

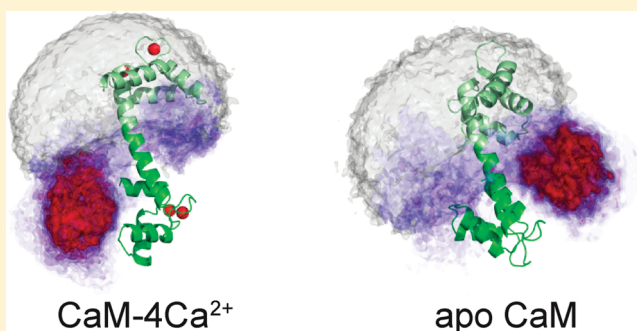
# Transient, Sparsely Populated Compact States of Apo and Calcium-Loaded Calmodulin Probed by Paramagnetic Relaxation Enhancement: Interplay of Conformational Selection and Induced Fit

Nicholas J. Anthis, Michaelleen Doucleff,<sup>†</sup> and G. Marius Clore\*

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States

**S** Supporting Information

**ABSTRACT:** Calmodulin (CaM) is the universal calcium sensor in eukaryotes, regulating the function of numerous proteins. Crystallography and NMR show that free CaM-4Ca<sup>2+</sup> exists in an extended conformation with significant interdomain separation, but clamps down upon target peptides to form a highly compact structure. NMR has revealed substantial interdomain motions in CaM-4Ca<sup>2+</sup>, enabled by a flexible linker. In one instance, CaM-4Ca<sup>2+</sup> has been crystallized in a compact configuration; however, no direct evidence for transient interdomain contacts has been observed in solution, and little is known about how large-scale interdomain motions contribute to biological function. Here, we use paramagnetic relaxation enhancement (PRE) to characterize transient compact states of free CaM that are too sparsely populated to observe by traditional NMR methods. We show that unbound CaM samples a range of compact structures, populated at 5–10%, and that Ca<sup>2+</sup> dramatically alters the distribution of these configurations in favor of states resembling the peptide-bound structure. In the absence of Ca<sup>2+</sup>, the target peptide binds only to the C-terminal domain, and the distribution of compact states is similar with and without peptide. These data suggest an alternative pathway of CaM action in which CaM remains associated with its kinase targets even in the resting state. Only CaM-4Ca<sup>2+</sup>, however, shows an innate propensity to form the physiologically active compact structures, suggesting that Ca<sup>2+</sup> activates CaM not only through local structural changes within each domain but also through more global remodeling of interdomain interactions. Thus, these findings illustrate the subtle interplay between conformational selection and induced fit.



## INTRODUCTION

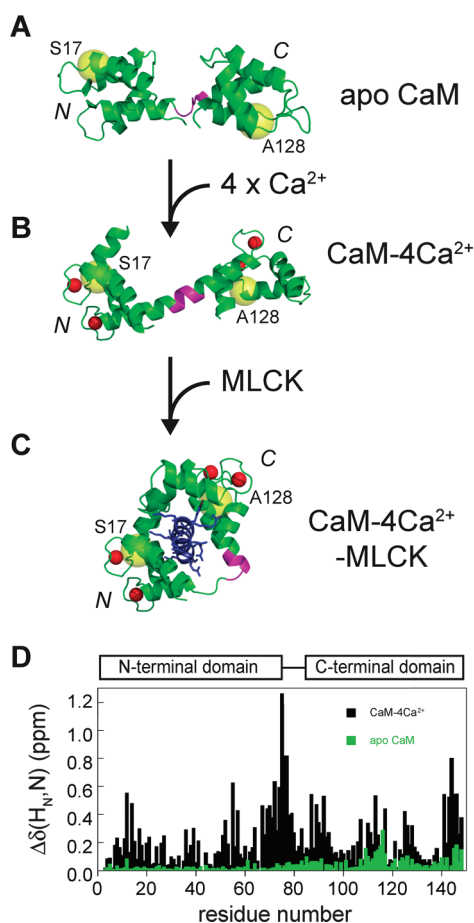
As the cell's prototypical Ca<sup>2+</sup> sensor, calmodulin (CaM) is responsible for amplifying and adapting this universal signal into specific downstream signals via interactions with hundreds of other proteins. This biological adaptability correlates with the significant structural plasticity that has been observed in CaM.<sup>1</sup> Figure 1 summarizes existing static structures of apo CaM,<sup>2</sup> CaM-4Ca<sup>2+</sup>,<sup>3</sup> and the complex between CaM-4Ca<sup>2+</sup> and the CaM binding domain (CBD) of myosin light chain kinase (MLCK).<sup>4</sup> Upon binding calcium, 25–50° shifts in the interhelical angles within the two domains of CaM expose methionine-rich hydrophobic patches,<sup>2,5</sup> thereby priming CaM to interact with a variety of binding partners that are self-inhibited by a CBD in the resting state (such as MLCK and other kinases). Target self-inhibition is relieved upon binding of CaM-4Ca<sup>2+</sup> to form a highly compact structure (CaM-4Ca<sup>2+</sup>-MLCK) in which the N- and C-terminal domains of CaM are clamped around an  $\alpha$ -helical CBD peptide and make contact with one another.<sup>4,6</sup> This is accomplished primarily through extensive bending of the central linker, with the two domains of CaM behaving largely as

rigid bodies, and through the concerted formation of an extensive network of hydrophobic and ion-pair interactions between the helical peptide and the two domains of CaM.<sup>7</sup> Although CaM can also form other types of complexes, with differing Ca<sup>2+</sup> dependencies,<sup>1,8</sup> we concern ourselves here specifically with the canonical Ca<sup>2+</sup>-dependent 1:1 CaM/peptide interaction, represented by CaM-4Ca<sup>2+</sup>-MLCK (Figure 1C).

Although all crystal structures of free CaM-4Ca<sup>2+</sup> (e.g., refs 3, 9, 10), with one exception,<sup>11</sup> appear as an extended dumbbell structure with a contiguous central helix connecting the two domains separated by a center-to-center distance of  $\sim 40$  Å (Figure 1B), fluorescence energy transfer experiments<sup>12</sup> and small-angle X-ray and neutron scattering<sup>13</sup> indicate that the average structure of CaM-4Ca<sup>2+</sup> in solution is somewhat more compact. NMR relaxation studies on both apo CaM<sup>14</sup> and CaM-4Ca<sup>2+</sup><sup>15,16</sup> have shown unambiguously that the central residues (77–81) of the linker are highly mobile (with  $S^2$  order parameters

**Received:** September 1, 2011

**Published:** October 17, 2011



**Figure 1.** CaM structures and peptide interactions. (A–C) Structures of CaM in various physiological states. Ca<sup>2+</sup> ions are shown as red spheres, and spin-labeling sites S17C and A128C are indicated by yellow spheres. The flexible linker (residues 77–81)<sup>15</sup> is shown in magenta. (A) NMR structure of apo CaM (PDB 1CFD).<sup>2</sup> (B) Crystal structure of CaM-4Ca<sup>2+</sup> (PDB 1CLL)<sup>3</sup> in the extended dumbbell conformation with a continuous central helix. (C) Crystal structure of the CaM-4Ca<sup>2+</sup>-MLCK complex (PDB 1CDL).<sup>4</sup> (D) Weighted combined <sup>1</sup>H/<sup>15</sup>N chemical shift perturbations, Δδ(H<sub>N</sub>, N), for the backbone amide groups of apo CaM (green) or CaM-4Ca<sup>2+</sup> (black) upon saturation with MLCK peptide.

of 0.5–0.6), and that the two domains reorient almost independently of one another in solution on a time scale of ~3 ns, undergoing restricted diffusion within a cone of semiangle ~30°. Studies based on lanthanide-induced residual dipolar couplings (RDC) and pseudo-contact shifts (PCS) suggest that an even larger cone of semiangle 50–80° can be sampled.<sup>17</sup> In one instance, CaM-4Ca<sup>2+</sup> has been crystallized in a compact conformation that resembles the peptide-bound structure,<sup>11</sup> but this specific conformation must have been largely selected by crystal packing forces since there is no conventional NMR evidence (in the form of interdomain nuclear Overhauser effects (NOE)) for any direct contacts between the N- and C-terminal domains of CaM-4Ca<sup>2+</sup>.<sup>18</sup> Thus, the population of such a state, if it exists, must be very low in solution. Ensemble molecular dynamics (MD) simulations<sup>19</sup> also suggest that compact states can potentially be sampled, but it is unclear whether the empirical force field employed is sufficiently accurate or the sampling sufficiently extensive to reliably characterize such states.

Our goal in the current work is to directly probe and visualize compact, sparsely populated conformational substates of CaM, characterized by interdomain contacts, using paramagnetic relaxation enhancement (PRE), to determine how these differ in the presence and absence of Ca<sup>2+</sup>, and how such states may contribute to CaM function. These sparsely populated states are invisible to conventional structural biology techniques, owing to their low occupancy and transient nature. In the fast exchange regime, however, the footprint of sparsely populated states can be observed on the PRE profile of the major species, provided that the distances between the paramagnetic label and the protons of interest are shorter in the minor species than the major one.<sup>20</sup> The exquisite sensitivity of the PRE in this regard arises from the large magnetic moment of the unpaired electron and the  $\langle r^{-6} \rangle$  dependence of the transverse PRE rate ( $\Gamma_2$ ), resulting in extremely large  $\Gamma_2$  values at short distances.<sup>21</sup> Since interdomain motion in CaM takes place on the nanosecond time scale,<sup>14–16</sup> any CaM conformations exhibiting close interdomain contacts should be observable by PRE, thereby offering novel insights over prior studies, which, by their nature, primarily characterized the major, more extended states of CaM-4Ca<sup>2+</sup>.

## MATERIALS AND METHODS

**Mutagenesis and Protein Expression.** The 148-residue CaM protein from *Xenopus laevis* (identical in amino acid sequence to human CaM) was cloned into the pET21a (Novagen) vector using PCR primers with *Bam*HI and *Nde*I restriction sites. Spin-label sites were created by making single cysteine point mutations (T5C, S17C, T117C, and A128C) with the QuikChange Site-Directed Mutagenesis Kit (Stratagene).

Uniformly <sup>2</sup>H/<sup>13</sup>C/<sup>15</sup>N-labeled CaM was expressed in BL21-Codon-Plus (DE3) RIPL cells (Stratagene) at 37 °C using standard methods. Briefly, cells were grown in 1 L of M9 medium (using <sup>2</sup>H<sub>2</sub>O, <sup>15</sup>NH<sub>4</sub>Cl, and [<sup>2</sup>H,<sup>13</sup>C]-glucose as isotope sources) and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at A<sub>600 nm</sub> ~ 0.6. Cells were harvested by centrifugation 6–9 h later.

**Protein Purification and Nitroxide Spin-Labeling.** Cells were resuspended in 40 mM Tris, pH 7.4, 2 mM ethylenediaminetetraacetic acid (EDTA), 2 mM dithiothreitol (DTT), and 1× Roche Complete Protease Inhibitor. Cells were lysed using a microfluidizer, cleared by centrifugation, and then loaded on a DEAE ion-exchange column (GE Life Sciences) pre-equilibrated with 20 mM Tris, pH 7.4, 80 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 1 mM MgCl<sub>2</sub>. Protein was eluted with a 0–100% gradient over 10 column volumes with a high-salt buffer containing 1.2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. Relevant fractions were pooled and loaded onto phenyl-sepharose resin (GE Life Sciences) pre-equilibrated with 50 mM Tris, pH 7.4, 2 mM CaCl<sub>2</sub>, and 100 mM NaCl. CaM was eluted with the addition of 2 mM ethylene glycol tetraacetic acid (EGTA). Protein was stored in the elution buffer plus 0.5 mM tris(2-carboxyethyl)phosphine (TCEP) and 1× Roche Complete Protease Inhibitor.

CaM was paramagnetically labeled by the addition of a 10-fold excess of the nitroxide spin-label (1-oxyl-2,2,5,5-tetramethyl-δ-3-pyrroline-3-methyl) methanethiosulfonate (“MTSL”; Toronto Research Chemicals) to ~200 μM CaM. The reaction was allowed to proceed in the dark for 2 h and tested for completion by mass spectroscopy. The diamagnetic control was labeled by the same method using (1-acetoxy-2,2,5,5-tetramethyl-δ-3-pyrroline-3-methyl) methanethiosulfonate (Toronto Research Chemicals).

Unreacted spin-label was removed with a HiPrep 26/10 Desalting Column (GE Life Sciences) and exchanged into NMR buffer, comprising 5% D<sub>2</sub>O/95% H<sub>2</sub>O, 100 mM KCl, 0.1× Roche Complete Protease Inhibitor, 25 mM HEPES, pH 6.5, and either 8 mM CaCl<sub>2</sub> (for Ca<sup>2+</sup>-

bound CaM) or 2 mM EDTA and 2 mM EGTA (for apo CaM). NMR samples were concentrated to 0.3 mM with Amicon Ultra Centrifugal Filter Units (3 kDa molecular weight cutoff). For isothermal titration calorimetry (ITC), NMR assignment experiments, or other work on wild-type protein, CaM was purified as above, except instead of spin-labeling, buffer exchange was performed immediately following phenyl-sepharose purification. Experiments were performed on CaM alone, or in the presence of a synthetic peptide (KRRWKKNFIAVSAANRFK-KISSSGAL; Anaspec), corresponding to the CaM-binding domain (CBD) of human skeletal muscle myosin light chain kinase (skMLCK). For Ca<sup>2+</sup>-loaded CaM, a 1.2-fold excess of peptide was added, and the sample was then concentrated to 0.3 mM CaM. For apo CaM, final NMR samples were made directly with 0.3 mM CaM and a 2-fold excess of peptide.

**Isothermal Titration Calorimetry.** ITC was performed on a MicroCal iTC<sub>200</sub> calorimeter with CaM in the instrument cell and MLCK peptide in the injection syringe. Experiments were carried out at 27 °C in NMR buffer. Sixteen 2.4-μL injections were performed (5 s each, separated by 180-s intervals). Experimental data were analyzed using MicroCal Origin software.

**NMR Spectroscopy.** NMR experiments were performed on uniformly <sup>2</sup>H/<sup>13</sup>C/<sup>15</sup>N-labeled CaM. Backbone resonance assignments for apo CaM, CaM-4Ca<sup>2+</sup>, and CaM-4Ca<sup>2+</sup>-MLCK were obtained by triple resonance through-bond scalar correlation spectroscopy<sup>22</sup> using HNCO, HNCACB, and CBCA(CO)NH pulse sequences. Data were processed using NMRPipe<sup>23</sup> and analyzed with the program XIPP (in-house software written by D. S. Garrett). Transverse <sup>1</sup>H<sub>N</sub>-Γ<sub>2</sub> PRE rates were obtained from the differences in the transverse <sup>1</sup>H<sub>N</sub>-R<sub>2</sub> relaxation rates between the paramagnetic and diamagnetic samples<sup>21</sup> using either of two different pulse sequences, which yielded identical results within experimental error. Some measurements used a 3D HNCO-based pulse scheme similar to that described by Hu et al.,<sup>24</sup> but without using transverse relaxation optimized spectroscopy (TROSY). Others used a TROSY version of a 2D pulse scheme similar to that depicted by Iwahara et al.,<sup>25</sup> but with the variable delay performed at the end of the pulse sequence, thus, measuring the TROSY <sup>1</sup>H<sub>N</sub>-R<sub>2</sub>. Data were recorded at 27 °C on a Bruker 600 MHz spectrometer equipped with a triple resonance z-gradient cryoprobe. Two time points (separated by 20 ms) were used for the <sup>1</sup>H<sub>N</sub>-R<sub>2</sub> measurements, and the errors in the <sup>1</sup>H<sub>N</sub>-Γ<sub>2</sub> PRE rates were calculated as described previously.<sup>25</sup> Samples for PRE experiments contained 0.3 mM <sup>2</sup>H/<sup>13</sup>C/<sup>15</sup>N-labeled spin-labeled (or control) CaM. Control experiments were also performed to detect any intermolecular PREs that might be present. In these cases, measurements were made on 0.3 mM <sup>2</sup>H/<sup>13</sup>C/<sup>15</sup>N-labeled CaM with no spin label in the presence of 0.3 mM CaM at natural isotopic abundance (either spin-labeled or wild type for diamagnetic control).

<sup>1</sup>H-<sup>15</sup>N heteronuclear single quantum coherence (HSQC) correlation spectroscopy was used to follow the <sup>1</sup>H<sub>N</sub>/<sup>15</sup>N chemical shifts of <sup>15</sup>N-labeled CaM (0.1 mM) upon addition of increasing amounts of unlabeled MLCK peptide. Weighted combined <sup>1</sup>H<sub>N</sub> and <sup>15</sup>N amide shifts, Δ(*H<sub>N</sub>N*), for CaM saturated with MLCK peptide were calculated using the equation: Δ(*H<sub>N</sub>N*) = [(Δ<sub>H</sub>W<sub>H</sub>)<sup>2</sup> + (Δ<sub>N</sub>W<sub>N</sub>)<sup>2</sup>]<sup>1/2</sup> where W<sub>H</sub> and W<sub>N</sub> are weighting factors for the <sup>1</sup>H and <sup>15</sup>N amide shifts, respectively (W<sub>H</sub> = 1, W<sub>N</sub> = 0.154, based on average variances<sup>26</sup> in the BioMagResBank; <http://www.bmrb.wisc.edu>), and Δ is the chemical shift difference between bound and free states, δ<sub>bound</sub> - δ<sub>free</sub>.

**PRE Back-Calculations and Ensemble Simulated Annealing.** Theoretical <sup>1</sup>H<sub>N</sub>-Γ<sub>2</sub> PREs were back-calculated for each nitroxide spin-label position from the coordinates of the structure of CaM in four different states: (i) the Ca<sup>2+</sup>-loaded extended “dumbbell” configuration (PDB 1CLL),<sup>3</sup> (ii) the Ca<sup>2+</sup>-loaded, peptide-bound complex with smooth muscle MLCK (PDB 1CDL),<sup>4</sup> (iii) apo CaM (PDB 1CFD),<sup>2</sup> and (iv) the Ca<sup>2+</sup>-loaded (peptide free) compact configuration of CaM (PDB 1PRW).<sup>11</sup> All of these structures are X-ray structures, with one

exception, apo CaM (1CFD), which was solved by NMR. <sup>1</sup>H<sub>N</sub>-Γ<sub>2</sub> rates were back-calculated in Xplor-NIH<sup>28,29</sup> using a five-conformer ensemble for the spin-label together with the Solomon-Bloembergen Model Free (SBMF) representation, optimizing the coordinate positions in torsion angle space by simulated annealing to minimize the difference between observed and calculated <sup>1</sup>H<sub>N</sub>-Γ<sub>2</sub> PRE rates, as described previously.<sup>30</sup> In addition to the PRE pseudopotential, the target function includes stereochemical restraints, a quartic van der Waals repulsion term to prevent atomic overlap between the spin-label and the protein, and a multidimensional torsion angle database potential of mean force.<sup>31</sup> (Note atomic overlap between spin-label ensemble members is allowed since the ensemble represents a distribution of states.) For the spin-labels at positions S17C and A128C, the entire molecule, with the exception of side chains adjacent to the spin-label (which were given torsional degrees of freedom), was kept rigid during the fitting procedure. For the spin-labels at positions T5C and T117C, residues 1–8 and 115–120, respectively, were also given torsional degrees of freedom and represented by the same five-member ensemble as the spin-label itself, to account for local structural perturbations induced by the spin-label (Supporting Information Figure S3). Unless stated otherwise, spin-label positions were set by fitting to intradomain PRE data only.

Because the electron relaxation rate τ<sub>s</sub> of the nitroxide spin-label is much longer than that of the protein rotational correlation time τ<sub>r</sub>,<sup>25</sup> the PRE correlation time τ<sub>c</sub> [= (τ<sub>r</sub><sup>-1</sup> + τ<sub>s</sub><sup>-1</sup>)<sup>-1</sup>] for the calculation of PRE rates was assumed to be approximately the same as τ<sub>r</sub> (9.9 ns) for CaM-4Ca<sup>2+</sup>.<sup>32</sup> Agreement between observed and calculated Γ<sub>2</sub> rates is given by the PRE Q-factor:<sup>30</sup>

$$Q = [\sum_i \{\Gamma_{2,i}^{\text{obs}} - \langle \Gamma_{2,i}^{\text{calc}} \rangle\}^2 / \sum_i (\Gamma_{2,i}^{\text{obs}})^2]^{1/2} \quad (1)$$

where Γ<sub>2,i</sub><sup>obs</sup> and ⟨Γ<sub>2,i</sub><sup>calc</sup>⟩ are the observed and ensemble average back-calculated transverse Γ<sub>2</sub> rates for residue *i*, respectively. All members of the ensemble are weighted equally.

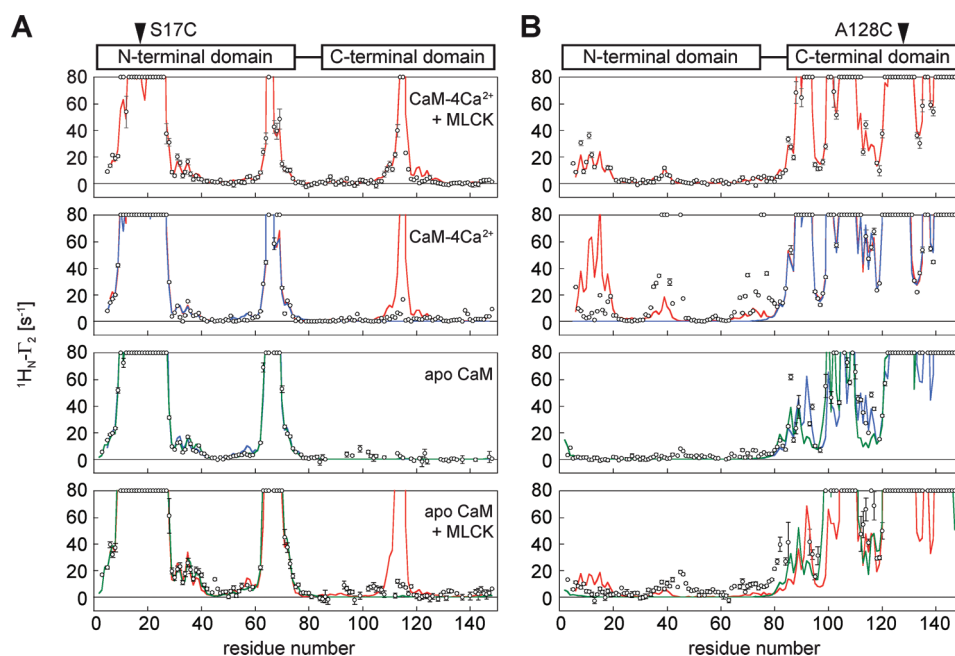
To optimize agreement with the interdomain PREs for all specified spin-labels simultaneously, ensembles of minor states were calculated using the ensembleSimulation module in Xplor-NIH.<sup>33</sup> Spin-label positions (represented by a five-member ensemble) were first set by fitting the intradomain PREs as described above. Conjoined rigid body/torsion angle simulated annealing<sup>34</sup> was then performed using an ensemble of minor states (consisting of 0–16 members and making up 0–40% of the total population): the two domains of CaM were treated as rigid bodies, and only the interdomain linker region (residues 77–81) was given torsional degrees of freedom. The target function was the same as that used to fit the intradomain PREs. The radial and angular order parameters for the SBF representation (that are calculated on the fly from the coordinates during simulated annealing) were computed separately for each ensemble member rather than averaged over the entire ensemble. This is because the time scale of motion of the spin-label (represented by a five-member configuration for every member of the minor state ensemble) is much faster (picosecond regime) and therefore decoupled from the slower interdomain reorientations (nanosecond regime).

Each simulated annealing calculation was run 100 times, and the results were represented as reweighted atomic probability density maps.<sup>35</sup> Images of the probability maps and protein structures were generated using PyMOL (<http://www.pymol.org>).

## RESULTS AND DISCUSSION

**CaM–Peptide Interactions, Spin Labeling, and PRE Measurements.** Large backbone <sup>1</sup>H<sub>N</sub>/<sup>15</sup>N chemical shift perturbations throughout CaM-4Ca<sup>2+</sup> are observed upon binding the MLCK peptide, particularly in the central linker (Figure 1D). Interestingly, we also detected binding of the MLCK peptide to





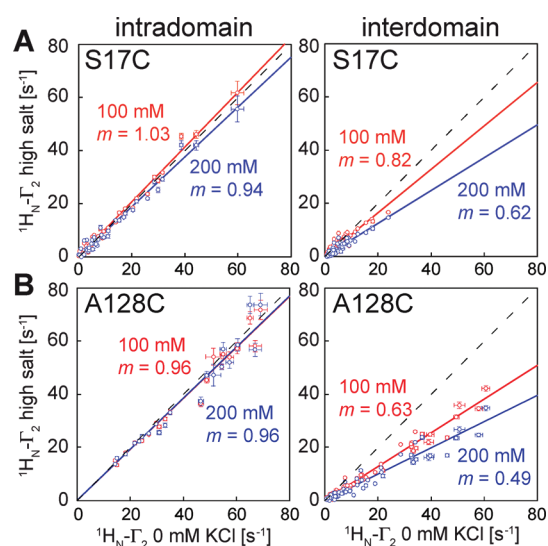
**Figure 2.** PRE profiles. (A and B) PRE profiles for CaM with a spin-label at (A) S17C or (B) A128C. Experimental PRE profiles for apo or  $\text{Ca}^{2+}$ -loaded CaM, in the presence or absence of MLCK peptide, are shown as circles (error bar, 1 SD). PREs too large ( $>80 \text{ s}^{-1}$ , roughly) to be accurately measured are plotted at the top of the charts. PRE profiles back-calculated from the structures of  $\text{CaM-4Ca}^{2+}$ -MLCK (red),  $\text{CaM-4Ca}^{2+}$  (blue), and apo CaM (green) are shown as solid lines. Only for the  $\text{CaM-4Ca}^{2+}$ -MLCK complex are the experimental interdomain PREs correctly predicted by the corresponding known structure. Note that slightly different PRE profiles can be back-calculated from the same structure depending on which experimental intradomain PRE data set is used to fit the spin-label position (and, in the case of  $\text{CaM-4Ca}^{2+}$ -MLCK, whether the peptide is present—as in the top panel—which slightly restricts the conformational space available to the spin-label). In the case of A128C, the apo NMR structure predicts some small PREs at the N-terminus; these are likely artifacts of the specific conformation adopted by the deposited average regularized structure (PDB 1CFD) since no interdomain NOEs were observed in that study.<sup>2</sup>

apo CaM, albeit 2 orders of magnitude weaker than in the presence of  $\text{Ca}^{2+}$  ( $K_D \sim 1 \mu\text{M}$  versus 10 nM; Supporting Information Figure S1). Exchange between free and peptide-bound states is fast on the chemical shift time scale for apo CaM but slow for  $\text{CaM-4Ca}^{2+}$  (Supporting Information Figure S2). The peptide-induced chemical shift perturbations in apo CaM are much smaller in magnitude than those observed for  $\text{CaM-4Ca}^{2+}$ , and are restricted primarily to the C-terminal domain (Figure 1D). That such an interaction is observed at all suggests that  $\text{Ca}^{2+}$ /CaM activation of MLCKs may be more complicated than the conventional mechanism illustrated in Figure 1A–C. Thus, to explore this mechanism, and the role of interdomain dynamics, PREs were measured for  $\text{Ca}^{2+}$ -loaded and  $\text{Ca}^{2+}$ -free CaM, in the presence and absence of the MLCK CBD peptide. Nitroxide spin-labels were conjugated via S–S linkages to engineered surface cysteines strategically placed at four positions: T5C, S17C, T117C, A128C (Figure 1A–C, and Supporting Information Figure S3B).  $\text{CaM-4Ca}^{2+}$  displays wild-type peptide binding activity for each spin-label (Supporting Information Figure S1). No evidence was observed for structural or dynamic effects induced by the spin-labels at S17C or A128C; however, labels at positions T5C and T117C result in local structural heterogeneity (Supporting Information Figure S3), likely due to a small degree of fraying at the N-terminal ends of helices I and VII, respectively, owing to disruption of an N-cap side chain-backbone hydrogen bond between the threonine hydroxyl oxygen atom and the backbone amide of the  $i + 3$  residue.<sup>3</sup> Although these effects are minor and can be accounted for (Supporting Information Figures S3 and S4), they complicate PRE analysis.

Hence, this work relies primarily on the S17C and A128C data, while the T5C and T117C data are used only for independent validation purposes (see below).

Interdomain PREs are observed for  $\text{CaM-4Ca}^{2+}$  and apo CaM (with or without bound peptide), which differ both in terms of their profiles and magnitude, with the latter (Figure 2, panels on lower two rows) being significantly smaller than the former (Figure 2, panels on second row). Thus, one can immediately conclude that sparsely populated states of  $\text{CaM-4Ca}^{2+}$  and apo CaM, characterized by close interdomain contacts and undetected by NOEs, are sampled in solution and that the distribution of these states is modulated by calcium. In addition, PRE experiments were also performed on  $\text{CaM-4Ca}^{2+}$  at various ionic strengths, varying  $[\text{KCl}]$  from 0 to 200 mM. While ionic strength has no significant effect on the intradomain PREs (Figure 3, left panels), the interdomain PREs decrease significantly as the ionic strength is increased (Figure 3, right panels). Thus, the interdomain contacts responsible for the interdomain PREs in  $\text{CaM-4Ca}^{2+}$  appear to be mediated by interactions that comprise a significant electrostatic component.

To discount and/or cancel out any intermolecular effects as the source of the interdomain PREs, control experiments were performed on samples comprising a mixture of  $^{15}\text{N}/^{13}\text{C}/^2\text{H}$ -labeled CaM and paramagnetically labeled CaM at natural isotopic abundance to exclusively detect intermolecular PREs occurring under these experimental conditions. No intermolecular PREs were detected for  $\text{CaM-4Ca}^{2+}$ -MLCK or apo CaM, and only very small intermolecular PREs were observed for apo CaM in the presence of MLCK (Supporting Information Figure S5).



**Figure 3.** Ionic strength dependence of CaM- $\text{Ca}^{2+}$  PREs. Experimental PREs (circles, error bars = 1 SD) were measured for CaM-4 $\text{Ca}^{2+}$  with the nitroxide spin label at (A) S17C and (B) A128C at various ionic strengths. Intradomain PREs are shown in the left panels and interdomain PREs in the right panels. The data are presented as scatter plots with the PRE value at 0 mM KCl shown on the x-axis and the PRE value at 100 mM (red) and 200 mM (blue) shown on the y-axis. In each case, the slope of a best-fit line through the origin is given. A dashed line corresponding to a slope of  $m = 1$  is also displayed for reference. The intradomain PREs do not depend on ionic strength, but increasing the salt concentration significantly decreases the magnitude of the interdomain PREs, indicating that the interdomain interactions explored here have a significant electrostatic component.

However, some more sizable intermolecular PREs were in fact detected for CaM-4 $\text{Ca}^{2+}$ , though still much smaller in magnitude than the intramolecular PREs. Thus, for maximal accuracy and relevance, intermolecular PRE effects were subtracted from the PREs measured for CaM-4 $\text{Ca}^{2+}$  elsewhere in this report (unless otherwise specified) to ensure that quantitative analysis was restricted to intramolecular PREs and not complicated by intermolecular effects. Interestingly, the intermolecular PREs that are observed appear to be a manifestation of the same intramolecular/interdomain interactions in CaM-4 $\text{Ca}^{2+}$  that are characterized below (Supporting Information Figure S5).

**Interdomain PREs Are Unaccounted for by Current CaM Structures.** Agreement between experimental PREs and those back-calculated from known structures was assessed by simulated annealing optimization of the spin-label coordinates using an ensemble representation to account for their inherent flexibility. There is excellent quantitative agreement for the intradomain PREs (Figure 2); the intradomain PRE Q-factors (Table 1; see eq 1 in Materials and Methods for definition) range from 0.1 to 0.2, with the exception of those for the C-terminal domain of apo CaM (with or without bound peptide) where agreement is a little worse ( $Q = 0.3\text{--}0.4$ ) owing to the presence of intermediate chemical exchange between two states within the C-terminal domain,<sup>2,36</sup> which causes line broadening and a concomitant decrease in the quality of the experimental PRE data. The fit of the intradomain PRE data for the C-terminal domain of apo CaM can be improved slightly by fitting to a mixture of apo and  $\text{Ca}^{2+}$  loaded conformations at an optimal ratio of  $\sim 9:1$  (Supporting Information Figure S6), in agreement with previous findings

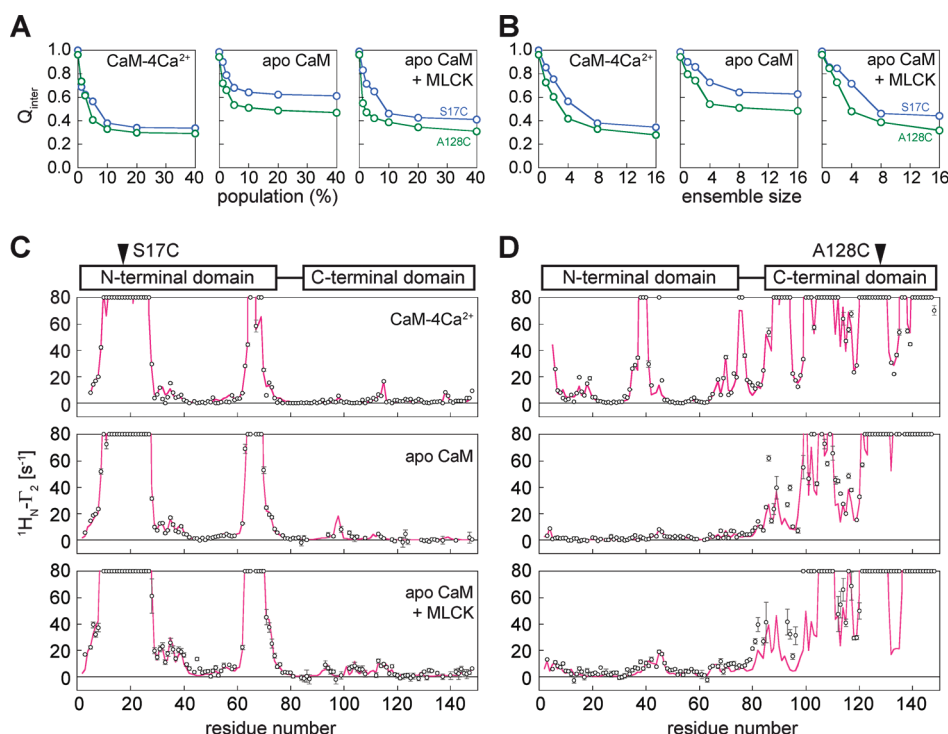
**Table 1.** Q-Factors for Fits of CaM PRE Data

	intradomain fit <sup>a</sup>			full molecule fit <sup>b</sup>		
	intra	inter	all	intra	inter	all
CaM-4 $\text{Ca}^{2+}$ -MLCK <sup>c</sup>						
S17C	0.16	0.41	0.22	0.17	0.21	0.18
A128C	0.13	0.34	0.17	0.13	0.25	0.15
CaM-4 $\text{Ca}^{2+}$ <sup>d</sup>						
S17C	0.17	0.99	0.29	0.17	0.99	0.29
A128C	0.13	0.99	0.49	0.14	0.99	0.49
Apo CaM <sup>e</sup>						
S17C	0.16	0.96	0.17	0.16	0.96	0.17
A128C	0.35	0.93	0.36	0.35	0.93	0.36
Apo CaM + MLCK <sup>e</sup>						
S17C	0.20	0.96	0.30	0.20	0.96	0.30
A128C	0.42	0.93	0.55	0.44	0.92	0.54

<sup>a</sup> Paramagnetic tag positions were fit to intradomain PREs only (residues 1–76 for the N-terminal domain; residues 83–148 for the C-terminal domain), and the Q-factors were then calculated for intradomain, interdomain, or all PREs. <sup>b</sup> Paramagnetic tag positions were fit to PREs for the full molecule, and Q-factors were then calculated for intradomain, interdomain, or all PREs. <sup>c</sup> Q-factors calculated for experimental PREs from CaM-4 $\text{Ca}^{2+}$ -MLCK, fit to crystal structure of CaM-4 $\text{Ca}^{2+}$ -MLCK complex (PDB 1CDL)<sup>4</sup> (Figure 1C). <sup>d</sup> Q-factors calculated for experimental PREs from CaM-4 $\text{Ca}^{2+}$ , fit to “dumbbell” crystal structure of CaM-4 $\text{Ca}^{2+}$  (PDB 1CLL)<sup>3</sup> (Figure 1B). <sup>e</sup> Q-factors calculated for experimental PREs from  $\text{Ca}^{2+}$ -free CaM (with or without peptide, as indicated), fit to NMR structure of apo CaM (PDB 1CFD)<sup>2</sup> (Figure 1A). Note that while the structure for the entire molecule of apo CaM was deposited in the PDB and is used in the calculations presented in this table, there were, in fact, no interdomain NOE restraints (or any other NMR restraints) to orient the two domains relative to one another.<sup>2</sup>

based on  $^{15}\text{N}$  relaxation data.<sup>36</sup> The full PRE profiles for CaM-4 $\text{Ca}^{2+}$ -MLCK fit extremely well to the known X-ray structure<sup>4</sup> of the complex (Figure 2A,B, top panels) with Q-factors of  $\sim 0.2$  (Table 1). For the other three states of CaM, however, agreement with the interdomain PRE data is extremely poor (Figure 2A,B, lower three panels) with Q-factors of  $\sim 1$  (Table 1).

The dumbbell X-ray structure of CaM-4 $\text{Ca}^{2+}$ <sup>3,9,10</sup> does not predict the existence of the interdomain PREs observed experimentally for CaM-4 $\text{Ca}^{2+}$  or apo CaM, and indeed, a very wide range of interdomain orientations, spanning a cone of semiangle  $\sim 90^\circ$ , is fully consistent with interdomain PRE values  $\leq 2\text{ s}^{-1}$  (see below). The interdomain PRE profiles back-calculated from the structure of CaM-4 $\text{Ca}^{2+}$ -MLCK<sup>4</sup> exhibit some qualitative similarities with the experimental PRE profiles for CaM-4 $\text{Ca}^{2+}$  with peaks in similar locations but yield poor quantitative fits to the experimental data (Figure 2A,B, second row panels): in particular, the back-calculated interdomain PREs for S17C are much larger, and there are regions with both smaller (residues 34–46 and 69–80) and larger (residues 7–20) back-calculated PREs for A128C. Similar results are obtained using the crystal structure of a compact state of CaM-4 $\text{Ca}^{2+}$ <sup>11</sup> (Supporting Information Figure S4). The MD ensemble for CaM-4 $\text{Ca}^{2+}$ <sup>19</sup> predicts some but not all features of the experimental PRE profiles (Supporting Information Figure S8): large intermolecular PREs



**Figure 4.** PRE-driven ensemble simulated annealing calculations characterizing the minor closed states of CaM with interdomain contacts. (A and B) Dependence of the interdomain PRE Q-factor for S17C (blue) and A128C (green) as a function of (A) minor state population (using an 8-member ensemble) and (B) ensemble size (for a minor state population of 10%) for CaM-4Ca<sup>2+</sup>, apo CaM, and apo CaM + MLCK peptide. (C and D) The resulting fits to the PRE profiles for (C) S17C and (D) A128C for a minor state population of 10% represented by an 8-member ensemble (calculated, magenta lines; experimental, black circles; error bars, 1 SD).

are predicted for residues 47–59 from A128C where none are observed, while other experimental intermolecular PREs such as those for residues 15–19 from A128C are not predicted (Supporting Information Figure S7).

The interdomain PREs observed for apo CaM (which are qualitatively similar in the presence and absence of bound peptide) are not predicted by any of the existing structures (Figure 2A,B, lower two panels). Moreover, the peak locations in the experimental PRE profiles of apo CaM do not coincide with those back-calculated from the structure of the CaM-4Ca<sup>2+</sup>-MLCK complex.

On the basis of the above observations, one can infer, without any further analysis, that a significant fraction of the sparsely populated compact states of CaM-4Ca<sup>2+</sup> samples a region of conformational space that likely overlaps with that occupied by the CaM-4Ca<sup>2+</sup>-MLCK complex; in contrast, the major fraction of sparsely populated compact states of apo CaM (both free and bound to peptide) does not significantly overlap with the CaM-4Ca<sup>2+</sup>-MLCK complex. Further, while the MD simulation ensemble<sup>19</sup> samples compact states of CaM-4Ca<sup>2+</sup>, the distribution and conformational space occupied by these states are not the same as those present in solution.

**Visualization of Sparsely Populated States by PRE-Driven Ensemble Simulated Annealing Calculations.** To obtain a quantitative description of the sparsely populated states sampled by CaM-4Ca<sup>2+</sup> and apo CaM, we carried out a series of PRE-driven simulated annealing calculations<sup>29</sup> in which the minor compact state was represented by an ensemble of varying size and population.<sup>37</sup> The two CaM domains were treated as rigid bodies, and only the linker residues (77–81) were given

torsional degrees of freedom. The major species was represented by the extended dumbbell conformation, but identical results were obtained using an ensemble generated from calculations in which the interdomain PREs were constrained to values  $\leq 2 \text{ s}^{-1}$ . From the dependence of the interdomain PRE Q-factor on ensemble size and population of the compact sparsely populated species, it is observed that the interdomain Q-factor decreases rapidly up to a population of  $\sim 10\%$  for CaM-4Ca<sup>2+</sup> and apo CaM bound to peptide (Figure 4A, left and right panels), and  $\sim 5\%$  for apo CaM without bound peptide (Figure 4A, middle panel), and up to an ensemble size of 8 (Figure 4B), leveling off thereafter. The apparently lower population of compact states for apo CaM is consistent with the NMR finding that the central linker region of apo CaM appears to adopt a helical conformation about 30% of the time,<sup>2</sup> whereas in CaM-4Ca<sup>2+</sup>, there is no NMR evidence for such transient helix formation.<sup>18</sup> A comparison of calculated and observed intermolecular PREs yield good agreement in all cases (Figures 4C,D; Table 2). Note that the higher Q-factors for the interdomain PREs of apo CaM simply reflect the low magnitude of the PREs, but the rms difference between observed and calculated values is small ( $1.5\text{--}1.6 \text{ s}^{-1}$ , see Table 2).

Reweight atomic probability maps,<sup>35</sup> generated from the calculated ensembles and describing the spatial distribution of the compact minor states of CaM, are shown in Figure 5. For reference, a superposition of the dumbbell extended structure of CaM-4Ca<sup>2+</sup><sup>3</sup> and the CaM-4Ca<sup>2+</sup>-MLCK complex<sup>4</sup> is displayed in Figure 5A, along with a probability map generated from a superposition of a large number of other peptide-bound CaM crystal structures.



**Table 2. Goodness of Fit for Interdomain PRE Data with and without Minor States Generated by Simulated Annealing Calculations**

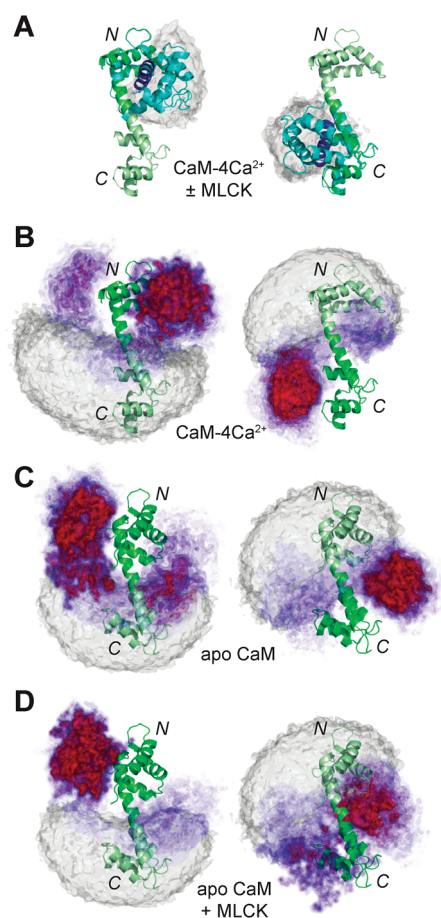
	Q-factor		rms difference ( $s^{-1}$ )	
	major alone <sup>a</sup>	major + minor <sup>b</sup>	major alone <sup>a</sup>	major + minor <sup>b</sup>
CaM-4Ca <sup>2+</sup>				
S17C	0.99	0.38	3.1	1.2
A128C	0.96	0.33	12.7	4.3
Apo CaM				
S17C	0.99	0.64	2.4	1.6
A128C	0.94	0.51	2.8	1.5
Apo CaM + MLCK				
S17C	0.99	0.46	4.7	2.2
A128C	0.96	0.39	7.9	3.2

<sup>a</sup> Paramagnetic tag positions were first fit to the intradomain PRE data (residues 1–76 for N-terminal domain, 83–148 for C-terminal domain), and then the PRE Q-factors and rms differences between observed and calculated PREs were determined for the interdomain PREs using the extended dumbbell crystal structure (1CLL)<sup>3</sup> for CaM-4Ca<sup>2+</sup>. Note that a very wide-range of interdomain orientations spanning a cone of semiangle  $\sim 90^\circ$  (cf. gray atomic probability density maps in Figure 5B–D and in Supporting Information Figure S8A) is fully consistent with interdomain PRE values  $\leq 2 s^{-1}$  and yields the same results as the extended dumbbell crystal structure. Thus, the major state comprises all states with no interdomain contacts. <sup>b</sup> Paramagnetic spin-label positions were first fit to intradomain PREs. PRE-driven ensemble simulated annealing calculations were then performed with an 8-member minor state ensemble encompassing 10% of the total population giving the linker region (residues 77–81) torsional degrees of freedom. Interdomain Q-factors and rms differences between observed and calculated PREs were then determined for the entire ensemble.

We first consider the conformational space compatible with the absence of intermolecular PREs (i.e., PRE rates  $\leq 2 s^{-1}$ ) and represented by the gray probability density maps in Figure 5B–D. These probability maps provide only a rough guide for the major species (population  $\sim 90\%$ ) for which there are no close interdomain contacts, since the calculated ensemble is defined by the absence of PREs rather than by any direct experimental measurements. In each case, the conformation of one domain relative to the other occupies roughly one full hemisphere (i.e., a cone of semiangle  $\sim 90^\circ$ ). This ensemble bears a very strong resemblance in terms of size and orientation to that recently calculated on the basis of PCS, RDC, and SAXS data.<sup>17,38–40</sup> Indeed, even the tilt of the hemisphere of the major species ensemble from the dumbbell structure is very similar in the two cases ( $\sim 45^\circ$  versus  $\sim 30^\circ$ <sup>17</sup>).

The atomic probability maps describing the conformational space sampled by the compact minor states (population  $\leq 10\%$ ) are plotted at multiple contour levels ranging from 0.1 (transparent blue) to 0.5 (opaque red) of maximum (Figure 5B–D). While all compact state ensembles encompass a large range of conformational space, the distribution for CaM-4Ca<sup>2+</sup> is clearly distinct from that of apo CaM (free or peptide-bound).

For CaM-4Ca<sup>2+</sup>, about half of the compact minor species ensemble samples a region of space (Figure 5B) that clearly overlaps with that occupied by the CaM-4Ca<sup>2+</sup>-peptide complexes



**Figure 5.** Visualizing the minor closed-state ensembles of CaM-4Ca<sup>2+</sup> and apo CaM. (A) CaM-4Ca<sup>2+</sup>-MLCK complex<sup>4</sup> (with CaM in cyan and MLCK peptide in blue) overlaid on the CaM-4Ca<sup>2+</sup> dumbbell structure<sup>3</sup> (green), best-fitted to either the N-terminal (left panel) or C-terminal (right panel) domains. Twenty-six additional peptide-bound structures were overlaid in the same manner, and these are represented by the atomic probability density maps<sup>35</sup> shown in gray for the C-terminal domain (left) and N-terminal domain (right). These structures (PDB codes 1CDL, 1CDM, 1MXE, 1NIW, 1QS7, 1QTX, 1VRK, 1YR5, 1ZUZ, 2BE6, 2F3Y, 2F3Z, 2FOT, 2HQW, 2OSG, 2O60, 2VAY, 3BXL, 3BYA, 3DVE, 3DVJ, 3DVK, 3DVM, 3EWT, 3EWW, 3GP2) represent a wide array of different 1:1 canonical CaM-peptide complexes, but it can be seen that the conformational space occupied by the two domains is very similar to that in the CaM-4Ca<sup>2+</sup>-MLCK complex. (B–D) Atomic probability density maps showing the conformational space sampled by the minor species ensemble for (B) CaM-4Ca<sup>2+</sup>, (C) apo CaM, and (D) the apo CaM-MLCK complex. The minor state atomic probability maps are derived from 100 independent PRE-driven simulated annealing calculations using an 8-member ensemble (i.e., 800 total structures) at a population of 10%, and plotted at multiple contour levels ranging from 0.1 (transparent blue) to 0.5 (opaque red). The gray probability density maps, plotted at a single contour level of 0.1 of maximum, show the conformational space consistent with interdomain PRE values  $\leq 2 s^{-1}$  and represent the major state ensemble characterized by no interdomain contacts and an occupancy of  $\sim 90\%$  (see also Supporting Information Figure S9A). In the left-hand panels, all ensemble members are best-fitted to the N-terminal domain (dark green) and the probability density maps are shown for the C-terminal domain. In the right-hand panels, all ensembles are best-fitted to the C-terminal domain (dark green) and the probability density maps are shown for the N-terminal domain. All panels are displayed in the same orientation, and the extended dumbbell structure<sup>3</sup> is displayed as a ribbon diagram for reference (with the structures of the individual domains in panels C and D replaced by those for apo CaM).

(Figure 5A). We note that the atomic probability maps obtained with the S17C and A128C data are very similar to those obtained with the T5C and T117C data, or with all four PRE data sets together (Supporting Information Figure S8B–D). This consistency is also seen for the interdomain PRE profiles back-calculated from different ensembles obtained with various combinations of PRE data sets (Supporting Information Figure S9), providing independent validation of the results. Moreover, even calculations carried out using any of the PRE data sets individually yield similar results (Supporting Information Figure S8E–H). The PRE data for CaM-4Ca<sup>2+</sup> provide direct and unambiguous evidence not only that the two domains of CaM-4Ca<sup>2+</sup> can come into close contact in solution for a small fraction of the time (~10%), but also that these compact states preferentially adopt peptide-bound-like conformations about half of the time (i.e., the population of peptide-bound-like conformations is ~5%).

Thus the PRE data presented here provide distinct, but fully complementary, information to that afforded previously by analysis of RDC, PCS, and SAXS data.<sup>17,38–40</sup> The latter provide direct experimental data to characterize the major species ensemble that has no interdomain contacts, but only allow an approximate estimate of the percentage (~5%) of a compact-like species that would not impact the RDC, PCS, and SAXS data.<sup>38</sup> In this regard, it should be stressed that the contribution of close-contact minor species to the PRE is hugely enhanced relative to that of PCS owing to the  $\langle r^{-6} \rangle$  versus  $\langle r^{-3} \rangle$  dependence on the paramagnetic center-nucleus distance, and the absence of an orientational dependence of the position of the nuclear spin with respect to the principal axes of the magnetic susceptibility tensor.<sup>21</sup>

The atomic probability maps calculated for the sparsely populated compact states of apo CaM in the absence (Figure 5C) and presence (Figure 5D) of peptide are similar to one another. In contrast to CaM-4Ca<sup>2+</sup>, the minor state ensemble does not cluster around the CaM-4Ca<sup>2+</sup>-peptide conformations, but preferentially samples structures on the opposing interface. (Note that the accuracy of the calculated apo CaM ensembles is not as high as that for CaM-4Ca<sup>2+</sup> as the magnitude of the interdomain PREs is smaller and the quality of the data lower due to intermediate exchange in the C-terminal domain.) Thus, apo CaM samples a range of compact (5–10%) and extended (90–95%) structures, and even in the presence of peptide does not clamp down into a single compact structure as in the case of CaM-4Ca<sup>2+</sup>-MLCK.

## CONCLUDING REMARKS

The present findings imply that Ca<sup>2+</sup> activates CaM not only through local structural changes within each domain,<sup>2,5</sup> but also through more global effects, with Ca<sup>2+</sup> allowing CaM to clamp down upon its peptide target in a manner not possible in the apo state. This does not necessarily imply a purely conformational selection mechanism of binding, but rather that CaM-4Ca<sup>2+</sup> has the predilection to form the relevant compact structures (unlike the apo state), priming it to bind peptide when present. This is fully consistent with the model proposed by Kranz et al.,<sup>7</sup> based on amide exchange rates, in which the initial interaction between a random coil peptide and CaM-4Ca<sup>2+</sup> results in a loosely packed initial complex, followed by coil-to-helix transitions within the peptide and concomitant formation of specific intermolecular ion-pair and hydrophobic interactions to generate the final compact complex, a process akin to coupled folding and binding<sup>41</sup> and induced fit. CaM-4Ca<sup>2+</sup> does not necessarily sample the

CaM-4Ca<sup>2+</sup>-MLCK structure directly (which would not be productive, since the binding surface would be largely inaccessible), but samples a range of nearby conformations, which are also near the conformations of other canonical CaM complexes (Figure 5A), thus, speaking to the universality of this process.

Earlier work has found that binding entropy varies considerably between different CaM-4Ca<sup>2+</sup>-target interactions: the binding entropy is strongly correlated with differences in changes in CaM conformational entropy upon binding, which arise from highly localized subnanosecond side chain motions.<sup>42,43</sup> In contrast, here we examine large-scale rigid-body domain motions in CaM; given the conserved overall conformation of CaM-4Ca<sup>2+</sup>-target complexes (Figure 5A), the decrease in conformational entropy due to the loss of these motions upon binding would not be expected to vary considerably between interactions with different targets.

The salt dependence of the interdomain PREs in CaM-4Ca<sup>2+</sup> (Figure 3) indicates that the formation of these compact structures is at least partially driven by electrostatic forces. The electrostatic surface potential of apo CaM is dominated by large areas of negative charge (Supporting Information Figure S10A–D). CaM-4Ca<sup>2+</sup> is similarly negatively charged; however, additional positively charged patches are present on its surface as a result of the charge of the four Ca<sup>2+</sup> ions and the structural rearrangements they induce (Supporting Information Figure S10E–H). One of these positively charged patches partially overlaps with the interdomain interaction site (Figure 5B) on the N-terminal domain, while the interaction site on the C-terminal domain remains highly negatively charged. Such interactions would be much less favorable in apo CaM due to electrostatic repulsion. The role of electrostatic contacts in protein–protein encounter complex formation has also been observed in other systems.<sup>37,44</sup>

The PRE data also imply an alternative pathway of CaM action; instead of the scheme illustrated in Figure 1A–C where CaM remains unbound until activated by Ca<sup>2+</sup>, CaM could remain associated with its kinase targets (via its C-terminal domain) even in the resting state, and then only activate the kinase when Ca<sup>2+</sup> loading causes it to clamp down upon the CBD, thus, relieving kinase self-inhibition. CaM is known to bind to some of its noncanonical targets in a Ca<sup>2+</sup>-independent manner,<sup>1</sup> and weak interactions between apo CaM and MLCK peptides have been reported.<sup>45,46</sup> Although the CBD of the kinase remains self-associated in the resting state,<sup>47</sup> the portion that interacts with the C-terminal domain of CaM is partially exposed, potentially allowing the CaM C-terminal domain to bind without causing activation (Supporting Information Figure S11). Such a model would make activation dependent on Ca<sup>2+</sup> loading in the N-terminal domain, which is reasonable, as the C-terminal domain is likely to be partially Ca<sup>2+</sup>-loaded even at the resting intracellular Ca<sup>2+</sup> concentration of ~100 nM due to its much higher Ca<sup>2+</sup> affinity.<sup>48</sup> Whether the C-terminal domain of apo CaM remains associated with its peptide targets or just interacts with them transiently, this is compatible with the idea that target binding and activation by CaM is initiated by the C-terminal domain,<sup>49</sup> with the N-terminal domain then acting as the primary Ca<sup>2+</sup> sensor.<sup>50</sup>

The current PRE data on CaM, in conjunction with previously published PRE data that revealed the existence of an open-to-partially closed transition for apo maltose binding protein,<sup>51</sup> also suggest an alternative perspective on conformational selection and induced fit for proteins where ligand binding results in



interdomain reorientations that effectively trap the ligand and bury it from solvent. In such systems, it is evident that the holo conformation cannot itself bind ligand directly, since the ligand-binding site is not accessible. However, the existence of sparsely populated states that are partially closed yet still accessible to ligand, and that occupy a region of conformational space in reasonably close proximity to that of the holo state, may facilitate the transition to the closed ligand-bound complex, highlighting the complementarity and interplay of conformational selection and induced fit.

## ■ ASSOCIATED CONTENT

**S Supporting Information.** Binding data and additional PRE data. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

mariusc@mail.nih.gov

### Present Addresses

<sup>†</sup>Cell, 600 Technology Square, Cambridge, MA 02139, U.S.A.

## ■ ACKNOWLEDGMENT

We thank Ad Bax, Dusty Baber, Dan Garrett, Charles Schwieters, Nico Tjandra, and Jinfa Ying for useful discussions, and Max Valenstein for technical support. This work was supported by the Intramural Program of the NIH, NIDDK, and the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the NIH (to G.M.C.)

## ■ REFERENCES

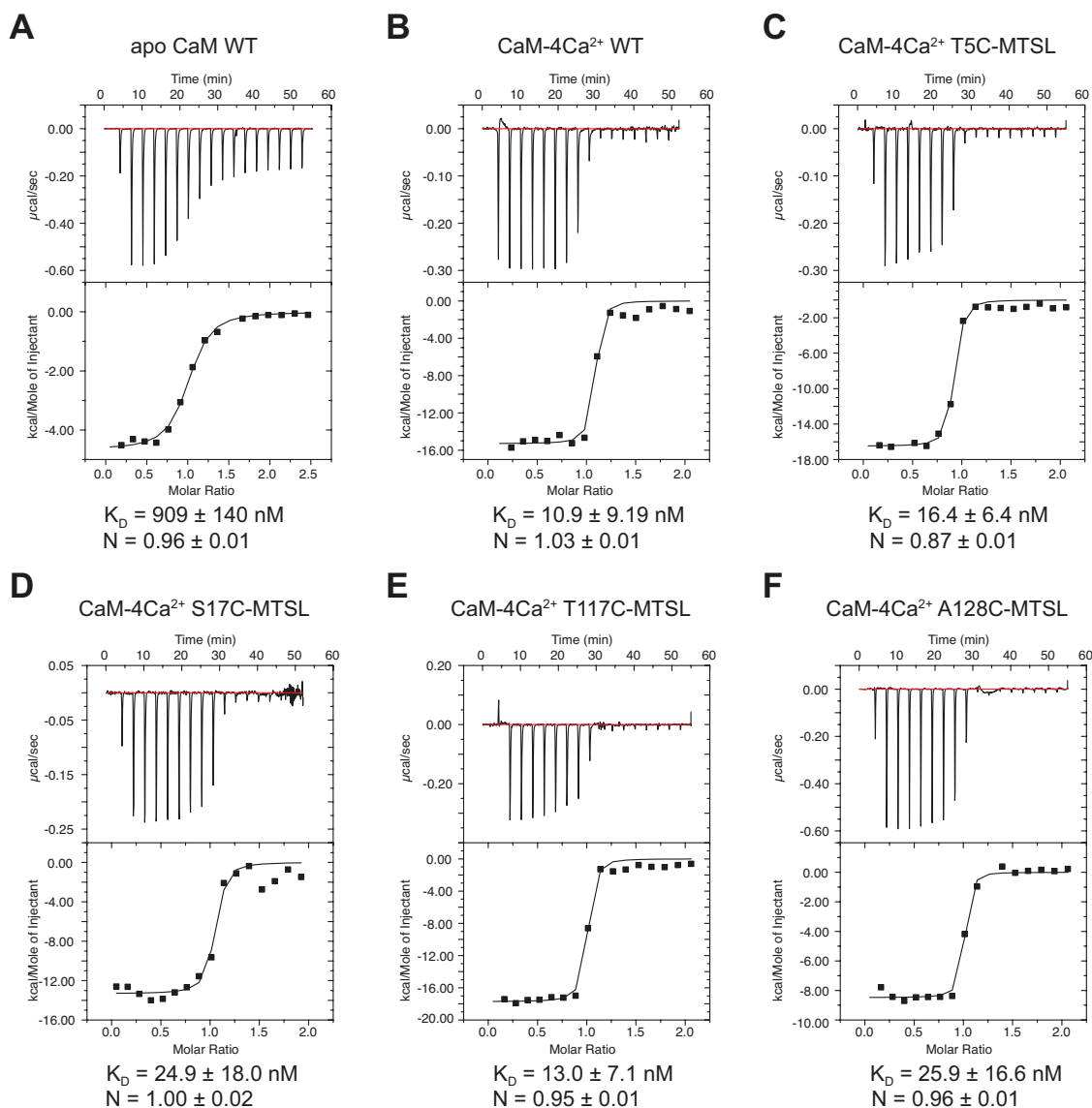
- (1) Yamniuk, A. P.; Vogel, H. J. *Mol. Biotechnol.* **2004**, *27*, 33–57.
- (2) Kuboniwa, H.; Tjandra, N.; Grzesiek, S.; Ren, H.; Klee, C. B.; Bax, A. *Nat. Struct. Biol.* **1995**, *2*, 768–776.
- (3) Chattopadhyaya, R.; Meador, W. E.; Means, A. R.; Quijcho, F. A. *J. Mol. Biol.* **1992**, *228*, 1177–1192.
- (4) Meador, W. E.; Means, A. R.; Quijcho, F. A. *Science* **1992**, *257*, 1251–1255.
- (5) Zhang, M.; Tanaka, T.; Ikura, M. *Nat. Struct. Biol.* **1995**, *2*, 758–767.
- (6) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax, A. *Science* **1992**, *256*, 632–638.
- (7) Kranz, J. K.; Flynn, P. F.; Fuentes, E. J.; Wand, A. J. *Biochemistry* **2002**, *41*, 2599–2608.
- (8) Hoefflich, K. P.; Ikura, M. *Cell* **2002**, *108*, 739–42.
- (9) Babu, Y. S.; Sack, J. S.; Greenhough, T. J.; Bugg, C. E.; Means, A. R.; Cook, W. J. *Nature* **1985**, *315*, 37–40.
- (10) Wilson, M. A.; Brunger, A. T. *J. Mol. Biol.* **2000**, *301*, 1237–1256.
- (11) Fallon, J. L.; Quijcho, F. A. *Structure* **2003**, *11*, 1303–1307.
- (12) Slaughter, B. D.; Bieber-Urbauer, R. J.; Johnson, C. K. *J. Phys. Chem. B* **2005**, *109*, 12658–12662.
- (13) Trewthella, J. *Cell Calcium* **1992**, *13*, 377–90.
- (14) Tjandra, N.; Kuboniwa, H.; Ren, H.; Bax, A. *Eur. J. Biochem.* **1995**, *230*, 1014–1024.
- (15) Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A. *Biochemistry* **1992**, *31*, 5269–5278.
- (16) Baber, J. L.; Szabo, A.; Tjandra, N. *J. Am. Chem. Soc.* **2001**, *123*, 3953–3959.
- (17) Bertini, I.; Del Bianco, C.; Gelis, I.; Katsaros, N.; Luchinat, C.; Parigi, G.; Peana, M.; Provenzani, A.; Zoroddu, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6841–6846.
- (18) Chou, J. J.; Li, S.; Klee, C. B.; Bax, A. *Nat. Struct. Biol.* **2001**, *8*, 990–997.
- (19) Gsponer, J.; Christodoulou, J.; Cavalli, A.; Bui, J. M.; Richter, B.; Dobson, C. M.; Vendruscolo, M. *Structure* **2008**, *16*, 736–746.
- (20) Iwahara, J.; Clore, G. M. *Nature* **2006**, *440*, 1227–1230.
- (21) Clore, G. M.; Iwahara, J. *Chem. Rev.* **2009**, *109*, 4108–4139.
- (22) Clore, G. M.; Gronenborn, A. M. *Methods Enzymol.* **1994**, *239*, 349–363.
- (23) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277–293.
- (24) Hu, K.; Doucleff, M.; Clore, G. M. *J. Magn. Reson.* **2009**, *200*, 173–177.
- (25) Iwahara, J.; Tang, C.; Marius Clore, G. *J. Magn. Reson.* **2007**, *184*, 185–195.
- (26) Mulder, F. A.; Schipper, D.; Bott, R.; Boelens, R. *J. Mol. Biol.* **1999**, *292*, 111–123.
- (27) Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; Nakatani, E.; Schulte, C. F.; Tolmie, D. E.; Kent Wenger, R.; Yao, H.; Markley, J. L. *Nucleic Acids Res.* **2008**, *36*, D402–D408.
- (28) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J. Magn. Reson.* **2003**, *160*, 65–73.
- (29) Schwieters, C. D.; Kuszewski, J. J.; Clore, G. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 47–62.
- (30) Iwahara, J.; Schwieters, C. D.; Clore, G. M. *J. Am. Chem. Soc.* **2004**, *126*, 5879–5896.
- (31) Clore, G. M.; Kuszewski, J. *J. Am. Chem. Soc.* **2002**, *124*, 2866–2867.
- (32) Chang, S. L.; Szabo, A.; Tjandra, N. *J. Am. Chem. Soc.* **2003**, *125*, 11379–11384.
- (33) Clore, G. M.; Schwieters, C. D. *J. Am. Chem. Soc.* **2004**, *126*, 2923–2938.
- (34) Schwieters, C. D.; Clore, G. M. *J. Magn. Reson.* **2001**, *152*, 288–302.
- (35) Schwieters, C. D.; Clore, G. M. *J. Biomol. NMR* **2002**, *23*, 221–225.
- (36) Malmendal, A.; Evenas, J.; Forsen, S.; Akke, M. *J. Mol. Biol.* **1999**, *293*, 883–899.
- (37) Tang, C.; Iwahara, J.; Clore, G. M. *Nature* **2006**, *444*, 383–386.
- (38) Bertini, I.; Giachetti, A.; Luchinat, C.; Parigi, G.; Petoukhov, M. V.; Pierattelli, R.; Ravera, E.; Svergun, D. I. *J. Am. Chem. Soc.* **2010**, *132*, 13553–13558.
- (39) Bertini, I.; Gupta, Y. K.; Luchinat, C.; Parigi, G.; Peana, M.; Sgheri, L.; Yuan, J. *J. Am. Chem. Soc.* **2007**, *129*, 12786–94.
- (40) Dasgupta, S.; Hu, X.; Keizers, P. H.; Liu, W. M.; Luchinat, C.; Nagulapalli, M.; Overhand, M.; Parigi, G.; Sgheri, L.; Ubbink, M. *J. Biomol. NMR* **2011**, *51*, 253–263.
- (41) Sugase, K.; Dyson, H. J.; Wright, P. E. *Nature* **2007**, *447*, 1021–5.
- (42) Marlow, M. S.; Dogan, J.; Frederick, K. K.; Valentine, K. G.; Wand, A. J. *Nat. Chem. Biol.* **2010**, *6*, 352–8.
- (43) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. *Nature* **2007**, *448*, 325–9.
- (44) (a) Volkov, A. N.; Worrall, J. A.; Holtzmann, E.; Ubbink, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18945–50. (b) Suh, J. -Y.; Tang, C.; Clore, G. M. *J. Am. Chem. Soc.* **2007**, *129*, 12954–12955.
- (45) Kleit, R. E.; Blumenthal, D. K.; Wemmer, D. E.; Krebs, E. G. *Biochemistry* **1985**, *24*, 8152–7.
- (46) Tsvetkov, P. O.; Protasevich, I. I.; Gilli, R.; Lafitte, D.; Lobachov, V. M.; Haiech, J.; Briand, C.; Makarov, A. A. *J. Biol. Chem.* **1999**, *274*, 18161–4.
- (47) Goldberg, J.; Nairn, A. C.; Kuriyan, J. *Cell* **1996**, *84*, 875–887.
- (48) Gilli, R.; Lafitte, D.; Lopez, C.; Kilhoffer, M.; Makarov, A.; Briand, C.; Haiech, J. *Biochemistry* **1998**, *37*, 5450–5456.
- (49) Bayley, P. M.; Findlay, W. A.; Martin, S. R. *Protein Sci.* **1996**, *5*, 1215–28.
- (50) Rodriguez-Castaneda, F.; Maestre-Martinez, M.; Coudeville, N.; Dimova, K.; Junge, H.; Lipstein, N.; Lee, D.; Becker, S.; Brose, N.; Jahn, O.; Carlomagno, T.; Griesinger, C. *EMBO J.* **2010**, *29*, 680–91.
- (51) Tang, C.; Schwieters, C. D.; Clore, G. M. *Nature* **2007**, *449*, 1078–1082.

**SUPPLEMENTARY INFORMATION**

**Transient, sparsely-populated compact states of apo and calcium-loaded calmodulin  
probed by paramagnetic relaxation enhancement: interplay of conformational selection  
and induced fit**

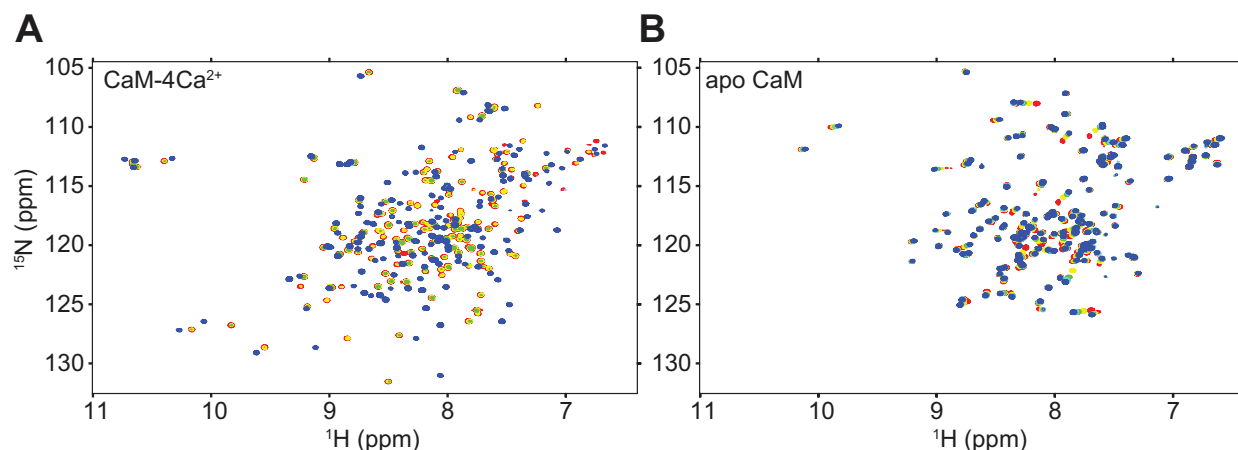
Nicholas J. Anthis, Michaelleen Doucleff<sup>1</sup>, G. Marius Clore\*

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney  
Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, U.S.A.

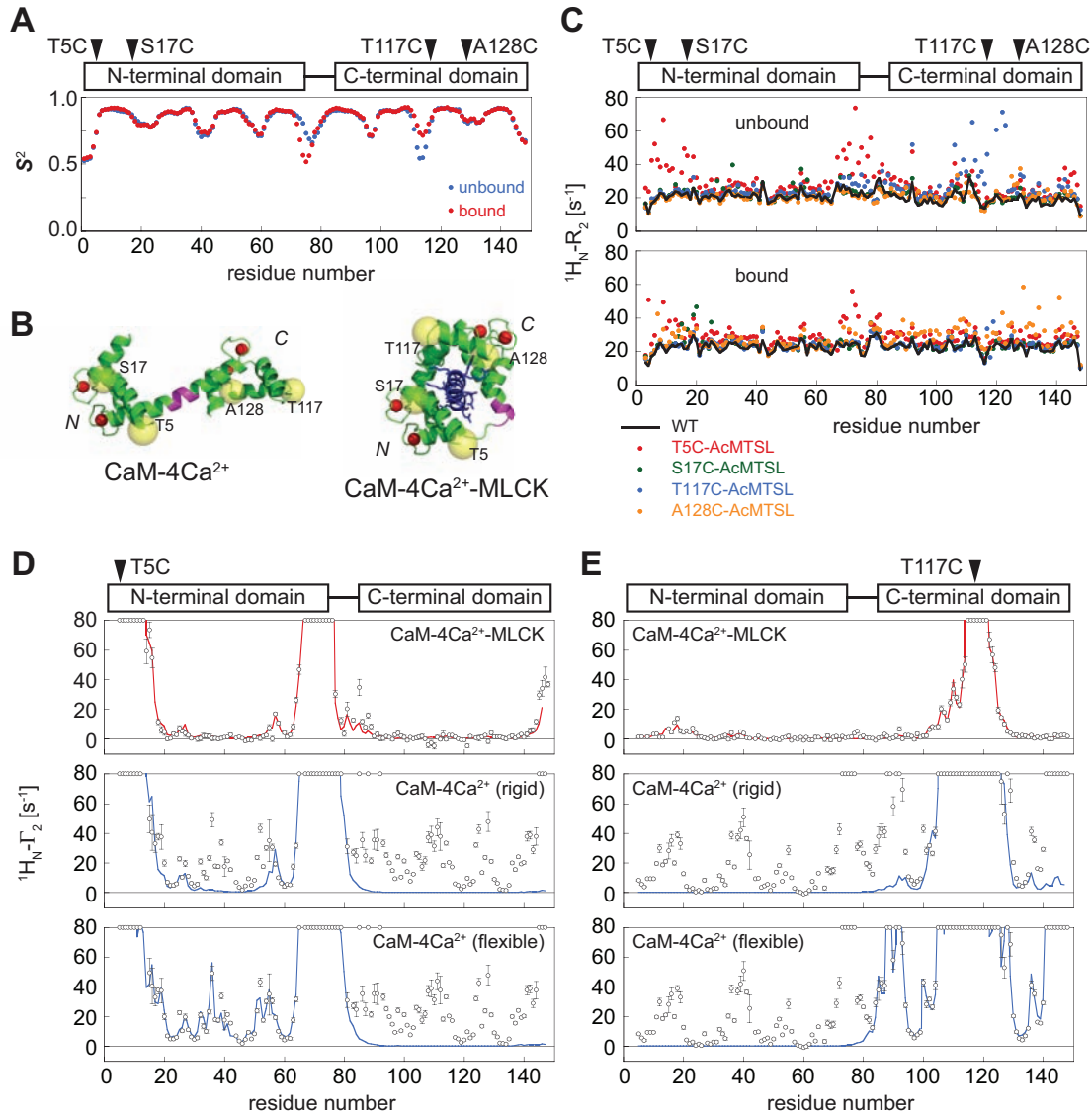


**Fig. S1. Affinity of CaM for the MLCK CBD peptide.** Isothermal titration calorimetry (ITC) measurements for the binding of the MLCK peptide to (A-B) wild type (WT) CaM and (C-F) nitroxide spin-labeled (“-MTSL”) mutants. ITC measurements were carried out in (A) the absence and (B-F) the presence of Ca<sup>2+</sup>. For the ITC experiments in the presence of Ca<sup>2+</sup>, the measurement cell contained 10  $\mu\text{M}$  CaM and the syringe 100  $\mu\text{M}$  MLCK peptide. For the experiment on apo CaM (i.e. no Ca<sup>2+</sup>) the corresponding concentrations were 50  $\mu\text{M}$  and 500  $\mu\text{M}$ , respectively (required because of the lower affinity of apo CaM for the MLCK peptide). The affinity of CaM for MLCK peptide increases by about two orders of magnitude (from  $K_D \sim 1 \mu\text{M}$  to  $\sim 10 \text{ nM}$ ) upon loading with Ca<sup>2+</sup>. Introduction of engineered surface cysteine mutations on CaM and subsequent spin label conjugation at these sites does not have any significant effect on the measured affinity, attesting to the suitability of the chosen sites. The stoichiometry of binding ( $N$ ) in all cases, including apo CaM, is one peptide molecule bound per molecule of CaM. The values of the equilibrium dissociation constants reported here are consistent with data in the literature.<sup>S1</sup>





**Fig. S2.  $^1\text{H}/^{15}\text{N}$  chemical shift perturbations in CaM upon binding MLCK CBD peptide.** (A)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -labeled  $\text{Ca}^{2+}$ -loaded CaM (100  $\mu\text{M}$ ) with 0 (red), 0.33 (yellow), 0.67 (green), 1 (cyan), and 1.2 (blue) equivalents of unlabeled MLCK peptide. (B)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -labeled apo CaM (100  $\mu\text{M}$ ) with 0 (red), 0.5 (yellow), 1 (green), 1.5 (cyan), and 2 (blue) equivalents of unlabeled MLCK peptide. The interaction between  $\text{Ca}^{2+}$ -loaded CaM and MLCK peptide is slow on the chemical shift time scale (i.e. two sets of cross-peaks are observed for unbound and peptide-bound CaM), whereas that with apo CaM is fast on the chemical shift time scale (i.e. a single set of cross-peaks with shifts at the weighted average of unbound and peptide-bound apo CaM). The slower timescale of the interaction for  $\text{Ca}^{2+}$ -loaded CaM and the greater magnitude of the chemical shift perturbations upon peptide binding (Fig. 1D main text), which are seen throughout the protein, are indicative of a larger binding interface and major structural rearrangements upon binding, consistent with the formation of the compact structure determined by NMR<sup>S2</sup> and crystallography.<sup>S3</sup> The much faster exchange kinetics and smaller chemical shift perturbations seen for apo CaM indicate that little structural change occurs upon peptide binding. That these perturbations are largely localized to the C-terminal domain indicates that the peptide only interacts primarily with the C-terminal domain and does not induce interdomain closing of apo CaM.

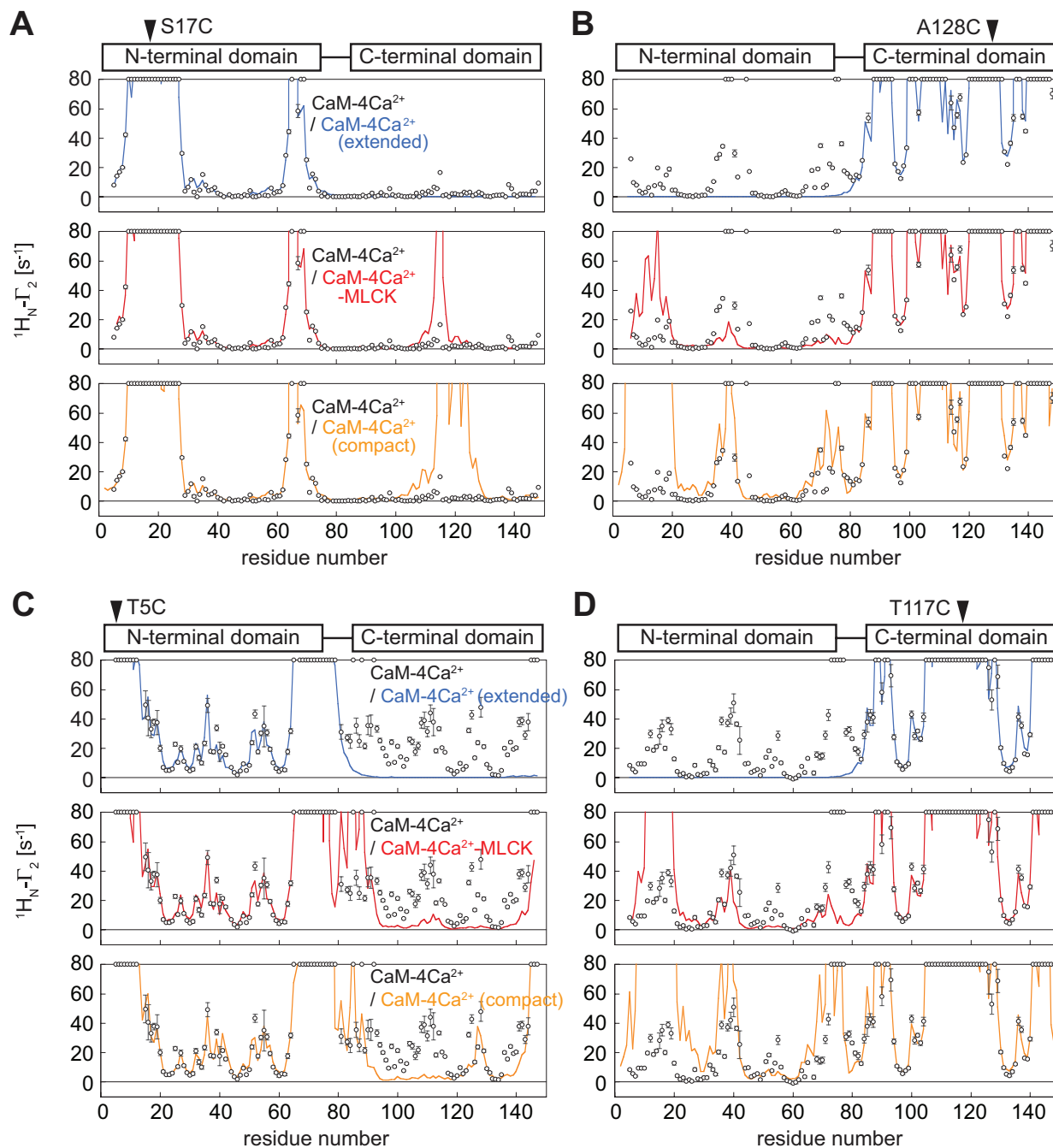


**Fig. S3. Local structural heterogeneity at T5C and T117C spin-labeling sites.** (A) Backbone chemical shift-derived order parameters ( $S^2$ ) calculated by TALOS+<sup>S4</sup> for Ca<sup>2+</sup>-loaded CaM in the presence (red) and absence (blue) of MLCK peptide. All tag positions are in regions with a high degree of rigidity, although T5C and T117C are located adjacent to mobile regions of the protein. (B) Structures of CaM-4Ca<sup>2+</sup> and CaM-4Ca<sup>2+</sup>-MLCK showing all four spin-label positions (yellow spheres). Ca<sup>2+</sup> ions are shown as red spheres, and spin-labeling sites are indicated by yellow spheres. The flexible linker (residues 77-81) is shown in magenta. (C)  $^1H_N-R_2$  values for Ca<sup>2+</sup>-loaded wild type CaM and CaM mutants conjugated to the diamagnetic control tag (“-AcMTSL”) in the absence (top) and presence (bottom) of MLCK peptide. Black, wild type; red, T5C; green, S17C; blue, T117C; and orange, A128C. Tagging at T117C causes local enhancement of  $R_2$  values, indicative of conformational exchange owing to local structural disruption; this effect disappears upon peptide binding. Tagging at T5C causes enhancement of  $R_2$  values both locally and near the linker region, which is close in space to the spin label site; this effect is less pronounced in the peptide-bound sample. Tagging at the S17C and A128C positions does not cause such effects, indicative of little to no structural perturbation. It is worth noting that the melting temperatures ( $T_m$ ) of all constructs (loaded with Ca<sup>2+</sup>), including T5C and T117C, determined by differential scanning calorimetry (data not shown) are similar with values of 115-120°C for the N-terminal domain and

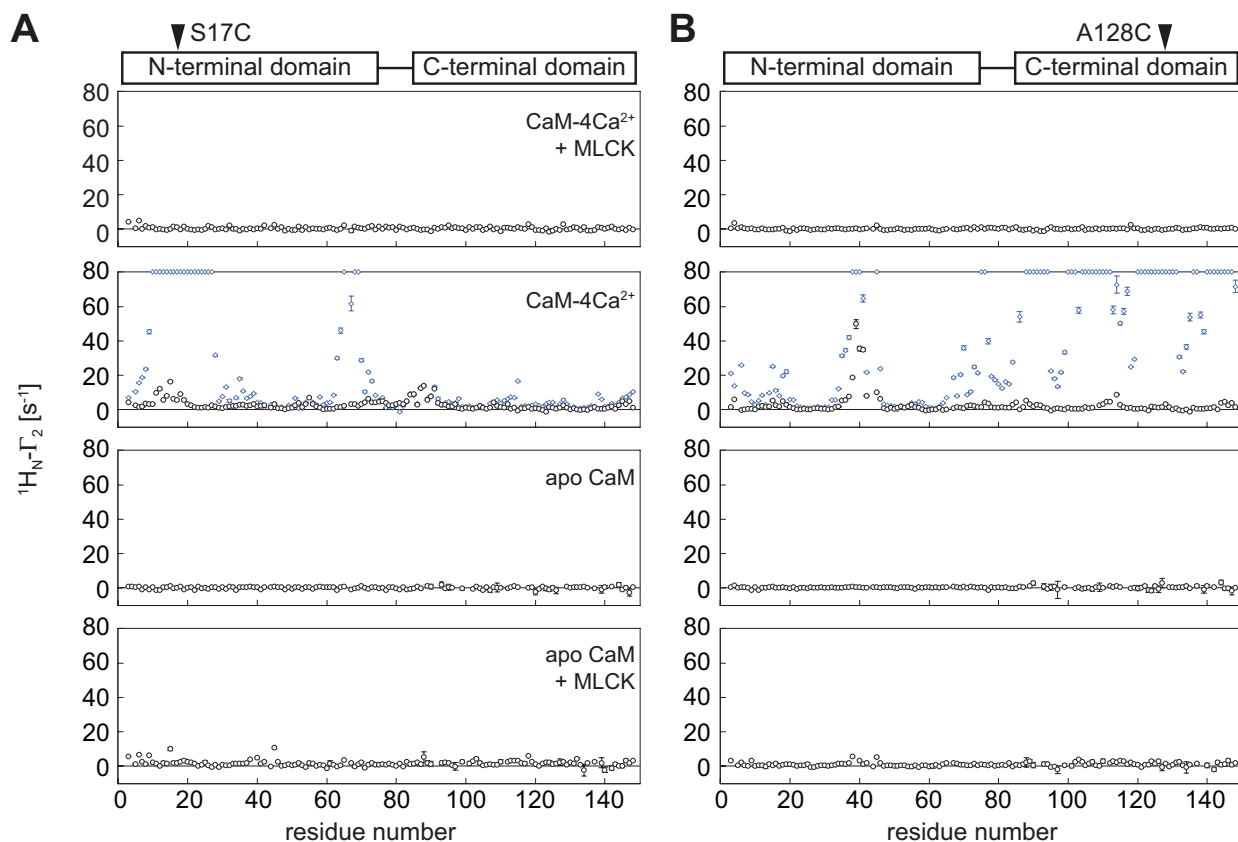
**Fig. S3 legend (cont.)**

95-100°C for the C-terminal domain, consistent with literature values of 113 and 93°C, respectively, for wild type CaM-4Ca<sup>2+</sup>.<sup>S1</sup> (*D,E*) PRE profiles for CaM with spin label at (*D*) T5C and (*E*) T117C in the presence (top panels) and absence (lower two panels) of the MLCK peptide. Experimental PREs are shown as circles (error bars, 1 s.d.) and back-calculated PREs are displayed as solid lines. Paramagnetic tag conformations (represented by a five-member ensemble) were obtained by fitting the intradomain PREs only; these positions were then used to back-calculate the PRE profiles for the remainder of the protein. The top panels show CaM-4Ca<sup>2+</sup>-MLCK PRE data fit to the peptide-bound structure, and the middle and bottom panels show the CaM-4Ca<sup>2+</sup> PRE data fit to the dumbbell structure. In the middle panel, the protein molecule was held rigid during the fit, but in the lower panel, residues 1-8 for T5C and residues 115-120 for T117C (including the respective paramagnetic labels) were represented by an ensemble of states and given torsional degrees of freedom to fit the data. These data are consistent with panel *C*, indicating that local structural heterogeneity in the immediate vicinity of the paramagnetic label needs to be taken into account for unbound CaM-4Ca<sup>2+</sup> conjugated at T5C or T117C, but not at S17C or A128C, or at any of the positions in the peptide-bound state. Thus, for this work we focus primarily on the PRE data from the S17C and A128C mutants, although the PRE data from T5C and T117C are still useful, when treated properly, for independent validation (cf. Figs. S8 and S9). Local heterogeneity arising from the introduction of paramagnetic labels at T5C and T117C, is likely due to fraying of the N-termini of helix I (in the N-terminal domain) and VII (in the C-terminal domain), respectively, as both T5 and T117 serve as helix N-capping residues with hydrogen bonds between the threonine O $\gamma$  atom and the backbone amide of the *i*+3 residue<sup>S5</sup>. In the complex, fraying is prevented by interactions with the peptide.<sup>S2,3</sup> Note that panels *D* and *E* show raw PRE profiles (see Fig. S5).

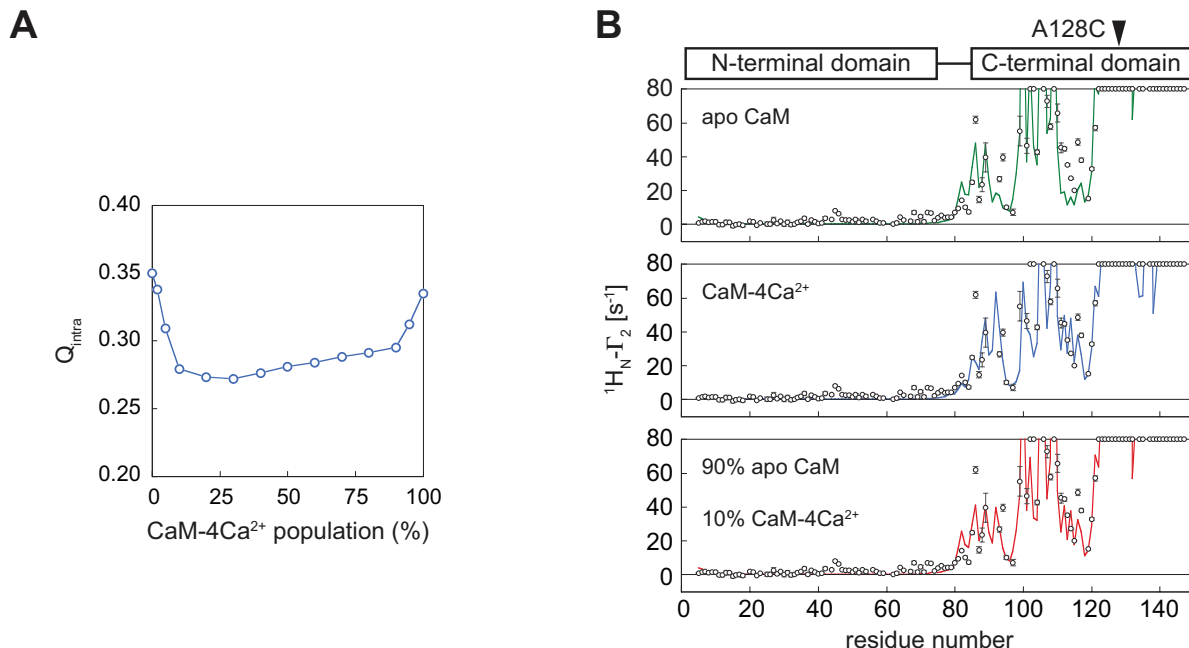




**Fig. S4. PRE profiles for Ca<sup>2+</sup>-loaded CaM.** Experimental PRE profiles (circles, error bars = 1 s.d.) for CaM-4Ca<sup>2+</sup> with the nitroxide spin label at (A) S17C, (B) A128C, (C) T5C, and (D) T117C. Where the PRE is too large to be accurately measured ( $\geq 80 \text{ s}^{-1}$ ), the experimental PREs are plotted at the top. PRE profiles back-calculated from known structures are shown as solid lines: CaM-4Ca<sup>2+</sup> in the dumbbell conformation (blue, PDB 1CLL<sup>S5</sup>), CaM-4Ca<sup>2+</sup>-MLCK (red, PDB 1CDL<sup>S3</sup>), and CaM-4Ca<sup>2+</sup> in the compact conformation (orange, PDB 1PRW<sup>S6</sup>). Paramagnetic tag conformations, represented by a five-member ensemble, were first obtained by fitting the intradomain PREs only; these positions were then used to back-calculate the PRE profiles for the rest of the protein. Note that panels C and D show raw PRE profiles (see Fig. S5).

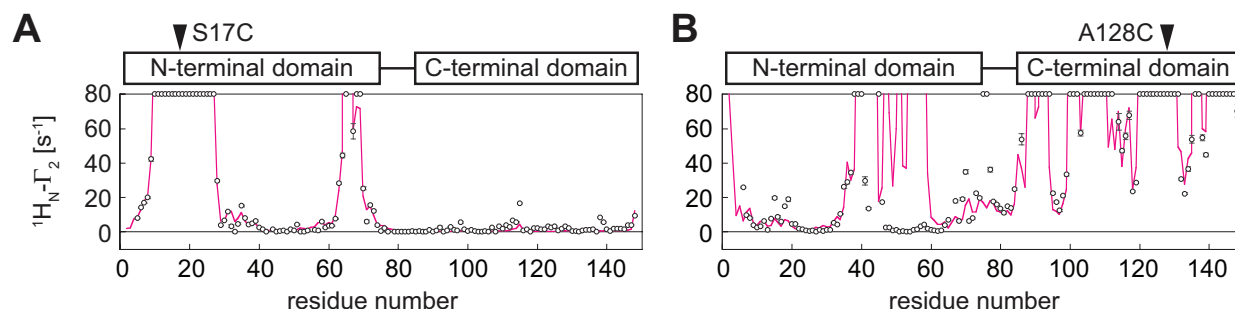


**Fig. S5. Controlling for intermolecular PRE effects.** Intermolecular PREs were exclusively observed using a sample comprising  $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled wild type CaM in the presence of CaM at natural isotopic abundance with a spin-label at (A) S17C or (B) A128C. Intermolecular PRE profiles for apo or  $\text{Ca}^{2+}$ -loaded CaM, in the presence or absence of MLCK peptide, are shown as black circles (error bar, 1 s.d.). Only for  $\text{CaM-4Ca}^{2+}$  were significant intermolecular PREs detected. For comparison, the raw experimental PREs measured for spin-labeled  $\text{U-}[^{15}\text{N}/^{13}\text{C}/^2\text{H}]\text{-CaM-4Ca}^{2+}$ , which reflect the sum of intermolecular and intramolecular PREs, are shown as blue diamonds in this case. Elsewhere in this report (unless otherwise specified), the PRE profiles displayed for  $\text{CaM-4Ca}^{2+}$  have had the intermolecular PREs subtracted out. No subtraction was necessary for CaM under the other conditions due to the absence of any significant intermolecular PREs. Although some minor intermolecular PREs were also observed for apo CaM in the presence of peptide, these were much smaller in magnitude and not large enough to affect the results of the calculations). With the exception of residues 84-91 for S17C, the intermolecular PREs for  $\text{CaM-4Ca}^{2+}$  follow a similar pattern to the intramolecular/interdomain PRE profiles, but with much smaller values, indicating that the intermolecular interactions between the N- and C-domains of CaM are likely very similar in nature to the corresponding intramolecular/interdomain interactions described in this report. The observation that the addition of MLCK peptide (or the removal of  $\text{Ca}^{2+}$ ) causes the intermolecular PREs to disappear supports this hypothesis.

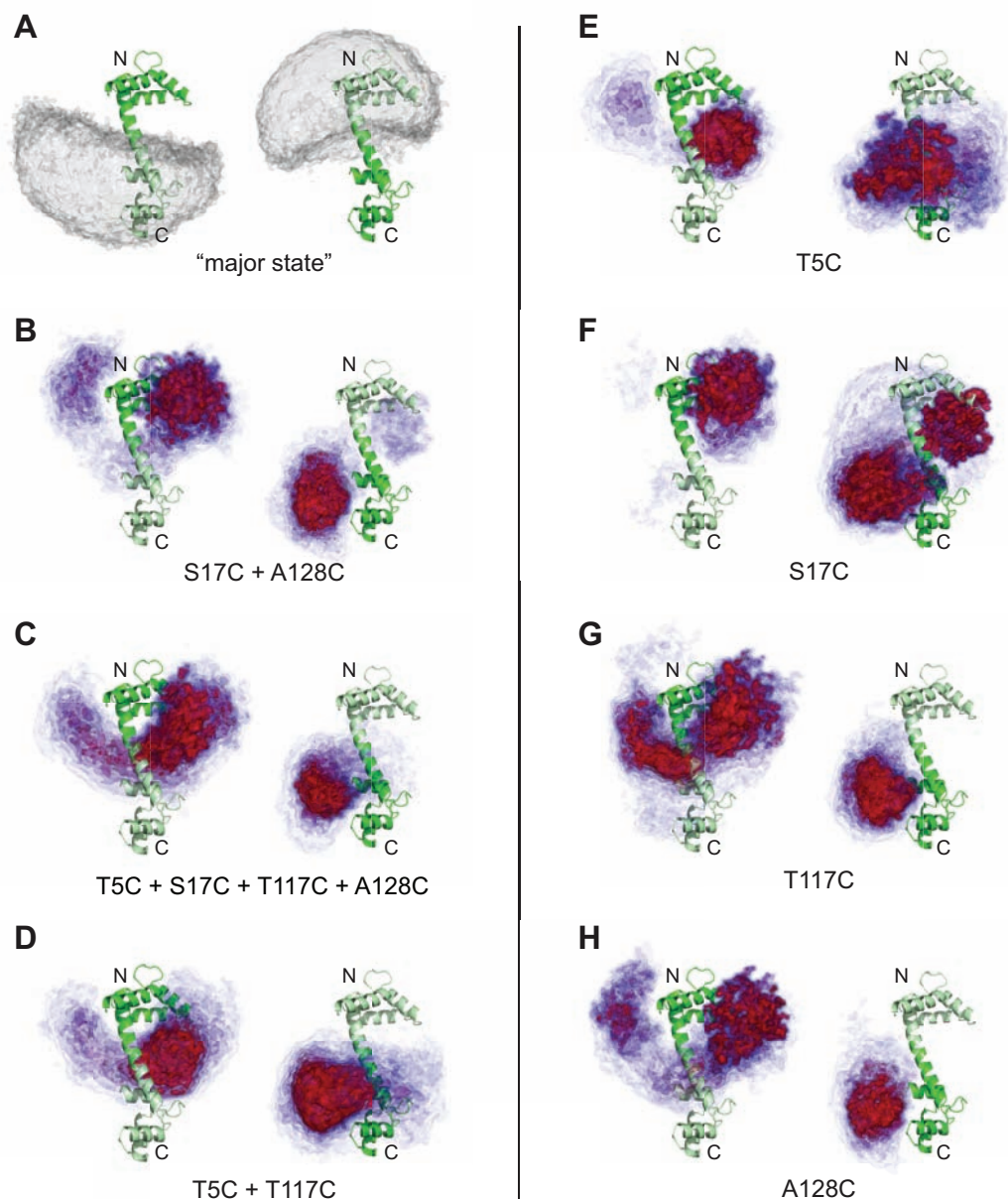


**Fig. S6. Fitting the intradomain A128C PRE data for apo CaM to a mixture of apo and Ca<sup>2+</sup>-loaded conformations.** Intradomain PREs for apo CaM tagged at A128C were fit to the NMR structure of apo CaM (PDB 1CFD), the crystal structure of CaM-4Ca<sup>2+</sup> (PDB 1CLL) or a mixture of the two. (A) Dependence of the intradomain PRE Q-factor as a function of the population of CaM-4Ca<sup>2+</sup> used in the fitting. The intradomain PRE Q-factor decreases sharply from 0.35 to 0.28 as the population of CaM-4Ca<sup>2+</sup> is increased from 0 to 10% (and as the population of apo CaM is concomitantly decreased from 100 to 90%), remains approximately constant up to a population of ~35% CaM-4Ca<sup>2+</sup>, and then increases at higher populations of CaM-4Ca<sup>2+</sup>. The C-terminal domain of apo CaM has been previously observed by <sup>15</sup>N relaxation studies to exchange between two species, with a minor state that resembles the Ca<sup>2+</sup>-loaded structure comprising up to 10% of the total population.<sup>S7,8</sup> Our PRE data are thus consistent with these earlier findings. (B) Experimental PRE profiles for apo CaM are shown as circles (error bar, 1 s.d.). PREs too large (> 80 s<sup>-1</sup>, roughly) to be accurately measured are plotted at the top of the charts. PRE profiles back-calculated from the structures of apo CaM (green), CaM-4Ca<sup>2+</sup> (blue), or a mixture consisting of 90% apo CaM and 10% CaM-4Ca<sup>2+</sup> (red) are shown as solid lines. The use of the mixture of conformations slightly improves the quality of the fit.





**Fig. S7. Comparison of experimental PREs for Ca<sup>2+</sup>-loaded CaM with those calculated from a molecular dynamics simulation.** The experimental PRE profiles (open circles) are presented as in Fig. 2 of the main text; PREs back-calculated from a published molecular dynamics (MD) ensemble (PDB 2K0E<sup>S9</sup>) are displayed as magenta lines. Paramagnetic tag conformations, represented by a five-member ensemble, were obtained by fitting the intradomain PREs only, and these positions were then used to back-calculate the PRE profiles for the remainder of the protein. The MD ensemble (which consists of 160 structures) predicts some, but not all, features of the experimental interdomain PRE profiles. The MD ensemble was generated using a complete empirical energy force field, in conjunction with restraints for intradomain NOEs and relaxation-derived backbone amide order ( $S^2$ ) parameters;<sup>S9</sup> thus there are no interdomain experimental restraints to drive the relative positioning of the two domains. The MD simulations show that the N- and C-terminal domains of CaM-4Ca<sup>2+</sup> can potentially sample a wide range of relative orientations, including some similar to the peptide-bound state, although not in any preferential manner as demonstrated by the experimental PRE data. The MD simulations also predict additional close-contact structures that are not compatible with the PRE data (particularly apparent for residues 47-59 in the A128C PRE data set shown in panel *B* and misses others (e.g. residues 138 and 139 in panel *A* and 15-19 in panel *B*).

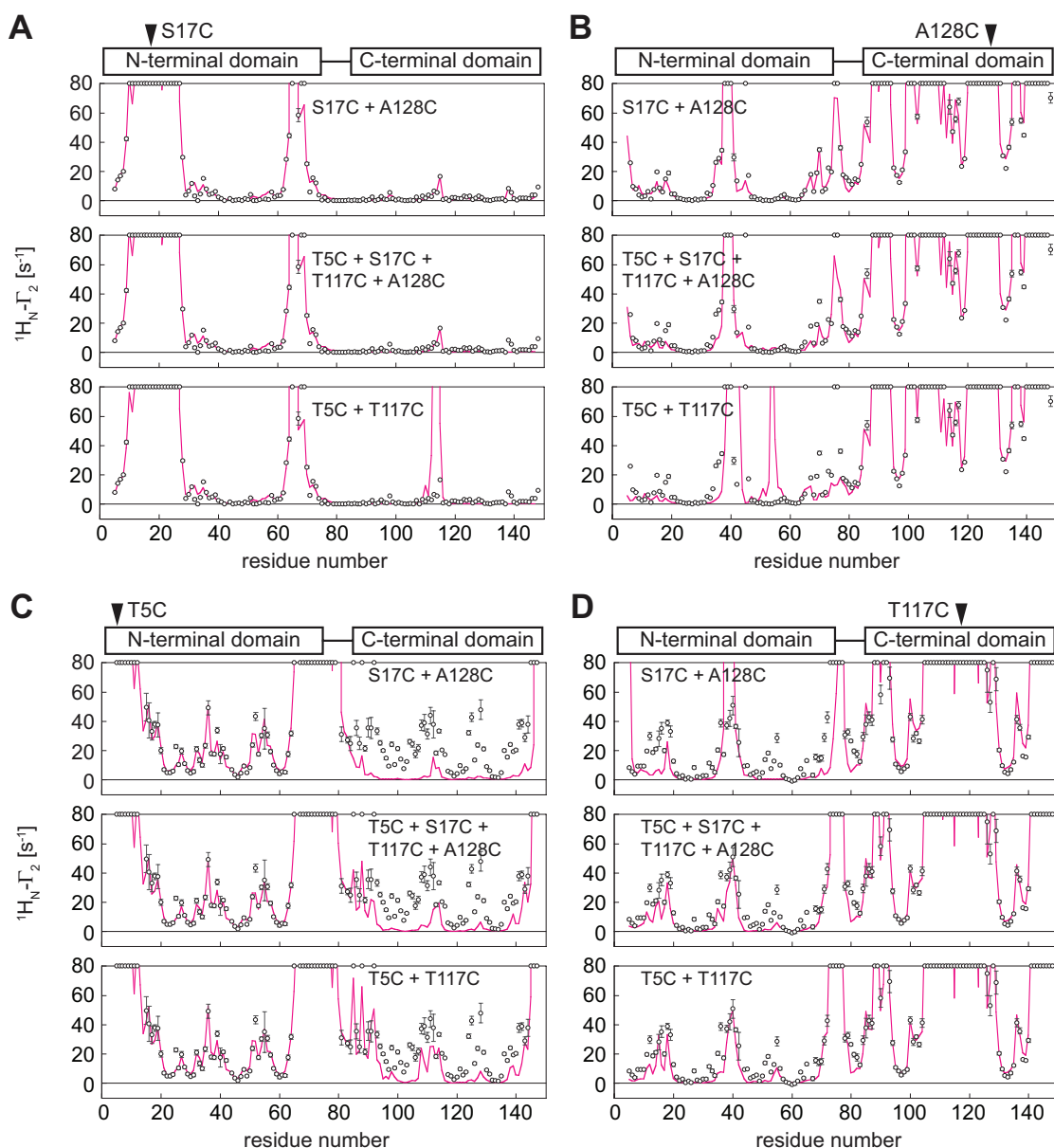


**Fig. S8. Reweighted atomic probability density maps for  $\text{Ca}^{2+}$ -loaded CaM derived from PRE-driven ensemble simulated annealing calculations.** (A) Atomic probability maps (gray, plotted at a contour level of 0.1 of maximum) representing the conformational space sampled by the "major" CaM-4 $\text{Ca}^{2+}$  extended state generated by calculations in which the interdomain PREs for S17C and A128C were restrained to be less than  $2 \text{ s}^{-1}$ . For this panel and all others, all ensemble members on the left side have been aligned to the N-terminal domain (dark green), and the probability density map is shown for the C-terminal domain; on the right side, all ensemble members have been aligned to the C-terminal domain (dark green), and the probability density map is shown for the N-terminal domain. The conformational space sampled by one domain relative to the other occupies roughly one full hemisphere (i.e. a cone with a semi-angle  $90^\circ$ ). While this is a greater degree of flexibility than that deduced from NMR relaxation data (cone of semi-angle  $\sim 30^\circ$ ),<sup>S10</sup> the magnitude and characteristics of the conformational space sampled in this calculation are very similar to those of the ensemble recently modeled from pseudocontact

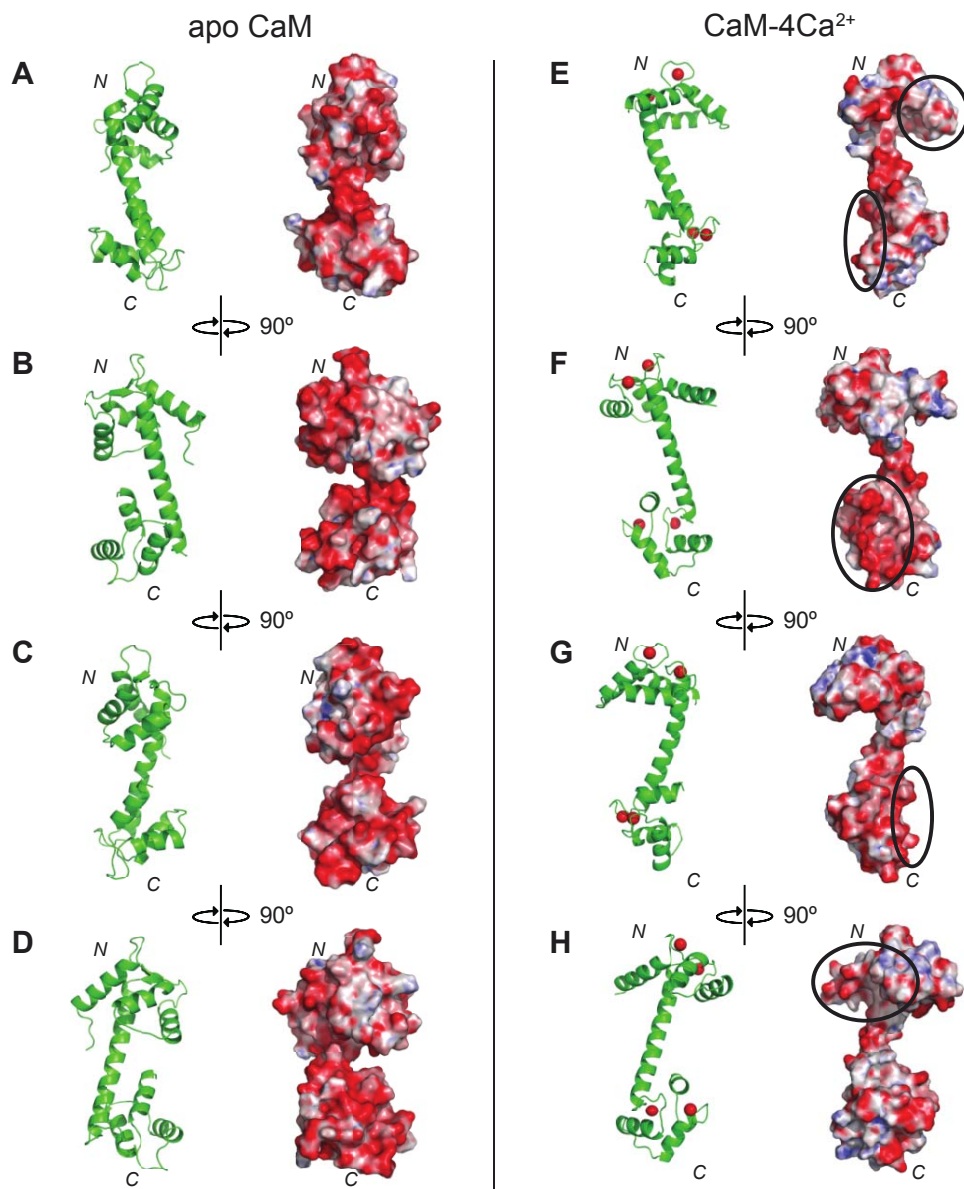
**Fig. S8 legend (cont.)**

shift (PCS), residual dipolar coupling (RDC), and small angle X-ray scattering (SAXS) data.<sup>S11</sup> Further, the tilt of our “major” state hemisphere from the dumbbell X-ray structure ( $\sim 45^\circ$ ) is similar to the tilt determined from analysis of the PCS and RDC data ( $\sim 30^\circ$ )<sup>S11,12</sup>. It should be emphasized, however, that the “major” state ensemble calculated here represents a rough guide, since it is not defined by direct measurements but by the absence of interdomain PREs. It should also be noted that the dumbbell X-ray structure fits within the cone depicted in panel *A*, and in the ensuing ensemble simulating annealing calculations, aimed at probing minor states with interdomain contacts from the experimental PRE data, it made no difference whether the “major” state was represented by such an ensemble or just by the dumbbell structure alone. (*B-H*) Atomic probability density maps of the CaM-4Ca<sup>2+</sup> “minor state” (i.e. states with close interdomain contacts) calculated from simulated annealing calculations using the indicated data set(s): (*B*) S17C and A128C; (*C*) T5C, S17C, T117C and A128C; (*D*) T5C and T117C; (*E*) T5C; (*F*) S17C; (*G*) T117C; (*H*) A128C. Each calculation included a minor state ensemble consisting of 8 members, encompassing 10% of the total population. Atomic probability density maps are plotted at multiple contour levels, from 0.1 (transparent blue) to 0.5 (opaque red) of maximum. It is apparent that the results of the simulated annealing calculations with individual PRE data sets (*E-H*) are approximately similar to those of the combined data sets. One feature that does emerge from the differences between the calculated ensembles is that spin labels located in the N-terminal domain more powerfully refine the positions of the C-terminal domain relative to the N-terminal domain—and vice versa for spin labels located in the C-terminal domain. Thus, for this sort of study, one nitroxide label in each domain represents the minimum amount of data needed to adequately determine such an ensemble with high confidence. It should be emphasized that the PRE data used in these calculations report directly upon close-contact interdomain conformations, offering novel descriptions of the compact CaM-4Ca<sup>2+</sup> minor states shown here as colored probability density maps. This is in stark contrast to earlier work, which primarily reports upon the details of the major extended state (i.e. similar to the grey probability density maps in panel *A*).<sup>S9-13</sup>

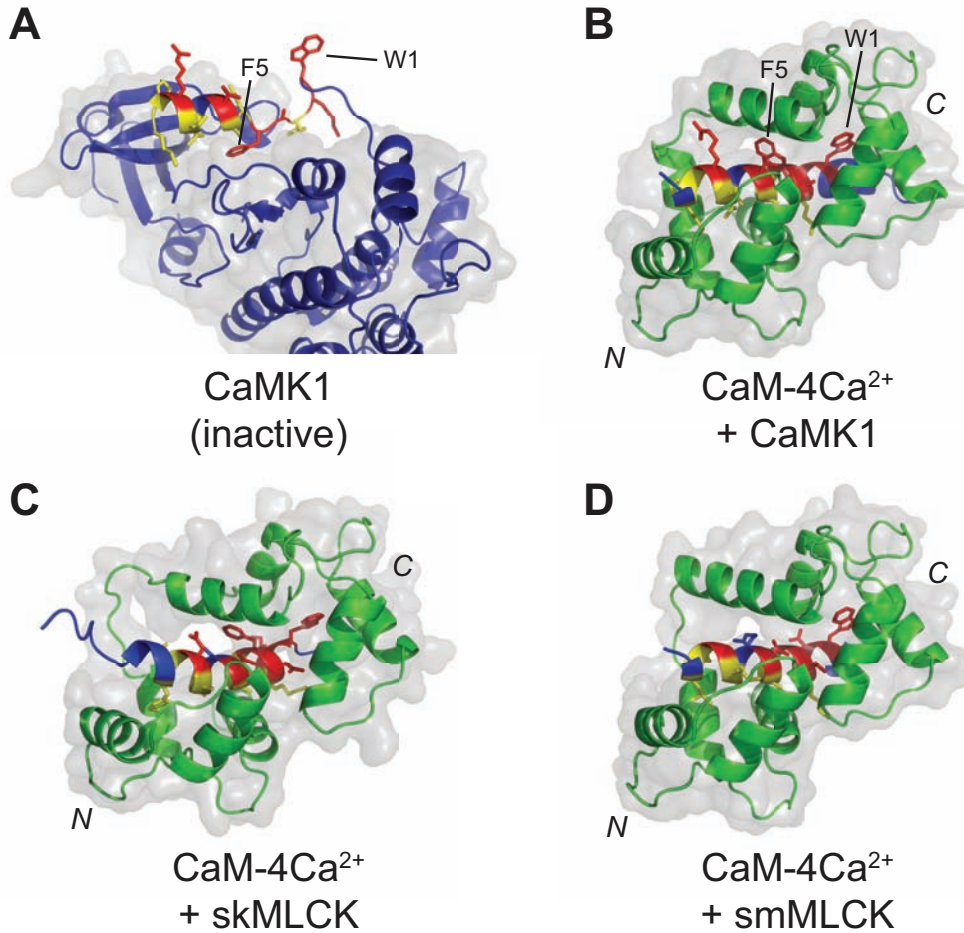




**Fig. S9. PRE-driven minor state ensemble simulated annealing calculations for CaM-4Ca<sup>2+</sup> using different data set combinations.** (A-D) PRE profiles for (A) S17C, (B) A128C, (C) T5C, and (D) A128C resulting from simulated annealing calculations with a minor compact state population of 10% represented by an 8-member ensemble (calculated, magenta lines; experimental, black circles; error bars, 1 s.d). Calculations were carried out using the indicated PRE data sets as follows: top row, S17C and A128C; middle row, T5C, S17C, T117C and A128C; bottom row, T5C and T117C. The interdomain PRE data can be fit for each spin-label position, although the fits are not as good for T5C owing to lower quality data and local dynamics in the vicinity of the spin-label (see Fig. S3). Although the inclusion of a given PRE data set in the simulated annealing calculations greatly improved the fit to that particular data set, there was a qualitative improvement in the fit even if that data set was *not* used to calculate the ensemble. For example, in the top row of panel D, the major features of the experimental T117C PRE data are reproduced by the PREs back-calculated from the ensembles obtained using only the S17C and A128C data. Thus, the individual PRE data sets have some predictive power over the others. This speaks to the consistency of the data, which is also highlighted by the ability to fit all four data sets simultaneously (middle row of each panel). Note that panels C and D show raw PRE profiles (see Fig. S5).



**Fig. S10. Electrostatic potential surface maps for CaM.** Electrostatic potentials were mapped onto structures of (A-D) apo CaM and (E-H) CaM-4Ca<sup>2+</sup>. The top row (A and E) is in the same orientation as Fig. 5 so that the electrostatic potentials can be compared with the interdomain interactions observed by PREs. The structures are shown in four different orientations related by 90° rotations about the y-axis. CaM is predominately negatively charged in both states, but CaM-4Ca<sup>2+</sup> displays additional patches of positive charge. The interdomain interaction surfaces in CaM-4Ca<sup>2+</sup> are indicated by black ellipses. The interaction surface on the C-terminal domain (which can be best seen in panel F) is highly negatively charged, while the interaction surface on the N-terminal domain (best seen in panels E and H) is partially positively charged. Thus, electrostatics may in part explain the formation of compact peptide-bound-like structures in the presence, but not absence, of Ca<sup>2+</sup>, consistent with the observed negative salt dependence of the interdomain PREs (Fig. 3). Electrostatic potential surface maps were calculated and visualized using APBS<sup>S14</sup> and PyMOL (<http://www.pymol.org>),



**Fig. S11. Exposure of C-terminal binding residues in the CBD of the inactive kinase.** Structure of the kinase CaM-binding domain (CBD), either (A) self-associated with the kinase active site or (B-D) bound to Ca<sup>2+</sup>-loaded CaM. (A) X-ray structure of inactive CaMK1 (PDB 1A06);<sup>S15</sup> (B) X-ray structure of CaMK1 CBD peptide bound to Ca<sup>2+</sup>-loaded CaM (PDB 1MXE);<sup>S16</sup> (C) NMR structure of skeletal muscle MLCK (skMLCK) CBD peptide (i.e. the peptide used in the experiments presented here) bound to Ca<sup>2+</sup>-loaded CaM (PDB 2BBM);<sup>S2</sup> (D) X-ray structure of smooth muscle MLCK (smMLCK) CBD peptide bound to Ca<sup>2+</sup>-loaded CaM (PDB 1CDL).<sup>S3</sup> CaM is shown in green and kinase in blue. Side chains of the 14-residue consensus region of the CBD are shown in stick representation, and Trp1 and Phe4 (numbered from the beginning of the CBD) are labeled. CBD residues buried by the C-terminal domain of CaM are colored red, and those buried by the N-terminal domain are colored yellow. The inactive kinase shown is CaM-dependent kinase 1 (CaMK1), because of the availability of structural information. However, panels (b-d) show that CaM interacts with CaMK1 in essentially the same manner as with the MLCKs (used in this study). Note that in the inactive kinase, residues that interact with the N-terminal domain of CaM are buried, but those that interact with the C-terminal domain—most notably Trp1—are largely accessible (with the exception of Phe5, which interacts directly with the ATP-binding region of the kinase). These structures, combined with the data presented here, support the hypothesis that the C-terminal domain of apo CaM could remain associated with the inactive kinase; activation would occur when Ca<sup>2+</sup> loading allows CaM to clamp down upon the CBD peptide, causing dissociation from the active site and kinase activation. This is consistent with the previously proposed hypothesis that CaM peptide binding is initiated by the C-terminal domain.<sup>S17</sup>

### Supplementary References

- (S1) Yamniuk, A. P.; Ishida, H.; Lippert, D.; Vogel, H. J. *Biophysical journal* **2009**, *96*, 1495-507.
- (S2) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax, A. *Science* **1992**, *256*, 632-638.
- (S3) Meador, W. E.; Means, A. R.; Quioco, F. A. *Science* **1992**, *257*, 1251-1255.
- (S4) Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. *J Biomol NMR* **2009**, *44*, 213-23.
- (S5) Chattopadhyaya, R.; Meador, W. E.; Means, A. R.; Quioco, F. A. *J. Mol. Biol.* **1992**, *228*, 1177-1192.
- (S6) Fallon, J. L.; Quioco, F. A. *Structure* **2003**, *11*, 1303-1307.
- (S7) Kuboniwa, H.; Tjandra, N.; Grzesiek, S.; Ren, H.; Klee, C. B.; Bax, A. *Nature Struct. Biol.* **1995**, *2*, 768-776.
- (S8) Malmendal, A.; Evenas, J.; Forsen, S.; Akke, M. *J. Mol. Biol.* **1999**, *293*, 883-899.
- (S9) Gsponer, J.; Christodoulou, J.; Cavalli, A.; Bui, J. M.; Richter, B.; Dobson, C. M.; Vendruscolo, M. *Structure* **2008**, *16*, 736-746.
- (S10) Baber, J. L.; Szabo, A.; Tjandra, N. *J. Am. Chem. Soc.* **2001**, *123*, 3953-3959.
- (S11) Bertini, I.; Giachetti, A.; Luchinat, C.; Parigi, G.; Petoukhov, M. V.; Pierattelli, R.; Ravera, E.; Svergun, D. I. *J. Am. Chem. Soc.* **2010**, *132*, 13553-13558.
- (S12) Bertini, I.; Del Bianco, C.; Gelis, I.; Katsaros, N.; Luchinat, C.; Parigi, G.; Peana, M.; Provenzani, A.; Zoroddu, M. A. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 6841-6846.
- (S13) Tjandra, N.; Kuboniwa, H.; Ren, H.; Bax, A. *Eur. J. Biochem.* **1995**, *230*, 1014-1024.
- (S14) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. *Proc Natl Acad Sci U S A* **2001**, *98*, 10037-41.
- (S15) Goldberg, J.; Nairn, A. C.; Kuriyan, J. *Cell* **1996**, *84*, 875-887.
- (S16) Clapperton, J. A.; Martin, S. R.; Smerdon, S. J.; Gamblin, S. J.; Bayley, P. M. *Biochemistry* **2002**, *41*, 14669-79.
- (S17) Bayley, P. M.; Findlay, W. A.; Martin, S. R. *Protein Sci* **1996**, *5*, 1215-28.