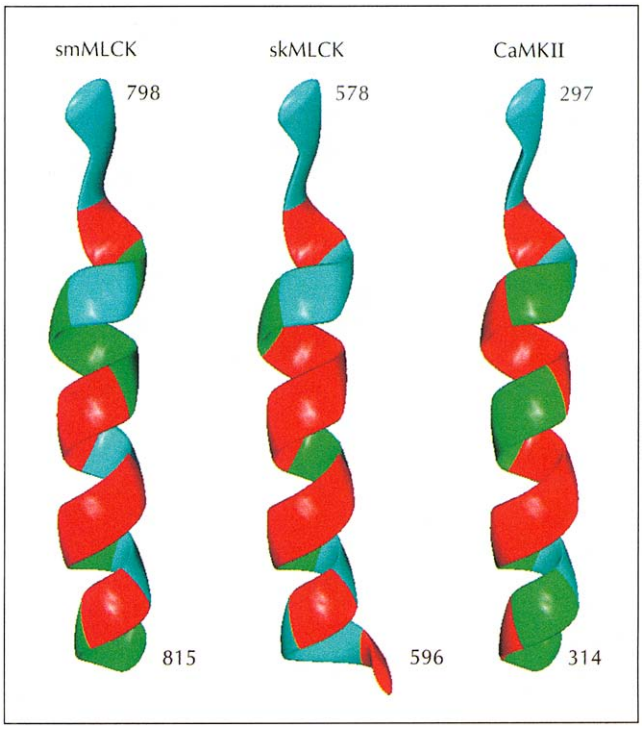


**Fig. 3.** Helical ribbon representations of the calmodulin-binding regions of three peptides whose atomic structures in complex with calmodulin have been determined: smMLCK, residues 796–815, ARRKWQKTGHHAVRAIGRLSS [22]; skMLCK, residues 577–602, KRRWKKNFIAVSAANRFKKISSSGA [19]; and CaMKII, residues 290–314, LKKFNARRKLKGAILTTMLATRNFS [23]. Underlining indicates the segments of the peptides which interact with calmodulin and which are shown in the figure. Hydrophobic residues are colored red, positively charged residues cyan, and polar residues green. All are reported to be helices and are based on the structure of skMLCK since the other two were not available from the Protein Data Bank.

Thus the apparent paradox between low sequence specificity and tight binding is the result of two factors: first,

the two domains can move relative to one another, and second, hydrophobic contacts dominate the interaction between $Ca^{2+}$–CaM and the peptides. The preponderance of long-chain hydrophobic (Met) and basic (Lys, Arg) residues allows for great side-chain flexibility within each half-site.

### $Ca^{2+}$-induced activation

The mechanism by which CaM is activated by $Ca^{2+}$ binding has been the subject of much investigation ever since the first X-ray crystal structure of $Ca^{2+}$–CaM was published. There has been much interest in what structure the inactive, apo-CaM might have and in what way $Ca^{2+}$ acts to convert apo-CaM to its active form. Unfortunately, apo-CaM has thus far defied all attempts to produce crystals of suitable quality for X-ray studies. The only clue offered by crystallographic studies was obtained from the structure of troponin-C mentioned above, in which only one domain was in the $Ca^{2+}$-activated form, while the other was in the inactive, apo-form [5]. The amino-terminal domain of apo-troponin C is more compact and its hydrophobic binding pocket is closed. For CaM, evidence from, for example, inhibitor-binding studies [24] lent support to the idea that CaM possesses similar features with respect to the inaccessibility of the hydrophobic binding sites in the apo-form. From this evidence and the troponin C structure, a model was presented for apo-CaM in which the two binding sites are closed but the orientation of the domains is maintained via the central helix [25]. In the absence of high resolution structural data, information on the precise structure of the apo-CaM form has been elusive, even leading some to speculate that there may be little or no stable structure in the apo-form. As the ability of NMR spectroscopy to deliver high-resolution data on proteins has increased, so too has the picture of apo-CaM come into focus. Building on a number of early studies on structural changes induced by $Ca^{2+}$ binding, partial NMR assignments of the individual domains of $TR_2C$ and $TR_1C$ gave indications of sequence-specific structural changes [26] (Fig. 4). More recently, complete $^1H$ assignment of the $TR_2C$ domain has been completed [27] and shows that the secondary structure is conserved in the apo-form but that the global fold is altered in a way that is consistent with the proposed model for apo-CaM [25]. These studies, along with heteronuclear assignments and preliminary structure calculations (B. Finn, unpublished data), indicate that the $TR_2C$ domain of apo-CaM possesses a distinct four-helix bundle structure with the helices closer to one another than in the $Ca^{2+}$-activated form and that the hydrophobic binding site is sequestered from solvent.

### Future directions

While great progress has been made in understanding the many structural transitions that calmodulin undergoes upon activation and target binding, there is much left to be done. Further studies on the apo-form of calmodulin and its domains promise to yield insights into the mechanism of activation for the entire family of EF-hand calcium-binding proteins. Studies of other calmodulin–target complexes, hopefully beyond the peptide level, would go a long way towards improving our understanding of the way in which the binding of helical segments is translated into the induction of activity in target proteins. Mutational studies could also shed light on these questions and many others.

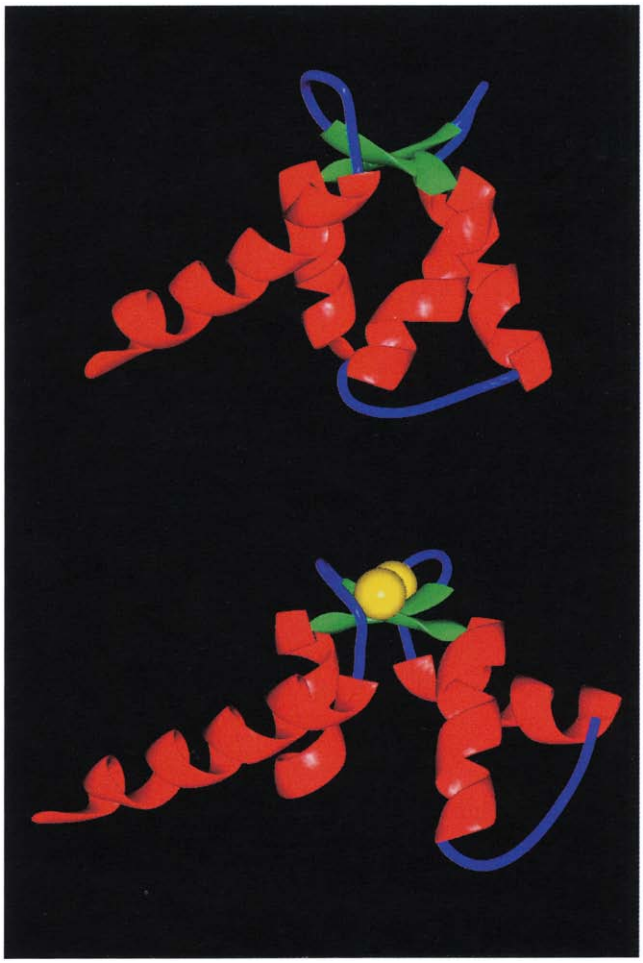

**Fig. 4.** A ribbon drawing showing the model for the $Ca^{2+}$-induced activation of the $TR_2C$ domain of CaM based on a model of apo-calmodulin [25] and NMR data [27]. The apo-form is shown above the $Ca^{2+}$-activated form in the same orientation. The binding site is the cleft at the bottom of the domain. The secondary structure is colored as in Fig. 1.

*Acknowledgement:* Molecular graphics images were produced using the MidasPlus program [28].

### References

1. Kawasaki, H. & Kretsinger, R. H. (1994). Calmodulin. *Protein Profile* **1**, 343–517.
2. Kretsinger, R.H. (1992). The linker of calmodulin - to helix or not to helix. *Cell Calcium* **13** ,363–376.
3. Babu, Y.S., Sack, J.S., Greenbough, T.C., Bugg, C.E., Means, A.R. & Cook, W.J. (1985). Three-dimensional structure of calmodulin. *Nature* **315**, 37–40.
4. 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.
5. Herzberg, O. & James, M.N.G. (1985). Structure of the calcium regulatory muscle protein troponin-C at 2.8 Å resolution. *Nature,* **313**, 653–659.

6. Sundaralingam, M., *et al.*, & Wang, B.C. (1985). Molecular structure of troponin C from chicken skeletal muscle at 3 Å resolution. *Science* **227**, 945–948.
7. Chattopadhyaya, R., Meador, W.E., Means, A.R. & Quiocho, F.A. (1992). Calmodulin structure refined at 1.7 Å resolution. *J. Mol. Biol.* **228**, 1177–1192.
8. Taylor, D.A., Sack, J.S., Maune, J.F., Beckingham, K. & Quiocho, F.A. (1991). Structure of a recombinant calmodulin from *Drosophila melanogaster* at 2.2-Å resolution. *J. Biol. Chem.* **266**, 21375–21380.
9. Rao, S.T., Wu, S., Satyshur, K.A., Ling, K.Y., Kung, C. & Sundaralingam, M. (1993). Structure of *Paramecium tetraurelia* calmodulin at 1.8 Å resolution. *Protein Sci.* **2**, 436–447.
10. Schutt, C. (1985). Hands on the calcium switch. *Nature* **315**, 15.
11. Seaton, B.A., Head, J.F. & Richards, F.M. (1985). Calcium-induced increase in the radius of gyration and maximum dimension of calmodulin measured by small-angle X-ray scattering. *Biochemistry* **24**, 6740–6743.
12. Heidorn, D.B. & Trewhella, J. (1988). Comparison of the crystal and solution structures of calmodulin and troponin C. *Biochemistry* **27**, 909–915.
13. Trewhella, J. (1993). The solution structures of calmodulin and its complexes with synthetic peptides based on target enzyme binding domains. *Cell Calcium* **13**, 377–390.
14. Ikura, M., Spera, S., Barbato, G., Kay, L. E., Krinks, M. & Bax, A. (1991). Secondary structure and side-chain $^{1}H$ and $^{13}C$ resonance assignment of calmodulin in solution by heteronuclear multi-dimensional NMR spectroscopy. *Biochemistry* **30**, 9216–9228.
15. Barbato, G., Ikura, M., Kay, L.E., Pastor, R.W. & Bax, A. (1992). Backbone dynamics of calmodulin studies by $^{15}N$ relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry* **31**, 5269–5278.
16. Persechini, A. & Kretsinger, R.H. (1988). Toward a model of the calmodulin–myosin light-chain kinase complex: implications for calmodulin function. *J. Cardiovasc. Pharm.* **12**, S1–S12.
17. Heidorn, D.B., Seeger, P.A., Rokop, S.E., Means, A.R., Crespi, & Trewhella, J. (1989). *Biochemistry* **28**, 6757–6764.
18. Persechini, A., Blumenthal, D.K., Jarrett, H.W., Klee, C.B., Hardy, D.O. & Kretsinger, R.H. (1989). The effects of deletions in the central helix of calmodulin on enzyme activation and peptide binding. *J. Biol. Chem.* **264**, 8052–8058.
19. M. Ikura, M., Clore, G.M., Gronenborn, A.M., Zhu, G., Klee, C.B. & Bax, A. (1992). Solution structure of a calmodulin-target peptide complex by multidimensional NMR. *Science* **256**, 632–638.
20. Szyja, W., Wrzosek, A., Brzeska, H. & Sarzäa, M.G. (1986). Interaction of calmodulin and its fragments with $Ca^{2+}$–ATPase and myosin light-chain kinase. *Cell Calcium* **7**, 73–88.
21. O'Neil, K.T. & DeGrado, W.F. (1990). How calmodulin binds its targets: sequence independent recognition of amphiphilic α-helices. *Trends Biochem. Sci.* **15**, 59–64.
22. Meador, W.E., Means, A.R. & Quiocho, F.A. (1992). Target enzyme recognition by calmodulin: 2.4 Å structure of a calmodulin-peptide complex. *Science* **257**, 1251–1255.
23. Meador, W.E., Means, A.R. & Quiocho, F.A. (1993). Modulation of calmodulin plasticity in molecular recognition on the basis of X-ray structures. *Science* **262**, 1718–1721.
24. LaPorte, D.C., Wierman, B.M. & Storm, D.R. (1980). Calcium-induced exposure of a hydrophobic surface on calmodulin. *Biochemistry* **19**, 3814–3819.
25. Strynadka, N.C.J. & James, M.N.G. (1988). Two trifluoroperazine-binding sites on calmodulin predicted from comparative molecular modeling with troponin C. *Proteins* **3**, 1–17.
26. Ikura, M., Minowa, O., Yazawa, M., Yagi, & Hikichi, K. (1987). Sequence-specific assignments of downfield-shifted amide proton resonances of calmodulin. *FEBS Lett.* **219**, 17–21.
27. Finn, B.E., Drakenberg, T. & Forsén, S. (1993). The structure of apo-calmodulin: a $^{1}H$ NMR examination of the carboxy-terminal domain. *FEBS Lett.* **336**, 368–374.
28. Ferrin, T.E., Huang, C.C., Jarvis, L.E. & Langridge, R. (1988). The MIDAS display system. *J. Mol. Graphics* **6**, 13–27.


Bryan E Finn and Sture Forsén, Department of Physical Chemistry 2, Chemical Center, Lund University, PO Box 124, S-22100 Lund, Sweden.