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, Michaeleen Doucleff[l], 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.

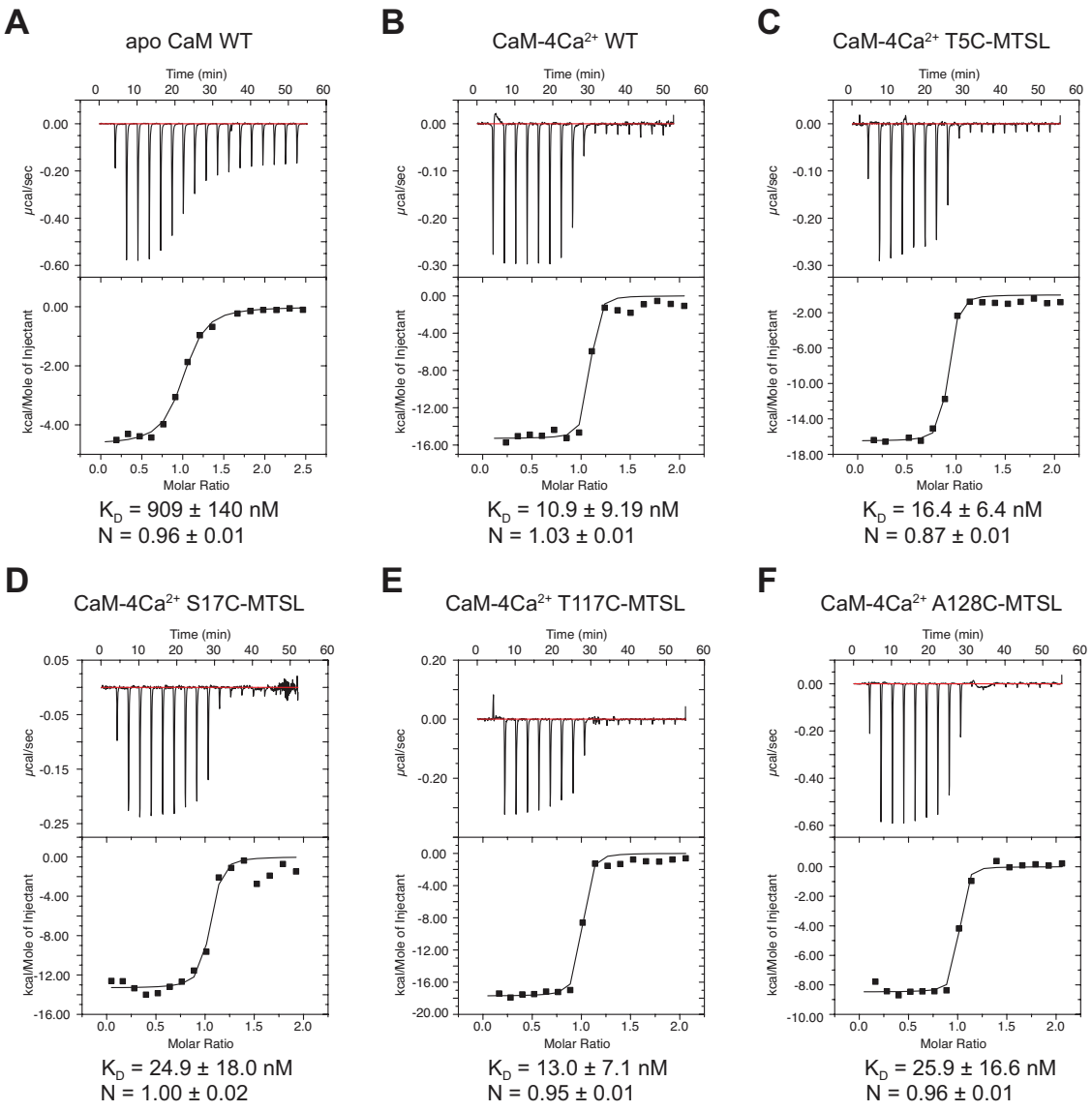

Fig. S1. Affinity of CaM for the MLCK CBD peptide. Isothermal titration calorimetry (ITC) measurements for the binding of the MLCK peptide to (*A*-*B*) wild type (WT) CaM and (*C*-*F*) nitroxide spin–labeled ("-MTSL") mutants. ITC measurements were carried out in (*A*) the absence and (*B*-*F*) the presence of Ca^{2+}. For the ITC experiments in the presence of Ca^{2+}, the measurement cell contained 10 µM CaM and the syringe 100 µM MLCK peptide. For the experiment on apo CaM (i.e. no Ca^{2+}) the corresponding concentrations were 50 µM and 500 µM, 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 ~1 µM to ~10 nM) upon loading with Ca^{2+}. 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.[S1]

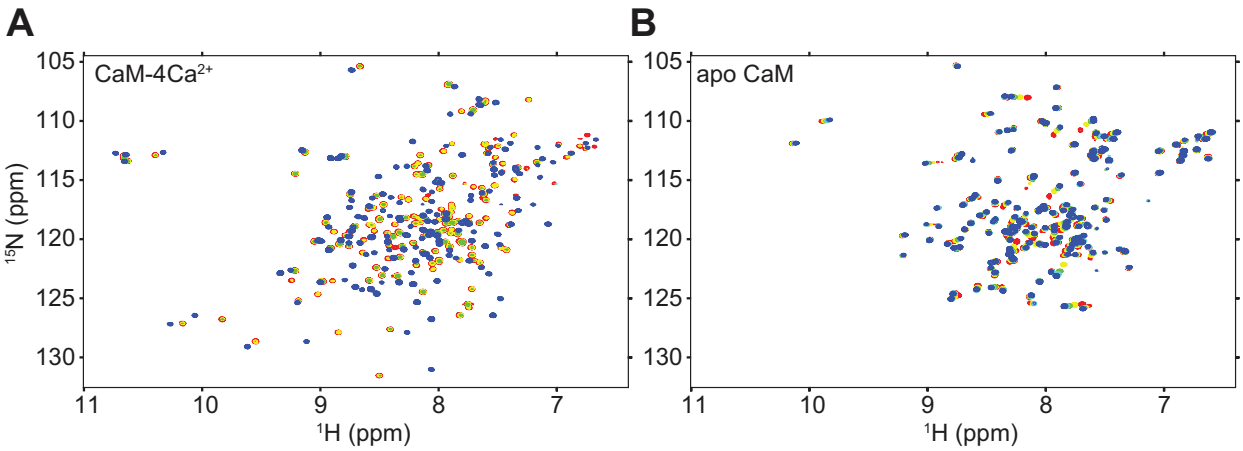

Fig. S2. $^{1}H_{N}/^{15}N$ chemical shift perturbations in CaM upon binding MLCK CBD peptide. (*A*) ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled Ca^{2+}-loaded CaM (100 µM) with 0 (red), 0.33 (yellow), 0.67 (green), 1 (cyan), and 1.2 (blue) equivalents of unlabeled MLCK peptide. (*B*) ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled apo CaM (100 µM) with 0 (red), 0.5 (yellow), 1 (green), 1.5 (cyan), and 2 (blue) equivalents of unlabeled MLCK peptide. The interaction between 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 Ca^{2+}-loaded CaM and the greater magnitude of the chemical shift perturbations upon peptide binding (Fig. 1*D* 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[S2] and crystallography.[S3] 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.

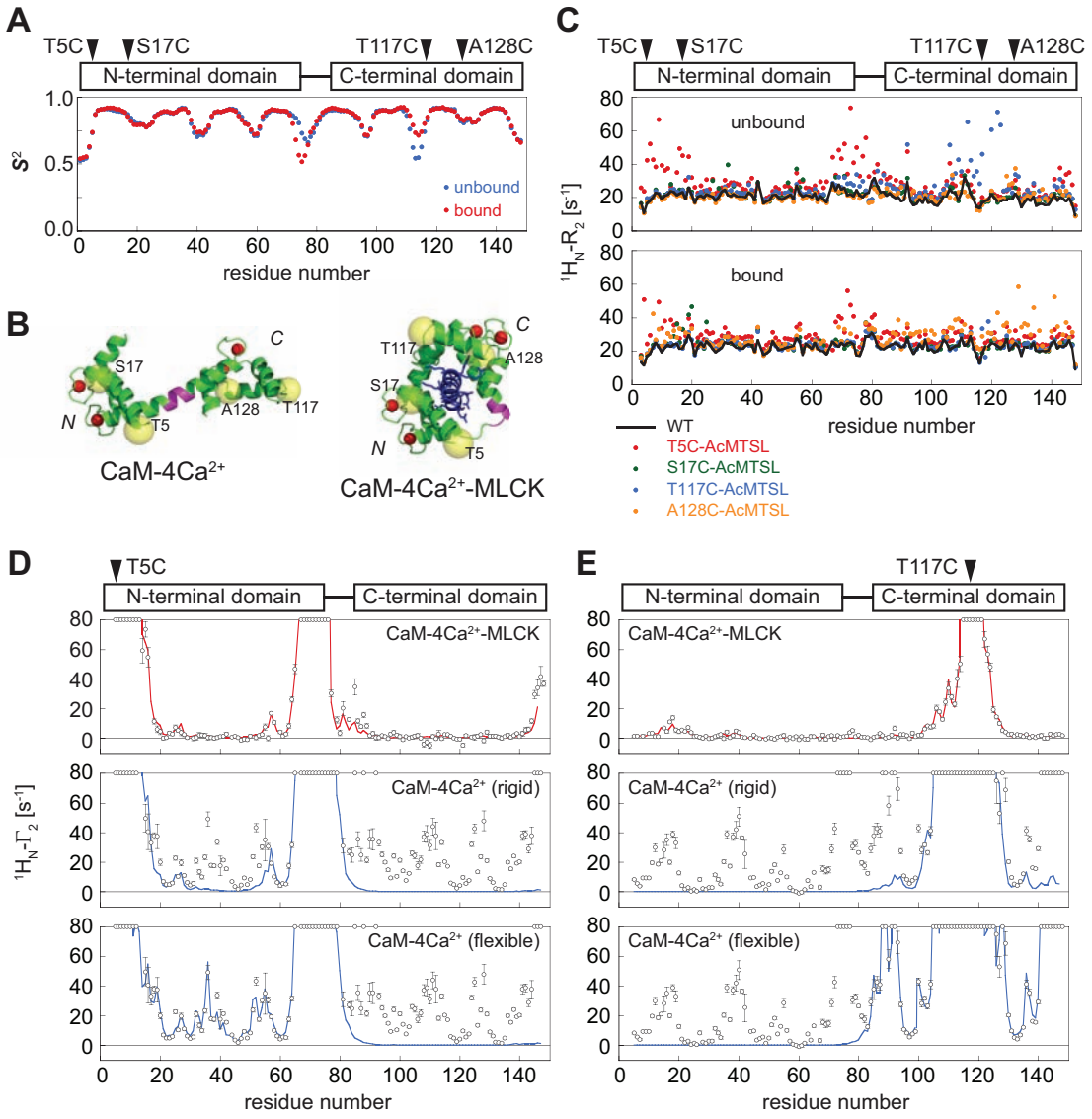

Fig. S3. Local structural heterogeneity at T5C and T117 spin-labeling sites. (*A*) Backbone chemical shift-derived order parameters ($\boldsymbol{S}^2$) calculated by TALOS+ [S4] for Ca^{2+}-loaded CaM in the presence (red) and absence (blue) of MLCK peptide. All tag positions are in regions with a high degree of rigidity, although T5C and T117C are located adjacent to mobile regions of the protein. (*B*) Structures of CaM-4Ca^{2+} and CaM-4Ca^{2+}-MLCK showing all four spin-label positions (yellow spheres). Ca^{2+} 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^{2+}-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^{2+}), 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^{2+}.[S1] (*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^{2+}-MLCK PRE data fit to the peptide-bound structure, and the middle and bottom panels show the CaM-4Ca^{2+} 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^{2+} 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 Oγ atom and the backbone amide of the *i*+3 residue[S5]. In the complex, fraying is prevented by interactions with the peptide.[S2,3] Note that panels *D* and *E* show raw PRE profiles (see Fig. S5).

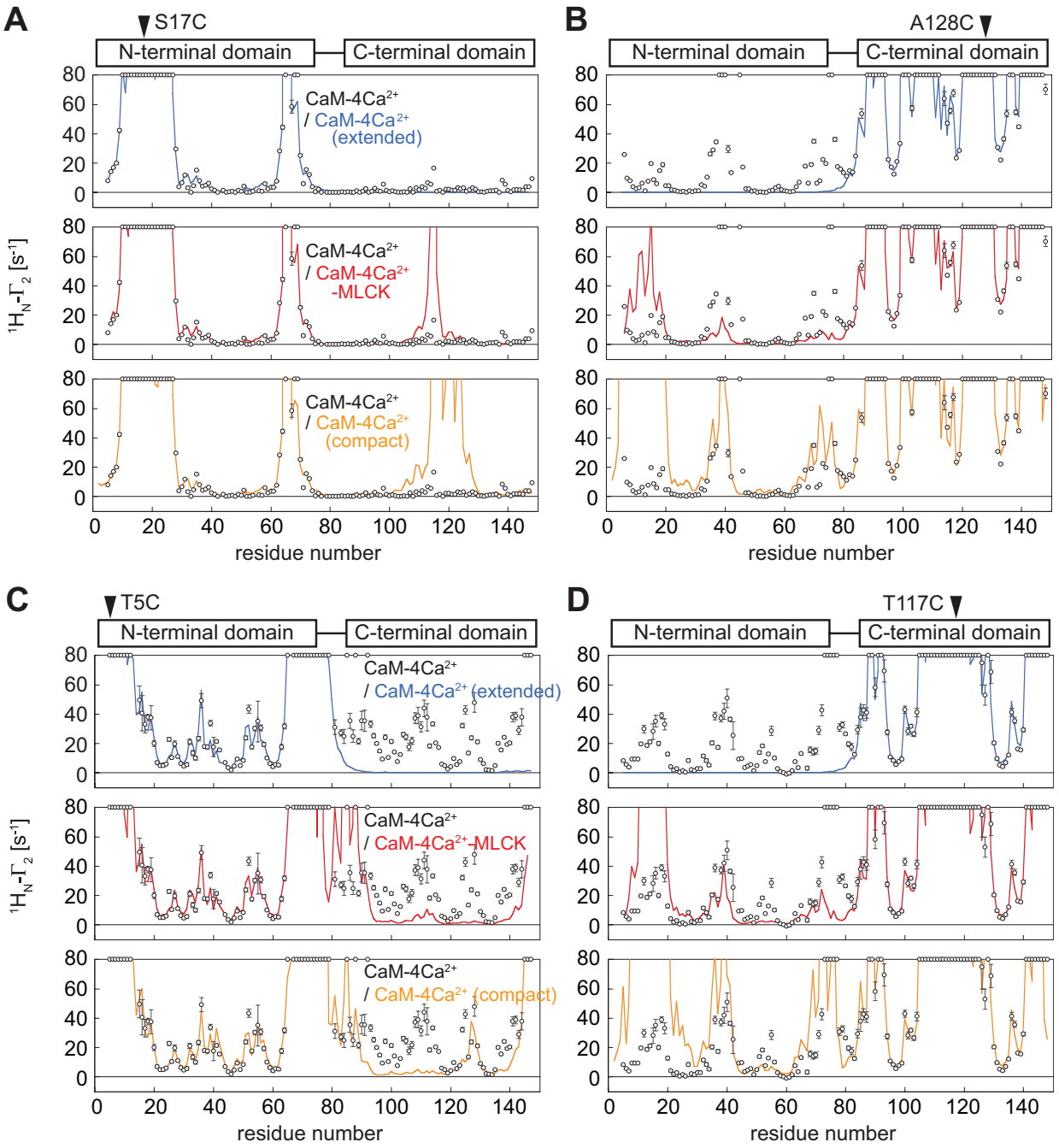

Fig. S4. PRE profiles for Ca^{2+}-loaded CaM. Experimental PRE profiles (circles, error bars = 1 s.d.) for CaM-4Ca^{2+} with the nitroxide spin label at (*A*) S17C, (*B*) A128C, (*C*) T5C, and (*D*) T117C. Where the PRE is too large to be accurately measured ($\geq$ 80 s^{-1}), the experimental PREs are plotted at the top. PRE profiles back-calculated from known structures are shown as solid lines: CaM-4Ca^{2+} in the dumbbell conformation (blue, PDB 1CLL[S5]), CaM-4Ca^{2+}-MLCK (red, PDB 1CDL[S3]), and CaM-4Ca^{2+} in the compact conformation (orange, PDB 1PRW[S6]). 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).

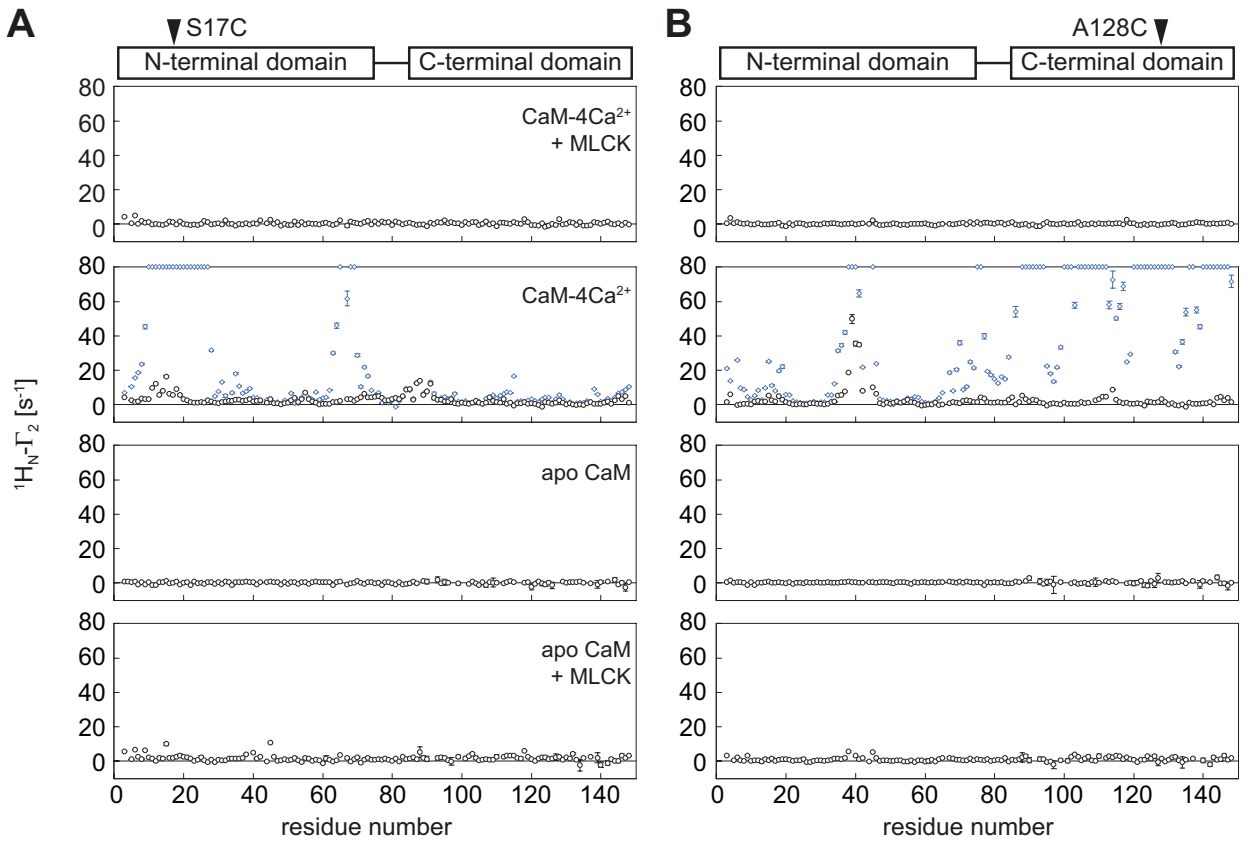

Fig. S5. Controlling for intermolecular PRE effects. Intermolecular PREs were exclusively observed using a sample comprising $^{2}H/^{13}C/^{15}N$-labeled wild type CaM in the presence of CaM at natural isotopic abundance with a spin-label at (*A*) S17C or (*B*) A128C. Intermolecular PRE profiles for apo or Ca^{2+}-loaded CaM, in the presence or absence of MLCK peptide, are shown as black circles (error bar, 1 s.d.). Only for CaM-4Ca^{2+} were significant intermolecular PREs detected. For comparison, the raw experimental PREs measured for spin-labeled U-[$^{15}N/^{13}C/^{2}H$]-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 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 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 Ca^{2+}) causes the intermolecular PREs to disappear supports this hypothesis.

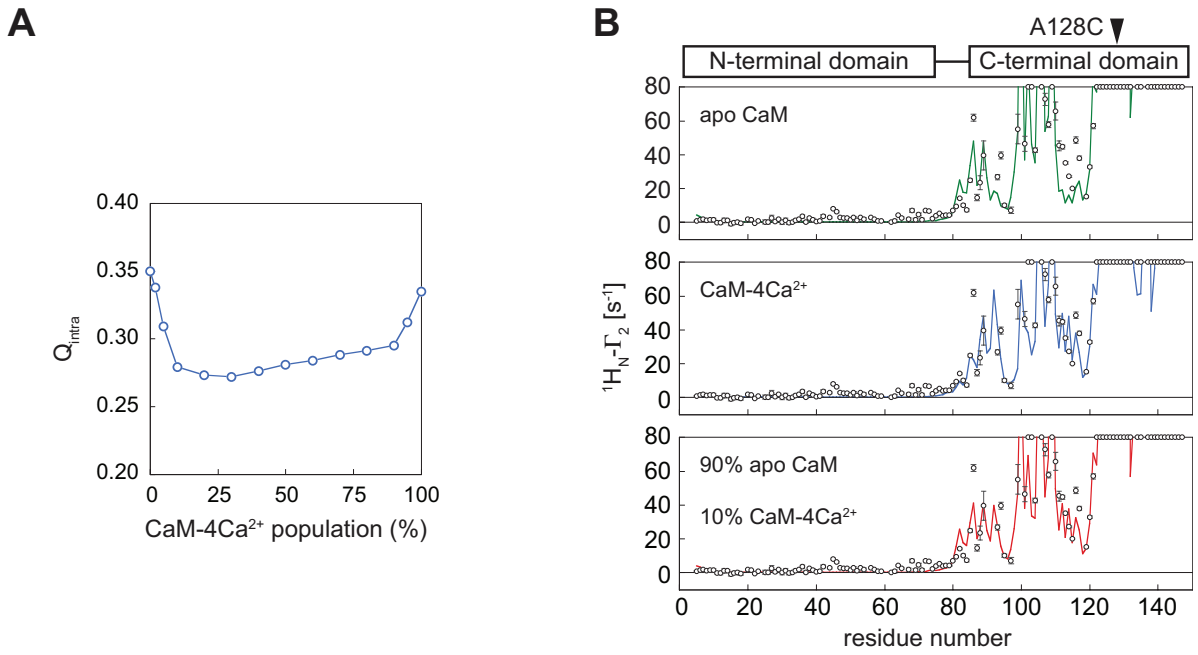

Fig. S6. Fitting the intradomain A128C PRE data for apo CaM to a mixture of apo and Ca^{2+}-loaded conformations. Intradomain PREs for apo CaM tagged at A128C were fit to the NMR structure of apo CaM (PDB 1CFD), the crystal structure of CaM-4Ca^{2+} (PDB 1CLL) or a mixture of the two. (*A*) Dependence of the intradomain PRE Q-factor as a function of the population of CaM-4Ca^{2+} used in the fitting. The intradomain PRE Q-factor decreases sharply from 0.35 to 0.28 as the population of CaM-4Ca^{2+} 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^{2+}, and then increases at higher populations of CaM-4Ca^{2+}. The C-terminal domain of apo CaM has been previously observed by ^{15}N relaxation studies to exchange between two species, with a minor state that resembles the Ca^{2+}-loaded structure comprising up to 10% of the total population.[S7,8] 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^{-1}, 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^{2+} (blue), or a mixture consisting of 90% apo CaM and 10% CaM-4Ca^{2+} (red) are shown as solid lines. The use of the mixture of conformations slightly improves the quality of the fit.

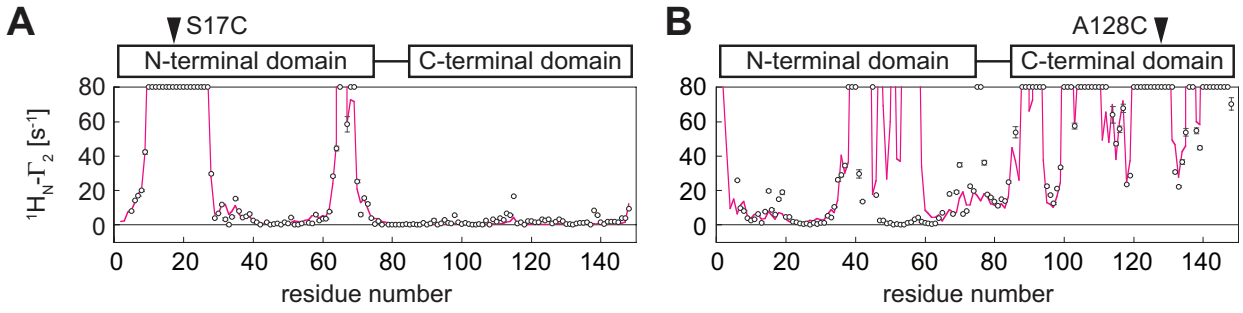

Fig. S7. Comparison of experimental PREs for Ca^{2+}-loaded CaM with those calculated from a molecular dynamics simulation. The experimental PRE profiles (open circles) are presented as in Fig. 2 of the main text; PREs back-calculated from a published molecular dynamics (MD) ensemble (PDB 2K0E[S9]) are 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 ($\boldsymbol{S}^2$) parameters;[S9] 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^{2+} 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*).

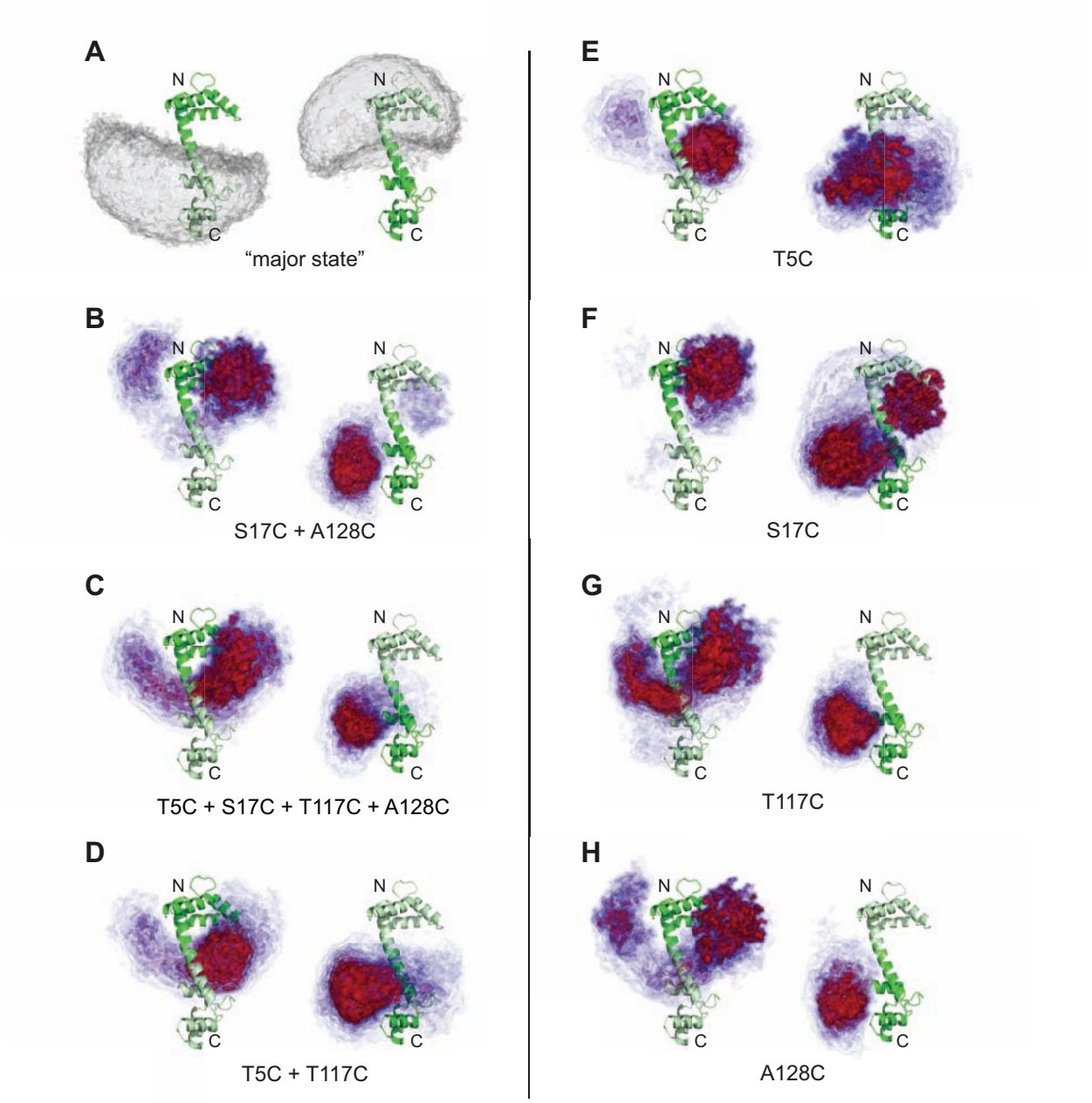

Fig. S8. Reweighted atomic probability density maps for Ca^{2+}-loaded CaM derived from PRE-driven ensemble simulated annealing calculations. (*A*) Atomic probability maps (gray, plotted at a contour level of 0.1 of maximum) representing the conformational space sampled by the "major" CaM-$4Ca^{2+}$ extended state generated by calculations in which the interdomain PREs for S17C and A128C were restrained to be less than 2 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°). While this is a greater degree of flexibility than that deduced from NMR relaxation data (cone of semi-angle ~30°),[S10] 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.[S11] Further, the tilt of our "major" state hemisphere from the dumbbell X-ray structure (~45°) is similar to the tilt determined from analysis of the PCS and RDC data (~30°)[S11,12]. 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^{2+} "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^{2+} 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*.[S9-13]

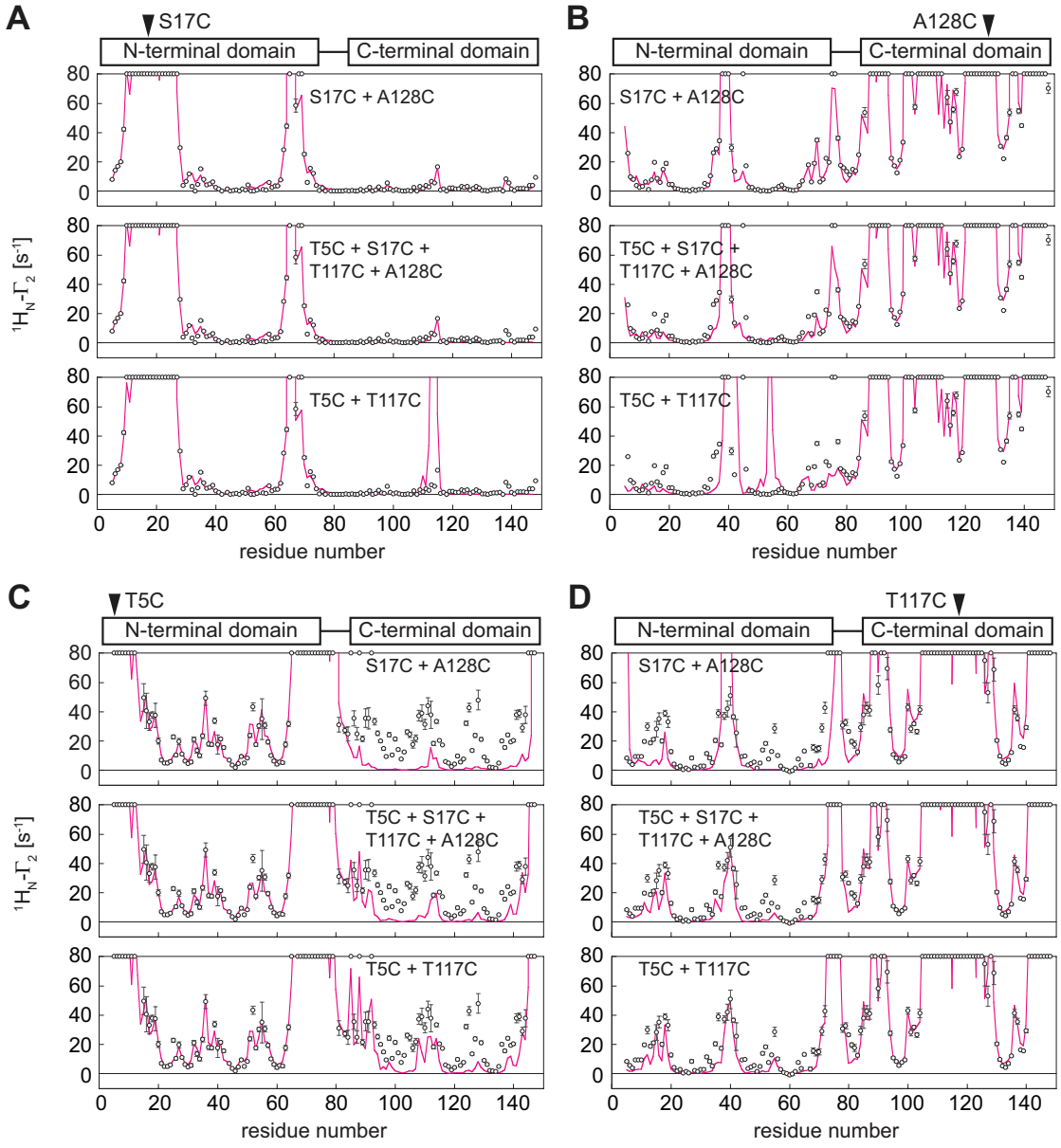

Fig. S9. PRE-driven minor state ensemble simulated annealing calculations for CaM-4Ca^{2+} using different data set combinations. (*A-D*) PRE profiles for (*A*) S17C, (*B*) A128C, (*C*) T5C, and (*D*) A128C resulting from simulated annealing calculations with a minor compact state population of 10% represented 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).

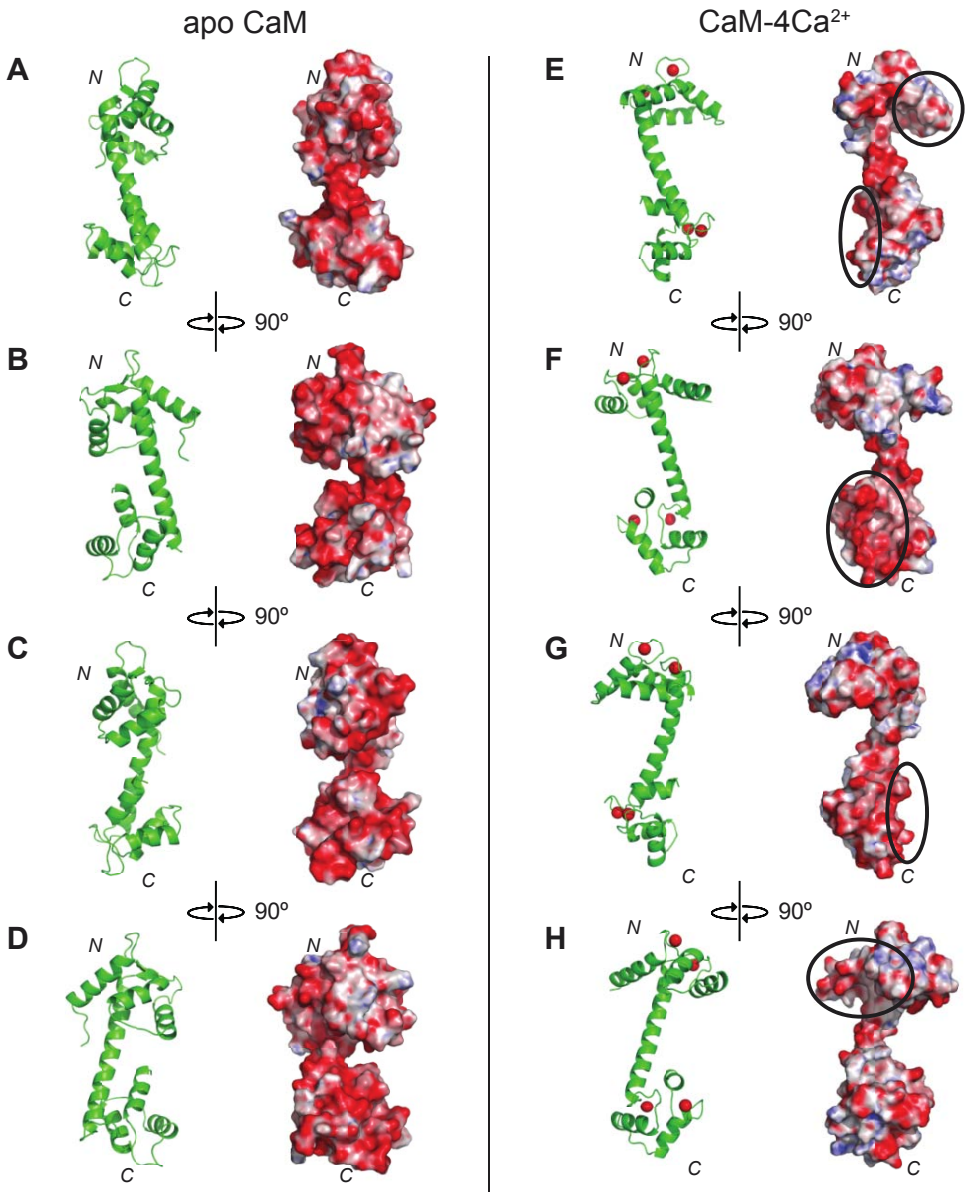

Fig. S10. Electrostatic potential surface maps for CaM. Electrostatic potentials were mapped onto structures of (*A-D*) apo CaM and (*E-H*) CaM-4Ca^{2+}. The top row (*A* and *E*) is in the same orientation as Fig. 5 so that the electrostatic potentials can be compared with the interdomain interactions 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^{2+} displays additional patches of positive charge. The interdomain interaction surfaces in CaM-4Ca^{2+} 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^{2+}, consistent with the observed negative salt dependence of the interdomain PREs (Fig. 3). Electrostatic potential surface maps were calculated and visualized using APBS[S14] and PyMOL (http://www.pymol.org),

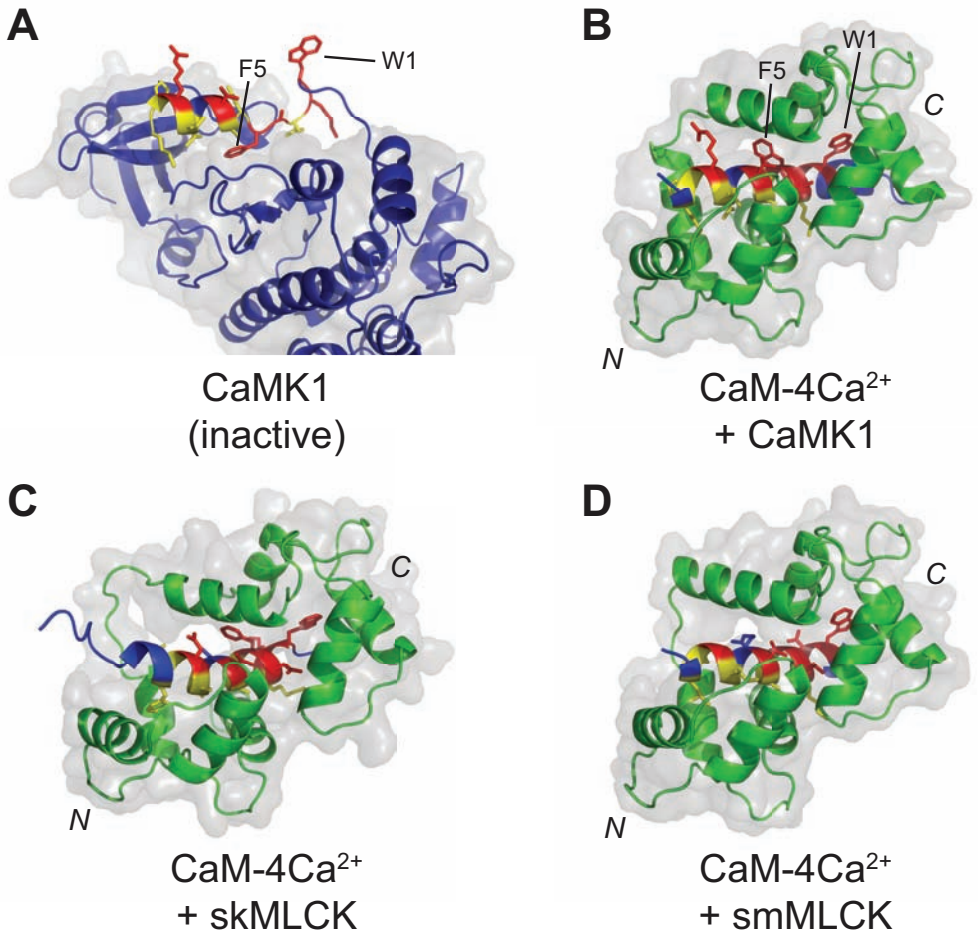

Fig. S11. Exposure of C-terminal binding residues in the CBD of the inactive kinase. Structure of the kinase CaM-binding domain (CBD), either (*A*) self-associated with the kinase active site or (*B-D*) bound to Ca^{2+}-loaded CaM. (*A*) X-ray structure of inactive CaMK1 (PDB 1A06);[S15] (*B*) X-ray structure of CaMK1 CBD peptide bound to Ca^{2+}-loaded CaM (PDB 1MXE);[S16] (*C*) NMR structure of skeletal muscle MLCK (skMLCK) CBD peptide (i.e. the peptide used in the experiments presented here) bound to Ca^{2+}-loaded CaM (PDB 2BBM);[S2] (*D*) X-ray structure of smooth muscle MLCK (smMLCK) CBD peptide bound to Ca^{2+}-loaded CaM (PDB 1CDL).[S3] 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^{2+} 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.[S17]

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.; Quiocho, 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.; Quiocho, F. A. *J. Mol. Biol.* **1992**, *228*, 1177-1192.
(S6) Fallon, J. L.; Quiocho, 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.