

Solution Structure of a Calmodulin-Target Peptide Complex by Multidimensional NMR

Mitsuhiko Ikura,* G. Marius Clore,† Angela M. Gronenborn,†
Guang Zhu, Claude B. Klee, Ad Bax†

The three-dimensional solution structure of the complex between calcium-bound calmodulin (Ca^{2+} -CaM) and a 26-residue synthetic peptide comprising the CaM binding domain (residues 577 to 602) of skeletal muscle myosin light chain kinase, has been determined using multidimensional heteronuclear filtered and separated nuclear magnetic resonance spectroscopy. The two domains of CaM (residues 6 to 73 and 83 to 146) remain essentially unchanged upon complexation. The long central helix (residues 65 to 93), however, which connects the two domains in the crystal structure of Ca^{2+} -CaM, is disrupted into two helices connected by a long flexible loop (residues 74 to 82), thereby enabling the two domains to clamp residues 3 to 21 of the bound peptide, which adopt a helical conformation. The overall structure of the complex is globular, approximating an ellipsoid of dimensions 47 by 32 by 30 angstroms. The helical peptide is located in a hydrophobic channel that passes through the center of the ellipsoid at an angle of $\sim 45^\circ$ with its long axis. The complex is mainly stabilized by hydrophobic interactions which, from the CaM side, involve an unusually large number of methionines. Key residues of the peptide are Trp⁴ and Phe¹⁷, which serve to anchor the amino- and carboxyl-terminal halves of the peptide to the carboxyl- and amino-terminal domains of CaM, respectively. Sequence comparisons indicate that a number of peptides that bind CaM with high affinity share this common feature containing either aromatic residues or long-chain hydrophobic ones separated by a stretch of 12 residues, suggesting that they interact with CaM in a similar manner.

Calmodulin (CaM) is a ubiquitous Ca^{2+} binding protein of 148 residues that is involved in a wide range of cellular Ca^{2+} -dependent signaling pathways, thereby regulating the activity of a large number of proteins including protein kinases, a protein phosphatase, nitric oxide synthase, inositol triphosphate kinase, nicotinamide adenine dinucleotide kinase, cyclic nucleotide phosphodiesterase, Ca^{2+} pumps, and proteins involved in motility (1). Binding

domains for CaM have been isolated from a number of CaM-dependent enzymes and have been shown to comprise peptide sequences with a high propensity for helix formation (2). In addition, both circular dichroism (3) and nuclear magnetic resonance (NMR) (4, 5) studies have shown that many naturally occurring CaM binding peptides, such as mellitin and mastoporan, as well as synthetic peptides corresponding to CaM binding domains adopt a helical

conformation upon binding to CaM. Although the crystal structure of Ca^{2+} -CaM has been solved (6), there is no crystal structure of a ternary complex of Ca^{2+} -CaM with the CaM binding domain of one of its target proteins yet available because of difficulties in obtaining suitable crystals. The crystal structure of Ca^{2+} -CaM is a dumbbell-shaped molecule with an overall length of ~ 65 Å that consists of two globular domains, each of which contains two Ca^{2+} binding sites of the helix-loop-helix type, connected by a long rigid central helix. In solution, on the other hand, NMR studies indicate that the central helix is disrupted near its midpoint and serves as a flexible linker between the two domains (7). A range of biophysical studies including cross-linking experiments (8) and small-angle x-ray and neutron scattering (9, 10) indicate that, upon complexation with certain target peptides, CaM adopts a globular conformation. The structural mechanism whereby Ca^{2+} -CaM recognizes its target sites has been the subject of considerable interest and debate, and a number of models for a Ca^{2+} -CaM-peptide complex have been proposed (11, 12). With the advent of a panoply of three-dimensional (3D) and 4D heteronuclear NMR experiments (13), it has now become possible to answer this question directly by solving the structure of a complex of this size (~ 19.7 kD) in solution (14). In this article, we present the structure determination of such a complex comprising recombinant CaM from *Drosophila melanogaster* and a 26-residue synthetic peptide (commonly referred to as M13) comprising the CaM binding domain of rabbit skeletal muscle myosin light chain kinase (MLCK).

Structure determination. All experiments were carried out at near-physiological conditions (100 mM KCl, pH 6.8, 35°C) on a 1:1 complex of uniformly ($>95\%$) ^{15}N - and ^{13}C -labeled Ca^{2+} -CaM and the synthetic 26-residue M13 peptide (KRRWKKNFIAVSAANRFKKISSGAL) (15) at natural isotopic abundance, with sample concentrations in the 1 to 1.5 mM range, prepared as described in (16). Structural information in the form of interproton distance restraints was derived from iso-

tope-edited and isotope-filtered 2D, 3D, and 4D nuclear Overhauser effect (NOE) experiments, on the basis of the previously obtained resonance assignments (4, 16, 17). Thus, 2D isotope-filtered nuclear Overhauser enhancement spectroscopy (NOESY) experiments (4, 18) were used to obtain interproton distance restraints for the bound peptide, suppressing the signal of ^{13}C - and ^{15}N -attached protons. The ^{13}C -separated, ^{12}C -filtered 3D NOESY spectra were used to specifically identify NOEs between protein protons and bound peptide protons. Finally, 3D ^{13}C - and ^{15}N -separated (19, 20) and 4D ^{13}C - ^{13}C -separated (21, 22) NOESY experiments were used to obtain NOEs between protein protons. At present, the 4D spectra have not as yet been extensively analyzed and have only been used to confirm NOE assignments derived from the 3D experiments.

An example of a small region of a $^1\text{H}(F_1)$ - $^1\text{H}(F_3)$ cross section through the 3D ^{13}C -separated, ^{12}C -filtered NOESY spectrum of the Ca^{2+} -CaM-M13 complex is shown in Fig. 1A. In this 3D spectrum, only NOE interactions between peptide protons and protein protons are observed. These NOE interactions are dispersed in an orthogonal dimension (F_2), which corresponds to the chemical shift of the ^{13}C atom attached to the protein proton. As a $^{13}\text{C}(F_2)$ spectral width of 23.8 ppm was used, this slice only contains NOE cross peaks between peptide protons and protein protons attached to ^{13}C

atoms resonating at $66.47 \pm (n \times 23.8)$ ppm (where $n = 0, 1, 2, \dots$). The side chain of the single Trp residue of M13 at position 4 shows intense NOE interactions to the methyl protons of Met¹⁴⁵ of CaM. The aromatic ring protons of Phe⁸ of M13 show interactions to the methyl group of Ala⁸⁸ and the αH proton of Ile⁸⁵, both located near the carboxyl-terminal end of the central helix. A $^{13}\text{C}(F_1)$ - $^{13}\text{C}(F_3)$ plane of the 4D ^{13}C - ^{13}C -separated NOESY spectrum of the Ca^{2+} -CaM-M13 complex at $\delta^1\text{H}(F_2)$, $\delta^1\text{H}(F_4) = 0.78, 1.23$ ppm is shown in Fig. 1B. Particularly noteworthy in this slice is the NOE between the methyl groups of Leu³⁹ and Val⁹¹, which indicates that they are separated by ≤ 5 Å. Thus, in the complex, the amino- and carboxyl-terminal domains are in close proximity. In contrast, in the crystal structure of Ca^{2+} -CaM, the distance between these two methyl groups is between 30 and 33 Å.

Approximate interproton distance restraints were obtained from 3D heteronuclear-separated and isotope-filtered NOE measurements at mixing times of 50 and 100 ms and were grouped into three classes, 1.8 to 2.7, 1.8 to 3.3, and 1.8 to 5.0 Å,

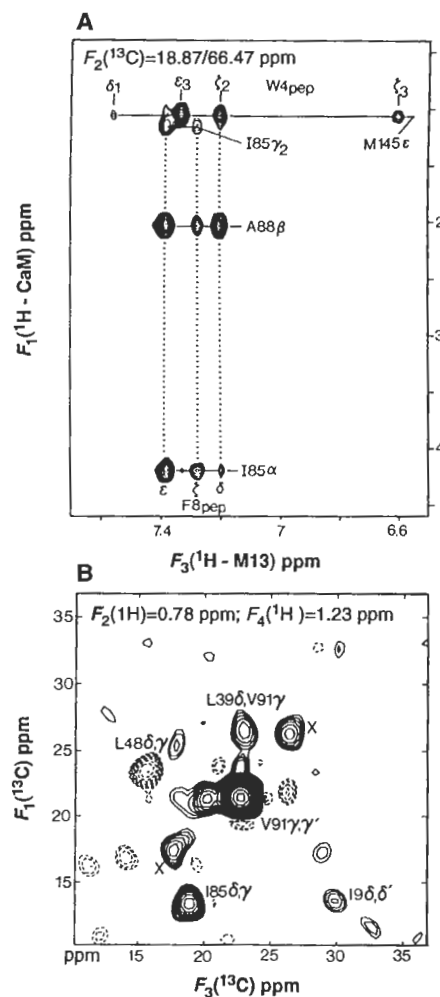


Fig. 1. Cross sections through (A) the 3D $^{13}\text{C}(F_2)$ -separated ^{13}C -filtered NOESY spectrum and (B) through the 4D ^{13}C - ^{13}C -separated NOESY spectrum of the Ca^{2+} -CaM-M13 complex. Spectrum A shows intermolecular NOE interactions exclusively between protons attached to ^{13}C (CaM protons) and protons attached to ^{12}C (M13 peptide protons) and displays among others NOEs between the CaM Met¹⁴⁵ methyl protons and the aromatic-ring protons of the M13 Trp⁴. Slice B displays intramolecular CaM NOE interactions. The ^{13}C - ^{13}C slice has been taken perpendicular to the NOESY cross peaks superimposed at $(F_2, F_4) = (0.78, 1.23)$ ppm, and reveals a long-range NOE cross peak (labeled L39 δ , V91 γ) between two methyl groups in the amino-terminal and carboxyl-terminal domains of CaM. Dashed contours correspond to negative intensity introduced by folding. Spurious diagonal peaks, marked "x," correspond to the tails of nearby intense diagonal methyl resonances. The 4D spectrum was recorded as described in (21). The 3D spectrum was recorded with a ^1H - ^{13}C HMQC-NOESY sequence (19) incorporating an isotope filter prior to the detection period (4). The spectra were processed with in-house routines for Fourier transformation (45) and linear prediction (46), together with the commercially available software package NMR2 (New Methods Research Inc., Syracuse, New York). Analysis of the spectra and peak picking was carried out with the in-house programs CAPP and PIPP (47).

M. Ikura, G. M. Clore, A. M. Gronenborn, G. Zhu, and A. Bax are members of the Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892. G. Zhu is also in the Chemical Physics Program, University of Maryland at College Park, MD 20740. C. B. Klee is a member of the Laboratory of Biochemistry, Building 37, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892.

*Present address: Division of Molecular and Structural Biology, Department of Medical Biophysics, Ontario Cancer Institute, University of Toronto, Toronto, Ontario, Canada M4X 1K9.

†To whom correspondence should be addressed.

which correspond to strong, medium, and weak NOEs, respectively (23, 24). The present structure calculations are based on 1787 interproton distance restraints, comprising 1490 and 164 restraints for CaM and M13, respectively, and 133 intermolecular

restraints between CaM and M13. These data were supplemented by 144 distance restraints for 72 hydrogen bonds identified on the basis of amide exchange and NOE data (25), 24 distance restraints between the four Ca^{2+} ions and protein groups

derived from the x-ray structure of Ca^{2+} -CaM (6), and 113 ϕ backbone torsion angle restraints derived from $^3J_{\text{HNH}}$ couplings (26). At present, no stereospecific assignments of C β H methylene protons or of methyl groups of Val and Leu have been obtained, and the structure calculations do not include any ψ or χ_1 torsion-angle restraints. Further, a large number of additional NOEs in the 4D spectrum have not as yet been interpreted. Thus, the present structure should be regarded as a low-resolution structure equivalent to what has been termed a second-generation NMR structure (13). A summary of the intermolecular NOEs observed between CaM and M13 is provided in Fig. 2.

Structure of the complex. A total of 24 simulated annealing (SA) structures (27) were calculated with the hybrid distance geometry-simulated annealing method of Nilges *et al.* (28, 29) (Table 1). The amino (residues 1 to 5) and carboxyl (residues 147 and 148) termini of CaM, the tether connecting the two domains of CaM (residues 74 to 82), and the amino (residues 1 and 2) and carboxyl termini (residues 22 to 26) of M13 are ill-defined by the present data and appear to be disordered. The atomic root-mean-square (rms) distribution about the mean coordinate positions for the rest of the structure (that is, residues 6 to 73 and 83 to 146 of CaM and residues 3 to 21 of M13) is 1.0 Å for the backbone atoms and 1.4 Å for all atoms. A stereoview showing a best-fit superposition of the 24 SA structures is shown in Fig. 3A.

The backbone atomic rms differences between the amino- (residues 6 to 73) and carboxyl- (residues 83 to 146) terminal domains of the Ca^{2+} -CaM-M13 complex and the corresponding residues of the rat Ca^{2+} -CaM x-ray structure (30) are 1.8 and 1.5 Å, respectively. Thus, within the errors of the present coordinates, there is no significant conformational change within each domain (residues 4 to 74 and 82 to 146) in the complexed and uncomplexed states. Further, the overall conformation of the two domains is similar and their backbone atoms can be superimposed with an atomic rms difference of 1.8 Å.

Table 1. The notation of the structures is as follows: <SA> are the 24 simulated annealing structures; $\overline{\text{SA}}$ is the mean structure obtained by averaging the coordinates of the individual SA structures best fitted to each other (the residues used in the best fitting are residues 6 to 73 and 83 to 146 of CaM and residues 3 to 21 of M13); and ($\overline{\text{SA}}$)_r is the restrained minimized (regularized) mean structure obtained from SA. The number of terms for the various restraints is given in parenthesis; rms, root-mean-square, and expt, experimental.

Parameter	<SA>	($\overline{\text{SA}}$) _r
Rms deviations from exp distance restraints (Å)*		
All (1995)	0.037 ± 0.006	0.039
Intramolecular CaM interproton distances		
Interresidue sequential ($ i - j = 1$) (393)	0.027 ± 0.006	0.023
Interresidue medium range ($1 < i - j \leq 5$) (253)	0.014 ± 0.006	0.023
Interresidue long range ($ i - j > 5$) (162)	0.030 ± 0.012	0.021
Intraresidue (682)	0.038 ± 0.006	0.040
Intramolecular M13 interproton distances		
Interresidue sequential ($ i - j = 1$) (51)	0.063 ± 0.037	0.064
Interresidue medium range ($1 < i - j \leq 5$) (33)	0.018 ± 0.009	0.037
Intraresidue (80)	0.066 ± 0.008	0.068
Intermolecular CaM-M13 interproton distances (133)	0.058 ± 0.013	0.062
CaM hydrogen-bond restraints (118)†	0.025 ± 0.008	0.023
M13 hydrogen-bond restraints (26)†	0.027 ± 0.020	0.037
Ca^{2+} distance restraints (24)‡	0.015 ± 0.015	0.000
Rms deviation from expt. ϕ restraints (degree) (113)*§	0.046 ± 0.071	0.126
Deviations from idealized covalent geometry		
Bonds (Å) (2715)	0.004 ± 0.000	0.006
Angles (degrees) (4894)	1.810 ± 0.004	2.267
Impropers (degrees) (1036)	0.510 ± 0.019	0.647
E_{LJ} (kcal mol ⁻¹)¶	-561.3 ± 37.3	-368

*None of the structures exhibit distance violations greater than 0.5 Å or dihedral angle violations greater than 1°. †Each hydrogen bond is characterized by two distance restraints: $r_{\text{NH}\cdots\text{O}} \leq 2.3$ Å and $r_{\text{N}\cdots\text{O}} = 2.5$ to 3.3 Å. All hydrogen-bonding restraints involve slowly exchanging backbone amide protons. ‡There are six distance restraints (1.5 to 3.3 Å) per Ca^{2+} to the interacting side-chain carboxylate and backbone carbonyl atoms of CaM, as deduced from the x-ray structure of Ca^{2+} -CaM (6). §The ϕ backbone torsion angles are restrained from -10° to -100° and from -70° to -170° for $^3J_{\text{HNH}} < 6$ Hz and > 8 Hz, respectively (24). ||The improper torsion restraints serve to maintain planarity and chirality. ¶ E_{LJ} is the Lennard Jones van der Waals energy calculated with the use of CHARMM (48) empirical energy function and is not included in the target function for simulated annealing or restrained minimization.

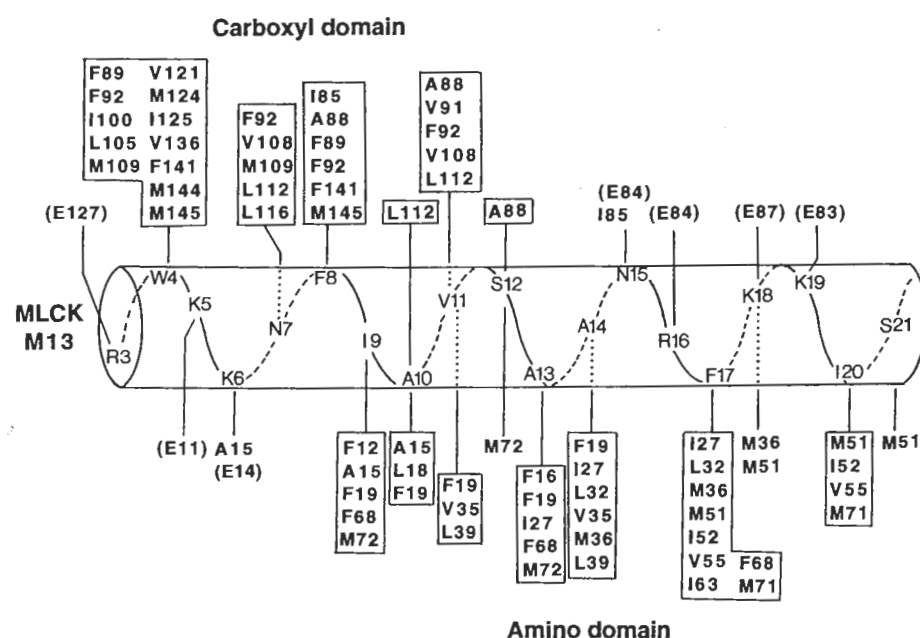


Fig. 2. Summary of residue pairs for which intermolecular NOEs between CaM and M13 are observed. CaM residues involved in hydrophobic interactions are boxed. Also included are potential electrostatic interactions between negatively charged Glu residues of CaM, shown in parentheses, and positively charged Lys and Arg residues of M13; it should be noted that the electrostatic interactions have been inferred purely from the structure as no NOEs have as yet been identified between these pairs of residues.

The major conformational change in Ca^{2+} -CaM that occurs upon binding M13 involves the dissolution of the long central helix (residues 65 to 93) observed in the x-ray structure into two helices (residues 65 to 73 and 83 to 93) connected by a long flexible loop (residues 74 to 82), thereby enabling the two domains to come together to grip the peptide rather like two hands capturing a rope. The hydrophobic channel formed by the two domains is complementary in shape to that of the peptide helix. This structure is illustrated by the schematic ribbon drawings shown in Fig. 4, which also serve to highlight the approximate twofold pseudo-symmetry of the complex. Thus, whereas the two domains of CaM are arranged in an approximately orthogonal manner to each other in the crystal structure of Ca^{2+} -CaM (6), in the Ca^{2+} -CaM-M13 complex they are almost symmetrically related by a 180° rotation about a twofold axis. A large conformational change also occurs in the M13 peptide upon complexation from a random coil state (4) to a well-defined helical conformation. Indeed, the helix involves all of the residues (3 to 21) of M13 that interact with CaM, whereas the amino (residues 1 and 2) and carboxyl (residues 22 to 26) termini of the peptide, which do not interact with CaM, remain disordered. Upon complexation there is a decrease in the accessible surface area of CaM and M13 of 1848 and 1477 $\text{\AA}^2$, respectively, which corresponds to a decrease in the calculated solvation free energy of folding (31) of 18 and 20 kcal mol^{-1} (32). This large decrease in solvation free energy of folding would account for the very tight binding (association constant $K_{\text{ass}} \approx 10^9 \text{ M}^{-1}$) of M13 to CaM (2). In addition, the accessible surface area of the portion of M13 (residues 3 to 21) in direct contact with CaM in the complex is only 494 $\text{\AA}^2$ compared to an accessible surface area of 3123 $\text{\AA}^2$ for a random coil and 2250 $\text{\AA}^2$ for a helix. Thus, more than 80 percent of the surface of the peptide in contact with CaM is buried.

In the views shown in Figs. 3 and 4, the roof of the channel is formed by helices II (residues 29 to 38) and VI (residues 102 to 111) of the amino- and carboxyl-terminal domains, respectively, which run antiparallel to each other; and the floor is formed by the flexible loop (residues 74 to 82) connecting the two domains and by helix VIII (residues 138 to 146) of the carboxyl-terminal domain. The front of the channel in Fig. 4A and the left wall of the channel in Fig. 4B are formed by helices I (residues 7 to 19) and IV (residues 65 to 73) and the mini-antiparallel β sheet comprising residues 26 to 28 and 62 to 64, all from the amino-terminal domain; the back of the channel in Fig. 4A and the right wall of the

channel in Fig. 4B are formed by helices V (residues 83 to 93) and VIII (residues 138 to 146) and the mini-antiparallel β sheet comprising residues 99 to 101 and 135 to 137, all from the carboxyl-terminal domain. The two domains of CaM are staggered with a small degree of overlap such that the hydro-

phobic face of the amino-terminal domain mainly contacts the carboxyl-terminal half of the M13 peptide, while the carboxyl-terminal domain principally interacts with the amino-terminal half of M13 (Fig. 4A).

The overall Ca^{2+} -CaM-M13 complex has a compact globular shape that is almost

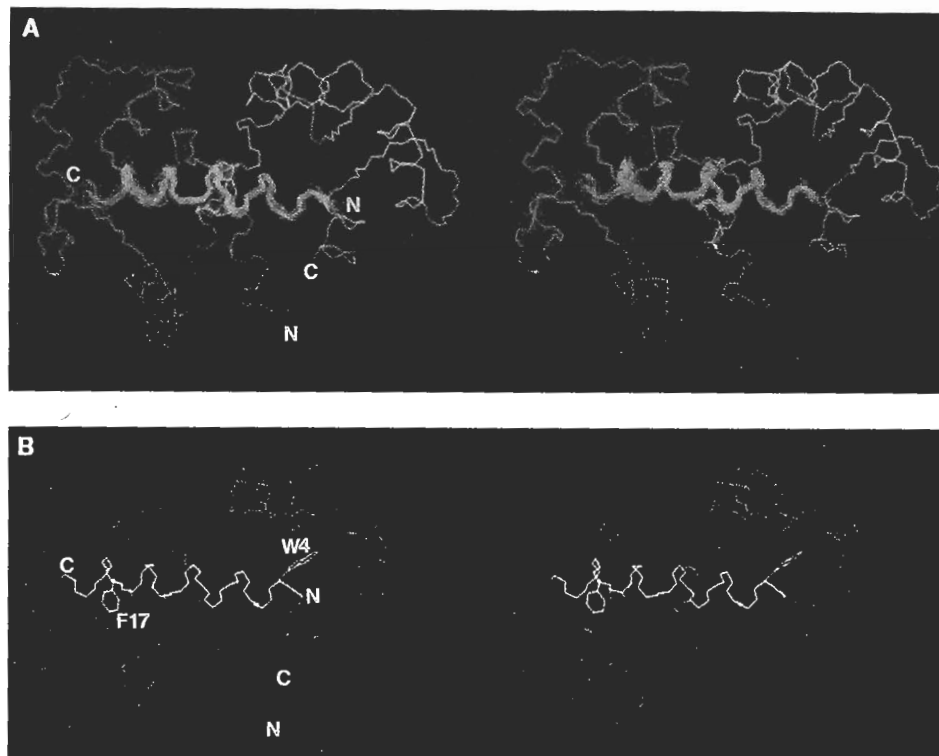


Fig. 3. (A) Stereoview showing a best-fit superposition of the backbone (N, $\text{C}\alpha$, C) atoms of the 24-SA structure of the Ca^{2+} -CaM-M13 complex with the amino-terminal domain in blue, the carboxyl-terminal domain in red, and the M13 peptide in green; the restrained minimized mean (SA)r structure is also included in the superposition and is highlighted. (B) Stereoview of the restrained minimized mean (SA)r structure illustrating the interaction of Trp⁴ and Phe¹⁷ of the peptide (shown in green) with hydrophobic side chains of the carboxyl- and amino-terminal domains of CaM, respectively (shown in red); the backbone atoms of the amino- and carboxyl-terminal domains are shown in blue and purple, respectively, while the backbone atoms of the peptide are shown in yellow. The residues shown are 6 to 146 of CaM and 3 to 21 of M13; the amino and carboxyl termini of both CaM and M13 are disordered.

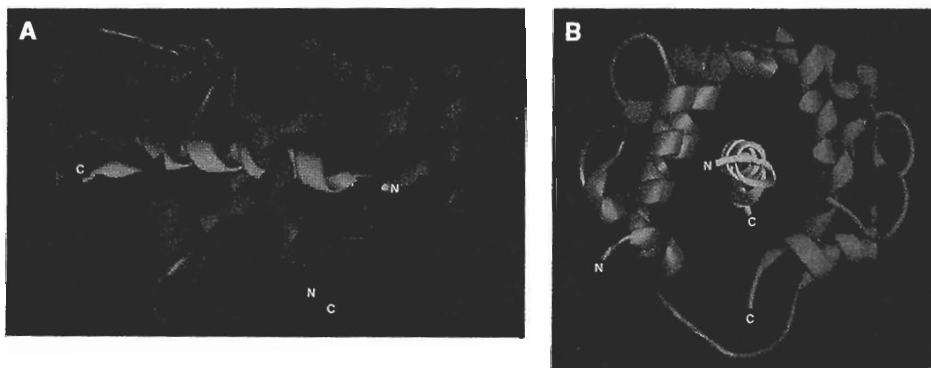


Fig. 4. Two orthogonal views of a schematic ribbon drawing representation of the restrained minimized mean (SA)r structure of the Ca^{2+} -CaM-M13 complex with the amino-terminal domain in blue, the carboxyl-terminal domain in red, and the M13 peptide in green. The residues shown are 5 to 148 of CaM and 1 to 22 of M13. The model was generated with the program Ribbon 2.0 (49).

ellipsoidal, with approximate dimensions of 47 by 32 by 30 Å. The helical M13 peptide passes through the center of the ellipsoid at an angle of ~45° with its long axis. In contrast, the approximate dimensions of the Ca²⁺-CaM x-ray structure are 65 by 30 by 30 Å (6). This change in dimensions is reflected by a change in the ratio of the three principal components of the inertia tensor from 1.00:0.22:1.07 for the x-ray structure of Ca²⁺-CaM to 1.00:1.44:1.58 for the solution structure of the Ca²⁺-CaM-M13 complex. In addition, the calculated radius of gyration for Ca²⁺-CaM-M13 is ~17 Å, which is completely consistent with the decrease in the radius of gyration from ~21 Å to ~16 Å observed by both small-angle x-ray and neutron scattering upon complexation of Ca²⁺-CaM with M13 (10).

The Ca²⁺-CaM-M13 complex is stabilized by numerous hydrophobic interactions which are summarized in Fig. 2. Particularly striking are the interactions of Trp⁴ and Phe¹⁷ of the peptide which serve to anchor the amino- and carboxyl-terminal halves of M13 to the carboxyl-terminal and amino-terminal hydrophobic patches of CaM, respectively (Fig. 3B). These interactions also involve a large number of Met residues, which are unusually abundant in CaM, in particular four methionines in the carboxyl-terminal domain (Met¹⁰⁹, Met¹²⁴, Met¹⁴⁴, and Met¹⁴⁵) and three methionines in the amino-terminal domain (Met³⁶, Met⁵¹, and Met⁷¹). As methionine is an unbranched hydrophobic residue extending over four heavy atoms (C β , C γ , S δ , and C ϵ), the abundance of methionines can generate a hydrophobic surface whose detailed topology is readily adjusted by minor changes in side-chain conformation, thereby providing a mechanism to accommodate and recognize different bound peptides (12).

In addition to hydrophobic interactions, there are a number of possible electrostatic interactions that can be deduced from the calculated SA structures. Putative interactions exist between the Arg and Lys residues of M13 and the Glu residues of CaM, and these are also included in Fig. 2. Glu¹¹ and Glu¹⁴ in helix I are within 5 Å of Lys⁵ and Lys⁶ of M13; Glu⁸³, Glu⁸⁴, and Glu⁸⁷ in helix V of CaM are close to Lys¹⁹, Arg¹⁶, and Lys¹⁸, respectively, of M13; and Glu¹²⁷ in helix VII of CaM is close to Arg³ of M13.

Correlation with biochemical data. The present structure explains a number of interesting observations previously reported. Studies of backbone amide exchange behavior have shown that upon complexation with M13, the amide exchange rates of residues 75 to 79 are substantially increased (33). Prior NMR studies on Ca²⁺-CaM indicate that the long central helix is already disrupted near its middle (from Asp⁷⁸ to Ser⁸¹) in solution (16) and that large variations in the orientation of one domain relative to the other occur randomly with time (7). The further disruption of the central helix upon complexation seen in the present structure is manifested by the increased amide exchange rates and supports the view of the central helix serving as a flexible linker between the two domains. Similarly, the present structure explains the finding that as many as four residues can be deleted from the middle of the central helix without dramatically altering the stability or shape of the Ca²⁺-CaM-M13 complex (34), as the long flexible loop connecting the two domains can readily be shortened without causing any alteration in the structure (compare with Fig. 4). The observation from photoaffinity labeling studies that the two domains of CaM interact simultaneously with opposite ends of the peptide, such that residue 4 of the peptide (current

numbering) can be cross-linked to Met¹²⁴ or Met¹⁴⁴ of the carboxyl-terminal domain and that residue 13 of the peptide can be cross-linked to Met⁷¹ of the amino-terminal domain (13, 35), is readily explained by the structural finding that the amino-terminal half of the peptide interacts predominantly with the carboxyl-terminal domain while the carboxyl-terminal half of the peptide interacts predominantly with the amino-terminal domain (Figs. 2 to 4). The observation that at least 17 residues of the M13 peptide from either skeletal muscle or smooth muscle are necessary for high-affinity binding (36, 37) is readily explained by the intimate interactions of the carboxyl-terminal hydrophobic residue (that is, Phe¹⁷) with the amino-terminal domain of CaM by which the peptide is anchored. Finally, the structure accounts for experiments in which cross-linking of residues 3 and 146 of CaM, mutated to Cys, has no effect on the activation of MLCK, even if the central helix is cleaved proteolytically at Lys⁷⁷ by trypsin (9). Thus, whereas the C α carbons of the residues 3 and 146 are 37 Å apart in the x-ray structure of Ca²⁺-CaM, they are only ~20 Å apart in the solution structure of the Ca²⁺-CaM-M13 complex, which is close enough to permit cross-linking to occur.

Comparison with previous models. It is of interest to compare the present structure of the Ca²⁺-CaM-M13 complex with the class III model proposed by Persechini and Kretsinger (11) generated by introducing a kink in the central helix by simply changing the ϕ, ψ angles of Ser⁸¹ from -57°, -47° to -54°, +98°. The general features of the structure, namely, the two domains coming together in close proximity to form a hydrophobic channel occupied by a peptide in a helical conformation, are similar in the two cases. However, there are substantial differences. First, the extent of the M13 peptide in direct contact with CaM is a little longer in the model (residues 1 to 21) than in the solution structure (residues 3 to 21). Second, the peptide in the model is oriented at 180° relative to that in the solution structure such that the amino- and carboxyl-terminal halves of the peptide in the model interact with the amino- and carboxyl-terminal domains of CaM, respectively. (The possibility, however, of inverting the orientation of the peptide is stated in the description of the model.) Third, the relative orientation of the two domains is significantly different; the central helix remains essentially intact in the model except for the kink introduced at Ser⁸¹, whereas in the solution structure it is broken up and partitioned into two helices separated by a long nine-residue flexible loop, thereby permitting a more optimal arrangement of the two domains. Thus, when the carboxyl-terminal domains of the

	1	5	10	15	20	25
SK-MLCK M13	K R R	W K K N	F I A V	S A A N R	F K K I S S S G A L	
SM-MLCK M13	R R K	W Q K T	G H A V	R A I G R	L S S S	
Ca Pump C24W	Q I L	W F R G	L N R I	Q T Q I R	V N A F R S S	
C20W	L R R G	Q I L W	F R G L	N R I Q	T Q I K	
Calspermin	A R R K	L K A A	V K A V	V A S S R	L G S	
Calcineurin	A R K E V	I R W K	I R A I	G K M A R	V S F V L	
Mastaporan		I N L K	A L A A	L A K K I	L	
Mastaporan X		I N W K	G I A A	M A K K L	L	
Mellitin	G I G A V	L K V L	T T G L	P A L I S	W I K R K R Q Q	
Interacting domain of CaM		C	C	C	N	

Fig. 5. Alignment of tightly binding ($K_{\text{ass}} > 5 \times 10^7 \text{ M}^{-1}$) CaM binding sequences based on the structural role of Trp⁴ and Phe¹⁷ in anchoring the M13 peptide to the carboxyl- and amino-terminal domains of CaM, respectively. In all of the sequences, there is a pair of aromatic or long-chain aliphatic residues or both (boxed) separated by a stretch of 12 residues that correspond to Trp⁴ and Phe¹⁷ of M13. In addition, there are generally hydrophobic residues at positions equivalent to Phe⁸ (which interacts with the carboxyl-terminal domain) and Val¹¹ (which interacts with both domains). Sequences: skeletal muscle myosin light chain kinase (SK-MLCK) M13 (residues 342 to 367) (50), smooth muscle myosin light chain kinase (SM-MLCK) M13 (residues 494 to 513) (51), plasma membrane Ca²⁺ pump (37, 52), calspermin (11, 53), and calcineurin (mouse brain) (54). The mastaporans and mellitin are naturally occurring insect peptides (55).

model and solution structures are superimposed, the atomic rms displacement between the amino-terminal domains of the two structures is ~ 8 Å. As a consequence the width of the channel is a little wider and the two domains are slightly more separated in the model than in the solution structure.

An alternative model has also been proposed by O'Neil and DeGrado (12). As in the previous model, a kink was introduced by altering the ϕ, ψ angles of Ser⁸¹. Although the orientation of the peptide in this model, based on the results of photolabeling studies, was correct, only a relatively small distortion was introduced in the central helix so that the complex retained to some degree a dumbbell shape as opposed to the globular shape observed in the present structure, and does not possess a well-defined hydrophobic channel.

Implications of the structure. A large body of experimental data shows that CaM binds to numerous proteins whose binding domains exhibit a propensity for α -helix formation (2). A comparison of these sequences reveals little homology. Nevertheless, many of the very tightly binding peptides ($K_{\text{ass}} \geq 5 \times 10^7 \text{ M}^{-1}$) have the common property of containing either aromatic residues or long-chain hydrophobic residues (Leu, Ile, or Val) separated by 12 residues (Fig. 5). In the case of M13, these two residues are Trp⁴ and Phe¹⁷ which are exclusively in contact with the carboxyl- and amino-terminal domains of CaM, respectively (Fig. 2). Given that these two residues are involved in more hydrophobic interactions with CaM than any other residues of the peptide (compare Figs. 2 and 3B), it seems likely that this feature of the sequence can be used to align the CaM binding sequences listed in Fig. 5, thereby permitting one to predict their interaction with CaM. It is clear from this alignment that the pattern of hydrophobic and hydrophilic residues is in general comparable for the various peptides, suggesting that the mode of binding and the structure of the corresponding complexes with Ca²⁺-CaM are also likely to be similar. For example, there is, in general, conservation of hydrophobic residues at the positions equivalent to Phe⁸ which interacts with the carboxyl-terminal domain and Val¹¹ which interacts with both domains (compare with Fig. 2). In addition, there are no acidic residues present that would result in unfavorable electrostatic interactions with the negatively charged Glu residues on the surface of CaM (compare with Fig. 2). The minimum length of peptide required for high-affinity binding to Ca²⁺-CaM is defined by the 14-residue mastaporans, which comprise the two hydrophobic residues at the amino and carboxyl termini (Fig. 5) and have approximately the same equilibrium associ-

ation constant ($K_{\text{ass}} \approx 1 \times 10^9$ to $3 \times 10^9 \text{ M}^{-1}$) as M13 (2).

This structural alignment also predicts that a peptide stopping just short of the second hydrophobic residue of the pair (that is, the residue equivalent to Phe¹⁷) would only bind to the carboxyl-terminal domain and that the resulting complex would therefore retain the dumbbell shape of Ca²⁺-CaM. This is exactly what has been observed by small-angle x-ray scattering with two synthetic peptides, C24W and C20W (Fig. 5), which comprise different portions of the CaM binding domain of the plasma membrane Ca²⁺ pump (38). The complex with the C24W peptide that corresponds to residues 1 to 24 of M13 and contains a Trp at position 4 and a Val at position 17 has a globular shape similar to that of Ca²⁺-CaM-M13. The complex with the C20W peptide, on the other hand, which corresponds to residues -4 to 16 of M13 and therefore lacks the carboxyl-terminal hydrophobic residue of the pair, retains the dumbbell shape of Ca²⁺-CaM (38), suggesting that the peptide only binds to the carboxyl-terminal domain.

The present structure of the Ca²⁺-CaM-M13 complex reveals an unusual binding mode in which the target peptide is sequestered into a hydrophobic channel formed by the two domains of CaM with interactions involving 19 residues of the target M13 peptide (that is, residues 3 to 21). In addition, a key requirement appears to be the presence of two long-chain hydrophobic or aromatic residues separated by 12 residues in order to anchor the peptide to the two domains of CaM. By analogy, the rope (that is, the CaM binding domain of the target) has to be long enough and have two knots at each end for the two hands (domains) of CaM to grip it. This particular mode of binding is therefore only likely to occur if the CaM binding site is located either at an easily accessible carboxyl or amino terminus or in a long exposed surface loop of the target protein. An example of the former is MLCK and of the latter is calcineurin, and, in accordance with their location, the CaM binding sites are susceptible to proteolysis (37, 39). Clearly, other types of complexes between Ca²⁺-CaM and its target proteins are possible given the inherent flexibility of the central helix (7, 16, 33). For example, in the case of the γ subunit of phosphorylase kinase, it appears that there are two discontinuous CaM binding sites that are capable of binding to Ca²⁺-CaM simultaneously (40), and binding of a peptide derived from one of these sites causes elongation rather than contraction of Ca²⁺-CaM (41), indicating that the complex is of a quite different structural nature. Similarly, in the case of cyclic nucleotide phosphodiesterase (42) and CaM

kinase II (43), the CaM binding sequences do not have the same spacing of hydrophobic residues seen in M13 and the other sequences listed in Fig. 5, and, in addition, CaM kinase II is not susceptible to proteolysis in the absence of phosphorylation (44), suggesting that the mode of binding is different again. Thus, in all likelihood, the present complex represents one of a range of Ca²⁺-CaM binding modes achieving CaM-target protein interactions in an efficient and elegant manner.

REFERENCES AND NOTES

- C. B. Klee and T. C. Vanaman, *Adv. Protein Chem.* **35**, 213 (1982); S. H. Ryu, S. Y. Lee, S. G. Rhee, *FASEB J.* **1**, 388 (1987); S. Forsen, H. J. Vogel, T. Drakenberg, in *Calcium and Cell Function*, W. Y. Cheung, Ed. (Academic Press, New York, 1986), vol. 6, pp. 113-157; C. B. Klee, in *Molecular Aspects of Cellular Regulation*, P. Cohen and C. B. Klee, Eds. (Elsevier, New York, 1988), vol. 5, pp. 35-56; A. R. Means, *Recent Prog. Horm. Res.* **44**, 223 (1988); D. S. Bredt and S. H. Snyder, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 682 (1990); C. B. Klee, *Neurochemistry* **16**, 1059 (1991).
- S. R. Anderson and D. A. Malencik, in *Calcium Cell Function*, W. Y. Cheung, Ed. (Academic Press, New York, 1986), vol. 6, pp. 1-42; D. K. Blumenthal and E. G. Krebs, in *Molecular Aspects of Cellular Regulation*, P. Cohen and C. B. Klee, Eds. (Elsevier, New York, 1988), vol. 5, pp. 341-356; W. F. DeGrado, *Adv. Protein Chem.* **39**, 51 (1988).
- L. McDowell, G. Sanyal, F. G. Prendergarst, *Biochemistry* **24**, 2979 (1985); W. F. DeGrado, F. G. Prendergarst, H. R. Wolfe, J. A. Cox, *J. Cell Biochem.* **29**, 83 (1985); J. A. Cox, M. Comte, J. E. Fitton, W. F. DeGrado, *J. Biol. Chem.* **260**, 2527 (1985); R. E. Klevitt, D. K. Blumenthal, D. E. Wemmer, E. G. Krebs, *Biochemistry* **24**, 8152 (1985); T. Vorherr *et al.*, *ibid.* **29**, 355 (1990).
- M. Ikura and A. Bax, *J. Am. Chem. Soc.* **114**, 2433 (1992).
- S. H. Seeholzer and A. J. Wand, *Biochemistry* **25**, 4011 (1989); S. M. Roth *et al.*, *ibid.* **30**, 10078 (1991); H. Prêcheur, H. Munier, J. Mispelter, O. Bâzu, C. T. Craescu, *ibid.* **31**, 229 (1992).
- Y. S. Babu *et al.*, *Nature* **315**, 37 (1985); R. H. Kretsinger, S. E. Rudnick, L. J. Weissman, *J. Inorg. Biochem.* **28**, 289 (1986); Y. S. Babu, C. E. Bugg, W. J. Cook, *J. Mol. Biol.* **204**, 191 (1988); D. A. Taylor, J. S. Sack, J. F. Maune, K. Beckingham, F. A. Quiocho, *J. Biol. Chem.* **266**, 21375 (1991).
- G. Barbato, M. Ikura, L. E. Kay, R. W. Pastor, A. Bax, *Biochemistry*, in press.
- A. Persechini and R. H. Kretsinger, *J. Biol. Chem.* **263**, 12175 (1988).
- N. Matsushima *et al.*, *J. Biochem. (Tokyo)* **105**, 883 (1989); M. Kataoka, J. F. Head, B. A. Seaton, D. M. Engelman, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6944 (1989).
- D. B. Heidorn *et al.*, *Biochemistry* **28**, 6757 (1989).
- A. Persechini and R. J. Kretsinger, *J. Cardiovasc. Pharmacol.* (S5) **12**, 1 (1988).
- K. T. O'Neil and W. F. DeGrado, *Proteins* **6**, 284 (1988); *Trends Biochem. Sci.* **15**, 59 (1990).
- G. M. Clore and A. M. Gronenborn, *Science* **252**, 1390 (1991); *Annu. Rev. Biophys. Chem.* **20**, 29 (1991); L. E. Kay, G. M. Clore, A. Bax, A. M. Gronenborn, *Science* **249**, 411 (1990).
- G. M. Clore, P. T. Wingfield, A. M. Gronenborn, *Biochemistry* **30**, 2315 (1991).
- Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
- M. Ikura, L. E. Kay, M. Krinks, A. Bax, *Biochemistry* **30**, 5498 (1991).
- M. Ikura *et al.*, unpublished data. The assign-

- ments of CaM in the ternary Ca^{2+} -CaM-M13 complex were obtained in 3D double and triple resonance NMR experiments [M. Ikura, L. E. Kay, A. Bax, *Biochemistry* 29, 4569 (1990); A. Bax, G. M. Clore, A. M. Gronenborn, *J. Magn. Reson.* 88, 425 (1990); G. M. Clore, A. Bax, P. C. Driscoll, P. T. Wingfield, A. M. Gronenborn, *Biochemistry* 29, 8172 (1990); G. M. Clore and A. M. Gronenborn, *Prog. Nucl. Magn. Reson.* 23, 43 (1991)].
18. G. Wider, C. Weber, R. Traber, H. Widmer, K. Wüthrich, *J. Am. Chem. Soc.* 112, 9015 (1990).
19. M. Ikura, L. E. Kay, R. Tschudin, A. Bax, *J. Magn. Reson.* 86, 204 (1990); E. R. P. Zuiderweg, L. P. McIntosh, F. W. Dahlquist, S. W. Fesik, *ibid.*, p. 210.
20. S. W. Fesik and E. R. P. Zuiderweg, *ibid.* 78, 588 (1988); D. Marion, L. E. Kay, S. W. Sparks, D. A. Torchia, A. Bax, *J. Am. Chem. Soc.* 111, 1515 (1989); D. Marion *et al.*, *Biochemistry* 29, 6150 (1989).
21. G. M. Clore, L. E. Kay, A. Bax, A. M. Gronenborn, *Biochemistry* 30, 12 (1991).
22. E. R. P. Zuiderweg, A. M. Petros, S. W. Fesik, E. T. Olejniczak, *J. Am. Chem. Soc.* 113, 370 (1991).
23. G. M. Clore *et al.*, *EMBO J.* 5, 2729 (1986).
24. Upper distance limits for distances involving methyl and methylene protons were corrected appropriately for center averaging [K. Wüthrich, M. Billeter, W. Braun, *J. Mol. Biol.* 169, 949 (1983)]. In addition, 0.5 Å was added to the upper limit of distances involving methyl protons to account for the higher apparent intensity of methyl resonances [G. M. Clore, A. M. Gronenborn, M. Nilges, C. A. Ryan, *Biochemistry* 26, 8012 (1987)].
25. K. Wüthrich, *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).
26. A. Pardi, M. Billeter, K. Wüthrich, *J. Mol. Biol.* 180, 741 (1984).
27. The coordinates of the 24 SA structures and of the restrained minimized mean structure, (SA)_r, together with the experimental restraints, have been deposited in the Brookhaven Protein Data Bank.
28. M. Nilges, G. M. Clore, A. M. Gronenborn, *FEBS Lett.* 229, 317 (1988).
29. The hybrid distance geometry-simulated annealing protocol for (27) makes use of the program XPLOR [A. T. Brünger, G. M. Clore, A. M. Gronenborn, M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.* 83, 3801 (1986); A. T. Brünger, *XPLOR Version 3 Manual* (Yale University, New Haven, 1992)] incorporating a distance geometry module (J. Kuszewski, M. Nilges, A. T. Brünger, *J. Biomol. NMR* 2, 33 (1992)). The protocol involves first calculating an initial set of substructures incorporating only about one-third of the atoms by projection from *n*-dimensional distance space into cartesian coordinate space, followed by simulated annealing with all atoms. The target function that is minimized during simulated annealing (as well as in conventional Powell minimization) comprises only quadratic harmonic potential terms for covalent geometry (that is, bonds, angles, planes, and chirality), square-well quadratic potentials for the experimental distance and torsion angle restraints (22), and a quartic van der Waals repulsion term for the nonbonded contacts (26). All peptide bonds were restrained to be trans. There are no hydrogen bonding, electrostatic, or 6-12 Lennard-Jones potential terms in the target function.
30. The coordinates used have the Brookhaven Protein Data Bank accession number 3CLN [Y. S. Babu, C. E. Bugg, W. J. Cook, *J. Mol. Biol.* 204, 191 (1988)].
31. D. Eisenberg and A. D. McLaghlán, *Nature* 319, 199 (1986); L. Cliche, L. M. Gregoret, F. E. Cohen, P. A. Kollman, *Proc. Natl. Acad. Sci. U.S.A.* 87, 3240 (1990).
32. The changes in accessible surface area and solvation free energy calculated for M13 peptide are relative to that of the free peptide in the same helical conformation. Relative to a theoretical extended, unfolded structure, the decrease in accessible surface area and solvation free energy of M13 upon binding is 2039 Å² and 29 kcal mol⁻¹, respectively.
33. S. Spera, M. Ikura, A. Bax, *J. Biomol. NMR* 1, 155 (1991).
34. A. Persechini *et al.*, *J. Biol. Chem.* 264, 8052 (1989); M. Kataoka, J. F. Head, A. Persechini, R. H. Kretsinger, D. M. Engelman, *Biochemistry* 30, 1188 (1991).
35. K. T. O'Neill, S. Erickson-Viitanen, W. F. DeGrado, *J. Biol. Chem.* 264, 14571 (1989).
36. T. J. Lukas, W. H. Burgess, F. G. Prendergast, W. Lau, D. M. Watterson, *Biochemistry* 25, 1458 (1986).
37. D. K. Blumenthal and E. G. Krebs, *Methods Enzymol.* 139, 115 (1987).
38. M. Kataoka, J. F. Head, T. Vorherr, J. Krebs, E. Carafoli, *Biochemistry* 30, 6247 (1991).
39. M. J. Hubbard and C. B. Klee, *ibid.* 28, 1867 (1989); D. Guerini and C. B. Klee, *Adv. Protein Phosphatases* 6, 391 (1991).
39. B. Kemp, R. Pearson, V. Guerriero, I. Bagchi, A. Means, *J. Biol. Chem.* 262, 2542 (1987).
40. M. Dasgupta, T. Honeycutt, D. K. Blumenthal, *ibid.* 264, 17156 (1989).
41. J. Trehwella, D. K. Blumenthal, S. E. Rokop, P. A. Seeger, *Biochemistry* 29, 9316 (1990).
42. H. Charbonneau *et al.*, *ibid.* 30, 7931 (1991); J. P. Novack, H. Charbonneau, J. K. Bentley, K. A. Walsh, J. A. Beavo, *ibid.*, p. 7940.
43. M. K. Bennett and M. B. Kennedy, *Proc. Natl. Acad. Sci. U.S.A.* 84, 1794 (1987); R. M. Hanley *et al.*, *Science* 237, 293 (1987).
44. A. P. Kwiatkowski and M. M. King, *Biochemistry* 28, 5380 (1989).
45. L. E. Kay, D. Marion, A. Bax, *J. Magn. Reson.* 84, 72 (1989).
46. G. Zhu and A. Bax, *ibid.* 90, 405 (1990).
47. D. S. Garrett, R. Powers, A. M. Gronenborn, G. M. Clore, *ibid.* 95, 214 (1991).
48. B. R. Brooks *et al.*, *Comput. Chem.* 4, 187 (1983).
49. M. Carson, *J. Mol. Graphics* 5, 103 (1987).
50. D. K. Blumenthal *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 82, 3187 (1985).
51. R. B. Pearson *et al.*, *Science* 241, 970 (1988); C. J. Foster, *Arch. Biochem. Biophys.* 280, 397 (1990).
52. A. K. Verma *et al.*, *J. Biol. Chem.* 263, 14152 (1988).
53. T. Ono, G. T. Slaughter, R. G. Cook, A. T. Means, *ibid.* 264, 2081 (1989).
54. R. L. Kincaid, M. S. Nightingale, B. R. Martin, *Proc. Natl. Acad. Sci. U.S.A.* 85, 8983 (1988); C. B. Klee, unpublished data.
55. D. A. Malencik and S. R. Anderson, *Biochem. Biophys. Res. Commun.* 114, 50 (1983).
56. Supported by the AIDS Targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health (G.M.C., A.M.G., and A.B.). We thank K. Bechingham and J. Maune for providing us with the *Drosophila* calmodulin coding construct.

31 January 1992; accepted 27 March 1992