Calcium-Induced Interactions of Calmodulin Domains Revealed by Quantitative Thrombin Footprinting of Arg37 and Arg106†

Madeline A. Shea,* Amy S. Verhoeven, and Susan Pedigo

Department of Biochemistry, University of Iowa College of Medicine, Iowa City, Iowa 52242-1109

Received August 16, 1995; Revised Manuscript Received January 8, 1996⊗

ABSTRACT: Calcium-dependent conformational states of calmodulin (CaM) were probed by thrombin to determine quantitative differences in the susceptibility of two bonds: Arg37−Ser38 (R37−S38, near site I in the N-terminal domain) and Arg106−His107 (R106−H107, near site III in the C-terminal domain). Quantitative thrombin footprinting of a discontinuous equilibrium calcium titration of wild-type calmodulin showed that the R37−S38 bond of the apoprotein was cleaved at a barely detectable level while the R106−H107 bond was maximally susceptible. Calcium binding to sites III and IV monotonically protected R106−H107 from proteolysis; concomitantly, the susceptibility of R37−S38 increased. However, calcium binding to sites I and II protected R37−S38 from cleavage, yielding a peaked biphasic profile composed of equal and opposite transitions. Both bonds were fully protected when calmodulin was saturated with calcium. Susceptibility profiles resolved from the fractional abundance of primary cleavage products (peptides 1-37, 38-148, 1-106, 107-148) were interpreted as directly reflecting calcium-induced conformational changes in whole calmodulin; free energies of calcium binding and cooperativity were estimated. Secondary cleavage was never observed; both R37 and R106 were sites of thrombinolysis in whole calmodulin only. In studies of E140Q-CaM (having a mutation in site IV), the susceptibility of R37−S38 decreased monotonically. Thus, the biphasic character of cleavage of R37 in helix B was not intrinsic to that domain but depended on propagation of effects of calcium-induced changes in the C-terminal domain. The observed patterns of susceptibility indicated that partially saturated wild-type calmodulin adopts at least one intermediate conformation whose structure is determined by calcium-mediated interactions between the domains.

Calmodulin (CaM)[1] serves as a ubiquitous eukaryotic intracellular calcium receptor (Klee & Vanaman, 1982; Watterson *et al*., 1976; Vanaman, 1980; Weinstein & Mehler, 1994). The molecular mechanism of calcium-induced conformational switching of CaM is of paramount physiological importance to a panoply of enzymatic targets activated by this small acidic protein (Meador *et al*., 1993; Török & Whitaker, 1994). As indicated in Figure 1a, its sequence is comprised of four similar segments configured in an α-*β*-α′-*β*′ pattern; each segment contains a metal binding site of the EF-hand type (Kretsinger, 1976; McPhalen *et al*., 1991).

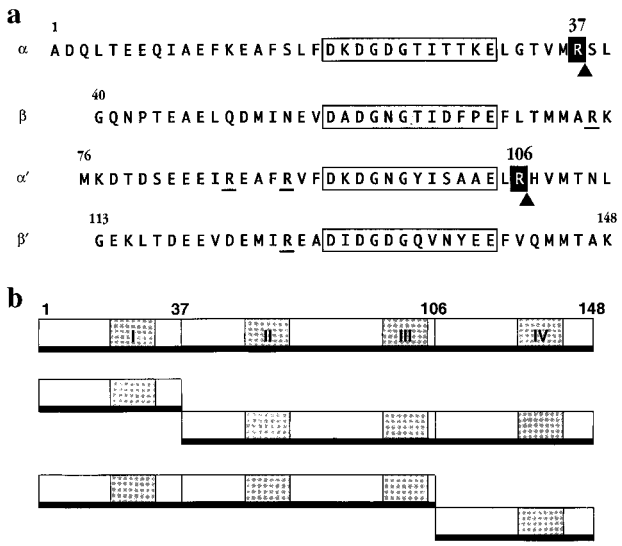


FIGURE 1: (a) Amino acid sequence of rat calmodulin. The 12 residues of each calcium binding site (boxed) are aligned. The sequence was divided into four segments: α (residues 1−39) includes site I; *β* (40−75) includes site II; α′ (76−112) includes site III; *β*′ (113−148) includes site IV. Arginine residues are underlined (no cleavage observed) or shaded (susceptible to thrombin). (b) Schematic representation of products isolated by HPLC after quantitative thrombin footprinting of calmodulin. Products were uncut calmodulin (residues 1−148) and four primary cleavage products (peptides 1−37, 38−148, 1−106, 107−148). The calcium binding sites are shaded.

Cooperative binding of calcium to the sites of vertebrate CaM (Crouch & Klee, 1980) induces large conformational

† These studies were supported by grants to M.A.S. from the American Heart Association (910148980), by a National Science Foundation Presidential Young Investigator Award (NSF DMB 9057157), and by the NIH Diabetes and Endocrinology Research Center (DK 25295). S.P. was supported by an NIH Predoctoral Traineeship in Biotechnology (PHS 1 T32 GM08365-04).

* Corresponding author. Telephone: (319) 335-7885. E-mail: madeline-shea@uiowa.edu.

⊗ Abstract published in *Advance ACS Abstracts,* February 15, 1996.

[1] Abbreviations: BAPTA, 1,2-bis(*O*-aminophenoxy)ethane-*N*,*N*,*N*′,*N*′-tetraacetic acid; CaM, calmodulin; CD, circular dichroism; EGTA, ethylene glycol bis(*β*-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid; EndoGluC, endoproteinase GluC; FFRCK, D-phenylalanylphenylalanylarginyl chloromethyl ketone; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; HPLC, reverse-phase high-performance liquid chromatography; IPTG, isopropyl *β*-D-thiogalactopyranoside; NMR, nuclear magnetic resonance; NTA, nitrilotriacetic acid; pCa, $-RT \ln [Ca^{2+}]_{free}$; SD, standard deviation; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; T-GPR-NA, tosylglycyl-prolylarginine-4-nitroanilide acetate; TFAA, trifluoroacetic acid; TFP, trifluoperazine.

changes (LaPorte *et al*., 1980; Klee, 1977). Both the number and position of calcium ions bound to CaM may regulate its interactions with targets (Haiech *et al*., 1981; Wang & Sharma, 1980; Klee, 1988). Functional evidence suggests that apo-CaM also may play an important role in some cells (Geiser *et al*., 1991) and engage in specific interactions with target enzymes [e.g., neuromodulin (Gerendasy *et al*., 1995)]. Structural and computational studies of CaM offer models of the conformational states that are populated upon calcium binding [cf. Weinstein and Mehler (1994)] and have suggested mechanisms for the transition between the two end states of CaM: apo and calcium-saturated ($Ca^{2+}{}_4$-CaM).

Crystallographic (Babu *et al*., 1988; Taylor *et al*., 1991; Chattopadhyaya *et al*., 1992) and NMR studies (Ikura *et al*., 1991; Barbato *et al*., 1992) of $Ca^{2+}{}_4$-CaM have shown that it has two distinct half-molecule domains (N-terminal and C-terminal) with nearly identical backbone structures; each has a contiguous pair of interacting Ca^{2+} binding sites. In the crystallographic studies, a long "central helix" was evident between sites II and III, giving CaM a dumb-bell shape as shown in Figure 2a (Babu *et al*., 1988). However, crystallization conditions have been shown to promote helix formation (Bayley & Martin, 1992; Török *et al*., 1992), and NMR studies indicated that residues 78−81 are generally disordered in solution (Ikura *et al*., 1991; Barbato *et al*., 1992). Other solution studies (Heidorn & Trewhella, 1988; Persechini & Kretsinger, 1988; Small & Anderson, 1988; Török *et al*., 1992; Yao *et al*., 1994; Persechini *et al*., 1993) and computational modeling (Pascual-Ahuir *et al*., 1991; see Figure 2b) indicate that $Ca^{2+}{}_4$-CaM can adopt conformations that are more compact than those observed in crystallagraphic studies, consistent with the interpretation that this region serves as a flexible tether (Persechini & Kretsinger, 1988; Kretsinger, 1992). Structural studies of co-complexes of CaM with a peptide (Ikura *et al*., 1992; Seeholzer & Wand, 1989; Meador *et al*., 1992; Rao *et al*., 1992; Porumb *et al*., 1994) or with a drug (Cook *et al*., 1994; Vandonselaar *et al*., 1994) indicate that in the presence of a bound target or antagonist, the two domains of CaM are in closer proximity than observed for CaM alone.

The structure of apo-CaM (and its isolated domains) has been characterized as being generally more flexible and less extended than that of fully saturated CaM (Hoffman & Klevit, 1991; Finn *et al*., 1993; Drabikowski & Brzeska, 1982; Heidorn & Trewhella, 1988; Török *et al*., 1992). In addition, recent NMR studies by Wand and co-workers (Urbauer *et al*., 1995) have shown that some critical structural features of both the secondary and tertiary structure of $Ca^{2+}{}_4$-CaM are missing in apo-CaM under low-salt conditions.

There are no high-resolution structural studies of intermediate (partially saturated) states of CaM. However, stoichiometric calcium binding studies monitored by NMR (Evans *et al*., 1988; Seamon, 1980; Starovasnik *et al*., 1992), CD and fluorescence (Klevit, 1983; Maune *et al*., 1992a; Wang *et al*., 1984; Yao *et al*., 1994; Anderson, 1991), and equilibrium binding studies using quantitative proteolytic footprinting (Pedigo & Shea, 1995a; Verhoeven & Shea, 1993) have provided evidence for at least one intermediate conformation. Studies of calcium binding using equilibrium dialysis (Crouch & Klee, 1980) and others monitoring the kinetics of calcium dissociation from CaM, its fragments, and mutants (Bayley *et al*., 1984; Martin *et al*., 1985; Martin *et al*., 1992) support this hypothesis. The kinetic studies of mutants defective in calcium binding also suggested that far-

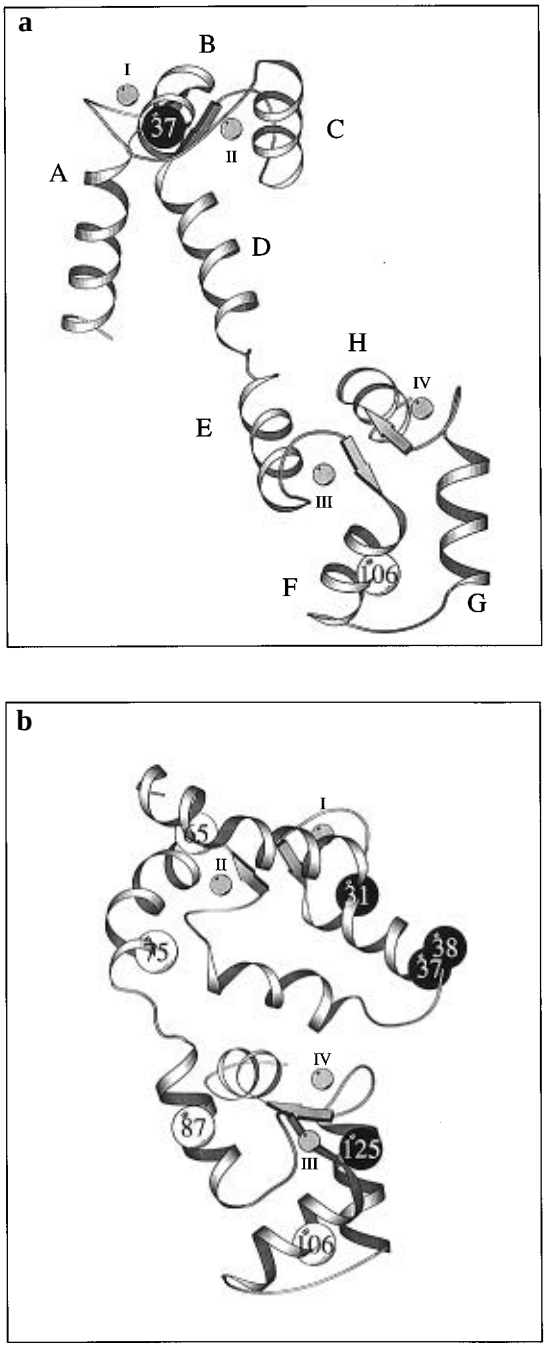
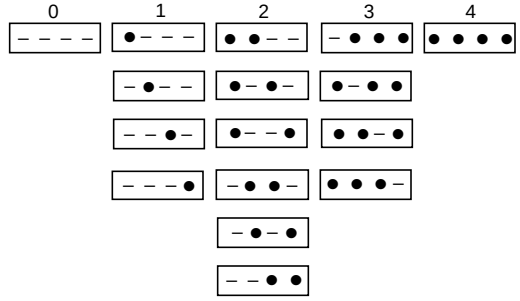


FIGURE 2: (a) Ribbon drawing (MolScript; Kraulis, 1991) of the α carbon backbone of a crystallographic structure of $Ca^{2+}{}_4$-CaM (Babu *et al*., 1988) in which residues 5−147 were resolved. Coordinates were taken from the Brookhaven Protein Data Bank file 3cln.pdb (Bernstein *et al*., 1977; Abola *et al*., 1987). Arginine residues 37 and 106, the only positions where thrombin cleaved calmodulin, are indicated; eight helices are designated with capital letters A−H. Secondary structure type of the D−E junction is based on NMR studies of $Ca^{2+}{}_4$-CaM in solution (Ikura *et al*., 1991). The center-to-center distance between the Cα carbon of R37 and the Ca^{2+} in site III is 37.4 Å; the distance between the Cα carbon of R37 and the Ca^{2+} in site IV is 36.6 Å. (b) Ribbon drawing (MolScript; Kraulis, 1991) of the α carbon backbone of CAM10, an energy-minimized model of $Ca^{2+}{}_4$-CaM (Pascual-Ahuir *et al*., 1991). Spheres indicate positions of calcium-dependent susceptibility to quantitative proteolytic footprinting [E31 (EndoGluC), R37 (thrombin), S38 (bromelain), F65 (chymotrypsin), K75 (bromelain), E87 (EndoGluC and bromelain), R106 (thrombin), I125 (bromelain)]. The shading of the sphere indicates the nature of the response: white (monotonic) or black (biphasic). The distance between the Cα carbon of R37 and the Ca^{2+} in site III is 24.7 Å, and between the Cα carbon of R37 and the Ca^{2+} in site IV is 15.8 Å.

ranging effects control the order of binding to the four sites.

These studies of CaM alone and in complexes with targets or antagonists answer many questions about molecular recognition by this important calcium receptor. However, they do not explain fully how calcium binding activates CaM to participate in such interactions. To understand the pathway of activation, it is essential to study the partially saturated species of CaM (having 1, 2, or 3 ligands) as well as the two end states (0 or 4 ligands). The intermediate ligation levels are comprised of 14 possible ligation forms differing in the number and position of vacant or occupied calcium binding sites. In the schematic diagram below, the top line indicates the number of calcium ions bound. For each species, − indicates a vacant site and ● indicates a site occupied by calcium; [− − − −] represents apo-CaM and [● ● ● ●] represents Ca^{2+}_4-CaM. Any of the 16 species may have additional conformational degeneracy.



In order to determine the relative abundance and functional significance of these intermediate ligation states, it is necessary to determine their energetic and structural properties directly. As has been demonstrated for several allosteric systems, development of approaches to monitor individual residues is essential for identifying the number and nature of intermediate states in any ligand binding pathway (Holt & Ackers, 1995; Heyduk & Lee, 1989; Ackers *et al*., 1983). It may be possible to dissect such a system into its component subunits or domains for comparison as reference states. Although pathways of cooperative interaction are severed in such analyses, observed differences may highlight the mechanistic role of altered interfaces or loops [cf. Powers *et al*. (1993), Chen *et al*. (1994), and LiCata and Ackers (1995)].

The focus of this study of CaM was to probe the molecular mechanism of calcium-induced switching and the nature of intermediate conformational states adopted by partially saturated CaM in the absence of targets. Previous studies of proteolysis of CaM by trypsin and clostripain indicated that several of the arginine residues in CaM underwent calcium-dependent conformational changes (Mackall & Klee, 1991; Walsh *et al*., 1977; Drabikowski *et al*., 1977; Drabikowski & Brzeska, 1982; Guerini & Krebs, 1985; Manalan *et al*., 1985). Studies of thrombin cleavage of apo-CaM showed that the single bond R106−H107 was susceptible (Wall *et al*., 1981). Thus, thrombin has been applied by many investigators to generate complemenatary peptides that contain either sites I, II, and III of CaM or only site IV (Figure 1a). These peptides have been used for structure−function analysis of CaM, particularly in comparison with the half-molecules generated by trypsinolysis near R74 (Tsalkova & Privalov, 1985; Gryczynski *et al*., 1991; Drabikowski & Brzeska, 1982). To our knowledge, there is no previous report of thrombin cleaving the bond between R37 and S38 of CaM.

To dissect the pathways of ligand-linked conformational changes, we have developed quantitative proteolytic footprinting methods for probing cooperative calcium binding to CaM under *equilibrium* conditions using several specific and nonspecific proteases. This study is the first to provide a quantitative simultaneous assessment of the calcium-dependent changes in susceptibility of the bonds R37−S38 (adjacent to site I) and R106−H107 (adjacent to site III) in CaM under equilibrium calcium binding conditions. For each position, the complementary primary cleavage products (Figure 1b) were monitored. The susceptibility profiles demonstrated the existence of an intermediate conformation induced by interactions between the two domains, consistent with other proteolytic footprinting studies using EndoGluC (Pedigo & Shea, 1995a) and bromelain (Verhoeven and Shea, data not shown). The free energy of calcium binding and intradomain cooperative interactions are estimated from these susceptibility profiles, and elements of method development are described.

EXPERIMENTAL PROCEDURES

Materials

Thrombin (EC 3.4.21.5) was purchased from Calbiochem (San Diego, CA; 605 190). The substrate T-GPR-NA (Boehringer Mannheim Biochemicals, Indianapolis, IN) was used for testing thrombin activity, and FFRCK (Calbiochem) was used as a protease inhibitor. SDS−PAGE molecular weight markers (SDS-17) were obtained from Sigma (St. Louis, MO). Other reagents were of the highest grade commercially available. Recombinant rat CaM was overexpressed bacterially, purified, characterized, and stored as described previously (Pedigo & Shea, 1995a). The mutant E140Q-CaM was prepared using cassette mutagenesis to substitute the codon GAG (glutamine) for CAG (glutamate) at residue 140 (position 12 in site IV) using standard procedures (Ausubel *et al*., 1991); bacterially expressed E140Q-CaM was purified by phenyl-Sepharose chromatography (Putkey *et al*., 1985) with 10 mM $CaCl_2$ in calcium-containing elution buffers (Hutchins, 1994).

Methods

Quantitative Footprinting Reactions. To obtain a reproducible measure of the calcium-induced change in the relative susceptibility of CaM to thrombin, exposure conditions were designed to provide an optimal abundance of products (i.e., high signal-to-noise ratio) without perturbing the equilibrium distribution of calcium bound to CaM. Thrombin activity was determined to be calcium-independent by monitoring the formation of the chromogenic portion of cleaved T-GPR-NA (change in absorbance at 405 nm per minute) over the range of free calcium concentrations used in the footprinting reactions.

FFRCK was tested for its ability to inhibit cleavage of T-GPR-NA by thrombin; addition of FFRCK at final concentrations of 0.02, 0.1, or 0.2 mM completely quenched the activity of thrombin (1.4 μM) within 10 min. In a separate experiment, FFRCK (0.1 mM) was used to quench proteolysis of CaM by thrombin after 15 min of proteolysis at pCa 6 or 7.7; products present at 0, 2, and 19 h after quenching were compared. No additional cleavage of CaM was apparent 2 h after quenching with FFRCK; there was a low degree ($\leq$5%) of additional cleavage of whole CaM at

19 h after quenching. In the thrombin footprinting reactions described below, products were separated on the HPLC within 10 h after addition of FFRCK; therefore, this small degree of additional proteolysis after 19 h was regarded to be a source of tolerable imprecision in the data.

The total time of exposure to thrombin was varied to determine its effect on the nature and abundance of proteolytic products; exposures over periods ranging between 15 and 60 min were found to yield the same four primary cleavage products at ratios that varied less than 10% (and were closer to 3% in most cases). Given that no secondary cleavage products were detected, the length of exposure to thrombin did not affect the relative abundance of complementary primary cleavage products; therefore, the 60 min time period was chosen to enhance the precision of peak integration.

A discontinuous calcium titration of CaM was conducted by preparing aliquots of CaM that each had been dialyzed exhaustively against one of a series of buffers (50 mM HEPES, 92 ± 3 mM KCl, 0.5 mM EGTA, 0.5 mM NTA, pH 7.40 ± 0.01 at 22.0 °C; variable $CaCl_2$) whose experimentally determined free calcium concentrations spanned a range from nanomolar to millimolar calcium. Samples of CaM used in this titration were aliquots of those studied previously by quantitative EndoGluC footprinting (Pedigo & Shea, 1995a) and 1-D ^{1}H-NMR (Pedigo & Shea, 1995b) of CaM. In the thrombin footprinting reactions (total volume of 140 μL), 6−12 μM CaM (14−28 μg, depending on the dialysate) was subjected to limited proteolysis by thrombin (stock solution 0.1 unit/μL; final concentration 0.005 unit/μL) at 22 °C. After 60 min of exposure to thrombin, proteolysis was quenched by addition of 5 μL of 10 mM FFRCK (final concentration 0.34 mM).

SDS−polyacrylamide gel electrophoresis (SDS−PAGE) was used for semiquantitative evaluation of the products of a thrombin footprinting study of calcium binding to wild-type CaM and E140Q-CaM under conditions similar to the discontinuous calcium titration described above. For these titrations, highly concentrated CaM stock solutions with no added calcium were diluted into buffers of defined pCa (50 mM HEPES, 100 mM KCl, 0.5 mM EGTA, 0.5 mM NTA, pH 7.40, 22 °C, variable $CaCl_2$). The duration of thrombin footprinting reactions was 15 min, and cleavage was quenched by addition of SDS-containing loading buffer. Peptide products of wild-type CaM were separated using a Tris−Tricine electrophoresis procedure (Schagger & von Jagow, 1987) with a resolving gel having a T:C ratio of 15.7%:6.3% and no spacer gel; they were detected by staining with silver nitrate. Peptide products of E140Q-CaM were separated on a Tris−glycine gel having a 17−20% gradient of acrylamide.

Two independent series of thrombin footprinting studies of equilibrium titrations of wild-type CaM were analyzed quantitatively using HPLC instrumentation as described previously (Pedigo & Shea, 1995a). For each titration, 2 volumes of 50 μL (10 μg of CaM) from a footprinting reaction at each pCa were injected onto an Aquapore RP-300 (C8) column (22 cm × 2.1 mm i.d.; Applied Biosystems, Foster City, CA) in solvent A (water with 0.06% TFAA) at a flow rate of 0.2 mL/min. Primary cleavage products were eluted with the following gradient: 35−42% B (80:20 acetonitrile:water ratio with 0.06% TFAA) in 9 min, followed by 42−56% B in 28 min. Proteolytic products were collected from the HPLC effluent; amino acid analysis was performed by the Protein Structure Facility at the University of Iowa College of Medicine.

Susceptibility profiles for both cleavage positions (bonds R37−S38 and R106−H107) were resolved from the fractional abundance of each pair of primary peptide products as measured by the integrated peak area monitored by A_{220} (absorbance at 220 nm). The dual-wavelength detector also monitored A_{280}, for corroboration of the identity of tyrosine-containing peptides and independent quantitation of relative peak areas. The quantitative conclusions reached in this study represent analysis of susceptibility profiles from duplicate analyses of two independent footprinting reactions. Equivalent susceptibility profiles were obtained in more than 20 titrations analyzed semiquantitatively by SDS−PAGE electrophoresis.

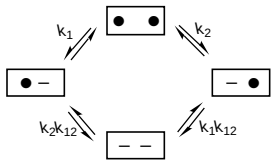
Resolution of Proteolytic Susceptibility Profiles. Chromatographic profiles were integrated manually, and the peak area (integrated absorbance at 220 nm) was determined for each CaM-derived peptide in the mixture of products from a footprinting reaction. The fractional area of each peptide, A_i, was determined by eq 1 where the denominator represents

$$A_i = \frac{\text{peak area of peptide}_i}{\text{peak area of CaM} + \sum_{j=1,4} \text{peak area of peptide}_j} \quad (1)$$

the sum of areas corresponding to whole CaM and all four proteolytic fragments (i.e., it represents the total mass of CaM in the footprinting reaction). For each titration, pairs of determinations of A_i for a given peptide resulted from analysis of replicate HPLC separations.

Susceptibility profiles were generated by treating A_i as proportional to the fractional abundance of each peptide. This was justified on the basis of Beer's Law (i.e., by assuming that the value of A_{220} was proportional to the number of peptide bonds in a fragment) and experimentally supported by comparing the observed ratio of peak areas to the calculated ratio of number of peptide bonds of primary cleavage products. For cleavage at R106−H107 (fragments of length 42 and 106), the expected ratio of the number of peptide bonds was 39% (41/105). Experimental values deviated by less than 1% from that value between pCa 9 and 5.5 (at lower pCa, noise in the integration of small peaks distorted the ratio). For cleavage at R37−S38, the absolute probability of cutting was lower than at R106−H107 (the maximal areas differed by more than 5-fold); thus, the signal-to-noise ratio was lower (less favorable) at all pCa values. The expected ratio of peak areas was 32.7% (36/110). The observed values were within 5% of that ratio between pCa 7 and pCa 5.

Data Analysis. As a simplifying assumption, the domains were treated as independent binding units, each having two sites for calcium. Normalized susceptibility was treated as being proportional to fractional saturation by calcium; assumptions and limitations inherent in this approach were discussed previously (Pedigo & Shea, 1995a). The isotherms were analyzed according to a standard linkage scheme for a macromolecule binding two ligands (Ackers *et al*., 1983), as given below. The microscopic association equilibrium constants are in lower case; k_1 and k_2 represent intrinsic binding constants for calcium, and k_{12} represents the cooperativity between two sites in a domain. In the diagram below, − indicates a vacant site, and ● indicates an occupied site.



The total fractional saturation of this two-site system ($\bar{Y}_t$ or fraction of moles of sites occupied) is given by the average of the fractional saturation of each of its sites:

$$\bar{Y}_t = \frac{\bar{Y}_1}{2} + \frac{\bar{Y}_2}{2} \tag{2}$$

where the individual-site isotherm $\bar{Y}_1$ represents the fraction of moles of site 1 occupied, and $\bar{Y}_2$ represents the fraction of moles of site 2 occupied [cf. Ackers *et al*. (1983)]. As given by eq 3, the individual-site isotherms may be expressed in terms of the intrinsic association binding constants (k_1 and k_2) and k_{12}, the cooperativity between the sites.

$$\bar{Y}_1 = \frac{k_1[X] + k_1k_2k_{12}[X]^2}{1 + (k_1 + k_2)[X] + k_1k_2k_{12}[X]^2} \tag{3a}$$

$$\bar{Y}_2 = \frac{k_2[X] + k_1k_2k_{12}[X]^2}{1 + (k_1 + k_2)[X] + k_1k_2k_{12}[X]^2} \tag{3b}$$

These relationships simplify in the case of a ligand binding to sites that are independent (noncooperative or $k_{12} = 0$) and also homogeneous (having equal intrinsic affinities for ligand or $k_1 = k_2$). In that case, the total average fractional saturation ($\bar{Y}_t$) of both sites in a domain is equivalent to both $\bar{Y}_1$ and $\bar{Y}_2$ and is given by a simple Langmuir binding isotherm (model I) in eq 4:

$$\bar{Y}_I = \frac{K[X]}{1 + K[X]} \tag{4}$$

In this model, K represents the intrinsic association constant ($k = k_1 = k_2$) for a calcium binding site and [X] is the concentration of free calcium in the dialysate. The total free energy (ΔG_2) of calcium binding to both sites in such a domain is given by $-2RT \ln K$.

Using NONLIN (Johnson & Frasier, 1985) as described previously (Pedigo & Shea, 1995a), susceptibility profiles were fit to both model I (eq 4) and model II (eq 5) which allowed heterogeneous (unequal intrinsic affinities) and cooperative binding to two sites in a domain (Adair equation for two sites).

$$\bar{Y}_{II} = \frac{K_1[X] + 2K_2[X]^2}{2(1 + K_1[X] + K_2[X]^2)} \tag{5}$$

The macroscopic equilibrium constant K_1 represents the the sum of two intrinsic microscopic equilibrium constants ($k_1 + k_2$) that are not necessarily equal. The parameter K_2 represents the equilibrium constant ($k_1k_2k_{12}$) for binding ligand to both sites; it accounts for any positive or negative cooperativity regardless of its source or magnitude. It is not possible to determine k_{12}, the intradomain cooperativity constant, analytically from these data alone; it may be estimated by assuming that the binding sites have equal intrinsic affinities ($k_1 + k_2 = 2k$). This leads to the expression of K_c, the apparent cooperativity constant, as given in eq 6; the value of K_c provides a lower limit for the value

$$K_c = \frac{4K_2}{K_1^2} \tag{6}$$

of k_{12}, the actual microscopic cooperativity constant. The compensatory contributions of cooperative free energy and heterogeneous binding energies are such that for the *same* cooperative interaction (i.e., k_{12} constant), the value of K_c will increase if there is a decrease in heterogeneity between sites. For example, if the ratio of k_1/k_2 is 10, K_c is $0.33k_{12}$, but for a k_1/k_2 of 2, K_c is $0.89k_{12}$ [see Pedigo and Shea (1995b)]. The value of K_c cannot indicate positive cooperativity if there is none and is necessarily <1 (i.e., indicates apparent anticooperativity) if the sites are noncooperative (i.e., $k_{12} = 1$) and also heterogeneous.

For ease of comparison of susceptibility profiles, data for each peptide were normalized first to the highest and lowest experimentally determined values of A_i. To account for finite variations in the asymptotes of the susceptibility profiles for different peptides, the function [$f(X)$] used for nonlinear least-squares anlaysis of the fractional abundance of each peptide was given by

$$f(X) = Y_{[X]_{low}} + \bar{Y}_t(\text{Span}) \tag{7}$$

where $\bar{Y}_t$ refers to the average fractional saturation as described by model I (eq 4) or II (eq 5). Note that the value of the parameter $Y_{[X]_{low}}$ corresponds to the value of the dependent variable (proteolytic susceptibility) at the lowest calcium concentration of the titration being fit and that the value of the parameter Span is negative for a monotonically decreasing signal. At the limit of saturating ligand concentration ($\bar{Y}_t = 1$), the experimental signal is equal to the sum of these two parameters. Values of both $Y_{[X]_{low}}$ and Span were determined simultaneously with free energies in most cases (see Tables 2 and 3). Data for complementary primary cleavage products (Figure 1b) were fit separately and as combined sets.

Due to the small separation between the median ligand concentrations of the two phases of the biphasic susceptibility profile for R37, it was not possible to fit both phases simultaneously. Thus, the data describing the susceptibility of R37 (abundance of peptides 1−37 and 38−148) were separated into two overlapping sets. Both sets included data for the dialysates at 1.9×10^{-6} M and 2.4×10^{-6} M free calcium, providing a constraint for continuity. The phase of induced susceptibility spanned a range from pCa 8.79 (1.6×10^{-9} M) to 5.61 (2.4×10^{-6} M) while the phase of induced protection ranged from pCa 5.71 (1.9×10^{-6} M) to 3.47 (3.4×10^{-4} M).

Multiple criteria were used for evaluating goodness-of-fit for the parameters reported for each case. As reported by NONLIN, these error statistics included (a) the value of the square root of variance, (b) the values of asymmetric 65% confidence intervals, (c) the distribution of residuals, (d) the magnitude of the span of residuals, and (e) the absolute value of elements of the correlation matrix. Values of ΔG_c ($-RT \ln K_c$) and its propagated confidence interval were estimated by NONLIN using the best-fit values for ΔG_1 and ΔG_2 determined in individual fits to model II (eq 5).

RESULTS

The motivation for these studies was to understand the molecular mechanism of calcium-induced switching of CaM.

Table 1: Amino Acid Analysis of Peptides of Calmodulin Generated by Thrombinolysis

amino acid[a]	A, 1−37[b]	A′, 38−148[c]	B′, 1−106[d]	B, 107−148[e]
Asx	4.0 (4)	19.2 (19)	16.2 (16)	6.3 (7)
Thr	4.8 (5)	7.1 (7)	8.7 (9)	3.0 (3)
Ser	1.1 (1)	2.7 (3)	3.8 (4)	0.1 (0)
Glx	6.9 (7)	20.7 (20)	18.4 (18)	9.1 (9)
Pro	0.0 (0)	1.9 (2)	2.0 (2)	0.0 (0)
Gly	3.1 (3)	8.3 (8)	8.1 (8)	3.2 (3)
Ala	3.0 (3)	8.4 (8)	9.1 (9)	2.3 (2)
Val	1.0 (1)	6.2 (6)	3.2 (3)	4.0 (4)
Ile	1.9 (2)	6.0 (6)	5.8 (6)	2.1 (2)
Leu	3.0 (3)	6.4 (6)	7.2 (7)	2.2 (2)
Tyr	0.0 (0)	1.9 (2)	1.0 (1)	0.9 (1)
Phe	2.9 (3)	5.2 (5)	6.9 (7)	1.1 (1)
His	0.0 (0)	0.8 (1)	0.0 (0)	0.5 (1)
Lys	3.0 (3)	4.7 (5)	5.4 (6)	1.2 (2)
Arg	4.0 (4)	4.6 (5)	4.6 (5)	0.7 (1)

[a] Standard three-letter abbreviations for amino acids except Glx represents the sum of glutamine and glutamic acid and Asx represents asparagine and aspartic acid; expected frequency is in parentheses. Raw data were normalized against expected frequency of several amino acids; average values are reported. [b] Normalized against Arg, Leu, Asx, Glx; SD ≤ 0.05. [c] Normalized against Tyr, Ile, Ala, Val, Lys; SD ≤ 0.2 except for Glx, Asx. [d] Normalized against Tyr, Ala, Val, Ser; SD ≤ 0.4 except for Glx, Asx. [e] Normalized against Tyr, Val; SD ≤ 0.03.

The quantitative thrombin footprinting method developed and applied here is similar to one reported previously using EndoGluC (Pedigo & Shea, 1995a) to probe calcium-dependent structural changes in CaM and indicated that the domains interact in the absence of target peptides or antagonists.

Cleavage by Thrombin. Thrombin cleaved selectively at the C-terminal side of two of six arginine residues in CaM (see Figure 2a). As indicated in Figure 1b, the bonds R37−S38 and R106−H107 were susceptible in whole CaM but not in its fragments (i.e., the single possible secondary cleavage product comprised of residues 38−106 was never detected). Cleavage after the other four arginine residues in CaM (positions 74, 86, 90, and 126; see Figure 1a) was not observed at any calcium concentration or period of proteolysis tested.

The four proteolytic products (two pairs of primary cleavage products) of thrombin proteolysis of CaM were identified by amino acid analysis of fractions separated by HPLC. In Table 1, the determined composition of each peptide is given as the average of experimental data normalized to the expected frequency of several amino acid residues (usually including tyrosine, isoleucine, alanine, and valine because these were quantified most reliably). As expected for this type of chemical analysis, the experimental evaluations of acidic and basic amino acids deviated the most from their expected abundance in each peptide. The amino acid compositions corresponded unambiguously to peptides containing residues 1−37, 38−148, 1−106, and 107−148 indicated schematically in Figure 1b.

Semiquantitative Analysis of Thrombin Footprinting of CaM. Figure 3a shows an SDS−PAGE analysis of the products of thrombin footprinting of a calcium titration of wild-type CaM. The frequency of cleavage at R106−H107 (as judged by the abundance of peptides 1−106 and 107−148) was reduced by calcium binding to CaM. The susceptibility of the bond R37−S38 was lower than that of R106−H107 at all calcium concentrations. Cleavage of R37 yielded two primary proteolytic products (1−37, 38−148) as quantified by HPLC and characterized by amino acid

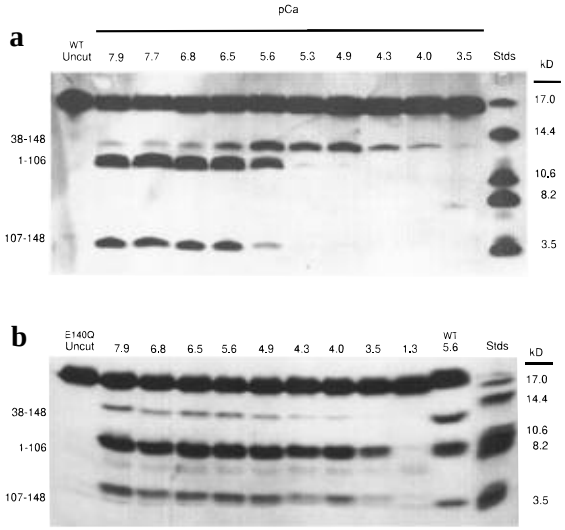


FIGURE 3: SDS−PAGE analysis of a thrombin footprinting of calcium titrations of calmodulin. Conditions were 50 mM HEPES, 100 mM KCl, 0.5 mM NTA, 0.5 mM EGTA, pH 7.40, 25 °C; pCa values are noted above lanes. (a) Wild-type CaM at 11.4 μM was exposed to thrombin at 0.01 unit/μL final concentration for 15 min; lanes 2−11 show products detected by staining with silver nitrate. Lane 1 shows uncut CaM; lane 12 represents molecular mass standards with sizes annotated. (b) E140Q-CaM in lanes 1−10; lane 1 shows uncut protein. Lane 11 contains products of thrombin footprinting of wild-type CaM at pCa 5.6 for comparison of the migration distance of peptides. Lane 12 represents molecular mass standards with sizes annotated.

analysis; however, peptide 1−37 was not resolved by SDS−PAGE. The abundance of fragment 38−148 was biphasic, indicating that R37−S38 was protected from cleavage at low (nanomolar) and high (millimolar) calcium concentrations but was susceptible at intermediate (micromolar) levels. Because of the nonlinear staining of peptides with silver nitrate and the lack of detection of peptide 1−37, the SDS−PAGE experiments were used only for semiquantitative assessment of the products of thrombin footprinting reactions of CaM. However, Figure 3a shows clearly that the pCa required for half-maximal protection of R106−H107 was close to that required for the half-maximal induced susceptibility of R37−S38. This indicated that calcium binding to the higher affinity sites (III and IV) affected the conformation of the N-terminal domain.

A mutant of *Drosophila* CaM (B4Q) containing a single substitution (Gln in place of the bidentate Glu at position 12 in site IV) has significantly diminished affinity for calcium at the C-terminal sites and enhanced affinity at the N-terminal sites (Martin *et al*., 1992; Maune *et al*., 1992b). A corresponding mutant of rat CaM (E140Q-CaM) was constructed to test whether the calcium-induced biphasic response of R37 (a) reflected intrinsic properties of the wild-type N-terminal domain in holo-CaM or (b) resulted from interactions between the two domains. On the basis of the size and calcium-dependent mobility shifts of the proteolytic products, thrombin was judged to cleave E140Q-CaM at the same two positions cleaved in wild-type CaM.

Thrombin footprinting studies of apo-E140Q-CaM (see Figure 3b) showed that the susceptibility of R37 was much lower than that of R106 as had been observed for wild-type CaM. However, both scissile bonds were protected monotonically by calcium binding. The relative values of the medians estimated from the susceptibility profiles suggested that sites I and II of E140Q-CaM have an affinity for calcium

that is equal to or higher than that of sites III and IV. This contrasts with the behavior of wild-type CaM and is consistent with NMR studies of the B4Q mutant of *Drosophila* CaM (Starovasnik *et al*., 1992) in which sites I, II, and III bound calcium with equivalent affinity.

The biphasic calcium-dependent response of the unmodified N-terminal domain of E140Q-CaM could not be detected using this semiquantitative SDS−PAGE analysis; therefore, the conformation of helix B was altered by mutating site IV in the C-terminal domain. This demonstrates that if the two domains interact in this mutant, they do so differently than in wild-type CaM. We conclude that the calcium-induced susceptibility of R37 does not appear to be an intrinsic property of all N-terminal domains of CaM having a wild-type sequence.

Quantitative Thrombin Footprinting. In the equilibrium titrations quantified rigorously by HPLC analysis, data at each pCa represent replicate analyses of an individual sample of CaM extensively dialyzed to equilibrate it at an experimentally determined level of free calcium (Pedigo & Shea, 1995a). The abundance of primary proteolytic products indicated calcium-dependent changes in the probability of cleavage of CaM by thrombin; the products at each pCa were resolved as shown in six representative HPLC chromatograms (Figure 4).

The abundance of peaks B′ (1−106) and B (107−148) resulting from cleavage of the R106−H107 bond was maximal for apo-CaM (see pCa 8.8) and decreased monotonically as calcium bound to CaM. In contrast to the monotonic response of R106, the abundance of peaks A (1−37) and A′ (38−148) formed from cleavage at R37 was low for apo-CaM, increased in magnitude at an intermediate fraction of calcium binding (see pCa 6.0), and then decreased in abundance for $Ca^{2+}{}_4$-CaM (see pCa 3.9). The maximal absolute abundance of products generated by cleavage at R106−H107 (pCa 8.8) was higher than the maximal abundance observed for R37−S38 (see pCa 6.0). Uncleaved CaM was present in every chromatogram at a retention time of approximately 40 min (not shown). The retention times for peptides varied slightly among chromatograms, but this variation did not interfere with assignments or integration of peak areas.

To emphasize the dramatic differences in the nature of the calcium-dependent susceptibilities of R37 and R106 (biphasic *vs* monotonic), Figure 5 shows the abundance of each peptide normalized to the range of its fractional area (A_i). Between nanomolar and micromolar calcium, responses at the two positions were opposite and slightly offset, such that the median ligand activity of the susceptibility of R37 was at a lower calcium concentration than that of the curve indicating susceptibility of R106. Between micromolar and millimolar calcium, both positions were protected in response to calcium binding.

Despite this evidence for structural interaction between the two domains, it was not possible to estimate the magnitude of the energy associated with these interactions. Therefore, the nonlinear least-squares analysis of quantitative proteolytic footprinting of CaM titrations treated each domain as a two-site binding system (eqs 4 and 5) in which susceptibility was proportional to average occupancy (Pedigo & Shea, 1995a). Two models of calcium binding were compared: model I forced the sites to have equal intrinsic affinities and independent (noncooperative) binding while model II allowed for heterogeneous intrinsic affinities and

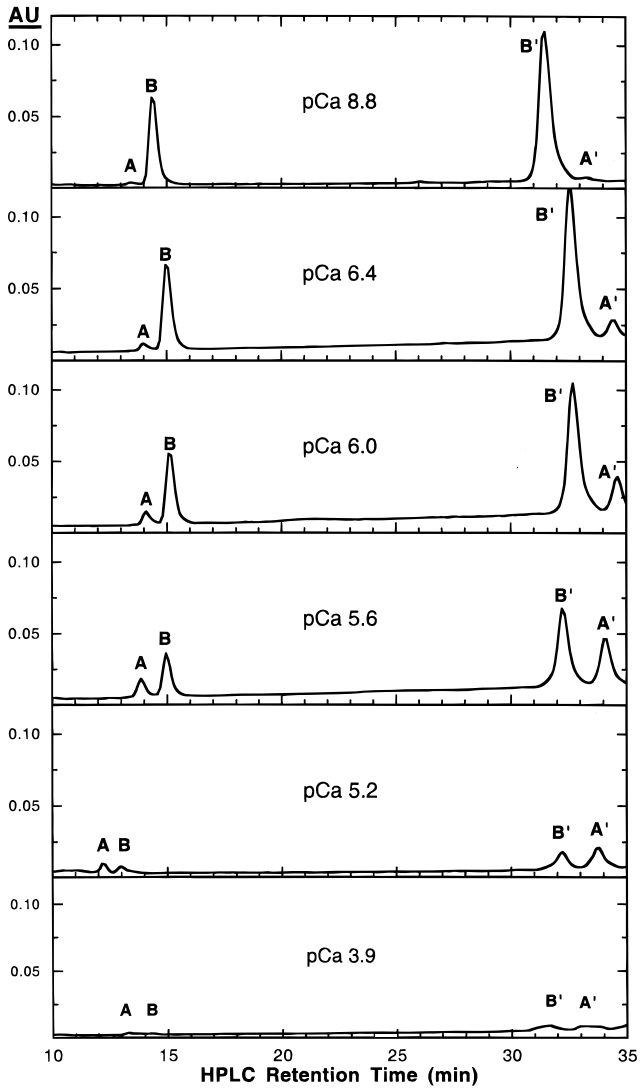


FIGURE 4: HPLC chromatograms of fragments generated by thrombin footprinting of a calcium titration of CaM. Comparison of absorbance units (AU) at 220 nm for proteolytic products with retention times between 10 and 35 min for six footprinting reactions; indicated pCa values (8.8, 6.4, 6.0, 5.6, 5.2, and 3.9) are rounded to nearest tenth. Peaks corresponding to proteolytic products are designated A (1−37), B (107−148), B′ (1−106), and A′ (38−148). Chromatographic and proteolytic conditions are described under Experimental Procedures.

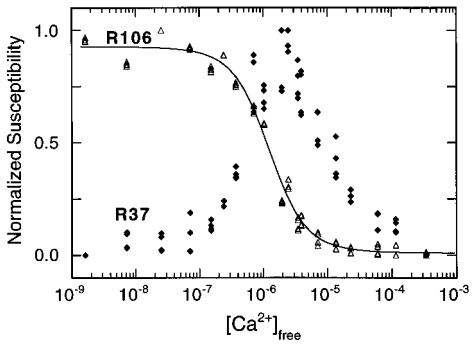


FIGURE 5: Normalized calcium-dependent susceptibility profiles for R37 and R106 of CaM. The normalized abundance of replicate analyses of peptides 1−37 and 38−148 (♦) and 1−106 and 107−148 (△) as a function of free calcium concentration in the dialysate. Smooth curve for 106 to guide the eye; simulated using values from Table 2.

cooperativity between paired sites. Analysis of the energetics of calcium binding inferred from the susceptibility of each domain is described below.

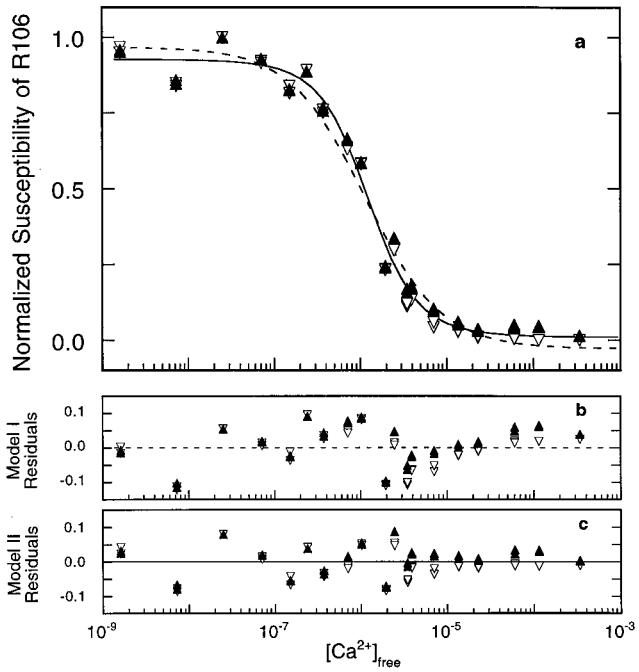


FIGURE 6: Analysis of the normalized susceptibility of R106 in CaM. (a) Abundance of peptides 1−106 (▲) and 107−148 (▽) compared to simulation using parameters (Table 2) resolved from nonlinear least-squares fit of both data sets simultaneously to model II. (b) Residuals of the data compared to nonlinear least-squares fit of both data sets simultaneously to model I; best-fit parameters were $\Delta G_1 = -8.02 \pm 0.06$ kcal/mol, $Y_{[X]low} = 0.967 \pm 0.025$, Span $= -0.995 \pm 0.024$ with square root of variance of 0.059. Values of ΔG_1 derived from simultaneous fit of data for both peptides differed by less than 0.01 kcal/mol from values resolved from individual analyses of the abundance of each peptide. (c) Residuals of the fit of both data sets simultaneously to model II (Table 2, Figure 6a).

Protection of R106−H107. In Figure 6a, the normalized fractional abundance of the peptide products 1−106 (▲) and 107−148 (▽) is shown as a function of the calcium concentration in each dialysate of CaM. For each peptide, there is close agreement between analysis of peak areas from replicate injections of thrombin footprinting reactions. As expected for primary cleavage products, the normalized fractional abundance for the two peptides is nearly identical at each pCa.

The titration curves resolved for each peptide were analyzed individually and in combination (Table 2). For binding models I and II (eqs 4 and 5), the values of the parameters resolved from fits to the data for individual peptides were nearly identical to each other and to those resolved from the combined data sets (e.g., ±0.02 kcal/mol for total free energies, ΔG_2). Resolved values were independent of starting guesses within a range of ±2 kcal/mol. Curves simulated using parameters resolved from simultaneous analysis of replicate determinations of both complementary peptides at each pCa are shown in Figure 6.

Residuals of fits to model I (Figure 6b) showed sinusoidal variation that is characteristic of fitting the fractional saturation of a cooperative system to a model limited to independent binding (eq 4). The systematic trend of the residuals decreased (Figure 6c) in fits to model II. This criterion, the relative values of the variance of the fit, and the values of ΔG_c calculated from the macroscopic free energies indicated that model II was more appropriate to describe these data; best-fit parameters are given in Table 2. Thus, calcium binding to sites III and IV in the C-terminal domain was cooperative and possibly heterogeneous.

Table 2: Free Energies of Calcium Binding to CaM Estimated from Analysis of Susceptibility of R106[a]

peptide	ΔG_1[b]	ΔG_2	$\sqrt{\text{var}}$[c]	ΔG_c[d]
1−106[e]	−7.75 ± 0.45	−15.95 ± 0.15	0.045	−1.27
107−148[f]	−7.68 ± 0.44	−15.97 ± 0.14	0.045	−1.42
both[g]	−7.71 ± 0.26	−15.96 ± 0.10	0.046	−1.35

[a] Parameters resolved from fits to model II which allowed heterogeneous and cooperative binding. [b] All free energies ($-RT \ln K$) expressed in kcal/mol (1 kcal = 4.184 J). To overestimate the uncertainty on each parameter, the greater of the positive or negative limit of each asymmetric 65% confidence interval was tabulated unless noted otherwise. [c] Correlation coefficients, magnitude, and distribution of residuals also were evaluated, but not shown. [d] ΔG_c was calculated by assuming that the intrinsic binding energies are equal and gives a lower limit of actual cooperativity ($\Delta G_c = \Delta G_2 - 2\Delta G_c - RT \ln 4$). [e] Fitted value of $Y_{[X]low}$ was 0.929 ± 0.031, and Span was −0.903 ± 0.044. [f] Fitted value of $Y_{[X]low}$ was 0.925 ± 0.031, and Span was −0.931 ± 0.044. [g] Simultaneous analysis of fractional abundance of both peptides; fitted value of $Y_{[X]low}$ was 0.927 ± 0.022, and Span was −0.917 ± 0.033.

Table 3: Free Energies of Calcium Binding to CaM Estimated from Analysis of Susceptibility of R37[a]

peptide	ΔG_1[b]	ΔG_2	$\sqrt{\text{var}}$[c]	ΔG_c[d]
Induced Susceptibility: pCa 8.79−5.61				
1−37[e]	−7.68 ± −1.13	−17.06 ± 0.23	0.081	−2.50
38−148[f]	−8.21 ± −0.55	−16.89 ± 0.21	0.081	−1.28
Induced Protection: pCa 5.71−3.47				
1−37[g]	−6.61 ± 0.50	−13.75 ± 0.15	0.071	−1.35
38−148[g]	−6.73 ± 0.42	−13.51 ± 0.18	0.082	−0.86

[a] Parameters resolved from fits to model II which allowed heterogeneous and cooperative binding. [b] All free energies ($-RT \ln K$) expressed in kcal/mol (1 kcal = 4.184 J). To overestimate the uncertainty on each parameter, the greater of the positive or negative limit of each asymmetric 65% confidence interval was tabulated unless noted otherwise. [c] Correlation coefficients, magnitude, and distribution of residuals also were evaluated, but not shown. [d] ΔG_c ($\Delta G_2 - 2\Delta G_c - RT \ln 4$) was calculated by assuming that the intrinsic binding energies are equal and gives a lower limit of actual cooperativity. [e] Fitted value of $Y_{[X]low}$ was 0.050 ± 0.043, and Span was fixed at 1.0. [f] Fitted value of $Y_{[X]low}$ was 0.002 ± 0.036, and Span was fixed at 1.0. [g] $Y_{[X]low}$ was fixed at 1.0, and Span was fixed at −0.95.

Induced Susceptibility of R37−S38. The most unexpected findings of these studies were that the R37−S38 bond was susceptible to thrombin and that its response to calcium binding was biphasic (Figure 5). Although R37−S38 was protected from cleavage at low and high calcium concentrations, it showed higher susceptibility to cleavage at an intermediate level of calcium saturation (occurring at micromolar calcium concentration). As described in detail previously for cleavage of CaM at E31 by EndoGluC (Pedigo & Shea, 1995a), the sharp biphasic transition of these data posed a significant challenge to nonlinear least-squares analysis. They were fit in a piecewise fashion as indicated under Experimental Procedures; resolved values are given in Table 3.

Data comprising the induced susceptibility phase of the R37 isotherms are shown in Figure 7. There was excellent agreement between the transitions defined by the abundance of complementary primary cleavage products (compare 1−37 in Figure 7a to 38−148 in Figure 7b). For analysis of both peptides, the square root of the variance decreased if either $Y_{[X]_{low}}$ or Span was held constant during simultaneous analysis of the binding free energies. The total free energy resolved

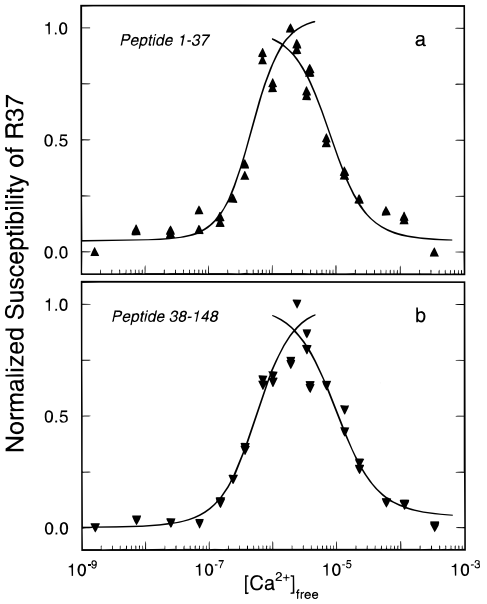


FIGURE 7: Piecewise analysis of the biphasic profile of the normalized susceptibility of R37 in CaM. Normalized fractional abundances of peptides (a) 1−37 (▲) and (b) 38−148 (▼) are indicated. Both phases of the R37 isotherms (induced susceptibility and protection) show curves simulated from parameters resolved from fits to model II (Table 3).

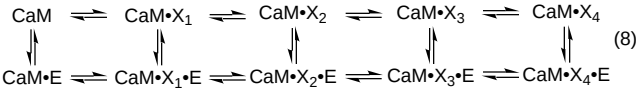
from fits of the individual peptides to model II with either fixed or floated end points differed by a maximum of 0.25 kcal/mol. Although this difference is greater than that found in the analysis of susceptibility at R106, it is insignificant and merely indicative of the difficulty in estimating parameters when the end points of a titration curve are poorly defined. Curves were simulated using parameters resolved from representative fits to model II (Table 3).

As was found in the analysis of the susceptibility of the R106−H107 bond, fits to model II were superior to those of model I as judged by several criteria as noted under Experimental Procedures. In particular, a sinusoidal pattern of systematic residuals for fits to model I (similar to those shown in Figure 6b) indicated that the transition of induced susceptibility represented a binding process with positive cooperativity. Errors associated with the calculation of ΔG_c were ~1 kcal/mol (reflecting the uncertainty in the determination of ΔG_1).

Protection of R37−S38. Calcium concentrations above micromolar induced protection of R37−S38 in the biphasic profiles shown in Figure 7. Parameters resolved from fits of the fractional abundance of each peptide to model II are reported in Table 3; the total energy of calcium binding to both sites in the N-terminal domain was between 2.5 and 3.5 kcal/mol less favorable than binding to the pair of sites in the C-terminal domain. Analyses of cooperativity were imprecise but suggested modest cooperative interactions (−1 kcal/mol) between sites I and II, consistent with NMR studies of these samples (Pedigo & Shea, 1995b) and other studies of CaM under similar conditions (Linse *et al*., 1991).

DISCUSSION

The molecular mechanism of calcium-induced conformational switching of CaM has pleiotropic physiological significance and serves as a model system for understanding basic issues in molecular recognition and energy transduction in cooperative ligand binding proteins. A complete thermodynamic linkage scheme between the binding of four calcium ions to CaM and its interactions with a target enzyme may be specified (eq 8; Wang & Sharma, 1980):



(8)

In accord with this scheme of chemical equilibria, calcium binding affects target recognition and *vice versa*; consistent with this is the finding that target binding differentially affects the calcium affinity of CaM (Olwin & Storm, 1985) and can restore calcium binding properties of calmodulin mutants (Haiech *et al*., 1991). It may be noted that the general form of this scheme is identical to that which describes the linked equilibria for other macromolecular associations such as hemoglobin binding four oxygen molecules and an organic phosphate (Ackers, 1979).

Although 10 macroscopic species are depicted above, most studies have treated activation of CaM as a 2-step process that requires (a) calcium-induced switching between two end states of CaM: an "inactive" (apo) form and an "active" (calcium-saturated) form (eq 9a):

$$CaM + 4Ca^{2+} \rightleftharpoons {Ca^{2+}}_4\text{-}CaM \qquad (9a)$$

and (b) conformational rearrangements that constitute recognition of target protein sequences by ${Ca^{2+}}_4$-CaM as indicated in eq 9b [cf. Wang *et al*. (1980)].

$${Ca^{2+}}_4\text{-}CaM + E_{free} \rightleftharpoons {Ca^{2+}}_4\text{-}CaM{\cdot}E \qquad (9b)$$

An increasing frequency of experimental reports indicate that apo-CaM may engage in specific target interactions; these are postulated to be reverse triggers such that calcium binding to CaM causes release of the target enzyme (Takeuchi *et al*., 1994, 1995; Gerendasy *et al*., 1995). Such a mechanism would be consistent with a two-state model of conformational states of CaM similar to that given in eq 9a; however, target interaction would be described by eq 9c.

$$CaM + E \rightleftharpoons CaM{\cdot}E \qquad (9c)$$

The biological role of intermediate ligation species (${Ca^{2+}}_1$-CaM, ${Ca^{2+}}_2$-CaM, ${Ca^{2+}}_3$-CaM) is less well understood. Some target proteins may be activated by partially saturated species of CaM or fragments (Kuznicki *et al*., 1981; Haiech *et al*., 1981; Wang & Sharma, 1980; Klee, 1988), and an ordered mechanism of domain interactions has been proposed to explain the distinction between CaM binding to smMLCK and its release of autoinhibition (VanBerkum *et al*., 1990). The interactions of CaM with caldesmon (Wang, 1988; Medvedeva *et al*., 1995) or nebulin (Root & Wang, 1994) may be qualitatively and quantitatively very different from those with smMLCK. Studies of enzyme activation by CaM mutants with altered calcium binding sites showed that site-specific effects were distinguishable (Gao *et al*., 1993). Recent studies of two classes of mutants of *Paramecium* exhibiting swimming defects showed a domain-specific division of effects of CaM mutations, suggesting that partially saturated species of CaM may have specific roles in regulation of sodium and potassium channels [see Saimi and Kung (1994) for a review]. Preliminary studies of their calcium binding properties and associated conformational changes (Harmon & Shea, 1995) support this view.

Inherent in the interpretation of these studies is that the four sites of CaM are not equivalent in their calcium binding properties. Heterogeneity among the calcium binding sites of CaM has been demonstrated experimentally but is not fully understood (Wang, 1985; Crouch & Klee, 1980; Drabikowski & Brzeska, 1982; Starovasnik *et al*., 1992; Klee, 1977, 1988; Martin *et al*., 1985; Kilhoffer *et al*., 1992; Haiech *et al*., 1981; Waltersson *et al*., 1993; Svensson *et al*., 1993; McPhalen *et al*., 1991). Mutagenic analyses combined with comparative metal binding studies have been the approaches most commonly applied to dissecting the rules governing ion selectivity of EF-hand sites in proteins (Falke *et al*., 1994; Wang *et al*., 1982, 1984; Forsén *et al*., 1993).

Motivated by an intrinsic interest in the cooperative switching mechanism of CaM and evidence that some of the intermediates have distinct physiological functions, the quantitative thrombin footprinting studies reported here were aimed at understanding the rules for transition between the partially calcium-saturated species of wild-type CaM. The presence of detectable intermediate(s) reflects the processes of propagated conformational change in CaM. Their abundance reflects the energetic costs of (a) binding calcium, including transferring the ion from water to a binding site, and (b) associated local and global responses by CaM, including the exposure of hydrophobic side chains and restriction of segmental motion. The quantitative thrombin footprinting method has revealed the presence of at least one intermediate structure and highlighted local features of the linkage between calcium binding and residue-specific conformational change under equilibrium conditions. It has been possible to estimate the magnitude of intrinsic calcium binding affinities and the nature of some pairwise cooperative interactions. The discussion below addresses the related issues of validation of the experimental method and new information obtained about the allosteric mechanism of calcium binding to CaM.

Thrombin Specificity and Lack of Secondary Cleavage. Classical methods (e.g., equilibrium or flow dialysis) that yield macroscopic equilibrium constants provide valuable, model-independent constraints on the energetics of calcium binding (Crouch & Klee, 1980; Haiech *et al*., 1980). However, many chemically distinct distributions of microscopic free energies may give rise to equivalent macroscopic equilibrium constants and degrees of fractional saturation by calcium (Ackers *et al*., 1983; Linse *et al*., 1991; Kilhoffer *et al*., 1992; Pedigo & Shea, 1995b). Thus, residue-specific methods must be developed to explore differences among intermediate states and ultimately allow resolution of the microscopic equilibrium constants that describe intrinsic binding and cooperative interactions for calcium binding. Proteolytic footprinting is especially advantageous for monitoring changes in the N-terminal domain where there are no naturally occurring spectral signals. Thrombin was chosen to monitor residue-specific responses to equilibrium calcium binding because it is more sensitive than many serine proteases to the tertiary structure of a protein substrate.

Thrombin preferentially cleaves only at the C-terminal side of arginine residues; however, its specificity is further restricted by residues in an exosite, an anionic region far removed from the scissile bond [cf. Aschner *et al*. (1992)]. A general advantage of highly selective cleavage is that it facilitates unambiguous identification and quantitation of products. Although several independent proteases must be applied to determine the generality of observed patterns, this minimizes the effects of possible systematic errors in a single study and facilitates comparisons when aliquots of CaM dialysates comprising a discontinuous titration can be studied by multiple methods.

Several proteases have been used to probe calcium-dependent conformational changes in CaM. Of the six arginine residues in CaM (R37, R74, R86, R90, R106, R126; Figure 1a), most are known to be substrates for trypsin (Kawasaki *et al*., 1986; Walsh *et al*., 1977; Drabikowski *et al*., 1977; Newton *et al*., 1984; Mackall & Klee, 1991) or clostripain (Manalan *et al*., 1985). The observed specificity of thrombin was expected to depend on both the solvent accessibility and conformation of peptide bonds of arginine. In this study, cleavage was observed at only two of six arginine residues: R37, in helix B leading out of site I, and R106, in helix F leading out of site III (Figure 2a). To our knowledge, this is the first report of the R37−S38 bond being susceptible to thrombin. The other four candidates (bonds R74−K75, R86−E87, R90−V91, and R126−E127) were not substrates for thrombin under the conditions tested. In apo-CaM, helix E (preceding site III and containing R86 and R90) is accessible to the action of diverse proteases (Walsh *et al*., 1977; Pedigo & Shea, 1995a; Mackall & Klee, 1991). The lack of cleavage at R90−V91 by thrombin emphasizes that target sites are defined by their tertiary as well as secondary structures and sequence.

The absence of detectable levels of secondary cleavage products provided additional evidence that thrombin specificity is related to the tertiary structure of its substrate and not determined exclusively by sequence. It was evident that removing segment α of CaM (see Figure 1) eliminated the action of thrombin at R106−H107 in the C-terminal domain (i.e., there was no cleavage at R106 in a peptide comprised of residues 38−148) and removing segment β' eliminated the susceptibility of R37−S38 in the N-terminal domain. This suggests that the native conformation of the bond R106−H107 depends on the presence of an intact N-terminal domain and, conversely, that the native conformation of R37−S38 depends on an intact C-terminal domain.

It is also possible that a primary cleavage event simply removes or disrupts a region of CaM that serves as an exosite for thrombin at each of the scissile bonds. This explanation would specify that an active site in one domain requires an exosite in the other. According to a crystallographic study of a complex of thrombin and hirudin (Rydel *et al*., 1990), the distance between the active site and residues in the exosite is in the range of 15−25 Å. This is possible in a crystallographic structure (Babu *et al*., 1988) in which the overall length of $Ca^{2+}{}_4$-CaM is greater than 60 Å. However, an energy-minimized model of $Ca^{2+}{}_4$-CaM such as CAM10 (Pascual-Ahuir *et al*., 1991; Figure 2b) shows a more compact configuration that seems more compatible with the suggestion that thrombin may interact with both domains simultaneously.

Conformational Change as a Measure of Binding. The lack of secondary cleavage allowed unambiguous interpretation of the abundance of the primary cleavage products as reflecting the calcium-dependent changes in the susceptibility of bonds within holo-CaM, rather than the properties of one of its fragments or a calcium-induced change in the enzymatic activity of thrombin. A remaining issue is the degree to which the conformational change correlates with the extent of calcium binding. As addressed extensively previously (Pedigo & Shea, 1995a,b), this is a complex issue

that is not unique to these quantitative proteolytic footprinting studies. It is intrinsic to the interpretation of all spectroscopic and exchange methods. However, these approaches have been very powerful for elucidating responses to calcium binding. The most persuasive argument in their favor is the agreement among conclusions obtained by applying orthogonal techniques to the same samples [cf. Pedigo and Shea (1995a,b)].

Calcium-Dependent Patterns of Susceptibility. Two distinct patterns of susceptibility were observed in footprinting studies of wild-type CaM (Figure 5). In the N-terminal domain, calcium binding caused a peaked biphasic transition of cleavage at R37 (six residues removed from E31, the bidentate glutamate of site I). In the C-terminal domain, calcium afforded monotonic protection of R106 (located two residues beyond E104, the bidentate glutamate of site III).

In apo-CaM, there was a marked difference between the intrinsic probability of cleavage of the two domains: R106−H107 in the C-terminal domain was a more efficient substrate than was R37−S38 in the N-terminal domain. The relative susceptibility of this pair of positions was the same as had been observed for (a) E87 *vs* E31 probed by EndoGluC and (b) E87 *vs* S38 probed by bromelain. As discussed extensively in the interpretation of quantitative EndoGluC footprinting studies of CaM titrations (Pedigo & Shea, 1995a), this relative difference was expected because the C-terminal domain of apo-CaM is known to be less stable thermodynamically (Tsalkova & Privalov, 1985) and its residues more reactive to chemical modification or spontaneous degradation than those of the N-terminal domain (Ota & Clarke, 1989, 1994; Potter *et al*., 1993; Giedroc *et al*., 1985, 1987). The greater proteolytic susceptibility of the C-terminal domain seen for wild-type CaM (Figure 3a) also was observed in studies of E140Q-CaM (Figure 3b). It appeared that the absolute probability of cleavage at R37 in the apo state was slightly greater in E140Q-CaM than in wild-type CaM. This would be consistent with a model in which a wild-type C-terminal domain contributes to protection of the N-terminal domain and lowers its affinity for calcium (Martin *et al*., 1992). However, this pattern must be verified quantitatively by chromatographic analysis.

The binding of calcium to a site on CaM induced protection of the thrombin-susceptible bond closest to the saturated calcium-binding site. The simplest example of this correlation was the calcium-induced change in susceptibility of R106 in both wild-type CaM (Figure 6) and E140Q-CaM (Figure 3); cleavage was maximal in the absence of [Ca^{2+}] but was barely detectable upon calcium saturation. This was expected if ligand binding stabilizes a folded conformation or causes burial of a region of the protein. It is consistent with observations that calcium binding to wild-type CaM raises the melting temperature by more than 40 °C (Tsalkova & Privalov, 1985) and causes rearrangement of helices that expose hydrophobic patches of each domain (LaPorte *et al*., 1980). Although the crippling effect of the E140Q mutation in site IV shifted the midpoint of the susceptibility profile of R106 to a higher calcium concentration than that observed for wild-type CaM, the qualitative nature of calcium-dependent susceptibility (i.e., monotonic protection) was the same. Binding of calcium to the sites having lower affinity in wild-type CaM (sites I and II) appeared to protect R37−S38 from cleavage in both proteins.

A monotonic response to calcium binding, such as that seen for R106, may be explained by the two-state model given in eq 9a; at each calcium concentration, the susceptibility could represent the sum of the fractional contributions of two end states (f_0 for apo and f_4 for $Ca^{2+}{}_4$-CaM; $f_0 + f_4 = 1$) having different intrinsic probabilities of cleavage (e.g., the susceptibility of $Ca^{2+}{}_4$-CaM given by S_4). The total abundance of a peptide (S_t) may be described by eq 10.

$$S_t = f_0 S_0 + f_4 S_4 \tag{10}$$

Because S_4 is effectively 0, S_t can be approximated as $f_0 S_0$, where f_0 is the fraction of apo-CaM in the population of 16 possible ligation species and is a function of calcium concentration. More complex models might give identical macroscopic behavior. For example, the presence of an intermediate that behaved as an additive hybrid of the two end states (Perrella *et al*., 1990) would not be distinguishable on the basis of a monotonic response such as that observed for R106.

The most remarkable feature of these quantitative footprinting studies was that the susceptibility of R37−S38 of wild-type CaM *increased* in response to calcium binding to sites III and IV and *decreased* upon calcium saturation. Helix B (see Figure 2a) is resistant to cleavage in the apo state and susceptible only after the C-terminal domain binds calcium. Thus, the biphasic response of R37 indicated the presence of a partially saturated species of CaM that cannot be an additive hybrid of the conformations of apo- and $Ca^{2+}{}_4$-CaM.

Although direct evidence for an intermediate structure had come from stoichiometric binding studies of CaM using NMR (Seamon, 1980; Klevit *et al*., 1984) and other techniques, such a finding could reflect two ordered and independent transitions. Our proteolytic footprinting studies (Pedigo & Shea, 1995a; Verhoeven & Shea, 1993) have shown unequivocally that the transition between the apo state and a partially saturated intermediate involves changing interactions between the two domains. The consensus among three quantitative footprinting studies of wild-type CaM identifies helix B of the N-terminal domain of CaM (see Figure 2a) as one that responds to calcium binding in the other domain. This pattern was not peculiar to titrations of CaM probed by thrombin but was also observed for EndoGluC cleavage of CaM at E31 and bromelain cleavage at S38, also in helix B (see Figure 2a).

Although it is possible that the conformational changes that lead to induced susceptibility are propagated from sites III and IV through an extended "central helix", this seems unlikely because the distances between the Cα carbon of R37 and each of the calcium ions in the C-terminal domains exceed 35 Å (see Figure 2a). It is more plausible that the two domains interact or that their relative orientation is changed: calcium binding to sites III and IV exposes or destabilizes the apo form of the N-terminal domain. The nature of this response may depend on the intrinsic stability of the N-terminal domain, differences in solvent accessibility afforded by the conformation of the C-terminal domain, and other propagated conformational changes. Although isolated domains self-associate weakly, the consequences of their interactions within holo-CaM may be similar qualitatively to the role of the $\alpha_1\beta_2$ interface of hemoglobin; alteration of residues in that interface dramatically affects cooperativity without necessarily altering intrinsic ligand binding affinity (Pettigrew *et al*., 1982).

The conclusion that the domains interact is supported by finding that (a) the biphasic susceptibility of wild-type CaM

consisted of two phases that were *opposite* in direction (presumably having different molecular origins), and (b) this pattern was modified in a mutant (E140Q-CaM) in which the N-terminal domain had a wild-type sequence but calcium binding to sites III and IV was compromised. In E140Q-CaM, monotonic protection was observed for the calcium-dependent susceptibility of R37. Evidently, the induced susceptibility seen in wild-type CaM was not an intrinsic property of the wild-type N-terminal domain of holo-CaM. It was retained by other mutants of the C-terminal domain such as R90A-CaM (Sorensen and Shea, data not shown) in which the calcium binding affinity of the C-terminal domain remained close to that of wild-type CaM. Studies of the N-domain of troponin C by NMR (Ding *et al*., 1994) have shown that helix B undergoes the largest reorientation upon binding calcium to sites I and II; however, such straightening of a helix in an isolated domain is caused exclusively by ligand binding rather than interdomain interactions.

Cooperative Calcium Binding. Quantitative analysis of susceptibility profiles of wild-type CaM indicated that sites III and IV have a higher affinity for calcium than do sites I and II and that cooperative interactions occur between paired binding sites within a domain (Tables 2 and 3).

An advantage of the discontinuous titration approach developed to study equilibrium binding properties of concentrated CaM solutions by quantitative EndoGluC footprinting (Pedigo & Shea, 1995a) and NMR (Pedigo & Shea, 1995b) is that the signal at each pCa represented the analysis of an aliquot of the same CaM dialysates as had been studied previously; thus, the binding parameters could be compared directly. The values of ΔG_2 resolved from calcium-dependent tyrosine fluorescence (−16.1 kcal/mol; Pedigo & Shea, 1995a) and 1-D ^{1}H-NMR studies of Y138 (−16.05 kcal/mol; Pedigo & Shea, 1995b) agree very closely with those resolved from calcium-induced protection of R106 (Table 2).

Although the profiles for protection of R106 and susceptibility of R37 change reciprocally as shown in Figure 5, they are not precisely complementary. Nonlinear least-squares analysis (Table 3) of the transition representing induced susceptibility of R37 (Figure 7) is slightly more negative (favorable) than that resolved for R106. The estimates of ΔG_2 for the C-terminal domain resolved from quantitative proteolytic footprinting of E87 using EndoGluC (Pedigo & Shea, 1995a) and bromelain (data not shown) and the value of ΔG_2 estimated from analysis of the induced susceptibility of E31 to EndoGluC cleavage are also more negative by ∼1 kcal/mol than the value of ΔG_2 resolved for R106. As discussed in detail previously (Pedigo & Shea, 1995a), we conclude that this small but consistent difference is a manifestation of heterogeneity in the intrinsic affinities of sites III and IV (i.e., $k_{III} > k_{IV}$). It may also reflect different contributions from the energy of interdomain interactions; however, this cannot be determined from these data alone. Although the calculated value of ΔG_c was less well-defined than ΔG_2 in all cases, the estimates from these studies are in good agreement with those reported previously (Linse *et al*., 1991; Pedigo & Shea, 1995a,b) and indicate positive cooperativity between calcium binding to pairs of sites in each domain.

There has been no previous study of the susceptibility of R37−S38 to thrombin; presumably the fragments generated by cleavage there escaped notice because that bond is only susceptible when CaM is partially saturated. The pattern of biphasic proteolytic susceptibility observed here is consistent with that observed for primary products of bromelain cleavage at the neighboring bond S38−L39 and the EndoGluC probe of nearby bond E31−L32. [It disagrees with studies showing that the R37−S38 bond was monotonically susceptible to trypsinolysis as a function of calcium binding to CaM (Mackall & Klee, 1991); however, effects of calcium on the activity of trypsin and significant secondary cleavage in that study prevent direct comparison.] The cooperativity resolved from the protection of R37 is less negative (less favorable) by ∼1 kcal/mol than that obtained for the C-terminal domain, in agreement with previous studies.

Calcium-Induced Switching of CaM. These quantitative thrombin footprinting studies of R37 and R106 have indicated that partially saturated CaM adopts at least one conformation distinct from apo or fully-saturated CaM and that calcium binding causes the two domains to interact. The postulate of interdomain interactions is consistent with solution studies of competition between calcium and terbium binding of CaM (Wang *et al*., 1984) and the pattern of calcium binding to mutants of *Drosophila* CaM (Maune *et al*., 1992b; Martin *et al*., 1992).

Results of our studies indicate that calcium binding at sites III and IV induces a change in the conformation of helix B (originating in site I; see Figure 2a). Quantitative proteolytic footprinting studies with three different proteases show that the responses of E31, R37, and S38 in helix B (see Figure 1a) all were found to be biphasic. A similar mechanism that entails calcium binding at site III affecting the conformation of site II has been postulated for recoverin, a calcium sensor in vision (Ames *et al*., 1994). Although all of these studies of CaM were conducted in the absence of target peptides and the family of conformations adopted is expected to differ, it is striking to note that in structures of co-complexes helices B and F are antiparallel and prominent as the "roof" of the peptide cavity (Ikura *et al*., 1992; Seeholzer & Wand, 1989; Meador *et al*., 1992; Rao *et al*., 1992).

One other position observed to have a biphasic response is I125 in the C-terminal domain as probed by bromelain (Verhoeven and Shea, data not shown). We have never observed or found a report of a position whose calcium-dependent susceptibility is biphasic but of opposite sign to that observed here for R37 (i.e., a bond that is protected when CaM is partially saturated CaM but susceptible in apo-CaM and Ca^{2+}_4-CaM), but this is not ruled out by any of our observations.

Monotonic protection in response to calcium binding to CaM has been observed in the C-terminal domain for a residue in helix E (E87 probed by both bromelain and EndoGluC) and helix F (R106, this study) and in the N-terminal domain, in site II/helix D (F65 probed by chymotrypsin). A monotonic increase in calcium-linked susceptibility has been observed at the junction of the two domains [K75 probed by bromelain, in contrast to the results of Mackall and Klee (1991) using trypsin]. This residue-specific sampling of the calcium-linked conformational responses of CaM suggests that a model of Ca^{2+}_4-CaM such as CAM10 (Pascual-Ahuir *et al*., 1991), in which residues R37 and R106 are in closer proximity than observed in a crystallographic determination (Babu *et al*., 1988), may pertain to the partially saturated species as well. However, any proposal of a series of conformational transitions induced by calcium binding must keep in mind that both of the structures in Figure 2 pertain to Ca^{2+}_4-CaM.

Specification of the number and properties of the intermediate states is critical for a determination of the molecular mechanism of ligand-induced structural transitions that control regulatory properties of CaM. Although the binding of divalent metals and their effects on CaM have been monitored using many methods, there is a paucity of optical spectroscopic probes in CaM. Protein engineering or chemical modification has been used to introduce spectroscopic reporter groups [cf. Kilhoffer *et al*. (1992) and Yao *et al*. (1994)] or create reduced valency mutants having fewer than four functional calcium binding sites (Maune *et al*., 1992b; Kosk-Kosicka *et al*., 1992). We engineered a mutation of the bidentate glutamate of site IV of the rat sequence analogous to B4Q (Maune *et al*., 1992b) as a test of whether the biphasic response of the N-terminal domain of holo-CaM reflects calcium-dependent properties attributable solely to that domain.

As was shown in semiquantitative analysis of a titration of E140Q-CaM (Figure 3b), the order of calcium binding to the two domains seemed to be equivalent or even the reverse of wild-type CaM: half-maximal protection of R37 occurred at a calcium concentration less than or equal to that required for half-protection of R106. Comparing the response of R37 in wild-type CaM to that of E140Q-CaM, the midpoints of protection were similar (between pCa 4.9 and 4.3). Thus, the calcium affinity of sites I and II was not affected dramatically, but the interactions between domains were disrupted. This is consistent with studies of isolated domains of wild-type CaM showing that calcium binding affinities are similar to those inferred for the domains within the holoprotein [Dalgarno *et al*., 1984; Aulabaugh *et al*., 1984; Ikura *et al*., 1984; reviewed in Klee (1988)] and their structural characteristics are similar (Dalgarno *et al*., 1984; Aulabaugh *et al*., 1984; Ikura *et al*., 1984; Finn *et al*., 1993). Although studies of isolated domains cannot elucidate interdomain interactions, they are very useful for defining reference properties. The findings of similar properties have led many investigators to conclude that there is little or no interaction between domains [cf. Weinstein and Mehler (1994)]. Results of our footprinting studies suggest that constraints imposed by interdomain interactions may be structurally significant, even if these conformational rearrangements have low energetic penalities.

We anticipate that the self-interactions of CaM found in this study are significant to the study of calcium activation of CaM and may be relevant to the diversity of more compact structures seen for $Ca^{2+}{}_4$-CaM interacting with targets or antagonist (Török & Whitaker, 1994; Meador *et al*., 1993). The rules that govern the specificity and affinity of calcium-regulated CaM−target interactions are an area of active study experimentally and computationally [see Crivici and Ikura (1995) for review]. Several structural studies of calcium-saturated CaM in complexes with target peptides (Meador *et al*., 1992; Ikura *et al*., 1992; Porumb *et al*., 1994; Seeholzer & Wand, 1989; Rao *et al*., 1992) have shown that both domains interact with the target. A prevalent interpretation of these studies has been that calcium binding opens hydrophobic clefts in each domain such that they act as independent clasps on the target.

A recent crystallographic study of a co-complex of $Ca^{2+}{}_4$-CaM with a single molecule of the anti-psychotic drug TFP showed that the drug interacted solely with the C-terminal domain (Cook *et al*., 1994). A remarkable feature was that the disposition of the two domains resembled that of CaM−peptide complexes described above; the central helix was no longer evident. It was noted that the only intra-CaM contacts <4 Å involved residues in helices A and B with helices E and F. The interpretation of this structure was that filling the cleft in the C-terminal domain induces a collapsed structure (Cook *et al*., 1994). However, another possibility is that a collapsed conformation is one of several that are energetically accessible to $Ca^{2+}{}_4$-CaM in solution and that at pH 5.8 such a form was stabilized by TFP binding so that crystals could form. The calcium-dependent pattern of susceptibility to thrombin observed here indicates that interdomain interactions occur upon binding calcium to sites in the C-terminal domain and are sufficient to cause rearrangement of the N-terminal domain. This supports the interpretation that the effects of TFP binding might be propagated between domains but also suggests that the two domains are closer on average than suggested by the dumbbell structure of $Ca^{2+}{}_4$-CaM (Figure 2a).

The co-complexes of $Ca^{2+}{}_4$-CaM with MLCK peptides show that the two domains of CaM appear to engage in similar numbers of contacts with the target and might serve as equal partners in recognition and binding (Clore *et al*., 1993). However mutations of CaM that selectively affect the activation of only one of many target enzymes have been identified (VanBerkum & Means, 1991; VanBerkum *et al*., 1990; Afshar *et al*., 1994). Understanding the molecular origins of these residue-specific effects will require a better understanding of the self-interactions of CaM as well as the nature of its specific contacts with different targets. Quantitative proteolytic footprinting methods allow us to explore residue-specific pathways of intraprotein communication and sample these linked processes comprehensively in the wild-type protein, without the filters that must be applied in protein engineering or that exist by definition when relying solely on naturally occurring spectroscopic probes.

SUMMARY

Despite the progress made in developing methods to explore the structural responses to calcium binding by calmodulin [see Kilhoffer *et al*., (1992), Weinstein and Mehler (1994), Cox (1988), and Wang (1985)], there are numerous questions remaining about the network of cooperative communication within this small protein. Using thrombin, it was possible to probe CaM specifically at two arginine positions: R37, near site I, and R106, near site III. For cleavage at each position, both primary proteolytic products were monitored and shown to give equivalent susceptibility profiles (i.e., primary cleavage products were observed in equimolar quantities). On the basis of proteolytic footprinting studies presented here and other studies using EndoGluC (Pedigo & Shea, 1995a), bromelain, and chymotrypsin (see Figure 2b), we conclude that calcium binding induces or alters interactions between the domains of CaM even in the absence of target peptides or drugs. The quantitative agreement between biphasic (equal and opposite) calcium-induced changes in susceptibility of residues in helix B toward cleavage by several proteases demonstrated that the changes reflect properties of CaM, not of the proteases used to probe its conformational changes.

These quantitative footprinting studies have shown that partially saturated CaM adopts at least one structure that is significantly different from apo- or $Ca^{2+}{}_4$-CaM and is not simply an average of those two end states. This interpreta-

tion emphasizes that self-interactions are an integral element of the calcium-induced switching mechanism and focuses attention on a region of the N-terminal domain that responds to calcium binding in both domains. In this report, we have estimated the free energy of the calcium binding events and cooperative interactions that induce these changes. Small differences between the protection of R106 and the induced susceptibility of R37 suggest that R106 reflects the average properties of the C-terminal domain while R37 is influenced disproportionately by the higher affinity binding sites in that domain. Consistent with other studies (Pedigo & Shea, 1995a,b), we conclude that site III binds calcium with higher affinity than site IV. The ultimate goal is to determine the pathway of conformational coupling between the two domains of this ubiquitous and complex protein.

ACKNOWLEDGMENTS

We thank H. Weinstein for sharing the coordinates of the CAM10 model structure; Gilson Medical Electronics for the donation of the autosampler unit of the HPLC used in this study; R. Mauer and P. Howard, University of Oregon, for the overexpression vector used for calmodulin; A. Bergold and K. Wright, University of Iowa College of Medicine Protein Structure Facility, for amino acid analysis and N-terminal sequencing; D. Moser, University of Iowa Molecular Biology Core Facility, for DNA sequencing; R. Rogers and S. Riahi, University of Iowa College of Medicine, Department of Pediatrics, for use of an atomic absorption spectrometer; P. Lashmit for technical assistance in construction of the overexpression plasmid for E140Q-CaM; M. U. Hutchins for purifying E140Q-CaM and for semiquantitative thrombin footprinting; and the reviewers for helpful suggestions.

REFERENCES

Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F., & Weng, J. (1987) in *Data Commission of International Union of Crystallography* (Allen, F. H., Bergerhoff, G., & Sievers, R., Eds.) pp 107−132.

Ackers, G. K. (1979) *Biochemistry 15*, 3372−3380.

Ackers, G. K.; Shea, M. A., & Smith, F. R. (1983) *J. Mol. Biol. 170*, 223−242.

Afshar, M., Caves, L. S., Guimard, L., Hubbard, R. E., Calas, B., Grassy, G., & Haiech, J. (1994) *J. Mol. Biol. 244*, 554−571.

Ames, J. B., Tanaka, T., Stryer, L., & Ikura, M. (1994) *Biochemistry 33*, 10743−10753.

Anderson, S. R. (1991) *J. Biol. Chem. 266*, 11405−11408.

Aulabaugh, A., Niemczura, W. P., & Gibbons, W. A. (1984) *Biochem. Biophys. Res. Commun. 118*, 225−232.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., & Struhl, K., Eds. (1991) *Current Protocols in Molecular Biology*, Vol. 1, Wiley Interscience, New York.

Babu, Y. S., Bugg, C. E., & Cook, W. J. (1988) *J. Mol. Biol. 204*, 191−204.

Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W., & Bax, A. (1992) *Biochemistry 31*, 5269−5278.

Bayley, P. M., & Martin, S. R. (1992) *Biochim. Biophys. Acta 1160*, 16−21.

Bayley, P. M., Ahlstrom, P., Martin, S. R., & Forsén, S. (1984) *Biochem. Biophys. Res. Commun. 120(1)*, 185−191.

Berliner, L. J., Ed. (1992) *Thrombin Structure and Function*, Plenum Press, New York.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Jr., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Brenowitz, M., Senear, D. F., Shea, M. A., & Ackers, G. K. (1986) *Methods Enzymol. 130*, 132−181.

Chattopadhyaya, R., Meador, W. E., Means, A. R., & Quiocho, F. A. (1992) *J. Mol. Biol. 228(4)*, 1177−1192.

Chen, J., Surendran, R., Lee, J. C., & Matthews, K. S. (1994) *Biochemistry 33*, 1234−1241.

Clore, G. M., Bax, A., Ikura, M., & Gronenborn, A. M. (1993) *Curr. Opin. Struct. Biol. 3*, 838−845.

Cohen, P., & Klee, C., Eds. (1988) *Calmodulin*, Elsevier, New York.

Cook, W. J., Walter, L. J., & Walter, M. R. (1994) *Biochemistry 33*, 15259−15265.

Cox, J. A. (1988) *Biochem. J. 249*, 621−629.

Crivici, A., & Ikura, M. (1995) *Annu. Rev. Biophys. Biomol. Struct. 24*, 85−116.

Crouch, T. H., & Klee, C. B. (1980) *Biochemistry 19*, 3692−3698.

Dalgarno, D. C., Klevit, R. E., Levine, B. A., Williams, R. J. P., Dobrowolski, Z., & Drabikowski, W. (1984) *Eur. J. Biochem. 138*, 281−289.

Ding, X.-L., Akella, A. B., Su, H., & Gulati, J. (1994) *Protein Sci. 3*, 2089−2096.

Drabikowski, W., & Brzeska, H. (1982) *J. Biol. Chem. 257*, 11584−15590.

Drabikowski, W., Kuznicki, J., & Grabarek, Z. (1977) *Biochim. Biophys. Acta 485*, 124−133.

Evans, J. S., Levine, B. A., Williams, R. J. P., & Wormald, M. R. (1988) in *Calmodulin* (Cohen, P., & Klee, C. B., Eds.) pp 57−82, Elsevier, New York.

Falke, J. J., Drake, S. K., Hazard, A. L., & Peersen, O. B. (1994) *Q. Rev. Biophys. 27*, 219−290.

Finn, B. E., Drakenberg, T., & Forsén, S. (1993) *FEBS Lett. 336*, 368−374.

Forsén, S., Kördel, J., Grundström, T., & Chazin, W. J. (1993) *Acc. Chem. Res. 26*, 7−14.

Gao, Z. H., Krebs, J., VanBerkum, M. F. A., Tang, W.-J., Maune, J. F., Means, A. R., Stull, J. T., & Beckingham, K. (1993) *J. Biol. Chem. 268*, 20096−20104.

Geiser, J. R., van Tuinen, D., Brockerhoff, S. E., Neff, M. M., & Davis, T. N. (1991) *Cell 65*, 949−959.

Gerendasy, D. D., Herron, S. R., Jennings, P. A., & Sutcliffe, J. G. (1995) *J. Biol. Chem. 270*, 6741−6750.

Giedroc, D. P., Sinhas, S. K., Brew, K., & Puett, D. (1985) *J. Biol. Chem. 260*, 13406−13413.

Giedroc, D. P., Puett, D., Sinha, S. K., & Brew, K. (1987) *Arch. Biochem. Biophys. 252*, 136−144.

Gryczynski, I., Steiner, R. F., & Lakowicz, J. R. (1991) *Biophys. Chem. 39*, 69−78.

Guerini, D., & Krebs, J. (1985) *Anal. Biochem. 150*, 178−187.

Haiech, J., Vallet, B., Aquaron, R., & Demaille, J. G. (1980) *Anal. Biochem. 105*, 18−23.

Haiech, J., Klee, C. B., & Demaille, J. G. (1981) *Biochemistry 20*, 3890−3897.

Haiech, J., Kilhoffer, M.-C., Lukas, T. J., Craig, T. A., Roberts, D. M., & Watterson, D. M. (1991) *J. Biol. Chem. 266*, 3427−3431.

Harmon, S. D., & Shea, M. A. (1995) *Biophys. J. 68*, A140.

Heidorn, D. B., & Trewhella, J. (1988) *Biochemistry 27*, 909−915.

Heyduk, T., & Lee, J. C. (1989) *Biochemistry 28*, 6914−6924.

Hoffman, R. C., & Klevit, R. E. (1991) *Tech. Protein Chem. 2*, 383−391.

Holt, J. M., & Ackers, G. K. (1995) *FASEB J. 9*, 210−218.

Hutchins, M. U. (1994) Honors Thesis, pp 1−23.

Ikura, M., Hiraoki, T., Hikichi, K., Minowa, O., Yamaguchi, H., Yazawa, M., & Yagi, K. (1984) *Biochemistry 23*, 3124−3128.

Ikura, M., Spera, S., Barbato, G., Kay, L. E., Krinks, M., & Bax, A. (1991) *Biochemistry 30*, 9216−9228.

Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., & Bax, A. (1992) *Science 256*, 632−638.

Johnson, M. L., & Frasier, S. G. (1985) *Methods Enzymol. 117*, 301−342.

Kawasaki, H., Kurosu, Y., Kasai, H., Isobe, T., & Okuyama, T. (1986) *J. Biochem. (Tokyo) 99*, 1409−1416.

Kilhoffer, M.-C., Kubina, M., Travers, F., & Haiech, J. (1992) *Biochemistry 31*, 8098−8106.

Klee, C. B. (1977) *Biochemistry 16*, 1017−1024.

Klee, C. B. (1988) in *Calmodulin* (Cohen, P., & Klee, C. B., Eds.) pp 35−56, Elsevier, New York.

Klee, C. B., & Vanaman, T. C. (1982) *Adv. Protein Chem. 35*, 213−321.

Klevit, R. E. (1983) *Methods Enzymol. 102*, 82−104.

Klevit, R. E., Dalgarno, D. C., Levine, B. A., & Williams, R. J. P. (1984) *Eur. J. Biochem. 139*, 109−114.
Kosk-Kosicka, D., Bzdega, T., Wawrzynow, A., & Watterson, D. M. (1992) *Biophys. J. 62*, 77−78.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kretsinger, R. H. (1976) *Annu. Rev. Biochem. 45*, 239−265.
Kretsinger, R. H. (1992) *Cell Calcium 13*, 363−376.
Kuznicki, J., Grabarek, Z., Brzeska, H., & Drabikowski, W. (1981) *FEBS Lett. 130*, 141−145.
LaPorte, D. C., Wierman, B. M., & Storm, D. R. (1980) *Biochemistry 19*, 3814−3819.
LiCata, V. J., & Ackers, G. K. (1995) *Biochemistry 34*, 3133−3139.
Linse, S., Helmersson, A., & Forsen, S. (1991) *J. Biol. Chem. 266*, 8050−8054.
Mackall, J., & Klee, C. B. (1991) *Biochemistry 30*, 7242−7247.
Manalan, A. S., Newton, D. L., & Klee, C. B. (1985) *J. Chromatogr. 326*, 387−397.
Martin, S. R., Andersson-Teleman, A., Bayley, P. M., Drakenberg, T., & Forsén, S. (1985) *Eur. J. Biochem. 151*, 543−550.
Martin, S. R., Maune, J. F., Beckingham, K., & Bayley, P. M. (1992) *Eur. J. Biochem. 205*, 1107−1114.
Maune, J. F., Beckingham, K., Martin, S. R., & Bayley, P. M. (1992a) *Biochemistry 31*, 7779−7786.
Maune, J. F., Klee, C. B., & Beckingham, K. (1992b) *J. Biol. Chem. 267*, 5286−5295.
McPhalen, C. A., Strynadka, N. C. J., & James, M. N. G. (1991) *Adv. Protein Chem. 42*, 77−144.
Meador, W. E., Means, A. R., & Quiocho, F. A. (1992) *Science 257*, 1251−1255.
Meador, W. E., Means, A. R., & Quiocho, F. A. (1993) *Science 262*, 1718−1721.
Medvedeva, M. V., Bushueva, T. L., Shirinsky, V. P., Lukas, T. J., Watterson, D. M., & Gusev, N. B. (1995) *FEBS Lett. 360*, 89−92.
Newton, D. L., Oldewurtel, M. D., Krinks, M. H., Shiloach, J., & Klee, C. B. (1984) *J. Biol. Chem. 259*, 4419−4426.
Olwin, B. B., & Storm, D. R. (1985) *Biochemistry 24*, 8081−8086.
Ota, I. M., & Clarke, S. (1989) *Biochemistry 28*, 4020−4027.
Ota, I. M., & Clarke, S. (1994) *J. Biol. Chem. 264*, 54−60.
Pascual-Ahuir, J.-L., Mehler, E. L., & Weinstein, H. (1991) *Mol. Eng. 1*, 231−247.
Pedigo, S., & Shea, M. A. (1995a) *Biochemistry 34*, 1179−1196.
Pedigo, S., & Shea, M. A. (1995b) *Biochemistry 34*, 10676−10689.
Perrella, M., Benazzi, L., Shea, M. A., & Ackers, G. K. (1990) *Biophys. Chem. 35*, 97−103.
Persechini, A., & Kretsinger, R. H. (1988) *J. Biol. Chem. 263*, 12175−12178.
Persechini, A., Jarrett, H. W., Kosk-Kosicka, D., Krinks, M. H., & Lee, H. G. (1993) *Biochim. Biophys. Acta 1163*, 309−314.
Pettigrew, D. W., Romeo, P. H., Tsapis, A., Thillet, J., Smith, M. L., Turner, B. W., & Ackers, G. K. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1849−1853.
Porumb, T., Yau, P., Harvey, T. S., & Ikura, M. (1994) *Protein Eng. 7*, 109−115.
Potter, S. M., Henzel, W. J., & Aswad, D. W. (1993) *Protein Sci. 2*, 1648−1663.
Powers, V. M., Yang, Y. R., Fogli, M. J., & Schachman, H. K. (1993) *Protein Sci. 2*, 1001−1012.
Putkey, J. A., Slaughter, G. R., & Means, A. R. (1985) *J. Biol. Chem. 260(8)*, 4704−4712.
Rao, U., Teeter, M. M., Erickson-Viitanen, S., & DeGrado, W. F. (1992) *Proteins: Struct., Funct., Genet. 14*, 127−138.
Root, D. D., & Wang, K. (1994) *Biochemistry 33*, 12581−12591.
Rydel, T. J., Ravichandran, K. G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C., & Fenton, J. W., II (1990) *Science 249*, 277−280.
Saimi, Y., & Kung, C. (1994) *FEBS Lett. 350*, 155−158.
Schagger, H., & von Jagow, G. (1987) *Anal. Biochem. 166*, 368−379.
Seamon, K. B. (1980) *Biochemistry 19*, 207−215.
Seeholzer, S. H., & Wand, A. J. (1989) *Biochemistry 28*, 4011−4020.
Small, E. W., & Anderson, S. R. (1988) *Biochemistry 27*, 419−428.
Starovasnik, M. A., Su, D.-A., Beckingham, K., & Klevit, R. E. (1992) *Protein Sci. 1*, 245−253.
Svensson, B., Jönsson, B., Thulin, E., & Woodward, C. E. (1993) *Biochemistry 32*, 2828−2834.
Takeuchi, Y., Birckbichler, P. J., Patterson, M. K., Jr., Lee, K. N., & Carter, H. A. (1994) *Z. Naturforsch. 49*, 453−457.
Takeuchi, Y., Birckbichler, P. J., & Patterson, M. K., Jr. (1995) *Experientia 51*, 339−342.
Taylor, D. A., Sack, J. S., Maune, J. F., Beckingham, K., & Quiocho, F. A. (1991) *J. Biol. Chem. 266(32)*, 21375−21380.
Török, K., & Whitaker, M. (1994) *BioEssays 16(4)*, 221−224.
Török, K., Lane, A. N., Martin, S. R., Janot, J.-M., & Bayley, P. M. (1992) *Biochemistry 31*, 3452−3462.
Tsalkova, T. N., & Privalov, P. L. (1985) *J. Mol. Biol. 181*, 533−544.
Urbauer, J. L., Short, J. H., Dow, L. K., & Wand, A. J. (1995) *Biochemistry 34*, 0000.
Vanaman, T. C. (1980) in *Calcium and Cell Function* (Cheung, W. Y., Ed.) Vol. I, pp 41−58, Academic Press, New York.
VanBerkum, M. F. A., & Means, A. R. (1991) *J. Biol. Chem. 266-(32)*, 21488−21495.
VanBerkum, M. F. A., George, S. E., & Means, A. R. (1990) *J. Biol. Chem. 265*, 3750−3756.
Vandonselaar, M., Hickie, R. A., Quail, J. W., & Delbaere, L. T. J. (1994) *Struct. Biol. 1*, 795−801.
Verhoeven, A. S., & Shea, M. A. (1993) *Biophys. J. 64*, A169.
Wall, C. M., Grand, R. J. A., & Perry, S. V. (1981) *Biochem. J. 195*, 307−316.
Walsh, M., Stevens, F. C., Kuznicki, J., & Drabikowski, W. (1977) *J. Biol. Chem. 252*, 7440−7443.
Waltersson, Y., Linse, S., Brodin, P., & Grundström, T. (1993) *Biochemistry 32*, 7866−7871.
Wang, C. A. (1988) *Biochem. Biophys. Res. Commun. 156(2)*, 1033−1038.
Wang, C. A., Aquaron, R. R., Leavis, P. C., & Gergely, J. (1982) *Eur. J. Biochem. 124*, 7−12.
Wang, C.-L. A. (1985) *Biochem. Biophys. Res. Commun. 130(1)*, 426−430.
Wang, C.-L. A., Leavis, P. C., & Gergely, J. (1984) *Biochemistry 23*, 6410−6415.
Wang, J. H., & Sharma, R. K. (1980) *Ann. N.Y. Acad. Sci. 356*, 190−204.
Wang, J. H., Sharma, R. K., & Tam, S. W. (1980) in *Calcium and Cell Function* (Cheung, W. Y., Ed.) Vol. I, pp 305−328, Academic Press, New York.
Watterson, D. M., Van Eldik, L. J., Smith, R. E., & Vanaman, T. C. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73(8)*, 2711−2715.
Weinstein, H., & Mehler, E. L. (1994) *Annu. Rev. Physiol. 56*, 213−236.
Yao, Y., Schöneich, C., & Squier, T. C. (1994) *Biochemistry 33*, 7797−7810.

BI951934G