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

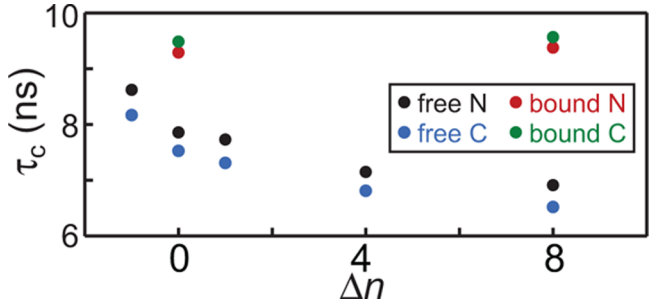

Figure 1. Dependence of CaM-4Ca^{2+} N- and C-terminal domain tumbling on the length of the linker connecting the two domains. The linker was lengthened by insertion of glycine residues and shortened by deletion of Thr79. The effective correlation time (τ_c), estimated from ^{15}N T_1/T_2, is plotted 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

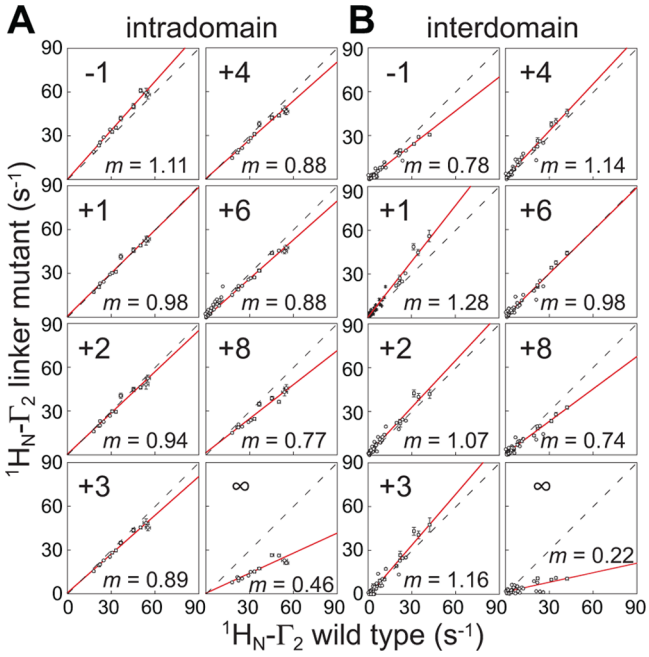

Figure 2. Dependence of CaM-4Ca^{2+} intra- and interdomain PREs on linker length. Experimental PREs (circles, error bars = 1 s.d.) were measured for CaM-4Ca^{2+} with the nitroxide spin label at A128C. (A) Intradomain PREs. (B) Interdomain PREs. The data are presented as scatter plots with the PRE value 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]

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

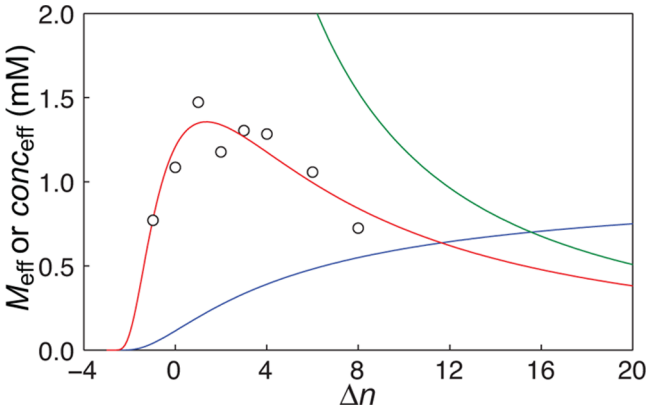

Figure 3. Dependence of CaM-4Ca^{2+} interdomain association on linker length can be largely explained by a simple random coil model. Empirical values of effective molarity (M_{eff}, circles) were calculated from interdomain PRE data for CaM-4Ca^{2+} with different linker lengths. These data were fit 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.