J|A|C|S
ARTICLES
Published on Web 04/27/2004

Designing Sequence to Control Protein Function in an EF-Hand Protein

Christopher G. Bunick,[†] Melanie R. Nelson,[‡,⊥] Sheryll Mangahas,[‡] Michael J. Hunter,[‡,#] Jonathan H. Sheehan,[†] Laura S. Mizoue,[†] Gerard J. Bunick,[§,∥] and Walter J. Chazin*[,†,‡]

Contribution from the Departments of Biochemistry and Physics and Center for Structural Biology, Vanderbilt University, 5140 BIOSCI/MRB III, Nashville, Tennessee 37232-8725; Department of Molecular Biology (MB-9), The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037; Department of Biochemistry, Cellular and Molecular Biology, and Graduate School of Genome Science and Technology, The University of Tennessee, Knoxville, Tennessee 37996; and Life Sciences Division, Oak Ridge National Laboratory, P.O. Box 2008, MS 6480, Oak Ridge, Tennessee 37831-6480

Received November 23, 2003; E-mail: walter.chazin@vanderbilt.edu

Abstract: The extent of conformational change that calcium binding induces in EF-hand proteins is a key biochemical property specifying Ca^{2+} sensor versus signal modulator function. To understand how differences in amino acid sequence lead to differences in the response to Ca^{2+} binding, comparative analyses of sequence and structures, combined with model building, were used to develop hypotheses about which amino acid residues control Ca^{2+}-induced conformational changes. These results were used to generate a first design of *calbindomodulin* (CBM-1), a calbindin D_{9k} re-engineered with 15 mutations to respond to Ca^{2+} binding with a conformational change similar to that of calmodulin. The gene for CBM-1 was synthesized, and the protein was expressed and purified. Remarkably, this protein did not exhibit any non-native-like molten globule properties despite the large number of mutations and the nonconservative nature of some of them. Ca^{2+}-induced changes in CD intensity and in the binding of the hydrophobic probe, ANS, implied that CBM-1 does undergo Ca^{2+} sensorlike conformational changes. The X-ray crystal structure of Ca^{2+}-CBM-1 determined at 1.44 Å resolution reveals the anticipated increase in hydrophobic surface area relative to the wild-type protein. A nascent calmodulin-like hydrophobic docking surface was also found, though it is occluded by the inter-EF-hand loop. The results from this first calbindomodulin design are discussed in terms of progress toward understanding the relationships between amino acid sequence, protein structure, and protein function for EF-hand CaBPs, as well as the additional mutations for the next CBM design.

Calcium (Ca^{2+}) signaling pathways are widely proliferated throughout the cell and participate in all basic cellular functions.[1] The ability to manipulate Ca^{2+} signaling pathways would provide a powerful tool for applications in therapeutic and biotechnological settings. One strategy to achieve this objective involves the rational re-engineering of EF-hand proteins, which perform the essential first step in converting the ion-based Ca^{2+} signal into biochemical cascades.

EF-hand proteins are a family of highly homologous Ca^{2+}-binding proteins (CaBPs) that have roles not only in intracellular Ca^{2+} signal transduction[2] but also in modulation of Ca^{2+} signals and Ca^{2+} homeostasis.[3,4] Although these proteins bind Ca^{2+} ions using a common helix−loop−helix structural motif known as the EF-hand,[5] they mediate diverse cellular processes.[6] This functional diversity results in part from the different types of Ca^{2+}-driven structural changes that the various EF-hand CaBPs undergo. EF-hand CaBPs involved in sensing and transducing Ca^{2+} signals undergo large Ca^{2+}-dependent conformational changes, whereas the response of those involved in Ca^{2+} signal modulation and homeostasis tends to be much more modest.[4,7] The key feature distinguishing these two classes is the generation of a hydrophobic target binding surface by the sensor proteins once they bind to calcium.[8−10] The molecular basis for the

[†] Vanderbilt University.
[‡] The Scripps Research Institute.
[§] The University of Tennessee.
[∥] Oak Ridge National Laboratory.
[⊥] Current address: Biomedical Information Solutions Division, SAIC, San Diego, California.
[#] Current address: Active-Sight, San Diego, California.

(1) Berridge, M. J.; Bootman, M. D.; Roderick, H. L. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 517−529.
(2) Van Eldik, L. J.; Watterson, D. M., Eds. *Calmodulin and Signal Transduction*; Academic Press: Boston, MA, 1998.
(3) Skelton, N. J.; Kördel, J.; Akke, M.; Forsen, S.; Chazin, W. J. *Nat. Struct. Biol.* **1994**, *1*, 239−245.
(4) Slupsky, C. M.; Sykes, B. D. In *Calcium as a Cellular Regulator*; Carafoli, E., Klee, C., Eds.; Oxford University Press: New York, 1999; pp 73−99.
(5) Kawasaki, H.; Kretsinger, R. H. *Protein Profile* **1994**, *1*, 343−517.
(6) Pochet, R., Ed. *Calcium: The molecular basis of calcium action in biology and medicine*; Kluwer Academic Publishers: Boston, MA, 2000.
(7) Nelson, M. R.; Chazin, W. J. *BioMetals* **1998**, *11*, 297−318.
(8) Finn, B. E.; Evenäs, J.; Drakenberg, T.; Waltho, J. P.; Thulin, E.; Forsén, S. *Nat. Struct. Biol.* **1995**, *2*, 777−783.

10.1021/ja0397456 CCC: $27.50 **© 2004 American Chemical Society**

diverse EF-hand CaBP functions is poorly understood because it is not known how subtle differences in the biophysical and structural properties of EF-hand CaBPs correlate with functional specificity. Furthermore, it is not known how differences in the primary sequences of the EF-hand CaBPs influence those important biophysical and structural properties, which, in turn, specify function.

The results described here represent an early step toward our ultimate goal to correlate sequence with structure and function for EF-hand CaBPs. Our strategy is to derive the requisite knowledge through an understanding of the inter-relation of EF-hand CaBP sequence, 3D-structure, and the Ca^{2+}-induced conformational response. Information of this type is a prerequisite for using protein engineering routinely in the manipulation of calcium signaling pathways and other biotechnology applications.[11,12]

This paper describes a method to develop and test hypotheses about the relationship between primary sequence and Ca^{2+}-induced structural response. The approach is intended to be iterative, involving successive design−characterize−analyze−redesign cycles. To focus our efforts, we have established the goal of reengineering a typical Ca^{2+} signal modulator, calbindin D_{9k}, to respond to the binding of Ca^{2+} in the manner of the typical Ca^{2+} sensor, calmodulin. We call this hybrid protein *calbindomodulin*.

Differences in the Ca^{2+}-Induced Responses of EF−Hand Proteins. Calbindin D_{9k} and calmodulin are representative examples of the modulator and sensor classes of EF-hand CaBPs, respectively. Calbindin D_{9k} (75 residues, ∼9 kDa) contains two EF-hands that are structurally organized into a single globular domain. Calmodulin (148 residues, ∼17 kDa) has two such globular domains (N- and C-terminal). Although the sequences of calbindin D_{9k} and calmodulin's N-terminal domain (CaM-N) are 25% identical and their structures in the apo state are very similar, there is a striking difference in their conformational response to Ca^{2+} binding: CaM-N undergoes a pronounced Ca^{2+}-induced conformational change while calbindin D_{9k} does not. CaM-N adopts an open conformation upon Ca^{2+} binding, characterized by an exposed hydrophobic patch.[8−10] Ca^{2+}-loaded calbindin D_{9k} remains in a closed conformation similar to its apo conformation.[3] Previous protein engineering studies of calbindin D_{9k} (e.g., refs 13−16) also provided a strong foundation from which to begin the studies described in this paper.

Calbindin D_{9k} is a member of the S100 subfamily of EF-hand CaBPs, which are characterized by a 14-residue S100-specific EF-hand binding loop in the N-terminal EF-hand. In contrast, canonical 12 residue loops are present in both of CaM-N's EF-hands.[17] Four of the seven Ca^{2+} ligands in S100-specific EF-hand binding loops are provided by backbone oxygen atoms. This contrasts to the Ca^{2+} coordination in canonical loops where the same four ligands are instead three monodentate side chain oxygens and only one backbone oxygen atom.[18] However, it has been shown that this difference in Ca^{2+} coordination is not the reason for the difference in Ca^{2+}-induced conformational changes: a mutant of calbindin D_{9k} in which the first EF-hand is mutated to a canonical-type Ca^{2+} binding loop (Ala14Δ + Ala15Asp + Pro20Gly + Asn21Δ) remains in the closed conformation upon binding calcium.[19]

One prominent proposal for why Ca^{2+}-loaded calbindin D_{9k} does not adopt an open conformation is the "preformation hypothesis." This hypothesis assumes calbindin D_{9k} does not undergo a CaM-like Ca^{2+}-induced conformational change because the positions of a limited number of critical side chain Ca^{2+} ligands are fundamentally different in the apo states (closed conformations) of Ca^{2+} signal modulators such as calbindin D_{9k} and Ca^{2+} sensors such as calmodulin and troponin C.[20,21] For instance, the Ca^{2+}-induced repositioning of the terminal bidentate glutamate in the binding loops of calbindin D_{9k} is clearly different from that seen in CaM-N or troponin C[3]. Glu27 in the N-terminal EF-hand of calbindin D_{9k} undergoes only a minor change in side chain conformation when Ca^{2+} binds. In contrast, Glu31 in the N-terminal EF-hand of CaM-N must undergo a significant change in backbone position upon Ca^{2+} binding.[20,21] The preformation hypothesis proposes that since this terminal bidentate glutamate is also part of the second helix in the EF-hand, the drastic repositioning of the Glu31 side chain in CaM-N requires a significant rearrangement in the backbone, which in turn leads to the altered interhelical positioning that is characteristic of the open conformation. This hypothesis has yet to be tested directly. Therefore, one goal of our research is to design mutations that disrupt the preformation in calbindin D_{9k} in order to determine how those mutated residues affect the protein's ability to undergo conformational change.

It is unlikely that the preformation hypothesis fully explains the differences in the Ca^{2+}-induced conformational response of signal modulators and signal sensors. The preformation hypothesis argues that CaM and troponin C adopt an open conformation in the Ca^{2+}-loaded state because the closed conformation is destabilized when Ca^{2+} binds, whereas in calbindin D_{9k} the closed conformation is not destabilized by Ca^{2+} binding. However, it is also possible that Ca^{2+}-loaded calbindin D_{9k} does not occupy an open conformation because the open conformation is destabilized by problems such as an excessive degree of exposed hydrophobic residues or packing conflicts in the hydrophobic core. These two possibilities are not mutually exclusive, and it seems likely that differences in both the open and closed conformations contribute to the observed differences in the Ca^{2+}-induced conformational response of signal modulator and signal sensor EF-hand CaBPs.

Mutation Design Strategy. The fundamental globular unit of EF-hand CaBPs is a four-helix domain containing a pair of

(9) Kuboniwa, H.; Tjandra, N.; Grzesiek, S.; Ren, H.; Klee, C. B.; Bax, A. *Nat. Struct. Biol.* **1995**, *2*, 768−776.
(10) Zhang, M.; Tanaka, T.; Ikura, M. *Nat. Struct. Biol.* **1995**, *2*, 758−767.
(11) Marshall, S. A.; Lazar, G. A.; Chirino, A. J.; Desjarlais, J. R. *Drug Discovery Today* **2003**, *8*, 212−221.
(12) Lazar, G. A.; Marshall, S. A.; Plecs, J. J.; Mayo, S. L.; Desjarlais, J. R. *Curr. Opin. Struct. Biol.* **2003**, *13*, 513−518.
(13) Forsén, S.; Kördel, J.; Grundstrom, T.; Chazin, W. J. *Acc. Chem. Res.* **1993**, *26*, 7−14.
(14) Julenius, K.; Thulin, E.; Linse, S.; Finn, B. E. *Biochemistry* **1998**, *37*, 8915−8925.
(15) Kragelund, B. B.; Jonsson, M.; Bifulco, G.; Chazin, W. J.; Nilsson, H.; Finn, B. E.; Linse, S. *Biochemistry* **1998**, *37*, 8926−8937.
(16) Linse, S.; Brodin, P.; Drakenberg, T.; Thulin, E.; Sellers, P.; Elmden, K.; Grundstrom, T.; Forsén, S. *Biochemistry* **1987**, *26*, 6723−6735.
(17) Strynadka, N. C. J.; James, M. N. G. *Curr. Opin. Struct. Biol.* **1991**, *1*, 905−914.
(18) Szebenyi, D. M. E.; Moffat, K. *J. Biol. Chem.* **1986**, *261*, 8761−8777.
(19) Johansson, C.; Ullner, M.; Drakenberg, T. *Biochemistry* **1993**, *32*, 8429−8438.
(20) Travé, G.; Lacombe, P.-J.; Pfuhl, M.; Saraste, M.; Pastore, A. *EMBO J* **1995**, *14*, 4922−4931.
(21) Gagné, S. M.; Li, M. X.; Sykes, B. D. *Biochemistry* **1997**, *36*, 4386−4392.

EF-hands. Previous work has shown that this four-helix domain should be treated as a single globally cooperative structural/functional unit.[22] Our approach to understanding the relationships among EF-hand CaBP sequence, structure, and function is based on the premise that within this globular unit a subset of residues govern essential molecular properties that provide fold and function and that these residues are primarily located in or directly contact the hydrophobic core of the four-helix domain. This parallels efforts to find such residues through identification of an evolutionarily conserved sequence.[23]

The initial hypotheses about the relationship between primary sequence and Ca^{2+}-induced structural response in calbindin D_{9k} and calmodulin were developed using a three-step process.[24]

1. Certain amino acids in the calbindin D_{9k} sequence were identified as preliminary candidates for mutation based on analyses of a sequence alignment of EF-hand CaBPs, analyses of a homology model of calbindin D_{9k} in the open conformation of CaM-N, and comparisons of all available three-dimensional structures of EF-hand CaBPs (see Materials and Methods for further details).

2. Each of these candidate residues was analyzed to determine whether it might contribute to the structural differences between calbindin D_{9k} and CaM. This candidate-by-candidate analysis also used the sequence alignments, the homology model of open calbindin D_{9k}, and the structural comparisons, as well as additional information from the literature about previously studied mutations at the residue site of interest.

3. Results of these analyses were combined with information about the biophysical properties of the amino acid side chains to formulate detailed hypotheses about how important a given residue site is for regulating Ca^{2+}-induced conformational change.

Four different classes of calbindin D_{9k} mutants (single-site, limited-site, multisite, calbindomodulin) were designed based on these hypotheses. Single- and limited-site mutants (typically 2−4 residues) test specific hypotheses about a single position or localized region in the EF-hand domain. Multisite mutants (typically 5−10 mutations) test broader hypotheses, for example, about the packing of helix pairs or triples. The more extensive calbindomodulin mutants are designed to achieve conversion of the conformational response upon Ca^{2+} binding. Our research program is using all of these approaches. Studies of one limited-site mutant (Phe36Gly) have been reported,[22] and analyses of additional single-site, limited-site, and multisite mutants are in progress to test specific hypotheses and the cooperative effects of multiple mutations on CaBP structure and function. This paper describes the design, production, and characterization of the first full calbindomodulin design (CBM-1).

Designing calbindin D_{9k} to undergo a Ca^{2+}-dependent conformational opening represents a stringent test of our understanding of the inter-relation of EF-hand CaBP sequence, structure, and function because it involves engineering in a new function. Desjarlais and co-workers have reported a highly complementary approach, with the goal of re-engineering CaM to remain closed upon Ca^{2+} binding.[25,26] Calbindomodulins could be produced either by the rational design of site-specific mutations or by a combination of random mutagenesis and selection. The extensive database of sequences, structures, and biophysical analyses for the EF-hand CaBP family led us to believe that a manual, rational design approach could be efficient and would also allow for direct testing of specific sequence-structure hypotheses we and others had already proposed. More high-throughput selection-based methods will ultimately be required to fully explore the inter-relation of sequence, structure, and function.

CBM-1 Design. There are 58 positions in the calbindin D_{9k} sequence that are not identical to structurally aligned residues in CaM-N (Figure 1a), and these were the prime candidates for CBM-1 mutations.

In all, 26 positions were identified as preliminary candidates that should be considered for further analysis in the design of CBM-1. Seven sites were identified from sequence information in the EF-Hand CaBP Data Library[24] (http://structbio.vanderbilt.edu/chazin/cabp_database/) because, although conserved, the Ca^{2+} sensors were different from the signal modulators. Three sites were identified from clear differences observed between the structures of apo calbindin D_{9k} and apo CaM-N and an additional site from a comparison of the closed and open conformations of classical signal sensors.[27] The remaining 15 sites were identified from the residue-by-residue scoring output from the MODELLER program used to generate the homology model of calbindin D_{9k} in the open conformation. From this group of 26 residues, 15 were selected as likely to play an important role in determining the conformational response to the binding of Ca^{2+}. Detailed hypotheses were formulated about the structural and functional roles of each of these residues, and mutations were designed for each of them.

The mutations were based on two general goals: destabilizing the closed conformation of calbindin D_{9k} by making it more like that of CaM and stabilizing the open conformation of calbindin D_{9k}. Eight mutations (Leu6Ile, Ile9Ala, Tyr13Phe, Leu23Ile, Leu31Val, Glu35Leu, Phe36Gly, Gln67Leu) were included relating to the closed conformation, built around the concept of breaking the preformation of calbindin D_{9k} by repacking the closed conformation and making it more CaM-like. These specific mutations were selected because the largest difference between apo calbindin D_{9k} and apo CaM is the interface between helices I and II, so repacking this particular interface is expected to be sufficient to break preformation of apo calbindin D_{9k}. The isolated multisite mutation designed to repack the helix I/II interface and specifically test the preformation hypothesis is currently under investigation in our laboratory. Mutations pertaining to the open conformation were made to improve the solvation properties of residues that become exposed at the surface and to relieve predicted steric conflicts (Figure 1b). Hence CBM-1 includes seven mutations (Lys12Val, Leu32Met, Leu49Met, Phe50Ile, Leu53Val, Leu69Met, Val70-Met) designed purely upon considerations of the stability of the open conformation. In the CBM-1 design, 13 of the calbindin D_{9k} residues were mutated to the CaM-N homologue and the other two were mutated to the CaM-C homologue (Figure 1a).

Four residues (Ile9Ala, Leu31Val, Phe36Gly, Gln67Leu) appear to be of particular importance in determining the response to the binding of Ca^{2+} because they are predicted to contribute

(22) Nelson, M. R.; Thulin, E.; Fagan, P. A.; Forsen, S.; Chazin, W. J. *Protein Sci.* **2002**, *11*, 198−205.
(23) Lockless, S. W.; Ranganathan, R. *Science* **1999**, *286*, 295−299.
(24) Nelson, M. R. Ph.D. Thesis, The Scripps Research Institute, 1999.
(25) Ababou, A.; Desjarlais, J. R. *Protein Sci.* **2001**, *10*, 301−312.
(26) Ababou, A.; Shenvi, R. A.; Desjarlais, J. R. *Biochemistry* **2001**, *40*, 12719−12726.
(27) Nelson, M. R.; Chazin, W. J. *Protein Sci.* **1998**, *7*, 270−282.

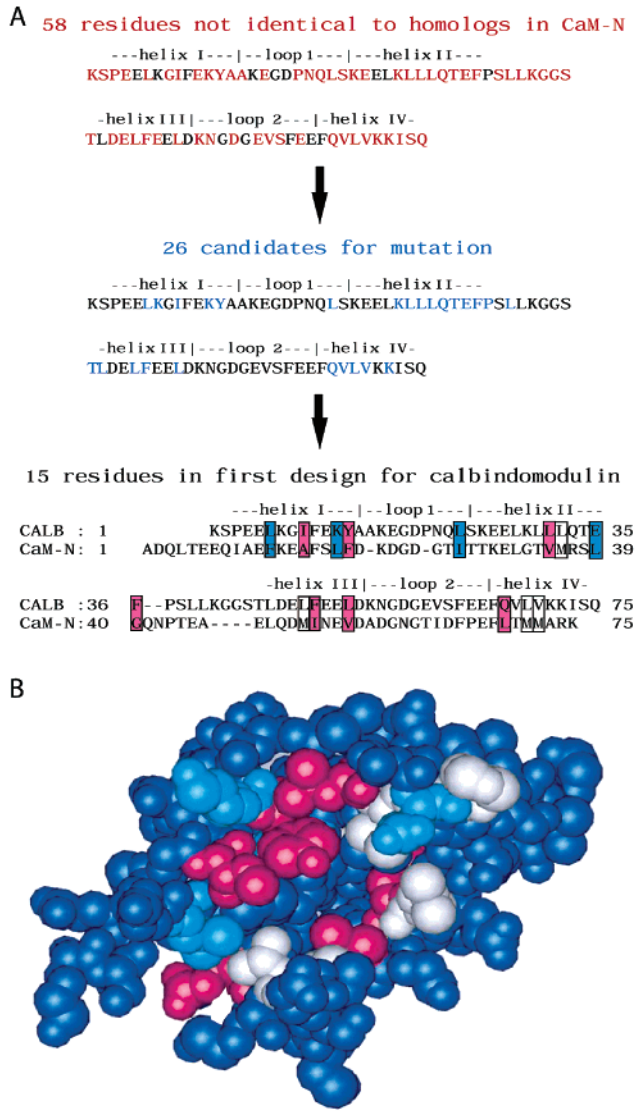

Figure 1. Mutations included in the first calbindomodulin design. (a) Selection of residues for mutation. Shown at the top in red are the locations of 58 nonidentical residues based on the alignment of the bovine calbindin D_{9k} and vertebrate CaM-N sequences. In the middle, the 26 potential candidates for mutation in calbindin D_{9k} are colored blue. At the bottom are the aligned sequences of bovine calbindin D_{9k} (CALB) and vertebrate CaM-N (CaM-N), with the 15 calbindin D_{9k} sites selected for mutation enclosed in boxes. The background color of a particular mutant site's box reflects the rationale for mutation: light blue, affect repacking of the closed conformation; white, affect solvation; magenta, affect steric hindrance. The corresponding CaM-N residues that the selected calbindin D_{9k} sites were mutated to are also enclosed in the box. Leu6 and Lys12 were the only calbindin D_{9k} residues mutated to the CaM C-terminal domain homologue (Ile and Val, respectively) instead of the CaM-N homologue (Phe and Leu, respectively). Analyses indicated that the Leu6Ile and Lys12Val mutations would not only repack the closed conformation but also better favor solvation in the open conformation (Leu6Ile) or cause less steric clash (Lys12Val). Three sites color coded here for steric hindrance (magenta) also were predicted to affect the closed conformation (Ile9Ala, Gln67Leu) or influence solvation (F36G). (b) Space-filling model of calbindin D_{9k} in the open conformation. The 15 sites of mutation are color coded based on the rationale for mutation, using the same color scheme described for the boxes in panel a.

significantly both to destabilization of the closed conformation and to stabilization of the open conformation. For example, Phe36 at the C-terminus of helix II was initially identified because its sequence homologue in Ca^{2+} sensors is a highly conserved glycine (Gly40 in calmodulin). In the open conformation, this residue is highly exposed, so Phe in this position would be highly destabilizing. In addition, glycine is a helix breaker, which implies the length and register of helix II would be altered. In fact, structure-based sequence alignment shows that the structural homologue of Phe36 in CaM-N is Leu39, not Gly40. These differences at the C-terminus of helix II would affect packing with other helices in the closed conformation and are proposed to contribute significantly to the preformation of calbindin D_{9k}. Support for this line of reasoning was obtained from the analysis of the single-site Phe36Gly mutation, which revealed perturbations of the apo protein structure that extend far beyond the site of mutation,[22] consistent with this residue being one of the key positions in the protein that control fold and function.

Calbindomodulin Characterization. The gene for CBM-1 was synthesized by one-pot shotgun ligation from a series of overlapping oligonucleotides.[28] The protein was expressed, purified, and characterized by standard methods. The dispersion of signals in the 1D ^{1}H NMR spectra of CBM-1 in both the absence and presence of Ca^{2+} indicated CBM-1 is folded (see Supporting Information). CBM-1 exhibits the increase in signal dispersion upon addition of Ca^{2+} characteristic of wild-type calbindin D_{9k} and virtually all other EF-hand CaBPs. Circular dichroism (CD) and fluorescence spectroscopy were used to monitor the Ca^{2+}-induced changes in protein conformation.

The prototypical EF-hand Ca^{2+} sensor CaM is distinguished by a substantial Ca^{2+}-dependent change in its CD spectrum.[29] In contrast, calbindin D_{9k} exhibits essentially no change in CD upon Ca^{2+}-loading.[30] Figure 2a shows the CD spectra of CBM-1, calbindin D_{9k}, and CaM in the absence and presence of Ca^{2+}. The CD spectra of all three proteins are characteristic of α-helical secondary structure. As expected, the spectra of calbindin D_{9k} in the apo and Ca^{2+}-loaded states are essentially identical, whereas the spectra of CaM are substantially different. The increase in ellipticity of CaM upon loading with Ca^{2+} reflects the reorganization of the helical packing within the globular domain, not an increase in helical content.[8,31] There is a similar Ca^{2+}-induced increase in ellipticity in the spectrum of CBM-1 upon Ca^{2+}-loading. This finding implies a Ca^{2+}-induced conformational reorganization in CBM-1 occurs similar to that in CaM.

To test whether the conformational change detected by CD results in an increased surface exposure of hydrophobic residues, the binding of the hydrophobic probe ANS to CBM-1 was monitored by fluorescence spectroscopy.[32,33] Control experiments were acquired for calbindin D_{9k}, which does not bind ANS, and for CaM, which binds ANS in a Ca^{2+}-dependent manner. A large 2.2-fold increase in fluorescence emission is observed upon addition of Ca^{2+} to CBM-1 (Figure 2b). This Ca^{2+}-induced increase is similar to the response seen with CaM (2.4-fold increase). CBM-1 also showed the ability to bind to phenyl-sepharose resin in a Ca^{2+}-dependent fashion, a property exhibited by Ca^{2+} sensors such as CaM,[34] but not calbindin

(28) Ye, Q. Z.; Johnson, L. L.; Baragi, V. *Biochem. Biophys. Res. Commun.* **1992**, *186*, 143−149.
(29) Martin, S. R.; Bayley, P. M. *Biochem J.* **1986**, *238*, 485−490.
(30) Chiba, K.; Ohyashiki, T.; Mohri, T. *J. Biochem* **1983**, *93*, 487−493.
(31) Gagne, S. M.; Tsuda, S.; Li, M. X.; Chandra, M.; Smillie, L. B.; Sykes, B. D. *Protein Sci.* **1994**, *3*, 1961−1974.
(32) Bayley, P.; Ahlstrom, P.; Martin, S. R.; Forsen, S. *Biochem. Biophys. Res. Commun.* **1984**, *120*, 185−191.
(33) Steiner, R. F. *Biopolymers* **1984**, *23*, 1121−1135.

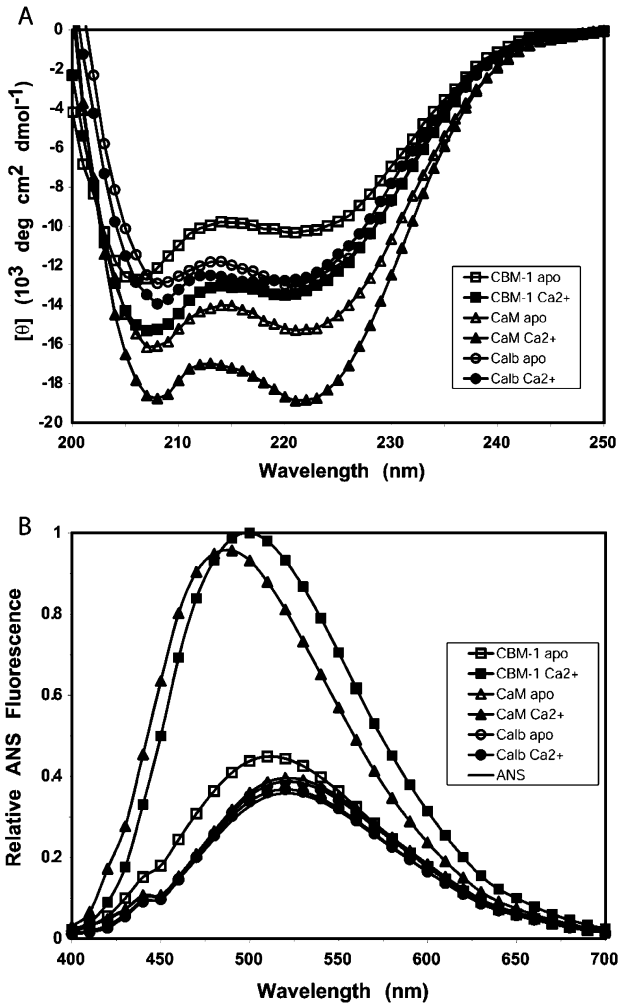

Figure 2. Ca^{2+}-dependent conformational changes in calbindomodulin detected by circular dichroism and ANS fluorescence. (a) Changes in circular dichroism (CD) intensity for CBM-1 indicate a Ca^{2+}-dependent conformational response. Ca^{2+}-CBM-1 (■) showed an increase in ellipticity compared to apo-CBM-1 (□). A similar increase was seen between apo-CaM (△) and Ca^{2+}-CaM (▲). In contrast, little change was seen between apo-calbindin D_{9k} (○) and Ca^{2+}-calbindin D_{9k} (●). (b) Changes in ANS fluorescence emission for CBM-1 indicate a Ca^{2+}-dependent increase in hydrophobic residue exposure. ANS showed increased fluorescence emission in the presence of Ca^{2+}-CBM-1 (■) compared to apo-CBM-1 (□). A similar increase in ANS fluorescence distinguished Ca^{2+}-CaM (▲) from apo-CaM (△), but no such increase was seen with calbindin D_{9k}. ANS fluorescence in the presence of apo (○) and Ca^{2+}-loaded (●) calbindin D_{9k} does not differ from that of ANS alone (solid black line).

D_{9k}. These data, combined with those from CD, indicate CBM-1 undergoes a conformational change upon binding Ca^{2+} that leads to an increase in the exposed hydrophobic surface.

In addition to monitoring Ca^{2+}-induced conformational change, the CD and ANS experiments along with ^1H NMR confirmed that the structural integrity of CBM-1 is maintained in the absence and presence of Ca^{2+} despite CBM-1 containing 15 mutations in the hydrophobic core, including some that are nonconservative such as Lys12Val, Glu35Leu, and Gln67Leu. This finding is significant, since several protein engineering studies have found a protein's core to be highly sensitive to mutations. For example, mutations in the central β core of the TIM barrel often caused losses in protein function,[35] core mutations greatly affected the stability and fold of the bacteriophage P22 Arc repressor,[36,37] and mutations to the T4 lysozyme core had affects on stability, cavity formation, and residue packing in the protein interior.[38−41]

Calbindomodulin Structure and Its Comparison to Calbindin D_{9k} and Calmodulin. To fully assess the CBM-1 design and continue the iterative design process, it was essential to determine the three-dimensional structure of Ca^{2+}-loaded CBM-1. A broad crystallization screen yielded promising results when zinc was added to the crystallization buffer and crystals diffracting to 1.44 Å resolution were obtained. The initial phase problem was solved using multiple anomalous dispersion (MAD) phasing, made possible by the presence of two zinc ions in the crystal asymmetric unit. The asymmetric unit also contained two CBM-1 molecules, which were refined independently without the use of noncrystallographic symmetry averaging. The structure was refined to a final *R*-factor of 0.158 (R_{free} = 0.194), and a summary of structural statistics is provided in Table 1.

CBM-1 is seen to adopt the two-EF-hand, four-helix domain fold characteristic of EF-hand CaBPs and coordinates two Ca^{2+} ions with the typical pentagonal bipyrimidal geometry (Figure 3).

The critical element in calbindomodulin design is to introduce opening of the globular domain upon Ca^{2+} binding. The CBM-1 design can be evaluated by comparing its structural features to calbindin D_{9k} and calmodulin. Since previous distance difference analyses showed that helices I and IV are the least variant among all EF-hand CaBPs,[7] these were used for all superpositions. The Ca^{2+}-induced opening of the domain can be described as the N-terminus of helix I shifting away from the C-terminus of helix II and the N-terminus of helix III shifting away from the C-terminus of helix IV, resulting in exposure of a large hydrophobic patch. In CBM-1, the helix I/II and helix III/IV interfaces have opened, though not to the extent to which they are open in CaM-N (Figure 4a). Comparison of the interhelical angles shows that the helical orientations in CBM-1 are between those of Ca^{2+}-calbindin D_{9k} and Ca^{2+}-CaM-N (Table 2).

Analysis of the accessible hydrophobic surface area (AHSA) of CBM-1 shows that CBM-1 has a significant exposed hydrophobic surface (Table 2). Corresponding values for calbindin D_{9k} and CaM-N are included for comparison. Because the sequences and corresponding surface areas of these proteins are not identical, the most insightful measure of relative AHSA is the ratio of AHSA to total accessible surface area, which is 10.4% for Ca^{2+}-calbindin D_{9k}, 16.8% for Ca^{2+}-CBM-1, and 18.7% for Ca^{2+}-CaM-N. Figure 4b illustrates the AHSA mapped onto the molecular surface of CBM-1. This shows not only that CBM-1 has a larger accessible hydrophobic surface compared

(34) Gopalakrishna, R.; Anderson, W. B. *Biochem. Biophys. Res. Commun.* **1982**, *104*, 830−836.
(35) Silverman, J. A.; Balakrishnan, R.; Harbury, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3092−3097.
(36) Milla, M. E.; Brown, B. M.; Sauer, R. T. *Nat. Struct. Biol.* **1994**, *1*, 518−523.
(37) Brown, B. M.; Sauer, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 1983−1988.
(38) Baldwin, E.; Xu, J.; Hajiseyedjavadi, O.; Baase, W. A.; Matthews, B. W. *J. Mol. Biol.* **1996**, *259*, 542−559.
(39) Mooers, B. H.; Datta, D.; Baase, W. A.; Zollars, E. S.; Mayo, S. L.; Matthews, B. W. *J. Mol. Biol.* **2003**, *332*, 741−756.
(40) Matthews, B. W. *Adv. Protein Chem.* **1995**, *46*, 249−278.
(41) Xu, J.; Baase, W. A.; Baldwin, E.; Matthews, B. W. *Protein Sci.* **1998**, *7*, 158−177.

Table 1. X-ray Data Collection and Refinement Statistics for Calbindomodulin

	native	peak	edge	remote
		Data Collection		
wavelength (Å)	1.0000	1.2818	1.2823	1.2703
resolution range[a] (Å)	50−1.44 (1.49−1.44)	50−1.85 (1.93−1.85)	50−1.85 (1.93−1.85)	50−1.85 (1.93−1.85)
number of reflections				
measured	149 862	96 734	96 894	99 659
unique	22 435	11 103	11 098	11 423
completeness[b]	94.7 (58.4)	97.7 (87.5)	97.8 (87.7)	98.9 (96.6)
$I/\sigma(I)$[b]	20.2 (6.1)	19.7 (7.3)	21.5 (7.0)	21.9 (5.4)
R_{sym} (%)[b]	7.5 (16.4)	8.3 (16.1)	7.7 (18.2)	7.9 (25.5)
		Refinement Statistics		
resolution range (Å)[a]	28.40−1.44 (1.48−1.44)			
reflections				
total	21 245			
test set[c]	1164			
R-factor[b] (%)	15.8 (17.3)			
R_{free}[b] (%)	19.4 (20.6)			
average *B*-factor[d] (Å^2)	10.8 (19.5)			
rms deviation from ideal values				
bond lengths (Å)	0.010			
bond angles (Å)	1.465			

[a] Values in parentheses show the highest resolution shell in Å. [b] Values for the highest resolution shell are given in parentheses. [c] Represents 5.2% of the reflections. [d] Wilson plot *B*-factor in parentheses.

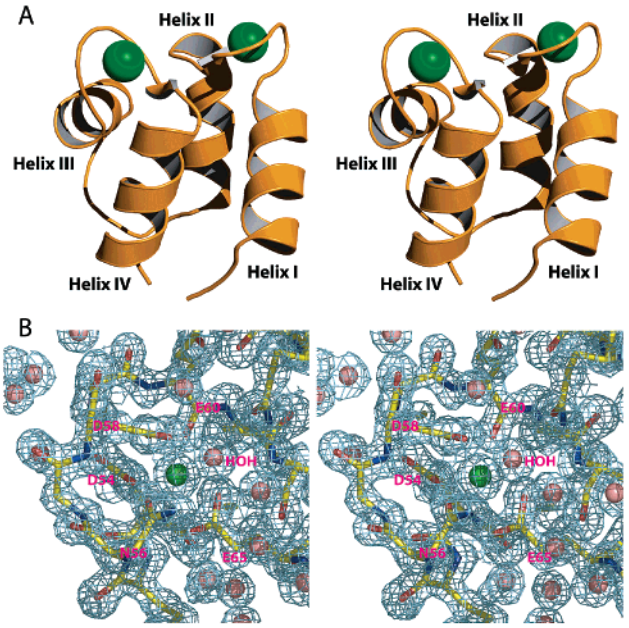

Figure 3. Calbindomodulin structure forming a four-helix domain characteristic of EF-hand calcium binding proteins. (a) Stereo representation of the global fold of CBM-1 (orange) and its bound calcium ions (green). (b) Stereo representation of the electron density at the calcium coordination site in the canonical EF-hand of CBM-1. The calcium ion is rendered as a green sphere, while water molecules are rendered as pink spheres. CBM-1 is colored according to atom type: carbon, yellow; oxygen, red; nitrogen, blue. Asp54, Asn56, and Asp58 provide monodentate Ca^{2+} coordination via their side chains. Glu60 coordinates Ca^{2+} with its main chain oxygen, whereas Glu65 is involved in bidentate coordination of the Ca^{2+} via its side chain. A water molecule (labeled HOH) provides the final direct ligand to the coordinated Ca^{2+}. The electron density is contoured at 1.0σ.

to calbindin D_{9k} but also importantly that significant hydrophobic patches are created.

The nature of the hydrophobic accessible surface of CBM-1 appears at first glance to be different than the deep hydrophobic pocket evident in CaM-N (Figure 4b). Detailed analysis of the CBM-1 structure indicates the difference is due to the positioning of the "linker" loop between the two EF-hands (connecting helices II and III), which makes significant contacts into the hydrophobic core in CBM-1 but not in CaM-N. The contacts in CBM-1 are primarily to side chains of residues in helix IV and result in the linker occupying a significantly different position than in CaM-N, one which occludes the hydrophobic patch. To understand if in fact the core of the protein has been altered in the manner of calmodulin, the linker can be deleted *in silico*. Figure 4b shows a surface representation of CBM-1 with the linker residues removed (CBM-1Δ_{36-45}), which reveals that CBM-1 does contain a deep hydrophobic binding pocket reminiscent of the CaM-N target binding surface. In contrast, removal of the linker in calbindin D_{9k} reveals clear differences in shape (convex versus concave) and size.

Conclusions

The S100 subfamily of EF-hand CaBPs, of which calbindin D_{9k} is a member, has a conserved inter-EF-hand linker motif known as the hydrophobic triad.[42] The consensus hydrophobic triad is Glu-H-X-X-H-H, where H represents a hydrophobic amino acid (typically Leu or Phe), and X, any nonhydrophobic amino acid. Phe36, Leu39, and Leu40 are the three hydrophobic amino acids in the triad motif of calbindin D_{9k}. The linker between the N- and C-terminal EF-hands of CaM-N differs from that of calbindin D_{9k} in that it is five residues shorter, does not have the consensus hydrophobic triad motif, and is much less hydrophobic overall. In the second iterative design of calbindomodulin (CBM-2), emphasis will be placed on disrupting the contacts observed in CBM-1 between the linker and residues in helix IV. This should release the linker so that it no longer occludes the exposed hydrophobic target binding surface. The packing of the linker into the hydrophobic core is also believed to prevent the EF-hands from fully opening. Hence, it is anticipated that the release of the linker will also promote increases in the I/II and III/IV interhelical angles. The selection of mutations will be directed to conversion of the hydrophobic triad residues to more polar, readily solvated residues.

The CBM-1 structure described in this paper represents an advance toward understanding the relationship between amino

(42) Groves, P.; Finn, B. E.; Kuznicki, J.; Forsen, S. *FEBS Lett.* **1998**, *421*, 175−179.

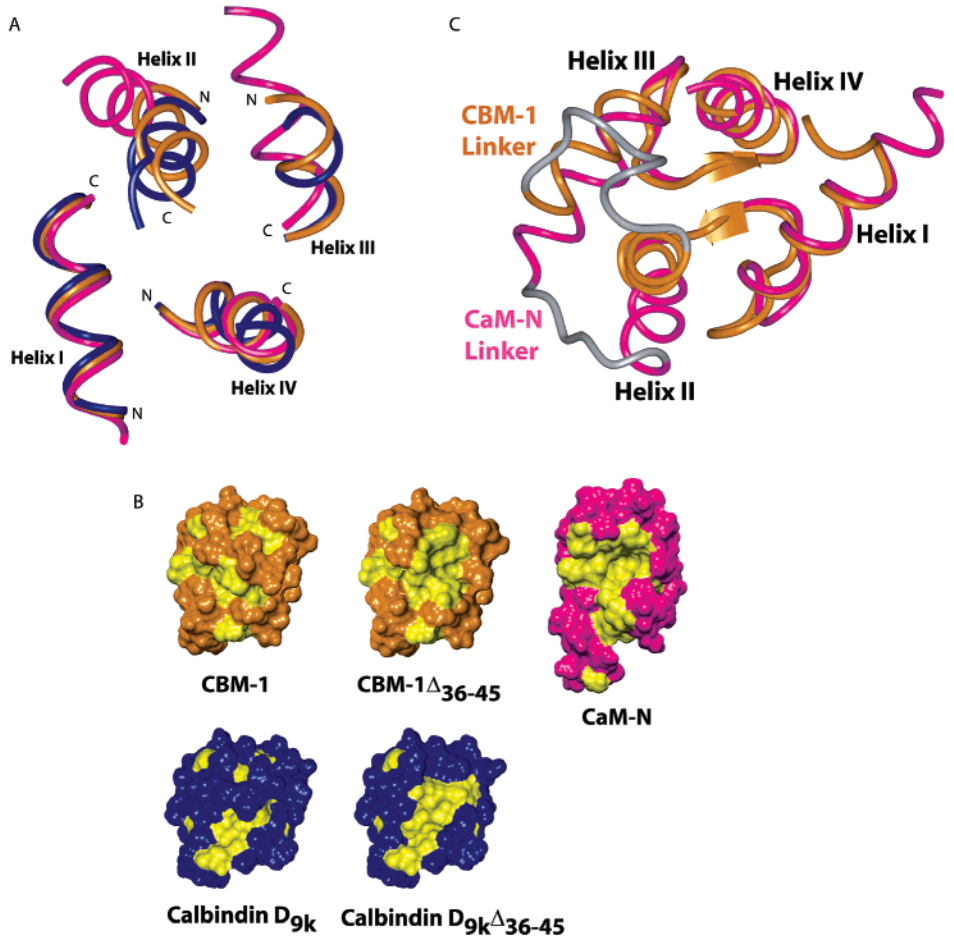

Figure 4. Structural comparison between the Ca^{2+}-bound states of calbindomodulin, calbindin D_{9k}, and the N-terminal domain of calmodulin. (a) Helical dispositions. CBM-1 (orange), calbindin D_{9k} (blue) (PDB Code 4icb), and CaM-N (magenta) (PDB Code 1cll) have been superposed on helices I and IV, and only the structured helical regions are shown. Helices II and III of CBM-1 have shifted conformation toward the positions of helices II and III in CaM-N. Also, helix IV of CBM-1 does not have the curvature seen in the fourth helix of calbindin D_{9k}. (b) Connolly and hydrophobic surface comparisons between calbindomodulin, calbindin D_{9k}, and CaM-N. From left to right, top row, there is intact calbindomodulin (CBM-1), calbindomodulin without the inter-EF-hand linker (CBM-1Δ_{36-45}), and intact CaM-N. From left to right, bottom row, there is intact calbindin D_{9k} and then calbindin D_{9k} without the inter-EF-hand linker (calbindin $D_{9k}\Delta_{36-45}$). The molecules have their helices in the same orientations as in panel a, providing a view into the hydrophobic pocket. The molecular surfaces were generated using InsightII after adding hydrogens to the structures (same PDB Codes as in panel a). The accessible hydrophobic surface is highlighted in yellow and defined by the residues Val, Met, Leu, Ile, and Phe. The rest of the molecular surface is highlighted in orange for CBM-1, in blue for calbindin D_{9k}, and in magenta for CaM-N. (c) Differences in the positioning of the inter-EF-hand linker between Ca^{2+}-CBM-1 and Ca^{2+}-CaM-N. Ca^{2+}-CaM-N (magenta, PDB Code 1cll) has been superimposed onto CBM-1 (orange) to contrast the positioning of the inter-EF-hand linkers (highlighted in gray) between the two.

acid sequence, protein structure, and protein function for EF-hand CaBPs. Several aspects of this first design cycle were successful: CD and ANS fluorescence detected a Ca^{2+}-dependent conformational change in CBM-1; structure determination of CBM-1 revealed helical orientations and AHSA that increased upon Ca^{2+} binding; and the structure showed the presence of a nascent CaM-like hydrophobic target binding surface. Additional challenges remain to bring the helical orientations and AHSA values closer to CaM and expose the putative target binding surface. The remarkable fact that the protein remained a well-folded globular domain despite mutating 15 residues is presumably a reflection of the detailed knowledge of EF-hand proteins already available. More importantly, it is an important positive sign for the probability of ultimately being successful in our protein engineering efforts. The next design cycles will attempt to improve the inter-EF-hand linker loop position and the helical orientations to achieve a more complete conversion of an EF-hand Ca^{2+} signal modulator into a Ca^{2+} sensor. Once a calbindomodulin is successfully created, it will be essential to work backward and determine which of the mutations are absolutely required. Calbindomodulin studies will be complemented by ongoing studies of other limited-site and multisite mutants. Ultimately, high-throughput selection-based strategies will be needed to more fully explore sequence space. The rational design of site-specific mutants such as those described here will continue to be an essential step toward uncovering the physical explanations for how the EF-hand CaBP sequence determines its structure and function and will provide the foundation for using protein engineering to manipulate calcium signaling pathways.

Materials and Methods

Sequence Alignments. The EF-Hand Calcium-Binding Proteins Data Library[24] (http://structbio.vanderbilt.edu/chazin/cabp_database/) was designed to link EF-hand CaBP amino acid sequences with information about the structure and function of the proteins. Therefore, the sequence information and appropriate annotations were taken from this resource. The alignments in the EF-hand CaBP Data Library were originally extracted from multiple sequence alignments constructed with CLUST-

Table 2. Comparison of the Structural Characteristics of Ca^{2+}-Calbindin D_{9k}, Ca^{2+}-Calbindomodulin (CBM-1), and Ca^{2+}-Calmodulin N-Terminal Domain (CaM-N)

	calbindin D_{9k}	CBM-1 A	CBM-1 B	CaM-N
	Interhelical Angles[a]			
interface				
I/II	134	130	129	89
I/IV	126	120	120	106
II/III	110	106	100	109
II/IV	−24	−24	−21	−40
III/IV	122	114	108	87
	Accessible Surface Area[b]			
total (TASA, Å^2)	4683.5	4739.0	4768.7	5247.6
hydrophobic (AHSA, Å^2)	486.8	770.8	801.2	983.1
AHSA as % of TASA	10.4	16.3	16.8	18.7

[a] Helices are defined as in the PDB coordinate files. PDB accession codes: Ca^{2+}-calbindin D_{9k}, 4icb (X-ray); Ca^{2+}-CBM-1, 1qx2 (X-ray); Ca^{2+}-calmodulin-N, 1cll (X-ray). [b] PDB accession codes were the same as those for the interhelical angle calculations. Hydrogen atoms were modeled onto the crystal structures using InsightII prior to the surface area calculations. Hydrophobic residues were defined as Val, Ile, Leu, Met, and Phe. Abbreviations: TASA, total accessible surface area; AHSA, accessible hydrophobic surface area.

ALW[43] using the web interface provided by the Network Protein Sequence Analysis (NPSA) group at Pôle Bio-Informatique Lyonnais (http://pbil.ibcp.fr/). Multiple sequence alignments were generated for each subfamily of EF-hand CaBPs cataloged in the data library. All available sequences of the proteins in the subfamily were aligned using the NPSA website and the alignments used to define the EF-hands in the proteins. The proteins were then broken into two-EF-hand units, since this is the smallest functional unit observed in this protein family, and all of the two-EF-hand units in each subfamily were aligned with CLUSTALW.

Model of Calbindin D_{9k} in the Open Conformation. A model of the calbindin D_{9k} sequence forced into the open conformation was constructed to assist in the identification of residues in the protein that may destabilize the open conformation. The model was built by homology with the Ca^{2+}-loaded N- and C-terminal domains of calmodulin, using the Modeller portion of the Homology module of InsightII [Molecular Simulations Inc. (MSI), San Diego, CA]. This module provides a graphical interface to the MODELLER algorithm,[44] which constructs probability density functions (PDFs) for the features of the unknown (model) structure. The default (medium) optimization level was used for all calculations. To improve the local geometry, a short restrained molecular dynamics simulated annealing protocol was performed on each structure after optimization of the molecular PDF. The final calculation was run 6 times using the CaM-N structure (PDB Code 1cll) and generated a family of converged representative conformers. The conformer with the lowest value of the total PDF was used as the representative structure.

The PDFs describe how likely the target structure is to have a given value for a particular feature. Since MODELLER reported the violations of the feature PDFs by residue, these violations were used to identify "problem" residues (i.e., residues that could not satisfy the PDFs for the open conformation). Such "problem" residues were considered to be good candidates for mutation but had to meet one of three conditions: (1) high total PDF violations; (2) high violations of the PDFs that determine side chain conformation (χ angles, side chain−main chain distances, and side chain−side chain distances); (3) high violations of other important conformational restraints (ϕ/ψ angles, C_α−C_α distances, N−O distances, soft sphere repulsions).

The model was built using the P43G mutation of calbindin D_{9k}. The proline at this position is mutated to either a methionine or glycine in all forms of calbindin D_{9k} studied in this laboratory to eliminate the cis−trans isomerization of Pro43 that leads to undesirable doubling of peaks in the nuclear magnetic resonance (NMR) spectra.[45] These P43 mutations do not greatly affect the properties of the protein.[46] The glycine mutation was chosen for the modeling studies because it was least likely to influence the rest of the model.

Structural Analyses. Detailed comparisons of the available structures of calbindin D_{9k} and calmodulin were used to help identify candidate residues for mutation. The comparisons were made using the methods described previously[27] and included analysis of interhelical angles calculated with the INTERHLX program,[47] interresidue contacts calculated with CHARMM,[48] and distance difference matrices calculated with DISCOM.[49] Extensive graphics-based comparisons were also performed using InsightII (MSI, San Diego).

Protein Production and Purification. The calbindomodulin (CBM-1) gene was synthesized by combining a series of overlapping oligonucleotides, which defined the desired CBM-1 gene and 5′ and 3′ oligonucleotides that allowed for amplification of the complete gene in a single polymerase chain reaction (PCR) mix.[28] DNA of the CBM-1 gene was purified by agarose gel electrophoresis, its sequence verified by DNA sequencing, and then it was digested with NdeI and NotI and subcloned into the NdeI/NotI sites in pSV271 (a pET27 derivative that was a generous gift from Dr. Navin Pokala, U. C. Berkeley). pSV271 was transformed into *Escherichia coli* strain BL21 (DE3), and preinduction cell cultures were grown in Luria Broth at 30 °C. After induction with 1 mM isopropyl-D-thiogalactopyranoside (IPTG), the temperature was raised to 37 °C for protein expression. After 4 h, cells were harvested and resuspended in 20 mM imidazole buffer (pH 7.0) containing 20 mM NaCl and 0.2 mM phenylmethylsulfonylfluoride (PMSF). The cells were lysed by three freeze/thaw cycles using an ethanol/dry ice bath. To this cell extract was added 10 mM $MgCl_2$, 0.8 mM PMSF, and 40 units/mL DNAse I, and then the solution was incubated at 37 °C for 20 min. Solutions were adjusted to 15 mM ethylenediaminetetraacetic acid (EDTA) and centrifuged at 2×10^4 g, 4 °C, for 25 min. The remainder of the protocol for CBM-1 purification was modified from that of Thulin for calbindin D_{9k}.[50] The supernatant was adjusted to 7 mM $CaCl_2$ and heated at 90 °C for 10 min, after which it was centrifuged at 1.6×10^4 g, 4 °C, for 10 min. The resulting supernatant was syringe-filtered through a 0.45 micron Whatman (Ann Arbor, MI) filter and applied to a 40 mL DEAE Sepharose column. The column was washed with 20 mM imidazole buffer (pH 7.0) containing 20 mM NaCl and 1.0 mM EDTA, and CBM-1 eluted by stepping the salt concentration to 135 mM. Collected fractions were analyzed by SDS-PAGE, and those containing CBM-1 were pooled, concentrated, and applied to a Superdex-75 Hi-Prep 16/60 column (Amersham Biosciences, Piscataway, NJ) with 20 mM Tris-HCl buffer (pH 8.0) containing 150 mM NaCl and 1 mM $CaCl_2$. Appropriate CBM-1 fractions were pooled, concentrated, and then either used in crystallization trials or dialyzed at 4 °C against several liters of water treated with Chelex (Bio-Rad, Hercules, CA), first with 1.0 mM EDTA, then without, to prepare apo protein. Mass spectrometry analysis confirmed CBM-1 expressed as native protein, and by convention, the N-terminal methionine is designated residue number "0." CBM-1 concentrations were determined using amino acid analysis.

Circular Dichroism and ANS Fluorescence. Circular dichroism (CD) measurements were made using samples that were 0.2 mg/mL

(43) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Res.* **1994**, *22*, 4673−4680.
(44) Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779−815.
(45) Chazin, W. J.; Kordel, J.; Drakenberg, T.; Thulin, E.; Brodin, P.; Grundstrom, T.; Forsen, S. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 2195−2198.
(46) Kördel, J.; Forsén, S.; Drakenberg, T.; Chazin, W. J. *Biochemistry* **1990**, *29*, 4000−4009.
(47) Yap, K. L.; Ames, J. B.; Swindells, M. B.; Ikura, M. *Methods Mol. Biol.* **2002**, *173*, 317−324.
(48) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(49) Gippert, G. Ph.D. Thesis, The Scripps Research Institute, 1995.
(50) Thulin, E. In *Calcium-Binding Protein Protocols: Volume I, Reviews and Case Studies*; Vogel, H., Ed.; Humana Press: Totowa, NJ, 2002; Vol. 172, pp 175−184.

protein, 10 mM Tris-HCl (pH 7.5), 100 mM KCl, and either 10 mM $CaCl_2$ (Ca^{2+}-loaded sample) or 1 mM EDTA (apo sample). A Jasco J-810 CD spectrometer (Easton, MD) was used to scan samples in a 0.1 cm path length cuvette from wavelength 260 to 190 nm (100 nm/min at 1 nm increments, 3 acquisitions) at 20 °C.

Fluorescence measurements were made using samples containing 2 μM protein and 40 μM 1-anilinonaphthalene-8-sulfonic acid (ANS) in Chelex-treated water, in the presence of either 2 mM $CaCl_2$ (Ca^{2+}-loaded sample) or 2 mM EDTA (apo sample). A Fluoromax 3 spectrometer (Jobin Yvon, Edison, NJ) was used to measure the emission spectrum from 400 to 700 nm (bandwidth 4 nm) at 20 °C, using 380 nm as the excitation wavelength for ANS.

Crystallization and Data Collection. CBM-1 crystals were grown at 18 °C by the hanging drop vapor diffusion method using a reservoir solution containing 25% polyethylene gylcol 550 monomethyl ether, 100 mM MES (pH 6.5), and 10 mM zinc sulfate. Crystallization drops were prepared by mixing 2 μL of reservoir solution with 1 μL of sterile water and 1 μL of Ca^{2+}-loaded CBM-1 solution. The CBM-1 solution contained 2.5 mM protein, 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 7.25 mM $CaCl_2$.

Both MAD and native data (Table 1) were collected on a single crystal using the ID-22 SER-CAT beamline at APS. The crystal was flash-cooled directly in the 100 K nitrogen-gas stream using the mother liquor as cryoprotectant and then underwent a single round of in situ macromolecular crystal annealing[51,52] to reduce its mosaicity (from 1.3° to 0.7°) prior to complete data collection. MAD data were collected to 1.84 Å resolution using zinc as the anomalous scattering atom. The 1.44 Å resolution native data set was used for the structure refinement. The crystal belonged to the orthorhombic space group $C222_1$ (cell dimensions: $a = 59.54$ Å, $b = 62.17$ Å, $c = 69.46$ Å, $\alpha = \beta = \gamma = 90°$) and had two molecules per asymmetric unit and a unit cell solvent content of 34%. All diffraction data were processed using HKL2000.[53]

Structure Determination and Refinement. SOLVE[54] was used to locate the two zinc sites in the asymmetric unit and to estimate the initial experimental phases. Electron density modification was performed using RESOLVE[55] and DM.[56,57] The resulting electron density map was used by ARP/wARP[57,58] to build 5 chains containing 148 out of the 152 residues in the asymmetric unit, with the model having an overall connectivity index score of 0.96. The CBM-1 structure then underwent iterative rounds of model building and refinement. Model building was performed using O[59] and XtalView.[60] Refinement was done using Refmac5[57,61] and SHELX,[62] following a typical stepwise crystallographic refinement strategy. Noncrystallographic symmetry averaging was not used in order to preserve conformational differences in the two asymmetric unit molecules. The final model of the two CBM-1 molecules in the asymmetric unit contains 151 amino acids, 213 waters, 4 Ca^{2+} ions, and 2 Zn^{2+} ions (located at the CBM-1 surface where they help pack neighboring protein molecules). The structure shows good stereochemistry (Table 1), and the Ramachandran plot shows all of the backbone torsional angles are in the most favored or additionally allowed regions. Structural analyses of CBM-1 used the INTERHLX, CHARMM, and InsightII programs as described for the rational design purposes (see Structural Analyses), as well as GRASP[63] for the calculation of accessible surface areas (surface area probe density level 4, 1.4 Å sphere radius). Figures were prepared using PyMOL[64] (Figure 3) and InsightII (Figure 4).

Coordinates. CBM-1 coordinates and structure factors have been deposited in the Protein Data Bank (accession code 1qx2).

Acknowledgment. We thank Drs. Sture Forsén, Bryan Finn, Anders Malmendal, John Christodoulou, Haitao Hu, Sara Linse, Yonglin Hu, B. Leif Hanson, Joel Harp, Gerald Stubbs, Feng Wang, Jarrod Smith, and staff at the SER-CAT beamline for assistance with experiments and helpful discussions. We also thank members of the Chazin lab, especially Dr. Kevin Weiss and Brian Weiner, for their contributions to the project and the Scripps Metalloprotein Structure and Design Program for their interest and support in pursuing this research. This work was supported by operating grants from the National Institutes of Health (PO1 GM 48495, RO1 GM 40120 to W.J.C.; RO1 GM 29818 to G.J.B.) and NASA (NAG8-1826 to G.J.B.), as well as a graduate fellowship from the National Science Foundation (to M.R.N.) and awards from the National Institute of General Medical Sciences for the Vanderbilt Medical Scientist Training Program (T32 GM 07347 to C.G.B.) and Vanderbilt Molecular Biophysics Training Program (T32 GM 04120 to J.H.S.).

Supporting Information Available: 1D ^{1}H NMR spectra of CBM-1 in the absence and presence of calcium (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0397456

(51) Harp, J. M.; Timm, D. E.; Bunick, G. J. *Acta Crystallogr.* **1998**, *D54*, 622−628.
(52) Harp, J. M.; Hanson, B. L.; Bunick, G. J. *Acta Crystallogr.* **1999**, *D55*, 1329−1334.
(53) Otwinowski, Z.; Minor, W. *Processing of X-ray Diffraction Data Collected in Oscillation Mode*; Academic Press: New York, 1997; Vol. 276, Macromolecular Crystallography, part A.
(54) Terwilliger, T. C.; Berendzen, J. *Acta Crystallogr.* **1999**, *D55*, 849−861.
(55) Terwilliger, T. C. *Acta Crystallogr.* **2000**, *D56*, 965−972.
(56) Cowtan, K. D.; Zhang, K. Y. *Prog. Biophys. Mol. Biol.* **1999**, *72*, 245−270.
(57) *Acta Crystallogr.* **1994**, *D50*, 760−763.
(58) Perrakis, A.; Morris, R.; Lamzin, V. S. *Nat. Struct. Biol.* **1999**, *6*, 458−463.
(59) Jones, T. A.; Bergdoll, M.; Kjeldgaard, M. *O: A macromolecular modeling environment*; Springer-Verlag Press: New York, 1990.
(60) McRee, D. E. *J. Struct. Biol.* **1999**, *125*, 156−165.
(61) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr.* **1997**, *D53*, 240−255.
(62) Sheldrick, G. M. *Acta Crystallogr.* **1990**, *A46*, 467−473.
(63) Nicholls, A.; Sharp, K.; Honig, B. *Proteins: Struct., Funct., Genet.* **1991**, *11*, 281ff.
(64) DeLano, W. L. *PyMOL*, 0.90 ed.; DeLano Scientific LLC: San Carlos, CA.