JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

The Length of the Calmodulin Linker Determines the Extent of Transient Interdomain Association and Target Affinity

Nicholas J. Anthis 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

Supporting Information

ABSTRACT: Calmodulin (CaM), the prototypical calcium sensing protein in eukaryotes, comprises two domains separated by a short flexible linker, which allows CaM to assume a wide range of extended and compact conformations. Here we use NMR relaxation measurements to explore the role of the linker in CaM function and dynamics. Using paramagnetic relaxation enhancement (PRE) measurements, we examine the effect of changes in the length and rigidity of the linker on the transient association between the two domains of Ca^{2+}-bound CaM (CaM-4Ca^{2+}). We observe that transient interdomain association, represented by an effective molarity (M_{eff}), is maximal for a linker extended by one residue from the wild-type length and decreases for lengths longer or shorter than that. The results can be quantitatively rationalized using a simplified model of a random coil whose two ends must be a specific distance apart for an interaction to occur. The results correlate well with the affinity of CaM-4Ca^{2+} for a target peptide, suggesting that the transient compact states adopted by CaM-4Ca^{2+} in the absence of peptide play a direct role in facilitating target binding.

Intrinsically unstructured regions are found throughout the proteome, where they perform myriad functions, including serving as flexible linkers between structured protein domains. The primary calcium-sensing protein in eukaryotes, calmodulin (CaM), comprises two domains connected by a flexible linker. Each domain contains two Ca^{2+}-binding sites and undergoes a local structural rearrangement upon binding Ca^{2+}, exposing a hydrophobic peptide-binding patch within each; this forms the basis of Ca^{2+} signaling within the cell.[1] CaM displays remarkable plasticity, including the adoption of a wide range of structures over the course of its biological cycle and the ability to bind hundreds of unique peptide sequences.[2] Ca^{2+}-bound CaM (CaM-4Ca^{2+}) adopts a rigid, compact structure when bound to target peptides,[3,4] but assumes a more flexible, extended conformation in the absence of peptide,[5] with its two domains tumbling semi-independently of one another.[6]

Although flexible linkers often comprise regions of lower sequence complexity and conservation than more rigid protein segments, the CaM linker is highly conserved, and its length is invariant (Supporting Information [SI] Figure S1). Thus, we sought to elucidate how the length and flexibility of the linker influence the biophysical and biological properties of CaM-4Ca^{2+}. We made a range of mutations in the central linker (SI Figure S2) to change its length, including deletion of the central residue (Thr79) to decrease its length by one residue ($\Delta n = -1$) and insertion of 1 to 8 glycine residues between Asp78 and Thr79 to lengthen it ($\Delta n = +1$ to $+8$). We also replaced the five residues of the linker (Lys77 to Ser81) by five alanines ("AAAAA")—to make a more rigid linker—or by five glycine residues ("GGGGG")—to replace the linker by a different flexible sequence. Measurements were also made on CaM with the wild-type (WT) linker, as well as on the isolated N-terminal (residues 1–76) and C-terminal (residues 81–148) domains.

We measured ^{15}N T_1 and T_2 values for several linker mutants and calculated an estimated effective correlation time (τ_c) for each domain (SI Table S1). As seen in Figure 1, τ_c for CaM-4Ca^{2+} displays a clear trend, decreasing with increasing linker length. The more rigid AAAAA mutant exhibits even higher τ_c values than the −1 mutant, while the GGGGG mutant exhibits τ_c values similar to the +1 mutant (SI Table S1). A full

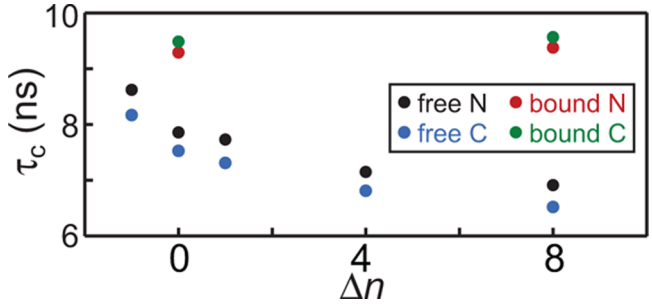

Figure 1. Dependence of CaM-4Ca^{2+} N- and C-terminal domain tumbling on the length of the linker connecting the two domains. The linker was lengthened by insertion of glycine residues and shortened by deletion of Thr79. The effective correlation time (τ_c), estimated from ^{15}N T_1/T_2, is plotted for different linker lengths ($\Delta n = 0$ for wild-type CaM). Values are shown for the N-terminal (black) and C-terminal (blue) domains of peptide-free CaM-4Ca^{2+} and for the N-terminal (red) and C-terminal (green) domains of MLCK-bound CaM-4Ca^{2+}. In the absence of peptide, the values of τ_c are inversely related to linker length, and the smaller C-terminal domain exhibits more rapid tumbling. When bound to peptide, however, τ_c is largely independent of linker length. The τ_c values of the isolated N- and C-terminal domains (not shown) are 4.5 and 3.9 ns, respectively. The uncertainties (±1 s.d.) in the values of τ_c lie within the circles and are given in Table S1 (SI).

Received: May 22, 2013
Published: June 19, 2013

dx.doi.org/10.1021/ja4051422 | J. Am. Chem. Soc. 2013, 135, 9648–9651

description of interdomain motion in CaM-4Ca^{2+} would require a more comprehensive extended model-free approach,[7] which is beyond the scope of the current work, although this has been done previously for WT CaM (apo[8] and Ca^{2+}-bound[9]). Our results show that the two domains of CaM-4Ca^{2+} tumble semi-independently of one another in solution (in agreement with previous studies[6,9]), and that as the length of the linker is increased the domains become less constrained and tumble more rapidly. The τ_c values of the isolated Ca^{2+}-bound domains, however, are significantly smaller than those of CaM-4Ca^{2+} with a +8 linker, indicating that the two domains are still constrained by this extended linker. This behavior disappears when CaM binds its target peptide from myosin light-chain kinase (MLCK), as the τ_c values of the CaM-4Ca^{2+}–peptide complex are similar for WT and +8 linkers. Thus, peptide-bound CaM-4Ca^{2+} forms a more rigid structure, with the two domains tumbling together as a single body.

We previously demonstrated that, in the absence of target peptide, CaM-4Ca^{2+} samples a wide range of conformational space, including a small population of states (5−10%) where the two domains are in close contact; some of these conformations resemble the peptide-bound structure and may play a direct role in target peptide binding.[10] Such sparsely populated states are invisible to conventional structural biology techniques, but they can be directly probed using paramagnetic relaxation enhancement (PRE).[11] Because CaM-4Ca^{2+} interdomain motion takes place in the fast exchange regime on the PRE time scale and because the distances between the domains are shorter in the minor species than in the major one, the footprint of these minor compact states are observed in the PRE profile of the major species of CaM-4Ca^{2+}.[10] The exquisite sensitivity of the PRE arises from the large magnetic moment of the unpaired electron and the $\langle r^{-6} \rangle$ dependence of the transverse PRE rate (Γ_2), resulting in extremely large Γ_2 values at short distances.[11]

Due to their unique ability to unlock these otherwise invisible states, we have used PREs as a direct measure of transient interdomain association in CaM-4Ca^{2+}. CaM was tagged with the paramagnetic nitroxide spin label MTSL at position 128 (as this position gave the largest interdomain PREs[10]), and PREs were measured for CaM-4Ca^{2+} with various linker mutations. PRE data for both interdomain PREs (originating from the tag in the C-terminal domain and observed on protons in the N-terminal domain) and intradomain PREs (originating from and measured on the C-terminal domain) were obtained. The intradomain PREs show the same trend as τ_c, increasing when the linker is shortened and decreasing when it is lengthened (Figure 2). In contrast, the interdomain PREs show a far more interesting trend. The interdomain PREs decrease when the linker is shortened by one residue and increase when the linker is lengthened by one residue; however, further increases in linker length show a general trend of decreasing interdomain PREs (Figure 2). (This trend is not completely smooth, as the +3 and +4 linkers give larger interdomain PREs than the +2 linker.) The equivalent PREs further decrease for a mixture of the isolated N- and C-terminal domains. The AAAAA mutation caused a marked decrease in interdomain PREs, indicating a more rigid linker, whereas the GGGGG mutation caused only a slight increase in interdomain PREs, probably because the WT linker already behaves as a random coil (SI Table S1).

To derive biophysical meaning from the above results, we converted the interdomain PREs into values of effective

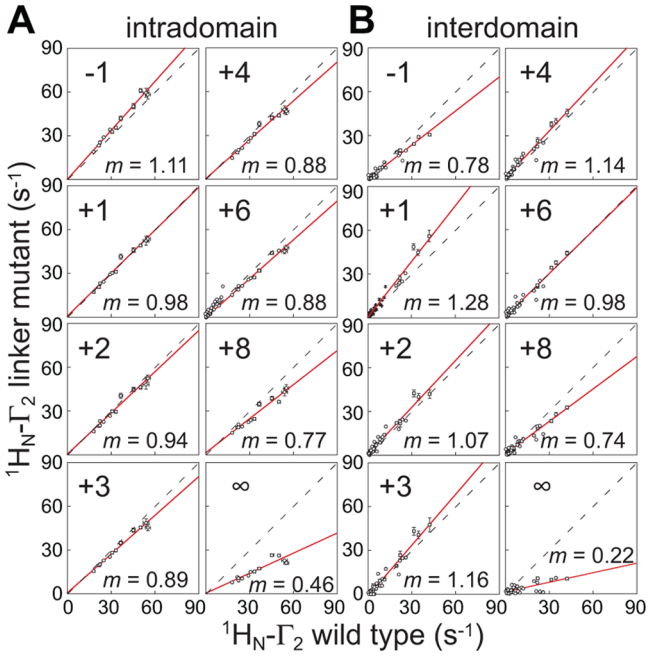

Figure 2. Dependence of CaM-4Ca^{2+} intra- and interdomain PREs on linker length. Experimental PREs (circles, error bars = 1 s.d.) were measured for CaM-4Ca^{2+} with the nitroxide spin label at A128C. (A) Intradomain PREs. (B) Interdomain PREs. The data are presented as scatter plots with the PRE value for CaM with the WT linker shown on the *x*-axis and the PRE value for CaM with various linker mutations shown on the *y*-axis. The infinity symbol (∞) indicates experiments performed on a mixture of the separate N- and C-domains without a linker (i.e., approximating CaM with an infinitely long linker). In each case, the slope of a best-fit line through the origin is given (red). The uncertainties in the values of the slope *m* are less than 7% in all cases except for those for the individual domains (SI Table S2). A dashed line corresponding to a slope of $m = 1$ is also displayed for reference. The magnitude of the intradomain PREs is inversely related to linker length, as longer linkers lead to faster tumbling domains (Figure 1). Interdomain PREs display a more complicated behavior.

molarity (M_{eff}) and compared these to values of effective concentration (conc$_{eff}$) calculated from a theoretical model. The full details of this procedure, which follows the framework described by Krishnamurthy et al.,[12] are described in SI. M_{eff} is a measure of the enhanced interdomain association brought about by connecting the two domains by a linker and is given by:

$$M_{eff} = M_{ind}(\Gamma_2^{tot}/\Gamma_2^{ind} - 1) \quad (1)$$

where Γ_2^{tot} is the total interdomain PRE (intermolecular + intramolecular) observed for CaM-4Ca^{2+} with a particular linker, and Γ_2^{ind} is the intermolecular PRE observed between the N- and C-terminal domains of CaM in different molecules. M_{ind} is the concentration of CaM (always 0.3 mM in this study). The PREs observed for CaM-4Ca^{2+} arise from both inter- and intramolecular interactions (SI Figure S3), which we assume are of the same nature on the basis of the correlation of the PREs (Figure 2 and SI Table S2).[10] By definition, M_{eff} follows the same trend as the interdomain PREs, first increasing to a maximum value of 1.47 mM at $\Delta n = +1$ and then gradually decreasing to 0 mM (for separate domains without a linker).

The empirical M_{eff} values were fit to an equation describing the effective concentration, conc$_{eff}$, of two points at the ends of

dx.doi.org/10.1021/ja4051422 | J. Am. Chem. Soc. 2013, 135, 9648−9651

a random coil, where an interaction only occurs when they are separated by a specific distance d (Figure 3 and SI Figure S4):[12]

$$\mathrm{conc}_{\mathrm{eff}} = \left(\frac{3}{2\pi}\right)^{3/2} \frac{p}{N_A(\langle r^2\rangle^{1/2})^3} \exp\left(\frac{-3d^2}{2(\langle r^2\rangle^{1/2})^2}\right) \quad (2)$$

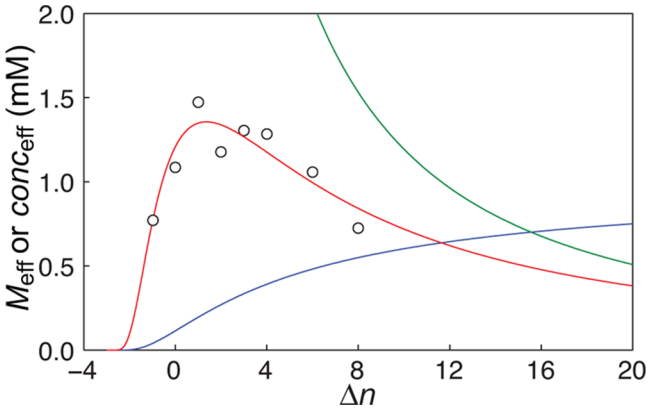

Figure 3. Dependence of CaM-4Ca^{2+} interdomain association on linker length can be largely explained by a simple random coil model. Empirical values of effective molarity (M_{eff}, circles) were calculated from interdomain PRE data for CaM-4Ca^{2+} with different linker lengths. These data were fit to eq 2, describing the theoretical values of effective concentration ($\mathrm{conc}_{\mathrm{eff}}$) of two points at the end of a random coil linker, where the interaction only occurs when the two points are separated by a defined distance. The best-fit line is shown in red (units of mM). A graph of the pre-exponential portion of the equation is shown in green (units of mM), and the exponential portion is shown in blue (dimensionless). The red line is the product of the green and blue lines.

where N_A is Avogadro's number and p is a constant that takes into account both the excluded volume occupied by the protein and the orientational requirements of the interaction. The root-mean-square distance (rmsd) between the ends of the random coil is given by:

$$\langle r^2\rangle^{1/2} = l\sqrt{C(n_0 + \Delta n)} \quad (3)$$

where l is the length of each unit of the chain (fixed to 3.8 Å for the length of the backbone of one amino acid residue). The number of units in the WT linker is given by n_0 (i.e., the linker of a given molecule contains $n = n_0 + \Delta n$ units). The characteristic ratio C is a measure of the rigidity of the linker and is fixed here to 2, the value for polyglycine.[13] The pre-exponential term of eq 2 (Figure 3, green line) describes the inverse cube dependence of $\mathrm{conc}_{\mathrm{eff}}$ on linker length. The exponential term (Figure 3, blue line) describes the requirement that the two ends of the random coil only interact at a distance d, causing a steep increase in $\mathrm{conc}_{\mathrm{eff}}$ as the linker length increases from zero, leveling off as it becomes much longer than d.

We obtained a reasonable fit of the PRE-derived M_{eff} data to eq 2 (Figure 3, red line) with $p = 0.016 \pm 0.002$, $d = 11.3 \pm 0.5$ Å, and $n_0 = 3.05 \pm 0.19$. The constant p acts primarily as a scaling factor, but its small value ($\ll 1$) likely represents the requirement of the two domains to be precisely oriented to interact ($p = 1$ if the ends of the linker are two discrete points with no orientational dependence).[12] The value of d represents the distance between the ends of the random-coil linker (and not, for example, the distance between the interacting domains). Since the value of d comes from a simplified model, one must be cautious about overinterpreting the result. However, the best-fit value of d is very similar to the distance of 11.5 Å between the backbone nitrogen atom of Lys77 and the carbonyl carbon atom of Ser81 (a reasonable interpretation for the two ends of the linker) in the NMR structure of CaM-4Ca^{2+} bound to the skeletal muscle MLCK peptide (PDB 2BBM),[3] although a shorter distance (7.3 Å) is observed in the crystal structure of CaM-4Ca^{2+} with the smooth muscle MLCK peptide (PDB 1CDL).[4] That n_0 was best fit to ~3, compared with the length of five residues observed by NMR,[6] could indicate that the linker is not as flexible as modeled (i.e., $C < 2$), that only the three central residues (Asp78, Thr79, Asp80) are totally flexible, or that additional factors are missing from the simple model. Regardless, the goodness of the fit and the reasonable values of the fitted parameters indicate that, despite its high level of simplicity, the random-coil model employed here describes the behavior of the CaM linker reasonably well.

To explore the role of transient CaM-4Ca^{2+} interdomain association further, we measured the affinity of wild-type and linker mutant CaM-4Ca^{2+} for the MLCK peptide using fluorescence anisotropy. These measurements were conducted under conditions of slightly higher temperature (37 °C vs 27 °C), lower pH (5.5 vs 6.5), and lower salt (0 mM vs 100 mM KCl) compared with the NMR conditions to lower the affinity from an equilibrium dissociation constant (K_D) of 10−30 pM (SI Table S1) to a K_D of ~400 pM (Table 1, SI Figure S5),

Table 1. Dependence of Target Peptide Affinity on the Length and Composition of the CaM Linker

CaM linker	K_D (pM)
−1	579 ± 27
0 (wild type)	426 ± 22
+1	320 ± 16
+8	746 ± 34
AAAAA	780 ± 37
GGGGG	387 ± 19

thereby permitting more accurate and reliable estimation of K_D values. The peptide contained an N-terminal 5-carboxy-X-rhodamine (5ROX) fluorescent label for greater sensitivity over intrinsic tryptophan fluorescence. Peptide affinity exhibited the same dependence on linker length as the interdomain PREs; the maximum affinity (smallest K_D) was measured for the +1 linker. Although the changes in affinity were small, they agree well with the interdomain PRE results.

The decrease in peptide affinity upon shortening the CaM linker agrees with previous findings. In one study,[14] the activity of CaM-dependent enzymes was tested in the presence of CaM-4Ca^{2+} with deletions of Glu84 ($\Delta n = -1$), Glu83-Glu84 (−2), and Ser81-Glu84 (−4) (deletions adjacent to, but not within, the flexible linker). These mutations caused an increase in the Michaelis−Menten constant K_M (i.e., a decrease in CaM activity) for the three enzyme variants tested; the effect ranged from 2- to 7-fold. In another study,[15] deletions of 2 or 3 residues within the flexible linker increased the K_M of some enzymes 2−18-fold, while another was unaffected. Larger deletions of 5 and 8 residues caused much larger increases in K_M, as well as significant decreases in maximum activity.

Our results, however, go further, demonstrating biphasic behavior. A small lengthening of the linker increases transient interdomain association, but this interaction progressively

dx.doi.org/10.1021/ja4051422 | J. Am. Chem. Soc. 2013, 135, 9648−9651

decreases as the linker is then further extended. Such behavior has been observed and modeled in other systems,[12,16,17] and here the overall trend can be explained by a simple random-coil model (Figure 3). More pertinently, the agreement shown here between transient interdomain association and target affinity suggests that compact sparsely populated states of CaM-4Ca^{2+} play a direct role in facilitating target binding.

■ ASSOCIATED CONTENT

Supporting Information

Experimental procedures, supplementary tables, and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

mariusc@mail.nih.gov

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Ad Bax for useful discussions. 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) Zhang, M.; Tanaka, T.; Ikura, M. *Nat. Struct. Biol.* **1995**, *2*, 758−767.
(2) Yamniuk, A. P.; Vogel, H. J. *Mol. Biotechnol.* **2004**, *27*, 33−57.
(3) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax, A. *Science* **1992**, *256*, 632−638.
(4) Meador, W. E.; Means, A. R.; Quiocho, F. A. *Science* **1992**, *257*, 1251−1255.
(5) Babu, Y. S.; Sack, J. S.; Greenhough, T. J.; Bugg, C. E.; Means, A. R.; Cook, W. J. *Nature* **1985**, *315*, 37−40.
(6) Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A. *Biochemistry* **1992**, *31*, 5269−5278.
(7) Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1990**, *112*, 4989−4991.
(8) Tjandra, N.; Kuboniwa, H.; Ren, H.; Bax, A. *Eur. J. Biochem.* **1995**, *230*, 1014−1024.
(9) Baber, J. L.; Szabo, A.; Tjandra, N. *J. Am. Chem. Soc.* **2001**, *123*, 3953−3959.
(10) Anthis, N. J.; Doucleff, M.; Clore, G. M. *J. Am. Chem. Soc.* **2011**, *133*, 18966−18974.
(11) Clore, G. M.; Iwahara, J. *Chem. Rev.* **2009**, *109*, 4108−4139.
(12) Krishnamurthy, V. M.; Semetey, V.; Bracher, P. J.; Shen, N.; Whitesides, G. M. *J. Am. Chem. Soc.* **2007**, *129*, 1312−1320.
(13) Cantor, C. R.; Paul, R. S. *Biophysical Chemistry: Part III: The Behavior of Biological Macromolecules*; W. H. Freeman: San Francisco, 1980.
(14) Persechini, A.; Blumenthal, D. K.; Jarrett, H. W.; Klee, C. B.; Hardy, D. O.; Kretsinger, R. H. *J. Biol. Chem.* **1989**, *264*, 8052−8.
(15) VanBerkum, M. F.; George, S. E.; Means, A. R. *J. Biol. Chem.* **1990**, *265*, 3750−3756.
(16) Robinson, C. R.; Sauer, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5929−5934.
(17) Shewmake, T. A.; Solis, F. J.; Gillies, R. J.; Caplan, M. R. *Biomacromolecules* **2008**, *9*, 3057−3064.

Supporting Information

The length of the calmodulin linker determines the extent of transient interdomain association and target affinity

Nicholas J. Anthis and G. Marius Clore*

Laboratories of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, USA

Experimental Procedures

Protein production, paramagnetic tagging, mutagenesis, and sample preparation

The full-length 148-residue human calmodulin (CaM) protein, and its N- (residues 1-76) and C- (residues 81-148) terminal domains were expressed and purified as described previously for full-length CaM.[S1] Proteins were either uniformly $^{2}H/^{13}C/^{15}N$-labeled or at natural isotopic abundance. After purification, CaM was exchanged into the "NMR buffer", which consisted of 5% D_2O/95% H_2O, 8 mM $CaCl_2$, 100 mM KCl, 1× Roche cOmplete Protease Inhibitor, 0.02% sodium azide, and 25 mM HEPES, pH 6.5. Samples were concentrated to 0.3 mM with Amicon Ultra Centrifugal Filter Units (3 kDa molecular weight cutoff). All NMR experiments were conducted in "NMR buffer". Fluorescence experiments were conducted in the "fluorescence buffer", which consisted of 100% H_2O, 3 mM $CaCl_2$, and 10 mM MES, pH 5.5. For fluorescence experiments, concentrated stocks of CaM in NMR buffer were diluted into the "fluorescence buffer". Concentrations were determined by absorbance at 280 nm (for CaM 1-148 and 81-148, ε_{280} = 2,980), except for CaM 1-76, which contains no groups that absorb at 280 nm. For the latter construct, concentrations mere measured by absorbance at 205 nm, using a molar absorptivity (extinction coefficient) of ε_{205} = 266,150 $M^{-1}\cdot cm^{-1}$, which was calculated from the amino acid sequence as previously described, using the server at http://spin.niddk.nih.gov/clore.[S2]

All mutations (including those in the CaM linker and the A128C mutation in the C-terminal domain for spin labeling) were made using the QuikChange Site-Directed Mutagenesis Kit (Agilent). CaM A128C 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 hours and tested for completion by mass spectrometry. 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) as CaM was exchanged into "NMR buffer".

For NMR experiments on peptide-bound CaM (CaM-4Ca^{2+}-MLCK), CaM was mixed with 1.2 equivalents of target peptide (skeletal muscle myosin light chain kinase, skMLCK, M13 peptide; commercially synthesized by Anaspec; KRRWKKNFIAVSAANRFKKISSSGAL). Fluorescence experiments were conducted on an M13 peptide with a 5-carboxy-X-rhodamine (5ROX) label conjugated to its N-terminus (custom-synthesized by Anaspec).

NMR spectroscopy

NMR experiments were performed on uniformly $^{2}H/^{13}C/^{15}N$-labeled CaM-4Ca^{2+}, including wild-type (WT) and mutant constructs. Data were recorded at 27°C on a Bruker 600 MHz spectrometer equipped with a triple resonance *z*-gradient cryoprobe. Data were processed using

NMRPipe[S3] and analyzed with the programs XIPP (in-house software written by D.S. Garrett) or Sparky (www.cgl.ucsf.edu/home/sparky).

Transverse $^{1}H_N$-Γ_2 PRE rates were obtained from the differences in the transverse $^{1}H_N$-R_2 relaxation rates between the paramagnetic and diamagnetic samples[S4] using a 2D TROSY pulse scheme with a variable delay performed at the end of the pulse sequence, thus measuring the TROSY component of $^{1}H_N$-R_2.[S1] Two time points (separated by 20 ms) were used for the $^{1}H_N$-R_2 measurements, and the errors in the $^{1}H_N$-Γ_2 PRE rates were calculated from spectral noise as described previously.[S5] Samples for PRE experiments contained 0.3 mM $^{2}H/^{13}C/^{15}N$-labeled CaM-4Ca^{2+}, tagged with a spin labeled or diamagnetic control tag at A128C. Experiments were also performed to measure intermolecular PREs; these measurements were made on 0.3 mM $^{2}H/^{13}C/^{15}N$-labeled CaM (1-76 or 1-148) with no attached spin label in the presence of 0.3 mM CaM (81-148 or 1-148, respectively) at natural isotopic abundance, spin-labeled at A128C. In this manuscript, the PREs for a given molecule are reported as the ratio of the PREs for that molecule to those of the WT. This ratio was determined by plotting the PREs of the molecule in question on the *y*-axis, plotting the PREs of the WT on the *x*-axis, and taking the slope of a best-fit straight line through the data.

^{15}N T_1 and $T_{1\rho}$ were measured with a TROSY readout using the pulse sequences described by Lakomek et al.[S6] on 0.3 mM $^{2}H/^{13}C/^{15}N$-labeled CaM-4Ca^{2+} (tagged at A128C with the diamagnetic control label, unless otherwise stated). Experiments were performed with 8 time points spanning from 0 to ~1.3 times T_1 or $T_{1\rho}$. T_1 and $T_{1\rho}$ values were calculated for each residue by fitting peak intensities to a single exponential decay. T_2 values were calculated from T_1 and $T_{1\rho}$ using the equation:

$$T_2 = \sin^2\theta / [(1/T_{1\rho}) - (1/T_1)\cos^2\theta] \qquad \text{(S1)}$$

where the angle θ is given by:

$$\theta = \tan^{-1}(\omega_1 / \delta\omega) \qquad \text{(S2)}$$

where ω_1 is the strength of the spin-lock field (1 kHz in this study) and $\delta\omega$ is the difference in frequency (in the ^{15}N dimension, in Hz) between the peak and the transmitter frequency. Overall rotational correlation times (τ_c) were estimated using the equation:

$$\tau_C = (1/2\omega_N)\sqrt{(6T_1/T_2 - 7)} \qquad \text{(S3)}$$

where τ_c is in units of seconds and ω_N is the ^{15}N frequency in radians·s^{-1}. The value of T_1/T_2 reported for each domain in Table S1 is the average for residues that are not in flexible loops or linkers and that do not exhibit T_1 or T_2 values outside of 1 standard deviation of the average for each domain. The value of τ_c for each domain was determined by optimizing a single value of τ_c that best fits all selected T_1/T_2 ratios within a given domain using Eq. S3. The uncertainties in the optimized values of τ_c are given in Table S1.

Fitting CaM linker data to a random coil model

To advance our understanding of the role of the CaM linker, a theoretical framework was applied to the analysis of the interdomain PRE data, interpreting this data through values of effective molarity (M_{eff}) and effective concentration ($conc_{eff}$). Specifically, we applied a framework based on a random coil model described by Krishnamurthy et al.[S7] To accomplish this, we converted the experimental interdomain PRE data for each linker mutant into values of M_{eff}. For each mutant, we took the ratio of the measured interdomain PREs of the mutant to those of WT CaM (i.e. the slope of the red lines in Figure 2); in this study, the term "PREs" generally refers to this ratio as opposed to the raw PRE data. We describe linker length using the variable Δn, which corresponds to the number of residues added (positive) or removed (negative) from the

WT linker (Δn = 0 for the WT linker). M_{eff} relates the strength of the intramolecular interaction for a given linker to that of the intermolecular interaction and is defined by the ratio of equilibrium constants:[S7]

$$M_{\text{eff}} = K_{\text{D}}^{\text{intermol}} / K_{\text{D}}^{\text{intramol}} \tag{S4}$$

$K_{\text{D}}^{\text{intermol}}$ is the bimolecular equilibrium constant for the intermolecular interaction, given in concentration units:

$$K_{\text{D}}^{\text{intermol}} = \frac{[\text{CaM}_{\text{N}}]_{\text{free}}[\text{CaM}_{\text{C}}]_{\text{free}}}{[\text{CaM}_{\text{N}}\text{CaM}_{\text{C}}]_{\text{associated}}} \tag{S5}$$

$K_{\text{D}}^{\text{intramol}}$ is the unimolecular equilibrium constant for the intramolecular interaction, which is dimensionless:

$$K_{\text{D}}^{\text{intramol}} = \frac{[\text{CaM}]_{\text{extended}}}{[\text{CaM}]_{\text{associated}}} \tag{S6}$$

Here, we consider three different populations of CaM-4Ca^{2+} in the absence of peptide (Figure S3). Population 1, the major population, comprises the conformations where the two domains are tumbling essentially independently of one another (and bound only by the linker, if present). Population 2 is a minor state that comprises the two linked domains interacting with one another. For the WT linker, we previously demonstrated that population 2 comprises 5-10% of the total population.[S1] The PREs from population 2 (interdomain, intramolecular) are denoted as $\Gamma_2^{\text{intramol}}$. Population 3 represents an even smaller population (~1-2%) of domains interacting intermolecularly. The PREs from Population 3 (interdomain, intermolecular) are denoted as $\Gamma_2^{\text{intermol}}$. Because the interdomain PREs in CaM display a similar profile to the intermolecular PREs (Figure 2), [S1] we can make the assumption that they are both due to the same type of interaction between an N- and C-terminal domain. Because these processes take place on a timescale that is fast on the PRE time scale, the total PRE (Γ_2^{tot}) for a given nucleus in the N-terminal domain when the C-terminal domain is tagged with a paramagnetic group is the sum of the intramolecular interdomain PREs and the intermolecular PREs:

$$\Gamma_2^{\text{tot}} = \Gamma_2^{\text{intramol}} + \Gamma_2^{\text{intermol}} \tag{S7}$$

Γ_2^{tot} and $\Gamma_2^{\text{intramol}}$ are functions of Δn. The value of $\Gamma_2^{\text{intermol}}$, which is derived in this study from the interdomain PREs measured between isolated domains, is constant and does not depend on Δn. This trend will be the same for other variables and constants in this study; all "intramol" and "tot" variables are dependent on Δn, whereas "intermol" will always signify a constant that is not affected by Δn. Γ_2^{tot} and $\Gamma_2^{\text{intermol}}$ are both dependent on the concentration of the spin-labeled domain (M_{ind}), which in this study is kept constant at 0.3 mM to simplify analysis and make all results directly comparable. $\Gamma_2^{\text{intramol}}$, on the other hand, is independent of protein concentration. We make the assumption that the populations of the transiently associated states are small enough that they do not directly compete with one another. Thus, for a given linker, the ratio of the intramolecular PREs to the universal intermolecular PREs is equal to the ratio of M_{eff} to M_{ind}:

$$\frac{\Gamma_2^{\text{intramol}}}{\Gamma_2^{\text{intermol}}} = \frac{M_{\text{eff}}}{M_{\text{ind}}} \tag{S8}$$

Combining Eq. S7 and S8, we can calculate M_{eff} for any CaM linker mutant as:

$$M_{\text{eff}} = M_{\text{ind}}\left(\Gamma_2^{\text{tot}}/\Gamma_2^{\text{ind}} - 1\right) \tag{S9}$$

Eq. S9 is presented in the main text as Eq. 1. These empirical values of M_{eff} can then be compared to theoretical values of effective concentration, $conc_{eff}$, calculated using a random coil model for the CaM linker (Figure S4); $conc_{eff}$ is a theoretical term related to the probability of the two ends of a molecule coming together.[S7] In this study, M_{eff} and $conc_{eff}$ are treated as functionally equivalent, with the distinction being that M_{eff} is an empirical observation and $conc_{eff}$ is a theoretical value calculated from a random coil model. Here we use a random coil model described by Krishnamurthy et al.[S7] for the $conc_{eff}$ of two points at opposite ends of random-coil linker that only interact when located at a defined distance d from each other. An illustration of this model is shown in Figure S4A. The alternative representation in Figure S4B shows how this model qualitatively relates to CaM. As this model here is intended to be a simplistic representation of CaM and its flexible linker, the values of the fitted parameters should be taken at face value. Using this model, $conc_{eff}$ is given by the following equation, which is also presented as Eq. 2 in the main text:

$$conc_{\mathrm{eff}} = \left(\frac{3}{2\pi}\right)^{3/2} \frac{p}{N_{\mathrm{A}}\left(< r^2 >^{1/2}\right)^3} \exp\left(\frac{-3d^2}{2\left(< r^2 >^{1/2}\right)^2}\right) \qquad \text{(S10)}$$

where N_A is Avogadro's number (6.0221×10^{23} mol^{-1}), and p is a dimensionless constant that takes into account both the excluded volume occupied by the protein domain and the orientational requirement of the interdomain interaction. The root-mean-square distance (rmsd) between ends of the random coil is given by:

$$< r^2 >^{1/2} = l\sqrt{C(n_0 + \Delta n)} \qquad \text{(S11)}$$

where l is the length of each unit of the chain (here fixed at 3.8 Å for the average length of the backbone of one amino acid residue). The number of units in the WT linker is given by n_0 (i.e. the linker of a given molecule contains $n = n_0 + \Delta n$ units). The constant C is the characteristic ratio of a random-coil polymer:

$$C = < r^2 > / nl^2 \qquad \text{(S12)}$$

C is described in detail by Cantor and Schimmel.[S8] In brief, C is a measure of the stiffness of the polymer, and $C = 1$ for a totally unconstrained random-coil polymer. In a biological polymer, however, C is greater than 1 and is dependent on n for small values of n, reaching an n-independent limit of C_∞ at large values of n. A polyglycine chain reaches C_∞ = 1.9-2.1 at $n \sim 6$. A polyalanine chain, however, exhibits a larger value of C for all values of n, but only approaches C_∞ = 9.0 near $n \sim 80$. For our purposes here, we used a fixed value of $C = 2$, roughly corresponding to the value expected for a polyglycine chain. (For example, a value of $C = 1.9$ is expected for a polyglycine chain of 5 units; even for a much stiffer polyalanine chain of 5 residues, C is only increased to 3.4.) Thus, $C = 2$ was deemed an appropriate value for this system.

The pre-exponential portion of Eq. S10 describes the inverse cube relationship between $conc_{eff}$ and the rmsd between the two ends of the linker. This is strictly based on the three-dimensional space sampled by the two ends of the linker, and it predicts a decreasing $conc_{eff}$ as the linker is lengthened. Depending on the desired units of concentration and the units of length used, a conversion factor may need to be included. The exponential portion of Eq. S10, on the other hand, describes the requirement of the interaction on the two ends of the linker being a specific distance d apart. Note that this distance d is not the distance between the surfaces of the two domains or the centers of mass of the two domains, but the distance between the ends of the linker when the two domains are in contact.

Eq. S11 was fit to the empirical data (M_{eff} vs. Δn), keeping l and C constant, and optimizing the values of d, p, and n_0 (Figure 3; fit in red, pre-exponential portion in green, exponential portion in blue). The fit was performed by minimizing the difference between the calculated $conc_{eff}$ values and the empirical M_{eff} values for each value of Δn. This fit was performed in MATLAB, and errors of the fitted parameters were estimated using Monte Carlo analysis, assuming an error in M_{eff} of 10%.

Fluorescence experiments

Fluorescence experiments were carried out at 37°C using a Jobin Ybon FluoroMax-3 fluorometer equipped with a Peltier temperature control unit. A 1 cm × 1 cm quartz cuvette was used (Starna Cells, Inc.). The fluorescence anisotropy of 5ROX-MLCK was monitored with excitation at 583 nm and emission at 607 nm. Measurements were acquired on 1 nM 5ROX-MLCK in the presence of 0-1 nM CaM 1-148 or 0-1.3 μM CaM 1-76 + CaM 81-148 (for the separated domains, a stock containing an equal concentration of the two domains was made and titrated into the fluorescent peptide).

Data were analyzed by fitting the fluorescence anisotropy versus CaM concentration to the following equation:

$$A = A_{min} + A_{max}\left(\frac{[F]_{tot} + [U]_{tot} + K_D - \sqrt{([F]_{tot} + [U]_{tot} + K_D)^2 - 4[F]_{tot}[U]_{tot}}}{2[F]_{tot}}\right) \quad \text{(S13)}$$

where A is the measured fluorescence anisotropy at each point (the dependent variable in the fitting); A_{max} the anisotropy of the MLCK peptide fully saturated with CaM; A_{min} the anisotropy of the free 5ROX-MLCK peptide; $[F]_{tot}$ the total (bound + unbound) concentration of the MLCK peptide ("F" for "fluorescent"); $[U]_{tot}$ the total (bound + unbound) concentration of CaM at each point (the independent variable in the fitting; "U" for "unseen"); and K_D the equilibrium dissociation constant, defined as:

$$K_D = \frac{[U][F]}{[UF]} \quad \text{(S14)}$$

Although this is the true K_D in most cases, for the titration of the MLCK peptide with the two separated domains this is an apparent K_D describing the concentration at half saturation. Data were fit using OriginPro 8.

The conditions used for the fluorescence experiments (low salt, low pH, high temperature) were chosen to weaken the CaM/target interaction to assist in accurate affinity measurement. Under the conditions used for NMR, the K_D of WT CaM-4Ca^{2+} for 5ROX-M13 is ~ 10-30 pM (Table S1), but it was difficult to determine an accurate value within the sensitivity limits of the experiment. The temperature was increased from 27ºC to 37ºC, which, due to the fact that the interaction is highly exothermic,[S1] was expected to decrease the affinity by a factor of ~2. In addition, based on results found in the literature on similar peptides,[S9,10] we predicted that lowering the pH from 6.5 to 5.5 and decreasing the ionic strength from [KCl] = 200 mM to 100 mM would decrease the affinity by approximately 10 to 100 fold. The ~20-to-40-fold reduction in affinity (to K_D ~ 400 pM) observed under these conditions was thus within the expected range. Despite the achieved decrease in affinity, low concentrations of CaM and 5ROX-M13 were still required. To prevent protein or peptide adsorption in the dilute samples, all dilutions were performed in Eppendorf Protein LoBind Tubes.

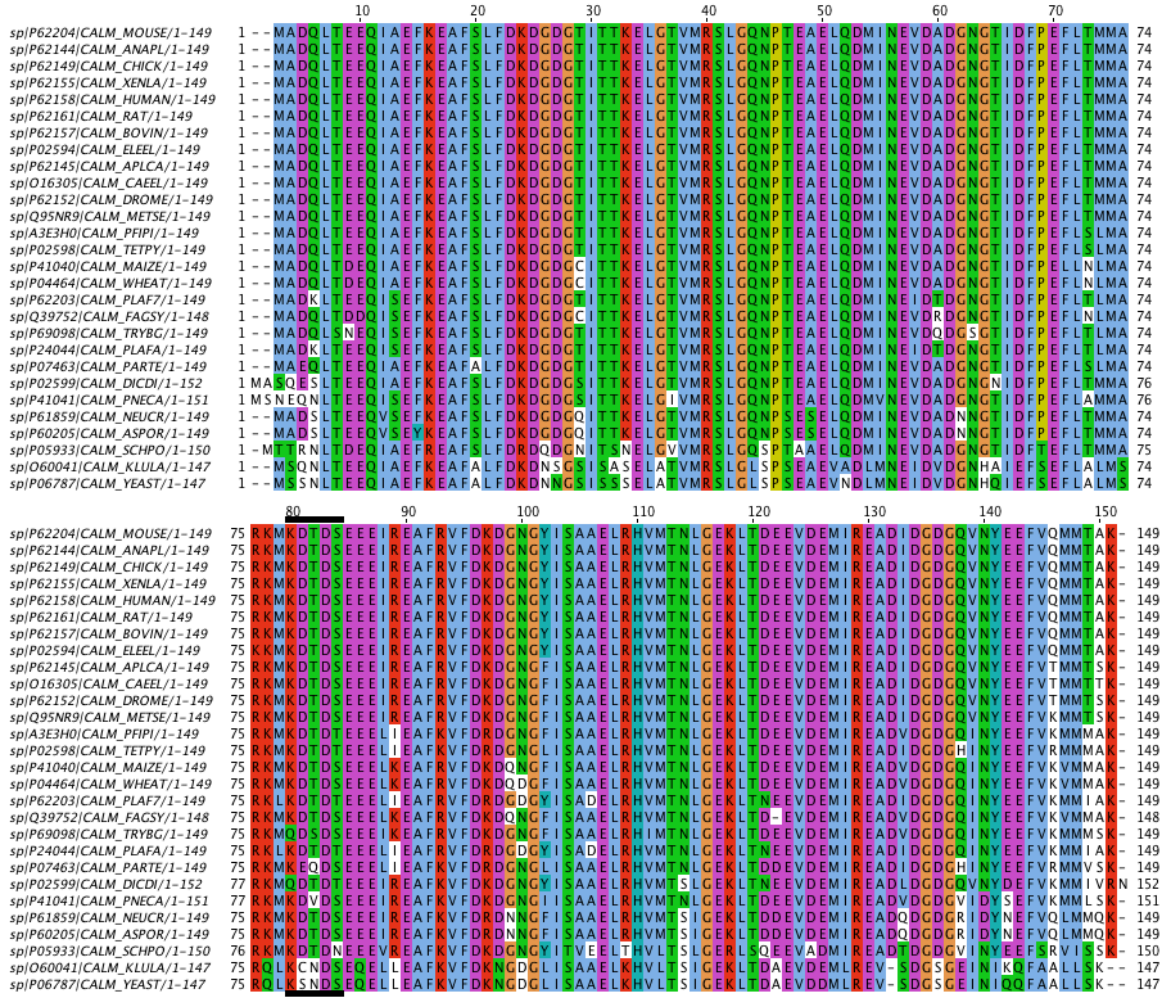

Figure S1. Alignment of a selection of calmodulin (CaM) sequences. Amino acid sequences of CaM from an assortment of organisms were aligned using the Clustal Omega[S11] web server on the EMBL-EBI website.[S12] Sequences were sorted and displayed in Jalview.[S13] Vertebrate sequences are located at the top and yeast sequences at the bottom. Others are found in between. Sequences are colored by amino acid type and identity. The UniProt[S14] accession number and entry name is given for each. The full UniProt entry is displayed for each, including the N-terminal methionine, which does not appear in the final processed protein sequence; thus, the numbering differs from that used in the rest of this study (and the numbers at the top correspond to the numbering of the longest sequence). The five-residue flexible linker is indicated by a black bar (beginning with number 80, but corresponding to residues 77-81 in the final processed human sequence). Note that the sequence of CaM is highly conserved, including that of the linker.

Figure S2. Mutations in the CaM linker used in this study. (A) Deletion and insertion mutations used to change the length of the CaM linker. (B) Other mutations used to alter the sequence and/or properties of the CaM linker. Additionally, the C-terminus of the N-terminal domain construct and the N-terminus of the C-terminal domain construct are shown. In both panels, CaM residues 70-90 are shown, and residues 77-81 (the flexible linker) are highlighted in red.

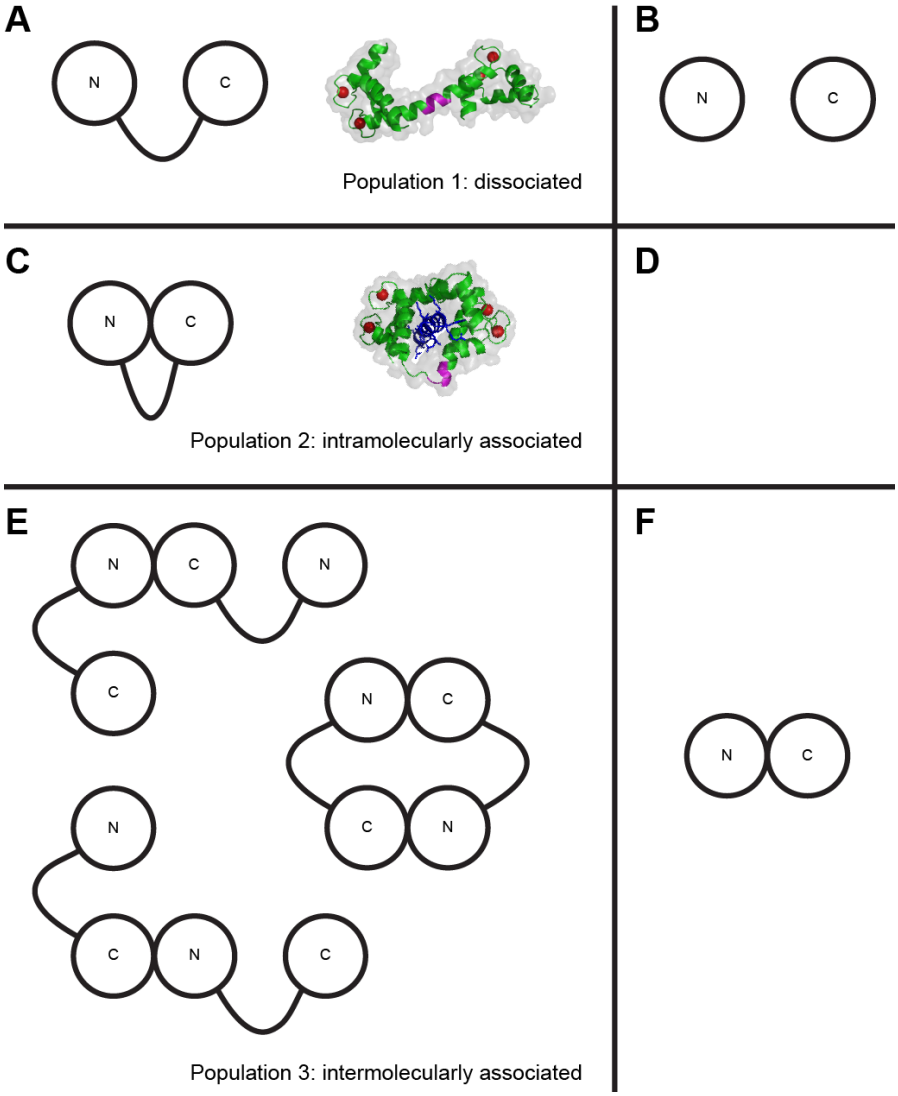

Figure S3. Summary of the different populations of CaM-4Ca^{2+}. CaM-4Ca^{2+} exists as an ensemble of states in equilibrium between three different populations. The panels on the left (A, C, E) show the populations for two CaM domains connected by a flexible linker (i.e. the natural state of CaM), and the panels on the right (B, D, F) show the populations for the isolated domains of CaM when the linker has been removed (i.e. an artificial situation). In population 1 (A, B), which is the major state, the two domains are dissociated from one another. In population 2 (C, D) and 3 (E, F), the N- and C-terminal domains are associated with each other. These are sparsely populated minor states (~5-10% for population 2; ~1-2% for population 3). Panel D is empty because intramolecular association is not possible for isolated domains. Although Panels C, E, and F, show different types of association between domains, it is assumed in this study that the interface between the N- and C-terminal domains is the same in each case, based on our previous results[S1] and the high correlation between intra- and intermolecular PREs (Figure 2), although panel E indicates that the intermolecular complex could take various forms. The extended dumbbell structure of CaM-4Ca^{2+} (PDB 1CLL)[S15] is shown in panel A, as a representation of one of the conformations that might be sampled in population 1. The compact structure of CaM-4Ca^{2+}-MLCK (PDB 1CDL)[S16] is shown in panel C, as an approximation of the interdomain conformation in population 2. The flexible linker is highlighted in magenta.

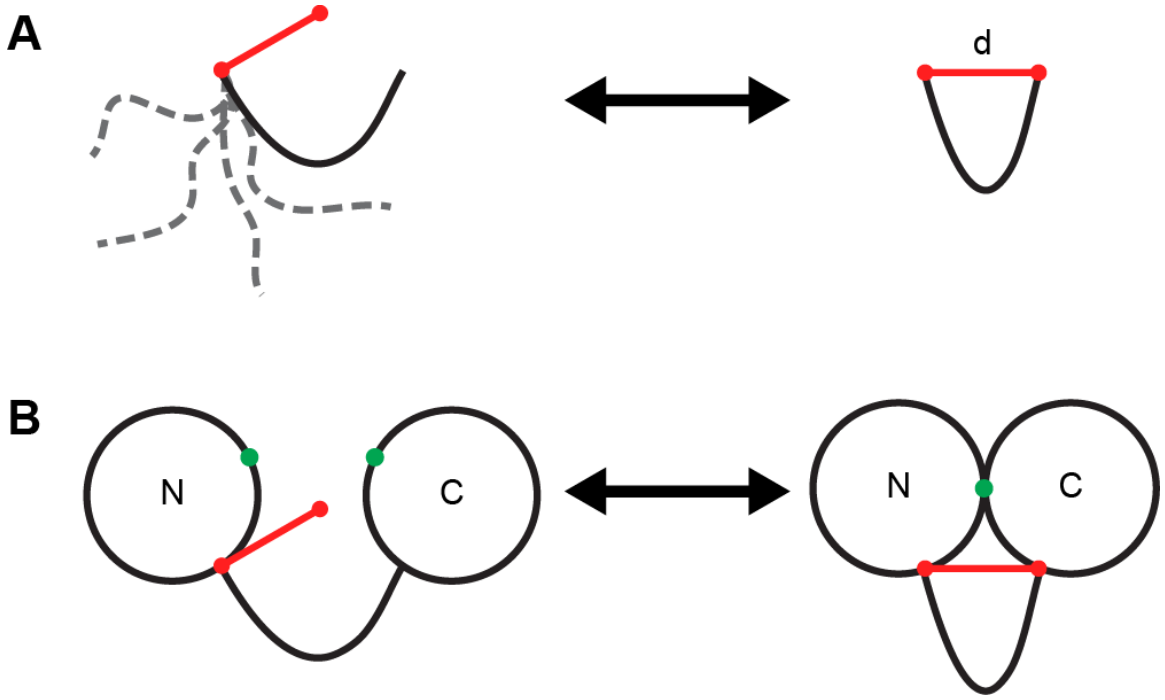

Figure S4. Illustration of the random-coil model used to model the CaM linker. (A) A random-coil chain (black) forms an interaction only when its two ends are located a certain distance (*d*, red) apart. The dotted grey lines represent the flexibility of the random-coil chain (left), which becomes restrained when forming its interaction (right). This model is the basis of Eq. 2 and Eq. S10, as described by Krishnamurthy et al.[S7] (B) A more realistic representation of CaM-4Ca^{2+} with its domains in the dissociated (left) and associated (right) states. The interaction surface is denoted by a green circle. The qualitative similarities between panels A and B support the use of the simple random-coil model to explain the behavior of the CaM linker.

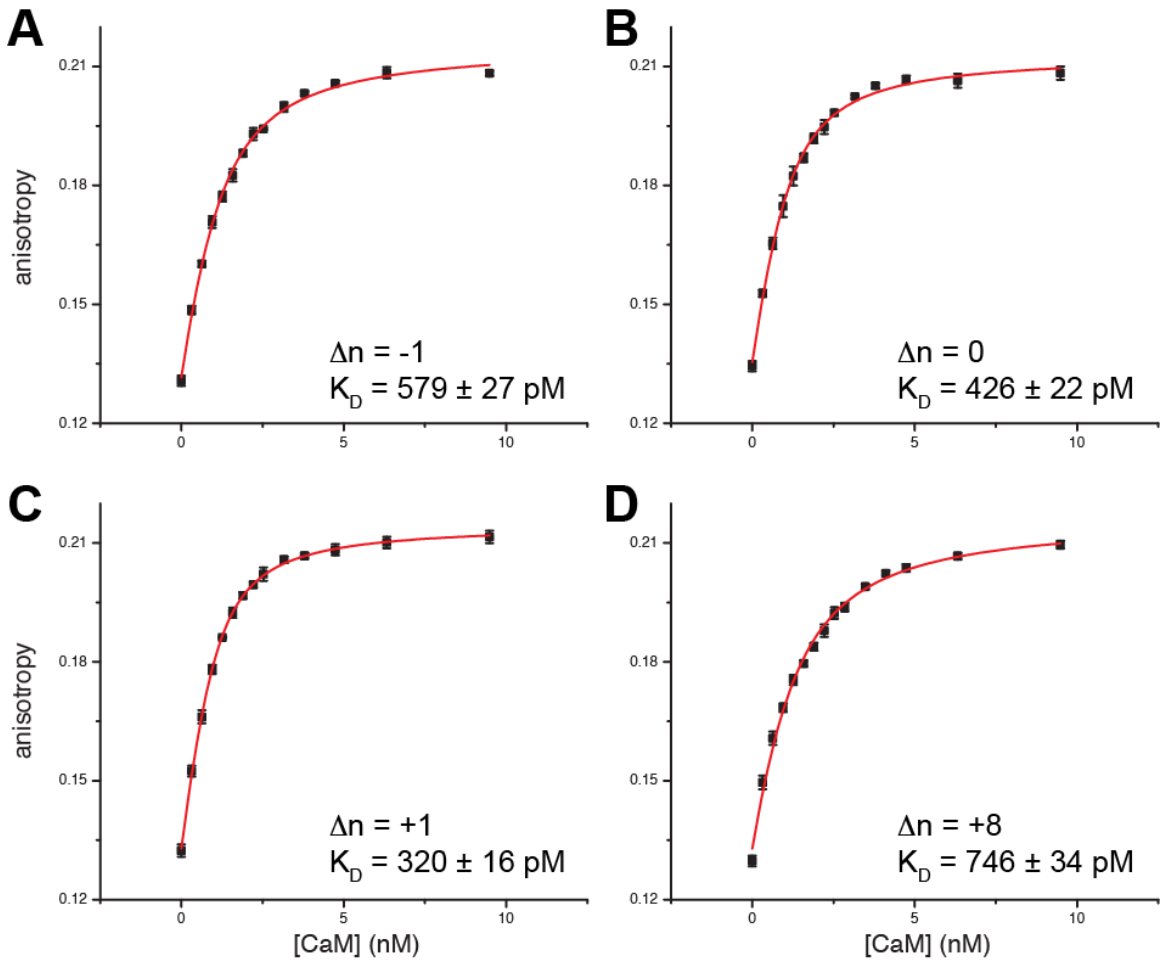

Figure S5. Sample fluorescence anisotropy titration curves. Fluorescence anisotropy was measured for 5ROX-M13 alone (1 nM) and in the presence of 0-10 nM CaM-4Ca^{2+}. Experimental data (average of three measurements) are plotted as filled-in squares, with error bars indicating one standard deviation. The best-fit line is shown in red. The linker length mutation (Δn) and K_D are indicated. Data are shown for linkers with Δn values of (A) -1, (B) 0, (C) +1, and (D) +8.

Table S1. Summary of experimental data

linker, condition[a]	^1H_N PRE		M_{eff} (mM)[d]	^{15}N relaxation								K_D (pM)[i]	
				T_1/T_2[e]				τ_c[h]					
	intradom[b]	interdom.[c]		free[f]		bound[g]		free		bound		low affinity[j]	high affinity[k]
				N-term	C-term	N-term	C-term	N-term	C-term	N-term	C-term		
-1	1.11	0.78	0.78	8.43 ± 1.04	7.71 ± 1.30			8.62 ± 0.05	8.17 ± 0.06			579 ± 27	
WT[l]	1[q]	1[q]	1.09	7.19 ± 0.68	6.72 ± 0.87	9.60 ± 0.64	9.95 ± 0.77	7.86 ± 0.04	7.53 ± 0.07	9.29 ± 0.07	9.49 ± 0.06	426 ± 22	13.6 ± 3.6
+1	0.98	1.28	1.47	7.00 ± 0.57	6.38 ± 0.52			7.73 ± 0.04	7.31 ± 0.07			320 ± 16	
+2	0.94	1.07	1.18										
+3	0.89	1.16	1.30										
+4	0.88	1.14	1.28	6.16 ± 0.55	5.70 ± 0.67			7.15 ± 0.05	6.81 ± 0.08				
+6	0.88	0.98	1.06										
+8	0.77	0.74	0.72	5.82 ± 0.45	5.32 ± 0.54	9.76 ± 0.86	10.11 ± 0.77	6.91 ± 0.05	6.52 ± 0.08	9.38 ± 0.09	9.57 ± 0.06	746 ± 34	
AAAAA	1.40	0.57	0.49	10.30 ± 0.86	9.72 ± 1.44			9.67 ± 0.07	9.34 ± 0.06			780 ± 37	
GGGGG	0.94	1.10	1.23	6.75 ± 0.59	6.17 ± 0.61			7.56 ± 0.05	7.15 ± 0.07			387 ± 19	
N-term[m]	[r]	0.22	0[q]	3.15 ± 0.23				4.51 ± 0.03				217,000 ± 11,000	
C-term[n]	0.46				2.66 ± 0.17				3.91 ± 0.13				
WT, intermol.[o]	1[q]	0.16	-0.08[s]										
WT, no tag[p]				7.32 ± 0.74	6.84 ± 0.91	9.17 ± 0.63	9.44 ± 0.52	7.94 ± 0.05	7.61 ± 0.07	9.06 ± 0.05	9.21 ± 0.06		23.6 ± 5.4[t]

[a]Data were measured and analyzed as described in Experimental Procedures. PREs were measured with CaM-4Ca^{2+} tagged with MTSL at position A128C in the C-terminal domain. Values of ^{15}N T_1 and T_2, unless otherwise stated, were measured on CaM-4Ca^{2+} conjugated to a diamagnetic control tag at A128C. Values of K_D, unless otherwise stated, were measured on CaM-4Ca^{2+} conjugated to MTSL at A128C.
[b]Ratio of the intradomain PREs (measured in the C-terminal domain for CaM paramagnetically tagged in the C-terminal domain) for CaM with the indicated linker relative to that of CaM with a WT linker.
[c]Ratio of the interdomain PREs (measured in the N-terminal domain for CaM paramagnetically tagged in the C-terminal domain) for CaM with the indicated linker relative to that of CaM with a WT linker.
[d]Calculated from interdomain PRE data as described in Experimental Procedures.
[e]Ratio of backbone ^{15}N T_1 to ^{15}N T_2, given as the average ± 1 s.d. for structured residues within either the N- or C-terminal domain.
[f]CaM-4Ca^{2+} in the absence of peptide.
[g]CaM-4Ca^{2+} bound to the MLCK peptide.
[h]Overall rotational correlation time calculated from T_1/T_2, as described in Experimental Procedures, given as the best-fit value ± 1 s.d..
[i]Affinity of CaM-4Ca^{2+} for 5ROX-MLCK, measured by fluorescence anisotropy, given as the best-fit value ± 1 s.d..
[j]Affinity measured under lower affinity conditions: higher temperature and "fluorescence buffer" (lower salt, lower pH).
[k]Affinity measured under higher affinity conditions: lower temperature and "NMR buffer" (higher salt, higher pH).

Footnotes to Table S1 (cont.)

[l]Here, WT (wild type) refers only to the CaM linker ($\Delta n = 0$), because this CaM sample has an A128C mutation, where MTSL or a diamagnetic control tag is conjugated.

[m]CaM 1-76. Values of M_{eff} and K_D were determined in concert with the C-terminal domain of CaM.

[n]CaM 81-148. Values of M_{eff} and K_D were determined in concert with the N-terminal domain of CaM.

[o]For this sample, PREs were measured on $^{2}H/^{13}C/^{15}N$ CaM-4Ca^{2+} in the presence of natural-abundance (NMR-invisible) CaM-4Ca^{2+} A128C-MTSL.

[p]For this sample, ^{15}N relaxation and affinity measurements were made on CaM-4Ca^{2+} with the full wild-type (WT) sequence (A128C, no tag).

[q]Equal to 0 or 1 by definition.

[r]Blank cells indicate measurements that are either not applicable or were not determined.

[s]This negative value is a result of the interdomain PRE ratio being slightly smaller for the intermolecular PRE measurements on full-length CaM, compared with those measured on the individual domains.

[t]Compare to $K_D = 50 \pm 50$ pM, determined by less-sensitive tryptophan fluorescence in our previous study.[S17]

Table S2. Error analysis for PRE correlation plots

linker	intradomain[a]				interdomain[b]			
	m			*R*	*m*			*R*
-1	1.11	±	0.047	0.99	0.78	±	0.023	0.98
+1	0.98	±	0.041	0.99	1.28	±	0.033	0.98
+2	0.94	±	0.053	0.98	1.07	±	0.033	0.98
+3	0.89	±	0.046	0.98	1.16	±	0.038	0.97
+4	0.88	±	0.038	0.99	1.14	±	0.032	0.98
+6	0.88	±	0.038	0.99	0.98	±	0.032	0.97
+8	0.77	±	0.054	0.97	0.74	±	0.029	0.96
AAAAA	1.40	±	0.077	0.98	0.57	±	0.027	0.95
GGGGG	0.94	±	0.038	0.99	1.10	±	0.043	0.97
N-term					0.22	±	0.028	0.72
C-term	0.46	±	0.057	0.92				

[a]The intradomain PREs (measured in the C-terminal domain for CaM paramagnetically tagged in the C-terminal domain) for CaM with the indicated linker were plotted on the *y*-axis, and those of CaM with a WT linker were plotted on the *x*-axis. The data were fit to a straight line through the origin, and the slope (*m*; ± 1 s.d.) and the correlation coefficient (*R*) are reported here.
[b]The same procedure as in footnote *a* was performed for interdomain PREs (measured in the N-terminal domain for CaM paramagnetically tagged in the C-terminal domain).

Supporting References

(S1) Anthis, N. J.; Doucleff, M.; Clore, G. M. *J. Am. Chem. Soc.* **2011**, *133*, 18966-18974.
(S2) Anthis, N. J.; Clore, G. M. *Protein Sci* **201**, *22*, 851-858.
(S3) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277-293.
(S4) Clore, G. M.; Iwahara, J. *Chem. Rev.* **2009**, *109*, 4108-4139.
(S5) Iwahara, J.; Tang, C.; Clore, G. M. *J.Magn. Reson.* **2007**, *184*, 185-195.
(S6) Lakomek, N. A.; Ying, J.; Bax, A. *J. Biomol. NMR* **2012**, *53*, 209-221.
(S7) Krishnamurthy, V. M.; Semetey, V.; Bracher, P. J.; Shen, N.; Whitesides, G. M. *J. Am. Chem. Soc.* **2007**, *129*, 1312-1320.
(S8) Cantor, C. R.; Paul, R. S. *Biophysical Chemistry: Part III: The Behavior of Biological Macromolecules*; W. H. Freeman, San Francisco, 1980.
(S9) Andre, I.; Kesvatera, T.; Jonsson, B.; Linse, S. *Biophys. J.* **2006**, *90*, 2903-2910.
(S10) Andre, I.; Kesvatera, T.; Jonsson, B.; Akerfeldt, K. S.; Linse, S. *Biophys. J.* **2004**, *87*, 1929-1938.
(S11) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Soding, J.; Thompson, J. D.; Higgins, D. G. *Mol. Syst. Biol.* **2011**, *7*, 539.
(S12) Goujon, M.; McWilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R. *Nucl. Acids Res.* **2010**, *38*, W695-W699.
(S13) Waterhouse, A. M.; Procter, J. B.; Martin, D. M.; Clamp, M.; Barton, G. J. *Bioinformatics* **2009**, *25*, 1189-1191.
(S14) The_UniProt_Consortium *Nucl. Acids Res.* **2012**, *40*, D71-75.
(S15) Chattopadhyaya, R.; Meador, W. E.; Means, A. R.; Quiocho, F. A. *J. Mol. Biol.* **1992**, *228*, 1177-1192.
(S16) Meador, W. E.; Means, A. R.; Quiocho, F. A. *Science* **1992**, *257*, 1251-1255.
(S17) Grishaev, A.; Anthis, N. J.; Clore, G. M. *J. Am. Chem. Soc.* **2012**, *134*, 14686-14689.