Calmodulin/Calcineurin Interaction Beyond the Calmodulin-binding Region Contributes to Calcineurin Activation

Bin Sun[1], **Darin Vaughan**[1], **Svetlana Tikunova**[3], **Trevor P. Creamer**[2], **Jonathan P. Davis**[3], **PM Kekenes-Huskey**[*,4,5]

[1]Department of Chemistry, University of Kentucky, Lexington, KY, USA 40506

[2]Center for Structural Biology and Department of Molecular & Cellular Biochemistry, University of Kentucky, Lexington, KY, USA 40536

[3]Department of Physiology and Cell Biology, Ohio State University, Columbus, Ohio, USA, 43210

[4]Department of Chemical and Materials Engineering, University of Kentucky, Lexington, KY, USA 40506

[5]Department of Cell and Molecular Physiology, Loyola University Chicago, Maywood, IL, USA 60153

Abstract

Calcineurin (CaN) is a calcium-dependent phosphatase involved in numerous signaling pathways. Its activation is in part driven by the binding of calmodulin (CaM) to a CaM-recognition region (CaMBR) within CaN's regulatory domain (RD). However, secondary interactions between CaM and the CaN RD may be necessary to fully activate CaN. Specifically, it is established that the CaN RD folds upon CaM binding and a region C-terminal to CaMBR, the 'distal helix', assumes an a-helix fold and contributes to activation (1). We hypothesized in (1) that this distal helix can binds CaM in a region distinct from the canonical CaMBR. To test this hypothesis, we utilized molecular simulations including replica-exchange molecular dynamics, protein-protein docking and computational mutagenesis to determine potential distal helix binding sites on CaM's surface. We isolated a potential binding site on CaM (site D) that facilitates moderate affinity inter-protein interactions and predicted that mutation of site D residues K30 and G40 on CaM would reduce CaN distal helix binding. We experimentally confirmed that two variants (K30E and G40D) indicate weaker binding of a phosphate substrate p-nitrophenyl phosphate (pNPP) to the CaN catalytic site by a phosphatase assay. This weakened substrate affinity is consistent with competitive binding of CaN auto-inhibition domain (AID) to the catalytic site, which we suggest is due to the weakened distal helix/CaM interactions. This study therefore suggests a novel mechanism for CaM regulation of CaN that may extend to other CaM targets.

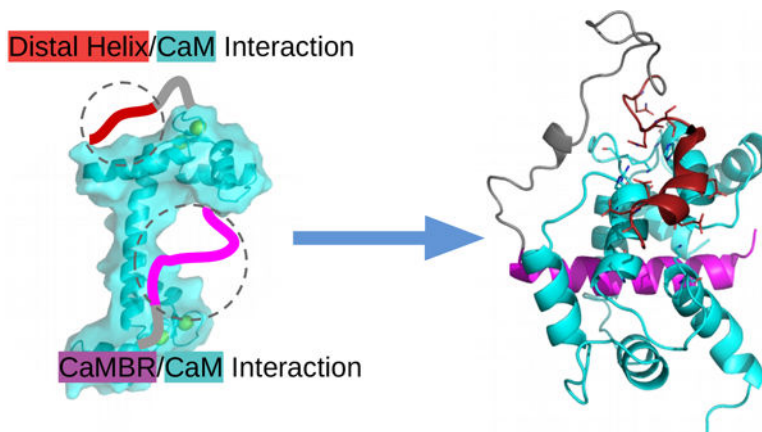[*]Corresponding authors. pkekeneshuskey@luc.edu (P. Kekenes-Huskey).

Graphical Abstract

2 Introduction

Calcineurin (CaN) is a phosphatase that contributes to gene expression in response to changes in Ca^{2+} homeostasis. It plays integral roles in processes including neurological development and maintenance, immune responses and tissue remodeling (2, 3). CaN is a heterodimeric protein consisting of two domains: chain A (57–61 kDa) contains the protein's catalytic site, while chain B (19 kDa) contributes to enzyme regulation (3). CaN is activated by rising intracellular Ca^{2+} levels. While it presents modest catalytic activity in response to Ca^{2+} alone, optimal phosphatase activity occurs upon binding of Ca^{2+}-saturated calmodulin (CaM) to the CaN regulatory domain. At depressed Ca^{2+} levels, the enzyme is inhibited by its auto-inhibitory domain (AID) that directly binds to the phosphatase's catalytic site.

Our current understanding of the phosphatase's activation and enzymatic activity has been shaped by a number of atomic resolution structures of CaN determined by X-ray crystallography (4–9) and nuclear magnetic resonance spectroscopy (10). Of the many CaN structures that have been deposited to the Protein Data Bank are examples that have revealed the phosphatase's auto-inhibited state (PDB ID: 1aui (4)), a 2:2 CaM/CaN stoichiometric configuration (7, 11, 12), complexes of the enzyme with immunosuppressants (5, 8), and transcription factors (6, 9). However, much less is known about the structural basis of CaM-dependent regulation of CaN, as atomic resolution CaM/CaN complexes are limited to intact CaM bound to small peptides comprising the CaM binding region (CaMBR) of the CaN regulatory domain (13). From those structures, while it is clear that the CaMBR assumes a-helical secondary structure when bound to CaM, the paucity of structural information inclusive of complete CaM and CaN proteins leaves critical details of CaM-dependent CaN regulation unresolved.

It is increasingly understood that CaM-dependent CaN activation depends on structural properties of the 95-residue (≈10 kDa) CaN regulatory domain (14). This segment is intrinsically disordered (4, 13–15), which signifies that it does not assume a well-defined fold in solution. Probes of its conformational properties in the absence and presence of

Ca^{2+}-activated CaM have revealed important clues about the mechanism of CaN regulation. It was first observed via circular dichroism (CD) by Rumi-Masante *et al* (14) that upon CaM binding, nearly fifty residues of the RD folded into α-helices, of which only half could be accounted for by the CaMBR region. By using hydrogen exchange mass spectrometry (HXMS), they further identified a region C-terminal to the CaMBR that formed an α-helix upon CaM binding (14). Dunlap *et al* (1) confirmed the observation in a mutagenesis study of that region. Namely, they revealed that single-point mutations of three alanines within the distal helix region (DH, residues K441 - I458) into glutamic acids disrupted helix formation. Importantly, these mutations reduced CaN's apparent affinity for a substrate, p-nitrophenyl phosphate (pNPP), that competes with the AID for the CaN catalytic site (1).

Simulations of CaN have helped bridge experimental probes of its phosphatase activity (3, 16, 17) with static, atomistic-resolution structural data. Li *et al* reported conformational changes of the CaN B domain following Ca^{2+} binding via molecular dynamics (MD) simulation and proposed that the similarity of the apo- to the holo-CaN B-domain conformations enables the former to regulate CaN activity independent of Ca^{2+} (18). Harish *et al* utilized virtual screening and MD simulations to design inhibitory peptides of CaN using the native AID peptide as a template (19). Simulations have also been used to study the involvement of CaN residues outside of its catalytic domain in the binding and anchoring of inhibitory immunosuppressant drugs and analogs thereof (20–23). Similarly, computational studies examining structural mechanisms of CaM-dependent regulation of targets have emerged recently, including myelin basic protein (MBP) (24) and myosin light chain kinase (MLCK) (25, 26). In complement to these studies, we have additionally shown via molecular dynamics and Brownian dynamics simulations that the CaMBR is highly dynamic in solution in the absence of CaM, that CaM binding to the CaMBR is diffusion-limited, and that the corresponding association rates are tuned by the charge density of the CaN peptide (27). Despite these contributions, the sequence of molecular events that follow CaMBR binding and culminate in relief of CaN auto-inhibition remain unresolved.

Observations in (1, 14, 28) formed the basis of a working model of CaN activation whereby the folding of the intrinsically-disordered distal helix into an α-helix-rich structure is coupled to relieving CaN autoinhibition. However, it was still unclear whether the distal helix directly bound to CaM, and if so, where they might share protein-protein interaction (PPI) interfaces or how those putative PPIs are stabilized. In large part, the challenge in identifying potential PPI sites arises because such interaction sites generally assume large, flat surfaces lacking specific interaction patterns (29), such as grooves formed between α-helical bundles (30, 31). Computational protein-protein docking engines have begun to address this challenge, including ZDOCK (32) and RosettaDOCK (33), which have been used to successfully elucidate structural details of intrinsically disordered peptide-involved regulation. For example, Hu *et al* utilized ZDOCK to successfully predict the binding modes between the disordered Yersinia effector protein and its chaperone partner (34). Schiffer *et al* explored the molecular mechanism of ubiquitin transfer starting from the top-ranked ZDOCK predicted binding pose between ankyrin repeat and SOCS box protein 9 (ASB9) and creatine kinase (CK) (35). Bui *et al* reported that phosphorylation of the intrinsically disordered fragment of transcription factor Ets1 leads to more binding-competent structures to its coactivator as evident by MD and RosettaDOCK (36). Our studies have therefore used

vetted protein-protein docking techniques and extensive MD to uncover and validate plausible sites for the secondary interaction between CaN's distal helix motif and CaM.

In this study, we harmonized a physical model of CaM-dependent CaN activation with phosphatase activity. This entailed using computational methods including protein-protein docking, enhanced sampling and classical MD simulations to identify potential interaction sites between the distal helix and CaM. The protein-protein docking yielded several candidate interaction sites that we defined as sites A through D. Of these, site D on the CaM solvent-accessible surface appears to stabilize the distal helix by moderate-affinity intermolecular interactions. Among the intermolecular interactions stabilizing this putative PPI are two residues, lysine (K30) and glycine (G40) found on the 'back-side' of CaM distal to where CaMBR is known to bind. Their mutation to K30E and G40D were found to abolish enzyme activity (37) in another globular CaM target, Myosin Light Chain Kinase (MLCK), that apparently relies on still unresolved secondary interactions to initiate catalysis (38, 39). Analogously, our simulations of CaM K30E and G40D variants indicate that the mutations substantially impair distal helix binding at site D. In complement to these simulations, we demonstrate that the distal helix A454E variant also destabilizes the distal helix/site D interaction in agreement with reduced phosphatase activity shown by Dunlap *et al* (1). Our data strongly suggest that the site D and CaN distal helix region are important to CaN activation, as we confirmed via experiment that the site directed variants at site D residues K30 and G40 reduces CaN-dependent dephosphorylation of pNPP. Based on these results, we provide an updated structural model of CaN activation by CaM that reflects specific CaM/distal helix interaction sites (see Fig. 1) beyond the classical CaM-peptide binding motif. We qualitatively rationalize that this mechanism controls the effective concentration of the auto-inhibitory domain (AID) near CaN's catalytic site, and by extension, CaN's catalytic activity.

3   Methods

Our simulation protocol consisted of four primary steps: 1) replica-exchange molecular dynamics (REMD) simulations to generate trial conformations of the isolated CaN distal helix region, 2) ZDOCK protein-protein docking to yield initial poses for putative CaM/CaN interaction sites, 3) refinement of poses using extensive, microsecond-length molecular dynamics simulations, and 4) Molecular Mechanics-Generalized Born and Surface Area continuum solvation (MM-GBSA) to rank-order distal helix/CaM pose interaction scores. We further challenged the predicted structural models by introducing mutations in the distal helix and putative interaction site D.

3.1   Replica-exchange molecular dynamics (REMD) sampling of the isolated distal helix

In accordance with our approach in (27), we performed replica-exchange molecular dynamics (REMD) simulations of the distal helix region (K441-I458) in the absence of CaM to exhaustively sample likely conformations that are in equilibrium. The distal helix peptide was constructed by the auxiliary TLEAP program in Amber16 (40) in an extended configuration and parameterized using the Amber ff99SBildn (41) force field. The peptide was then minimized via SANDER (42) *in vacuo* until convergence of the energy gradient

(drms   0.05) or the number of steps 1×10^5 (with first 50 steps of steepest descent and rest steps of conjugate gradients algorithm) was satisfied. The minimized structure was then used as the starting structure for REMD simulations coupled with the Hawkins, Cramer, Truhlar pairwise generalized born implicit solvent model (43) via the *igb = 1* option in Amber. The monovalent 1:1 salt concentration was set to 0.15 M and a non-bonded cutoff of 99 Å was chosen. Ten replicas were created with temperature ranges spanning 270–453 K. The temperature of each replica was calculated via the Patriksson *et al* webserver (44) to ensure the exchange probability between neighbouring replicas was approximately 0.4, as recommended in (45, 46). Each replica was first subjected to 1×10^5 steps of energy minimization via PMEMD with the first 50 steps via steepest descent and remaining via conjugate gradients. The minimized systems were subsequently heated from 0 to their respective target temperatures over an 800 ps interval using a timestep of 2 fs with a Langevin thermostat. The equilibrated replicas were then subjected to 100 ns of production REMD simulations under target temperature with a Langevin thermostat. The SHAKE (47) algorithms were also used for REMD simulations. Clustering analyses with a hierarchical agglomerative (bottom-up) approach using CPPTRAJ were conducted on the 300 K REMD trajectory to divide the trajectory into ten clusters; the average root mean squared deviations (RMSD) between each cluster was approximately 6 Å.

3.2.   Docking of distal helix to CaM/CaMBR complex via ZDOCK

The protein-protein docking webserver ZDOCK 3.0.2 (32) was used to determine probable binding poses for the REMD-generated distal helix conformations on the CaMBR-bound CaM complex. The CaM/CaMBR complex configuration was obtained from the Protein Data Bank (PDB ID: 4q5u (28)). It has been reported that 62% of experimentally-resolved PPIs are characterized by the binding of an a-helical peptide within grooves formed between adjacent a-helical on the target protein surface (31); therefore we narrowed the ZDOCK search to four a-helical-containing regions on the CaM solvent-exposed surface. These sites are shown in Fig. 2(a), from which we determined a list of probable amino acid contacts as input to ZDOCK (see Table S1). During the ZDOCK calculations, the receptor (CaM/CaMBR complex) was kept fixed while grids were constructed around the receptor with dimensions as $80 \times 80 \times 80$ and spacing as 1.2 Å. The ligand (distal helix) was then docked via the fast fourier transform (FFT) algorithm on the 3D grids. The scoring function consists of interface atomic contact energies (IFACE) (48), shape complementarity, and electrostatics with charge adopted from CHARMM19 force field (49). The generated 2×10^3 poses were subjected to a culling process to eliminate those having no contacts with residues we specified in Table S1. After culling, there were zero, two, eighty-eight and three poses left at sites A-D, respectively. The pose with the highest score at each site was chosen for further refinement using molecular dynamics.

3.3.   Conventional molecular dynamics (MD) simulations of ZDOCK-generated distal helix/CaM poses

Explicit-solvent MD simulations were performed on the ZDOCK-predicted distal helix/CaM complexes to further refine the distal helix binding poses. The amino acid sequence from the CaMBR to the distal helix is shown at the bottom of Fig. 1 and the sequence definition of CaMBR and distal helix are the same as (1). We first inserted peptide linkers for each pose

between the CaMBR C-terminus (R414) and the N-terminus (K441) of the distal helix via TLEAP. The initial linker was generated via TLEAP and energy-minimized as done in Sect. 3.1. The minimized structures were subsequently simulated *in vacuo* to heat the systems from 0 to 300 K. The last frame of the short equilibration run was subject to additional energy minimization *in vacuo* to facilitate its compliance with the distal helix and CaMBR terminus. The top poses from ZDOCK presented distal helix orientations that were all compatible with the CaMBR and linker configurations. The optimized linker was placed adjacent to the CaMBR and the distal helix; TLEAP was then used to link the peptide components. The resulting structures were then subjected to energy minimization, followed by a 100 ps heating process to raise the system temperature to 300 K, for which all atoms except the linker were fixed via the *ibelly* function in the SANDER MD engine of Amber. This minimization and heating was performed *in vacuo* to further relax the linker in the presence of the distal helix and CaM/CaMBR complex. The last frame of the heating stage was used as input configuration for explicit-solvent molecular dynamics simulations.

Each *in vacuo* starting configuration was solvated in a TIP3P (50) waterbox with 12 Å boundary margin. K^+ and Cl^- ions were added to neutralize the protein and establish a 0.15 M salt concentrations. After parameterizing the system using the ff14SB force field (51) via TLEAP, the system was subjected to energy minimization, for which all atoms except hydrogens, water and KCl ions were constrained by the *ibelly* function. The cutoff value for non-bonded interactions was set to 10 Å. A 2 fs timestep was chosen, as SHAKE (47) constraints were applied on bonds involving hydrogen atoms. Two heating procedures were performed to heat the system from 0 to 300 K using the Amber16 SANDER.MPI engine (42). In the first heating stage, the *ibelly* function was used to keep the protein fixed and the surrounding solvent unrestrained. The water box was heated to 300 K over a 100 ps interval under the NVT ensemble. For the second heating stage, the entire system was heated from 0 to 300 K over 500 ps under the NPT ensemble, for which the backbone atoms of CaM, CaMBR and distal helix were constrained by a harmonic potential (force constants of 3 kcal mol^{-1} $Å^{-2}$). Thereafter, an additional 1 ns equilibrium stage was conducted at 300 K under the same constraints, but with a reduced force constant of 1 kcal mol^{-1} $Å^{-2}$. These equilibrium simulations were followed by 100 ns production-level MD simulations. The weak-coupling thermostat (52) was used during the simulation. Clustering analysis was performed on the production trajectory using the same strategy in Sect. 3.1. The average root mean squared deviations (RMSD) between each cluster was approximately 6 Å. Based on the rationale that extending simulations using less-frequently sampled structures provides greater overall sampling of the conformational space (53), we identified 5–6 low-probability states as inputs for subsequent MD simulations. Approximately 1 μs of trajectory data were simulated in total for each site.

3.4 MD simulations of CaM (K30E and G40D) and CaN distal helix variant (A454E)

Clustering analyses were performed on the production-level MD trajectories of the distal helix-CaM configurations that yielded the most favorable binding scores by Molecular Mechanics-Generalized Born and Surface Area continuum solvation (MM-GBSA). The binding free energy between distal helix and CaM was estimated via MM-GBSA (54, 55) as follows,

$$\Delta G = \langle G_{DH-CaM} \rangle - \langle G_{CaM} \rangle - \langle G_{DH} \rangle \tag{1}$$

where $\langle G_{DH-CaM} \rangle$, $\langle G_{CaM} \rangle$ and, $\langle G_{DH} \rangle$ are ensemble-averaged free energies of distal helix-CaM complex, CaM and distal helix, respectively. A representative structure of the most populated cluster was selected as an input for *in silico* mutagenesis in order to validate the model against experiment. Namely, the CaM K30E and G40D variants, as well as the CaN A454E variant, were built by replacing and regenerating the amino acid side chains using TLEAP. Since the predicted A454E distal helix poses appeared to be inferior to those of the wild-type (WT) variant, we refined only the WT poses and thereafter introduced A454E mutations to the refined conformations. The resulting structures were energy minimized with a stop criterion of (drms <= 0.05) for the energy, during which all atoms except the mutated residues were fixed via the *ibelly* function in Amber. The energy-minimized structure was then solvated and simulated according to the same procedure in Sect. 3.3. All simulation cases in this study are listed in Table S3. In the MM-GBSA calculations, the trajectories of these three components in Eq. 2 were extracted from MD trajectories via CPPTRAJ at a 2 ns frequency. The generated sub-trajectories were used as input of MMPBSA.py in Amber16 to calculate the free energies of each part. The salt concentration was set as 0.15 M with the generalized Born model option setting as *igb = 5*. No quasi-harmonic entropy approximation was made during the calculation.

3.5 Structural Analyses

Clustering analysis, root mean squared deviations (RMSD)/root mean squared fluctuations (RMSF) calculations, hydrogen bonds, and secondary structure analysis were performed via CPPTRAJ (56) in Amber16. The reference structure used for these analyses was the CaM/CaMBR crystal structure (PDB ID: 4q5u (28)). Secondary structure for each residue was calculated using CPPTRAJ with the Define Secondary Structure of Proteins (DSSP) algorithm (57). The COLVAR module (58) within VMD was used to assess the total α-helix content of REMD-generated distal helix and DH_{A454E} conformation. The *hbond* command within CPPTRAJ was used to analyze hydrogen bonds between distal helix and CaM/CaMBR. During the hbond analysis, the angle cutoff for hydrogen bonds was disabled while the default 3 Å cutoff between acceptor and donor heavy atoms was used. Scripts and CPPTRAJ input files used for above analyses will be publicly available at https://bitbucket.org/pkhlab/pkh-lab-analyses/src/default/2018-CaMDH.

3.6 Calcineurin phosphatase assay using para-nitrophenyl phosphate (pNPP) substrate

Materials.—pNPP was obtained as the bis(tris) salt (Sigma), dithiothreitol reducing agent (Sigma), assay buffer (80 mM Tris pH 8, 200 mM KCl, 2 mM $CaCl_2$), and 50 mM $MnCl_2$ used as a CaN-activating cofactor. *Preparation of Enzymes and Proteins.* The CaM wild-type, K30E, and G40D variants were generated, expressed and purified as previously described (37). CaN was expressed from the pETagHisCN plasmid (from Addgene, Cambridge MA) in *E. coli* BL21 (DE3) Codon-Plus RIL cells (Agilent, La Jolla CA). The enzyme was unified via a Ni-NTA column followed by a CaM-sepharose column (GE Healthcare, Piscataway NJ) as described in (14). *Enzyme Assay.* Phosphatase assays were performed using 30 nM CaN, and 90 nM CaM in 96-well Corning Costar microtiter plates

with a reaction volume of 120 μL. Assays proceeded in the manner of (1) with each CaM assayed in triplicate and over three plates to account for technical variation. Control reactions absent CaN were added to the end of each lane with 200 mM pNPP to determine the rate of enzyme-independent substrate hydrolysis. *Kinetic Analysis.* The pNPP substrate reactions were varied over 11 concentrations, increasing from 0 mM to 200 mM for each column. 60 minute UV-Vis recordings were obtained on a Molecular Devices FlexStation 3 plate reader using Softmax Pro 7 software at 405 nm with 10 minute read intervals. The resulting data were inspected for appropriate Michaelis-Menten kinetics by plotting absorbance against substrate concentration. Readings were linearized to produce the double reciprocal Lineweaver-Burk plot for estimation of V_{max} and K_M based on the following equation:

$$\frac{1}{V} = \frac{K_M}{V_{max}}[pNPP] + \frac{1}{V_{max}} \qquad (2)$$

4 Results

Prior studies (14, 59) have indicated that CaM binding to CaN's canonical CaM-binding region requires secondary interactions beyond that region to fully activate the phosphatase. A study by Dunlap *et al* (1) suggested that a distal helix region spanning residues K441 to I458 comprise a secondary interaction that is likely involved in CaM binding. However, it was unclear which region(s) of the CaM solvent-exposed surface would contribute to a potential PPI. We therefore used molecular dynamics and protein-protein docking simulations to identify plausible wild-type CaN interaction sites on CaM, and challenge these predictions with mutagenesis. Our predicted site was validated using a CaN pNPP phosphatase assay.

4.1 Regulatory domain (RD)-construct propensity for secondary structure formation in absence of CaM

Circular dichroism (CD) and HXMS analysis in (14) suggest that there exists α-helical structure beyond the canonical CaMBR region after CaM's binding. We therefore sought to assess α-helicity in the REMD-simulated distal helix peptides. Previously (27), we found that extensive MD simulations of the isolated CaMBR yielded a small population of α-helical structures suitable for binding CaM in its canonical binding pose (60). We therefore applied a similar REMD procedure (see Sect. 3.1) to the proposed distal helix segment of the CaN regulatory domain to assess the propensity for the spontaneous formation of secondary structure in the absence of CaM. Here, we performed 100 ns of REMD simulations on the WT distal helix as well as an A454E variant. The latter was considered as it has been reported to exhibit reduced α-helical content in the presence of CaM (1), which is suggestive of abolishing the distal helix/CaM interaction. Following the REMD simulations, we performed clustering analysis to identify the predominant conformations of the two peptide configurations. Interestingly, we observed that both the WT distal helix and its A454E mutant partially fold into an α-helix in the absence of CaM. As shown in Fig. 3(a), representative structures of the most populated clusters of the distal helix and A454E mutant (83.8% and 85.3% of the total trajectory, respectively) both contain helical fragments. While

the overall α-helix content (≈45%) of these two fragments were statistically indistinguishable, a contiguous helix was formed in the WT distal helix, whereas it was fragmented in the mutant. Our calculations indicate approximately five amide hydrogen bonds present in WT distal helix (Fig. S11), that maintain its moderate helical content. We speculate that the spontaneous helix formation of isolated distal helix region is a consequence of the peptides's balance of charge and uncharged amino acids. Our rationale is based on the finding that fraction of charged residues (FCR) and net charge per residue (NCPR) of intrinsically disordered peptide correlates with ensemble properties (61, 62): when both FCR and NCPR are less than 0.25, collapsed globules are preferred, otherwise, expanded coil-like structures are preferred. The A454E mutation was located in the hydrophobic surface of distal helix region (1) and this apolar to polar mutation disrupted the hydrophobic surface and thus increased the peptide's charge density. After mutation, the FCR and NCPR was increased from 0.27 to 0.33 and 0.16 to 0.22, respectively. Ultimately this altered its secondary structures, thus leading a fragmented helix for the isolated peptide. These helicity features are further quantified as a residue's α-helix structural probability shown in Fig. 3(b); the distal helix region has a maximum probability in middle region while the A454E has maxima near the two terminus. Both the simulated distal helix and its variant therefore adopt α-helix content in the absence of CaM, but it remains to be determined whether the dominant structures are capable of binding the CaM surface. We note that experimental assays of the complete regulatory domain (RD) do not detect significant secondary structure; this discrepancy may be a result of using substantially different RD lengths (S374 to Q522 residues in Rumi-Masante *et al* (14) and K441-I458 in this study). We discuss this difference in further detail in the Limitations (Sect. 5.5).

4.2 Protein-protein interactions between RD-construct and peptide-bound CaM

The overwhelming majority of CaM-containing complex structures resolved to date include only limited fragments of the bound target protein (60). CaM-bound CaN is no exception, as the mostly likely physiological conformation (28) consists of a monomeric CaM in a canonical 'wrapped' conformation about a target region in CaN(A391-R414) (13); however, it is evident that secondary interactions beyond this domain play a role in CaN activity, yet atomistic-level structural details of these interactions have not yet been resolved. Therefore, in order to resolve potential binding regions for the distal helix region, we seeded a protein-protein docking engine, ZDOCK (32), with candidate α-helical structures identified through REMD simulations. The docking simulations were performed in regions that included grooves formed between α-helices we identified at the CaM solvent-accessible surface. We selected these regions, since such secondary structures are believed to nucleate protein-protein interactions (63). Furthermore, a thorough examination of protein-protein complex structures in the Protein Data Bank in 2011 suggested that α-helices contribute to 62% of all PPI interaction surfaces (31) between binding partners. Narrowing the search region on CaM to those containing α-helical regions yielded four candidate sites (A-D) that spanned nearly the entire CaM solvent-exposed surface (see Fig. 2(a)).

The most energetically-favorable distal helix-CaM poses predicted via ZDOCK at sites A-D are summarized in Fig. S1. The docked poses reflect significant interactions of at least the distal helix C-terminal loop with loops bridging adjacent α-helices on the CaM surface. At

site A, polar residues near N97, Y99 and D133 from two of the C-terminal CaM domain's loops interact with the distal helix, compared with just one EF-hand motif loop at site B (D129, D133 and D135). The site C poses were primarily stabilized by hydrophobic interactions formed from CaN residues L444/I458 and F16/L4 on CaM, in addition to a loop-loop interaction via CaM D64. The site D poses reflected distal helix C-terminal loop interactions with CaM EF-hand loop residues near N42 and K94. Most poses were unsurprisingly parallel to a-helical/a-helical 'grooves' on the CaM solvent-exposed surface and were evidently anchored through interactions between the proteins' loop regions.

In contrast, we found that the A454E variant docked poorly at sites A–D (see Fig. S2), as assessed by the proximity of docked poses to the designated CaM sites. In fact, most predicted poses tended to localize toward site A, albeit with weak interactions. Moreover, we speculate that the impaired binding of DH_{A454E} may arise from its fragmented a-helical structure, in contrast to the contiguous regions for the WT variant (see Table S2 for docking scores and Fig. S1/Fig. S2 for docking poses). Although docking scores were provided by the ZDOCK algorithm to rank order potential poses, we did not analyze these scores in detail as we later refined these structures using more detailed simulations and energy expressions. This refinement corrects for artifacts from the ZDOCK algorithm, which assumes rigid conformations for both proteins that would ordinarily be expected to relax in the bound complex. Hence, in the following section we pursue extensive microsecond-scale all-atom MD simulations to refine and assess the predicted poses.

The docked CaN/CaM configurations from the previous section were intended as inputs for MD-based refinement of nearly intact CaN regulatory domain complexes with CaM. Subsequent refinement using microsecond-length MD simulations were used to relax the rigid protein conformations assumed in ZDOCK. We first assess the integrity of the predicted binding modes based on Molecular Mechanics-Generalized Born and Surface Area continuum solvation (MM-GBSA). MM-GBSA scoring of the MD-generated configurations provides a coarse estimate of binding affinity without significantly more expensive free energy methods. We reported the binding free energy of the distal helix between CaM as well as between the CaMBR and CaMin Fig. 4. Significantly, we found that binding of WT distal helix at the CaM site D yielded a more pronounced favorable average binding free energy ($G = -27.7 \pm 2.3$ kcal mol^{-1}) than sites A, B and C (-3.3 ± 2.8 kcal mol^{-1}, -17.4 ± 2.6 kcal mol^{-1}, -22.6 ± 2.2 kcal mol^{-1}) with P-values (1×10^{-4}, 2.8×10^{-3} and 1.144×10^{-1}, respectively) confirming that the means are significant compared to the null hypothesis. Notably, these thermodynamically favorable scores are also suggestive of the potential for the distal helix to bind multiple regions on the CaM surface, although site D is the most favorable site. Additionally, the binding free energies of distal helix interactions were generally substantially weaker (-3.3 to -27.7 kcal mol^{-1}) than those between the CaMBR and CaM ($G < -1.20 \times 10^2$ kcal mol^{-1}).

We supplement the energy scores with structural indicators of stability, namely contacts and RMSF. We report in Fig. 5 the corresponding root mean squared deviations (RMSD) and root mean squared fluctuations (RMSF) of the peptide backbone atoms from CaM and the CaN CaMBR. We additionally include two CaM variants with mutations at site D to challenge our predicted pose. We found that the average RMSD values of the MD-predicted

conformations relative to the experimentally-determined CaM/CaMBR structure were at or below 2 Å; we attribute these small fluctuations to stable CaM/CaMBR interactions that were insensitive to the distal helix docking. Similar to the RMSD data, the CaM and CaMBR RMSF values are comparable in amplitude and nearly indistinguishable between distal helix docking poses, with most residues presenting values below 1.5 Å. The prominent peaks in excess of 5.0 Å correspond to the CaM terminus and the N-terminus of the CaMBR. We additionally observe a variable region midway along the CaM sequence, which corresponds to the labile linker between its globular N- and C- domains that is implicated in allosteric signaling (64).

The small and statistically indistinguishable RMSF values for the CaM/CaMBR in Fig. 5 suggest that distal helix binding had negligible impact on binding the CaM recognition motif. This is an important observation, as viable binding poses for the distal helix are expected to preserve the binding between the CaMBR and CaM. We base this assumption on CD data collected in (65) that indicated substantial alpha helical character in the CaM/CaN complex following dissociation of the distal helix domain. Therefore, we then assessed the integrity of the distal helix poses using RMSF analyses and measurements of inter-protein contacts. In Fig. 6 we report representative configurations of the distal helix region (red) in complex with CaM (cyan), as well as their corresponding per-residue RMSF values in Fig. 7. To guide interpretation, we hypothesized that RMSF values above 5 Å were indicative of poorly stabilized residues. At site A, both the distal helix/CaMBR linker and the distal helix reflect RMSF values in excess of ~10 and ~15 Å, respectively. These large fluctuations arise from the breadth of binding orientations evident in Fig. 6(a), which we interpreted as poorly-stabilized configurations. Similarly, the site B configurations also appeared to be loosely bound, based on linker and distal helix RMSF values beyond 10 Å. In contrast, the distal helix RMSF values at sites C and D were below 5 Å, with the latter site reporting the smallest values among the sites we considered, which is indicative of a stable binding configuration.

As has been shown in other proteins regulated by disordered protein domains (66–68), there are often multiple poses that contribute to regulation. We therefore assessed the most significant inter-protein contacts contributing to the ensemble of distal helix binding poses at sites A-D. Among these poses, the distal helix configurations at site D presented the lowest distal helix RMSF values among the considered sites. Significantly, the site D distal helix configuration exhibited several hydrogen bond-facilitated interactions with CaM, including two long-duration (37% and 55% of sampled configurations) interactions between Q445 and CaM residues R37/K94, pairing of CaM K21 with glutamic acids E453 and E450, as well as E456 with CaM residues K30 and R37. Contacts between CaM and CaN, as well as their duration (as assessed by the percentage of MD frames satisfying a hydrogen bond contact cutoff of 3 Å between oxygen and nitrogen atoms) are additionally quantified in Fig. S4 (specific values are listed in Table S4). The latter data indicate a modestly greater degree of hydrogen bonding of the distal helix at site D (10 h-bonds were above 10%) versus site B (9), and a significantly greater degree relative to sites A (1) and C (3). Furthermore, the site D pose appears to be stabilized by both the N- and C-domains of CaM (residues D20-S38 and R90-N111, respectively). We speculate that this bi-dentate interaction could improve CaMBR binding by locking CaM into its collapsed configuration and thereby preventing

disassembly. Although during the simulation the distal helix at site D maintained significant α-helix content (see Fig. S3 and Fig. S5), we note that some of the predicted structures exhibited beta sheet character in the linker region (see Fig. S6) that was not observed in the CD spectra collected by Rumi-Mansante *et al* (14). This persistent secondary structure was limited to a few residues (see Fig. S6 and Fig. S7) and thus may be beyond the limits of detection in earlier CD experiments. We comment on this further in the Limitations section (see Sect. 5.5). Meanwhile, site B reflected interactions with both CaM terminal domains that were attenuated relative to site D, while sites A and C were mostly bound by interactions of their linker regions with the CaM N-domain. Interestingly, we observed that the distal helix poses originating at site B migrated toward site D (see Fig. S3), which likely explains the higher hydrogen bonding in site B versus sites A and C.

As a result of HXMS conducted by Rumi-Masante *et al* (14) of the RD construct CaN in solution with CaM, it is apparent that residues R414 through E456 are within a stretch of residues that are somewhat protected from solvent, which suggest that relief of CaN autoinhibition entails binding at least the distal helix region. We note that the HXMS data could not precisely distinguish which residues were protected, as proteolysis and mass spectrometry were conducted on short peptides. Further, HXMS data detects only bonds involving backbone amide protons, thus we speculate that the CaN side chain interactions with CaM may stabilize the distal helix alpha helical structure. Hence, we suggest that CaM/CaN configurations that stabilize the distal helix region likely contribute to CaN activation. Based on this rationale, the small RMSF values and extensive hydrogen bonding of the CaN distal helix with the CaM site D relative to other ZDOCK identified regions suggest that CaN is most stabilized at site D.

4.3 Effects of distal helix/CaM site D mutagenesis

MD simulations of the WT CaN CaMBR-distal helix sequence suggest that CaM site D is a probable binding region for the CaN regulatory domain. To challenge this hypothesis, we performed MD simulations of CaN distal helix and CaM site D variants to test whether the distal helix/CaM interaction was impaired. Namely, we introduced the CaN A454E and CaM K30E and G40D mutations into the MD-optimized WT structures in accordance with prior experimental studies of CaN (1) and the Myosin Light Chain Kinase (MLCK) (37). We elected to mutate the WT CaMBR/distal helix complexes with CaM, as the WT complex appeared to have favorable stability, whereas repeating the REMD/ZDOCK steps with the mutants may not have yielded viable configurations. The proposed A454E CaN variant was based on CD data collected by Dunlap *et al* (1) that demonstrated reduced α-helical content upon binding CaM relative to the WT with impaired CaN activation. The CaM variants we examined in this study were based on studies (37) of the CaM-dependent MLCK activation, for which secondary interactions beyond the canonical CaM binding motif were implicated in enzyme activation (38, 39) (Fig. 8(a)). Although these secondary CaM interactions are involved in directly binding the MLCK catalytic domain in contrast to CaN (38), two residues (K30 and G40) implicated in binding (37) reside within the site D identified in our simulations.

We reported MM-GBSA-calculated binding free energies between the distal helix and CaM of the mutants from MD simulations of these variants in Fig. 4. While the WT distal helix at the CaM site D had the most stable binding with $G = -27.7 \pm 2.3$ kcal mol^{-1}, the three mutations K30E, G40D and A454E had less favorable G values of -21.8 ± 2.5 kcal mol^{-1}, -17.9 ± 2.6 kcal mol^{-1} and -14.4 ± 2.6 kcal mol^{-1} (P-values of 8.12×10^{-2}, 5.1×10^{-3} and 2×10^{-4}, respectively). The MM-GBSA-energies suggest that mutations would impair binding between the distal helix and CaM. Accordingly, we present the linker and distal helix RMSF data for the WT and mutants in Fig. 8(b). The distal helix RMSF values among the two CaM variants were moderately increased compared to the WT case. Specifically, for the WT system, the distal helix residues were entirely within 10 Å and as low as ~2.5 Å. In contrast, the K30E variant yielded RMSF values no smaller than approximately 5 Å, while the C-terminal half approaches values nearing 15 Å. This trend manifested in fewer long-lived hydrogen bond contacts between the distal helix and both CaM domains (see Fig. 8). Similarly, the G40D mutation appeared to significantly disrupt interactions with CaN, as the entire distal helix region was characterized with RMSF values over ~10 Å in amplitude, with corresponding decreases in hydrogen bond contacts. Among the mutations we considered, the A454E mutant had the most severe impact on RMSF values, as all residues comprising the linker and distal helix regions resulted in fluctuations above 8 Å. We also reported the α-helix probability of distal helix residue for variants in Fig. S8. It was found that all variants preserved a significant degree of overall helicity despite evidence of impaired interactions with CaM. However, the specific residues which formed the α-helix were different among the variants: the mutation of A454 to E454 shifted the helicity to the first half of the distal helix while the two CaM variants had the second half region being α-helical. Altogether, these simulation data suggest that: 1) the WT distal helix is stabilized at the site D CaM region, 2) site D residues K30 and G40D are implicated in distal helix binding, and 3) disruption of site D binding by CaN A454E is consistent with reduced helicity and enzyme activity measured experimentally.

4.4 Phosphatase assays of site-directed CaM mutants

To validate the simulation results, namely that the CaM site D stabilizes the distal helix and thereby promote CaN activity, we analyzed the kinetics of CaN mediated hydrolysis of pNPP. We hypothesized that disruption of site D/distal helix binding would reduce the accessibility of the catalytic site for pNPP binding by allowing the AID to bind to a greater extent. This could manifest as a reduced apparent substrate affinity. We therefore conducted CaN phosphatase assays with a pNPP substrate using two site D variants, K30E and G40D. We analyzed substrate turnover in a Michaelis-Menten model, as described in the Methods. Phosphatase assays performed on CaM variants strongly suggest a statistically significant reduction (p-values in Table 1) in catalytic activity by a substantial increase in K_M for K30E and G40D over the WT (46.0 ± 2.8 mM, and 35.5 ± 2.2 mM, 27.6 ± 1.3 mM respectively) indirectly indicating weaker binding of the distal helix peptide to the mutated CaM construct.

5 Discussion

5.1 Summary of Key Findings

We have used computational modeling and experiment to elucidate a potential mechanism for CaM-dependent regulation of CaN activity, whereby the binding of a 'distal helix' region of the regulatory domain relieves CaN auto-inhibition. Our microsecond-duration MD simulations indicate that the distal helix region remains bound to the solvent accessible CaM surface, which could decrease the ability of the AID to bind CaN's catalytic site (see Fig. 1). In contrast, we predict that an engineered variant (A454E) disrupts the domain's secondary structure and ability to competently bind CaM. Both predictions are in agreement with experimental probes of CaN regulatory domain structure and phosphatase activity (1). Namely, among the four potential regions on CaM's surface that were solvent-accessible after binding the CaMBR, our data suggest that an RD region spanning the CaMBR through the distal helix was best stabilized at a site nestled between the CaM N- and C-terminal domains. In silico mutagenesis of two N-terminal CaM residues (K30E and G40D), prevented distal helix binding in our model, which we suggest hinders CaN activation, similar to identical mutations in CaM that were found to inactivate another CaM target, Myosin Light Chain Kinase (MLCK). We confirmed the potential CaM site D binding site for the distal helix through site-directed K30E and G40D variants, which we found to weaken CaN binding as reflected by reduced (weakened) MM-GBSA scores and an increase in K_M (from 27.6 mM to 46.0 and 35.5 mM, respectively) in a pNPP phosphatase assay. Although our REMD simulations suggest that the isolated distal helix region spontaneously assumes moderate α-helical structure in absence of CaM, in contrast to trends observed in the complete RD domain observed experimentally (14), we do not believe this significantly impacts our suggestion that site D contributes to CaN activation. We discuss this further in Sect. 5.5.

5.2 Plausible binding modes for CaN distal helix with CaM

Previous studies suggest that 1) binding of regulatory domain residues beyond the CaMBR are involved in CaM-dependent relief of CaN autoinhibition (1, 14), based on increases in regulatory α-helical content reported upon binding CaM that could not be accounted for by the CaMBR alone 2) Alanine to glutamic acid mutations at RD positions (A451E, A454E and A457E) C-terminal to the CaMBR decreasing the α-helical content and CaN activity and 3) HXMS studies indicating reduced solvent accessibility for the distal helix relative to the entire RD for a complex formed between CaM and a CaN regulatory domain/AID/C-terminal domain construct (see Sect. S1.4 for further discussion). While we believe site D is the most probable site for distal helix binding, interactions with other potentially less-favorable sites could occur and contribute to the bound RD conformational ensemble. Such a diverse ensemble of strongly and weakly bound conformations is increasingly evident in complexes involving intrinsically disordered peptide (IDP)s and globular targets (27, 70) and may be adopted by CaN as well. It is also interesting that CD experiments in (1) suggested that the distal helix contact is abolished at temperatures above 38 degrees Celsius. It is tempting therefore to speculate that the comparatively larger RMSFs of the bound distal helix configurations relative to the CaMBR, in addition to the weaker interaction energies, may render the distal helix interaction susceptible to melting.

Strengthening the case for the involvement of the CaM site D in binding the CaN distal helix are our comparisons against two CaM variants with substantially impaired ability to relieve enzyme auto-inhibition in another CaM target, Myosin Light Chain Kinase (MLCK) (37). CaM appears to relieve MLCK auto-inhibition (71) through binding the kinase's regulatory domain (72) and adopts a similar conformation as the CaN/CaM complex with CaM 'wrapping' around an α-helical CaMBR motif (see also Fig. 8(a)) (1, 69). Importantly, both appear to utilize secondary interactions beyond the CaMBR motif and it was shown by Van Lierop *et al* (37) for MLCK that K30E and G40D mutations far from its CaMBR-binding domain prevented CaM-dependent kinase activity. These sites are localized to the site D region we identified for distal helix binding in our study. Although the secondary interactions in MLCK likely involve CaM binding directly adjacent to the enzyme's catalytic domain (73), we speculated that mutagenesis of these CaM residues could also impact CaN activation, but instead by disrupting distal helix interactions. We confirmed this hypothesis in our computational model by demonstrating less favorable distal helix binding scores, and validated these predictions via a pNPP assay.

5.3 Assessment of phosphatase activity

To challenge our hypothesis that impaired distal helix binding to CaM reduce CaN activity, we used kinetic phosphatase assays with the substrate pNPP on WT and the aforementioned CaM mutants. The Michaelis constant, K_M, obtained from these experiments informs on the ability of the catalytic site to bind and dephosphorylate pNPP. This substrate is specific to the catalytic site due to its low molecular weight, which allows for a probe of the extent to which CaM binding removes the AID. Mutations in the distal helix region that disrupt its folding and allow the AID to bind to the catalytic site would result in reduced pNPP binding (higher K_M). This explanation has been used by earlier authors studying the inhibitory properties of the AID as a peptide (28). We reported significantly higher K_M values for both K30E and G40D based on our pNPP assay, thus these mutants evidence weaker distal helix binding that impedes removal of the AID from the CaN catalytic site. As a result, the CaM variants reduce the CaN catalysis of the dephosphorylation reaction. This can be interpreted as the AID competing with pNPP at the catalytic site and yielding a reduced apparent substrate affinity. This loss in affinity coincides with 40% increases in K_M reported for CaN A454E relative to WT CaN (1), which were attributed to impaired distal helix formation. In contrast to common peptide-based dephosphorylation targets like RII (28) that bind to sites outside of the catalytic site (the LxVP site), therefore binding, and hence K_M, would be unaffected by mutations in the distal helix region.

5.4 Tether-model of CaM-dependent CaN activation

We recognize that a shortcoming of our modeling approach is that it is limited to simulations of CaM complexes with fragments of the CaN regulatory domain, whereas distal helix binding's effects on CaN activity are coupled to the entire regulatory domain and specifically, the AID. We therefore discuss a qualitative description of 'linker' dynamics of the regulatory domain appropriate for the AID-dependent inactivation of CaN. Specifically, we speculate that we can describe extents of CaN inactivation based on the AID's effective concentration at the CaN catalytic site as determined by the formation of distal helix/CaM interactions. This effective concentration is controlled by the tethering of the AID to CaN,

which effectively confines the AID to a smaller volume (than free diffusion) that results in a higher interaction probability with the catalytic site (74). We use this effective concentration perspective to qualitatively assess how distal helix interactions with CaM impact CaN activity, as explicit all-atom simulations of the complete RD are prohibitively expensive. Here we leveraged theoretical models of protein activation (75, 76) by describing AID binding to the CaN catalytic domain as an intra- PPI. This PPI leverages a molecular tether (the regulatory domain) to enhance the *local* effective AID (p) concentration near the catalytic domain.

To illustrate this principle in CaN, we provide a basic extension of a linker-dependent modulation model we recently applied to the calcium-dependent troponin I (TnI) switch domain binding to troponin C (TnC) (74). For this reaction, Ca^{2+} binding to TnC generates a conformation that can facilitate TnI binding:

$$\text{TnC} \xrightleftharpoons{\text{Ca}^{2+}} \text{Ca} \cdot \text{TnC} \xrightleftharpoons{\text{TnI}} \text{Ca} \cdot \text{TnC} \cdot \text{TnI};$$

(3)

hence, increasing the TnI concentration promotes the generation of TnC·TnI with fewer equivalents of Ca^{2+}. In the tethered state, we estimated that the *effective* switch peptide concentration was an order of magnitude greater near its TnC target than would be expected for a 1:1 stoichiometric ratio of untethered (free) switch peptide to TnC. Accordingly, we experimentally confirmed that formation of the TnC/TnI switch peptide occurred at lower Ca^{2+} concentrations for the TnC-tethered TnI compared to a cleaved system, in which both TnC and TnI were untethered (74).

In a similar vein, we created a hypothetical linker-based model of CaN activation, based on a polymer-theory based model for the probability distribution of the linker spanning the CaMBR and AID domains (see Fig. 9). We introduce this model with several assumptions. Firstly, we postulate the CaN inhibition is dependent on the free AID concentration, of which the latter is determined by the RD 'tether' length. This tether length can assume three distributions associated with the CaM-free, CaMBR-bound CaM and CaMBR+distal helix-bound CaM, respectively. Lastly, for simplicity we assume that the distal helix binds CaM independent of the AID's bound state, though in reality we recognize there will be a competition between these two events.

Under these assumptions, we describe the effective [AID] at the CaN catalytic domain, based on the RD linker length in its CaM-free, CaMBR-bound CaM and CaMBR+distal helix-bound CaM states. We based this on an effective concentration model for tethered ligands suggested by Van Valen *et al* (75),

$$[\text{AID}]^{eff} = \left(\frac{3}{4\pi\xi L}\right)^{3/2} \exp\left(-\frac{3D^2}{4\xi L}\right)$$

(4)

where D is the distance between CaMBR and catalytic site, L is linker length, and ξ is the persistence length. The units of $[\text{AID}]^{eff}$ in Eq. 4 was found via fitting to existing experimental data. Namely, experimental assays were reported to investigate the competitive inhibitory effect of isolated AID peptide on CaN phosphate activity on substrate peptide (77,

78). In the assays, the reduction of phosphatase activity was recorded as isolated AID peptide was added to intact CaN pre-incubated with CaM and substrate RII peptide. According to the experimental setup, there existed three competitive components that could bind the catalytic site of CaN: substrate RII peptide, isolated AID peptide and tethered AID from the intact CaN itself. Similar to P_{on} definition which represents the probability of switch peptide being on under the competitive binding of free ligand and tethered ligand to receptor in (75), we also defined a P_{on} that represents the percentage of CaN phosphate activity on substrate RII peptide under competitive binding from isolated AID peptide, and tethered AID:

$$P_{on} = \frac{1 + \frac{[\text{RII}]}{K_{d1}}}{1 + \frac{[\text{RII}]}{K_{d1}} + \frac{[\text{AID}]}{K_{d2}} + \frac{[\text{tAID}]}{K_{d2}}}$$

(5)

where [RII], [AID] and [tAID] are concentrations of substrate, isolated and tethered AID peptide, respectively. [RII] is set as 5 μM according to the experimental setup and the dissociation constant of substrate K_{d1} is assumed to be 10 μM. Tethered AID peptide is assumed to have the same dissociation constant as isolated peptide with an experimentally estimated K_{d2} of 40 μM (77, 78). The fitting of Eq. 5 to experimental data in (77) with [tAID] as free pameter is shown in Fig. 9(b). [tAID] was fitted as 2.07 μM and this value corresponds to [AID]^{eff} of 'CaMBR+distal helix-bound CaM' case in our tether model. In the following tether model analysis, the [AID]^{eff} from Eq. 4 was scaled by [tAID] to give meaningful units for the effective AID concentration.

We first provide a rough estimate for the linker length through simulations of residues E415-M490 C-terminal to the CaMBR (see Fig. 9(a)). Starting from WT/A454E site D simulations, an optimized fragment (residues K459 to M490) containing AID built by TLEAP was fused to the C-terminus of the distal helix in the representative structure of the first two most populated clusters. The complete structures were resolvated and simulated for 0.7μs as described in Sect. S1.1. These simulations indicate that the WT AID to CaM distance is approximately 23 Å, versus approximately 40 Å for the A454E variant that precludes distal helix binding.

Based on these data, in Fig. 9(c) we demonstrate the effective AID concentration over a range of ligand lengths (L), predicted from Eq. 4 assuming $D = 66$ Å for the distance between CaM and the CaN AID binding site and $\xi = 3$ Å (79). The black dot represents the CaMBR/distal helix (DH)-bound case, which has a tethered ligand length estimated from our simulation of approximately 23 Å or roughly 8 free amino acids. The blue dot represents free RD, which has a ligand length of 95 residues (M387 to E481). The red dot represents the CaMBR-bound (no distal helix interaction as for the A454E case, in this case, the tethered ligand length estimated from our simulation as 40 Å). Based on these linker lengths, the corresponding effective [AID] concentrations for CaMBR-bound (A454E) states were 6.76 μM versus 2.07 μM for the CaMBR/distal helix-bound case. For the free RD case, the effective [AID] is 3.20 μM. In other words, the distal helix/CaM interaction reduces the AID effective concentration near the catalytic site relative to the free (no CaM) RD or CaMBR-bound configurations. Hence, the distal helix-bound case yields more active enzyme. This

approximate model is therefore consistent with the experimental trends in activity data reported in the literature (1, 80), namely that maximal CaN activation requires CaM binding.

5.5 Limitations

We observed appreciable degrees of alpha helical and beta sheet character in the regulatory domain that were not evident in the CD data from (14). A primary distinction between the modeling and experimental studies is that we used a much smaller regulatory domain fragment (residue A391 to I458) than the full length domain in Rumi-Masante *et al* (14), owing to the computational expense. It is possible that there are different tendencies to form secondary structure, based on the regulatory domain length. Since we simulated only a small fragment of the RD domain, this might have increased the peptide's preference for alpha helical structure than would otherwise be observed in measurements of the entire RD. For instance, it has been shown that IDPs have length-dependent preference of residue compositions as a longer IDP has more enriched K, E and P than a short IDP (81), implying the conformational properties of IDPs which are determined by sequence charge distribution (61) are also length-dependent. As a concrete example, Lin *et al* (82) reported that the 40-residue disordered amyloid beta monomer has reduced β-hairpin propensity when compared to the longer 42-residue monomer.

We additionally recognize that differences in ionic strength or solvent composition might influence the percentage of alpha helical character, although this seemed to be a modest effect in our simulations of the CaMBR alone (27). Importantly, in that study, we reported negligible alpha helical character for that isolated CaMBR peptide, which suggests that our force field was not artificially stabilizing alpha helices, as had been an issue in earlier modeling studies of IDPs (83, 84). Nevertheless, the potential overestimate of alpha helical content for the isolated peptide is probably of little consequence, since the predicted bound distal helix was shown to exhibit significant alpha helical content consistent with experiment.

We utilized REMD to sample the distal helix sequence in the absence of CaM; although REMD has been shown to perform well in terms of qualitatively describing conformational landscape, chemical shifts, and α-helix stability for peptides of lengths comparable to the distal helix (85–87), we did not have the means to experimentally validate the predicted apo ensembles. Nevertheless, the simulations provide testable hypotheses in terms of the α-helical content. We additionally limited ourselves to subsets of the CaM surface for the docking search, which represented approximately 38% of the solvent-exposed surface area. However, given that the microsecond-length simulations were sufficient to reorient the site B configurations into the site D site, we anticipate the docked distal helix candidates reasonably sampled the thermodynamically-accessible regions of the CaM surfaces. Although it has been demonstrated that RD binding to CaM is diffusion-limited, it is also possible that the intermediate complexes could be further optimized to form a final bound state, which would perhaps lead to more accurate assessments of critical intermolecular contacts and energy estimates. For the latter, alchemical methods such as thermodynamic integration may provide more accurate affinity estimates, albeit at a substantially greater computational expense compared to 'end point' methods like MM-GBSA. In addition, more

detailed simulations of the RD ensemble in the presence of the complete CaM and CaN structures are needed to more accurately characterize the effective AID distribution controlling CaN (in)activation.

Lastly, there are several considerations that could improve the accuracy of the tethering model in the Discussion. These include assumptions that the linker follows a random-walk chain distribution, that the catalytic domain does not attract and thereby bias the AID distribution, and that the CaN molecule does not sterically clash with the linker chain. Further, precise knowledge of the CaM distribution relative to the CaN B-chain would be needed to refine the effective linker lengths. Despite these assumptions, the model provides a qualitative basis for how RD mutations or variations in RD length could influence the efficiency of CaN (in)activation, similar to the model systems with synthetic linkers, as in (88).

5.6 Additional considerations

There are several compelling directions to pursue that would provide essential clues governing CaM-dependent CaN activation. For one, we have predicted several contacts that appear to be involved in stabilizing the distal helix region; mutagenesis of these potential 'hotspots' on CaM and measurements of subsequent CaN phosphatase activity could help validate this site. In addition, more detailed characterization of the RD intrinsically-disordered conformation ensemble would benefit future modeling. Given the difficulty in probing ensemble properties of IDPs, it is likely that modeling and experiment, such as fluorescence resonance energy transfer (FRET) labeling, should work in tandem toward this goal. Furthermore, relating these RD ensemble properties to the propensity for AID and CaN catalytic domain interactions would comprise an essential step toward a complete model of CaM-dependent CaN activation.

We anticipate that the findings of this study could broadly expand to other classes of CaM-dependent targets. Namely enzymes that have autoinhibitory domains such as the CaM-dependent kinases (CaMKI, CaMKII) and MLCK (89). Similarly to CaN, these enzymes feature CaM-binding regions that are disordered in the absence of CaM (65, 90). In contrast to CaN, the CaM-binding regions are nearly adjacent to their autoinhibitory domains (89). As for CaN, the autoinhibitory domain and CaM-binding region are ~50 residue apart. For this reason, CaM essentially competes with the target for its regulatory domain, whereas for CaN, CaM-binding essentially controls the spatial probability distribution of the AID by modulating the 'tether-length' of the regulatory domain. Despite this distinction, at least for MLCK a secondary interaction between CaM and a region beyond CaMBR of MLCK might be necessary for latter's activation (39), thus raising the possibility that other enzymes share similar CaM -dependent activation mechanisms as we report for CaN.

6 Conclusions

We have developed a computational strategy to elucidate potential binding poses for a secondary interaction (the 'distal helix') between the CaN regulatory domain and CaM that is apparently essential for competent CaN activation. We combined REMD simulations of isolated distal helix peptides, protein-protein docking of the distal helix peptides to the

CaMBR-bound CaM surface, and microsecond-scale MD simulations of candidate poses to implicate a so-called CaM site D in binding the CaN distal helix. The predicted site D region is in part stabilized through direct interactions with K30 and indirectly through G40, which is consistent with experimental probes of a CaM-activated enzyme, MLCK. We confirmed the predictions via pNPP phosphatase assay in which mutations of K30E and G40D in CaM reduced CaN activity compared with WT CaM. With these data, we provide a qualitative model of AID-dependent CaN activation, which can be used to further refine potential molecular mechanisms governing the activation process and susceptibility to missense mutations. Importantly, our data suggest a potentially novel mechanism of CaM-dependent target regulation whereby interactions distal from the canonical CaM-peptide binding motif control target auto-inhibition. Given the broad range of physiological processes mediated by CaM binding to intrinsically disordered target proteins (60), the mechanistic details of CaN activation in this study may extend to other CaM targets, including CaM-dependent channels and cytoskeletal components (60, 91).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

References

(1). Dunlap TB, Cook EC, Rumi-Masante J, Arvin HG, Lester TE, and Creamer TP, (2013). The distal helix in the regulatory domain of calcineurin is important for domain stability and enzyme function. Biochemistry 52, 8643–8651. [PubMed: 24191726]

(2). Rusnak F, and Mertz P, (Jan. 2000). Calcineurin: Form and Function. Physiol. Rev 80, 1483–1521. [PubMed: 11015619]

(3). Klee CB, Ren H, and Wang X, (1998). Regulation of the Calmodulin-stimulated Protein Phosphatase, Calcineurin. J. Biol. Chem 273, 13367–13370. [PubMed: 9593662]

(4). Kissinger CR, et al. (1995). Crystal structures of human calcineurin and the human FKBP12 FK506 calcineurin complex. Nature 378, 641–644. [PubMed: 8524402]

(5). Griffith JP, Kim JL, Kim EE, Sintchak MD, Thomson JA, Fitzgibbon MJ, Fleming MA, Caron PR, Hsiao K, and Navia MA, (Aug. 1995). X-ray structure of calcineurin inhibited by the immunophilin-immunosuppressant FKBP12-FK506 complex. Cell 82, 507–522. [PubMed: 7543369]

(6). Li H, Zhang L, Rao A, Harrison SC, and Hogan PG, (June 2007). Structure of Calcineurin in Complex with PVIVIT Peptide: Portrait of a Low-affinity Signalling Interaction. J. Mol. Biol 369, 1296–1306. [PubMed: 17498738]

(7). Ye Q, Wang H, Zheng J, Wei Q, and Jia Z, (Apr. 2008). The complex structure of calmodulin bound to a calcineurin peptide. Proteins: Structure, Function and Genetics 73, 19–27.

(8). Jin L, and Harrison SC, (Oct. 2002). Crystal structure of human calcineurin complexed with cyclosporin A and human cyclophilin. Proc. Natl. Acad. Sci 99, 13522–13526. [PubMed: 12357034]

(9). Sheftic SR, Page R, and Peti W, (Dec. 2016). Investigating the human Calcineurin Interaction Network using the i LxVP SLiM. Sci. Rep 6, 38920. [PubMed: 27974827]

(10). Takeuchi K, Roehrl MH, Sun ZYJ, and Wagner G, (May 2007). Structure of the Calcineurin-NFAT Complex: Defining a T Cell Activation Switch Using Solution NMR and Crystal Coordinates. Structure 15, 587–597. [PubMed: 17502104]

(11). Ye Q, Li X, Wong A, Wei Q, and Jia Z, (2006). Structure of Calmodulin Bound to a Calcineurin Peptide: A New Way of Making an Old Binding Mode. Biochemistry 45, 738–745. [PubMed: 16411749]

(12). Majava V, and Kursula P, (Apr. 2009). Domain swapping and different oligomeric states for the complex between calmodulin and the calmodulin-binding domain of calcineurin A. PLoS One 4, ed. by Ratna, B., e5402. [PubMed: 19404396]

(13). Manalan AS, and Klee CB, (1983). Activation of calcineurin by limited proteolysis. Proc. Natl. Acad. Sci 80, 4291–4295. [PubMed: 6576338]

(14). Rumi-Masante J, Rusinga FI, Lester TE, Dunlap TB, Williams TD, Dunker AK, Weis DD, and Creamer TP, (Jan. 2012). Structural basis for activation of calcineurin by calmodulin. J. Mol. Biol 415, 307–317. [PubMed: 22100452]

(15). Shen X, Li H, Ou Y, Tao W, Dong A, Kong J, Ji C, and Yu S, (2008). The Secondary Structure of Calcineurin Regulatory Region and Conformational Change Induced by Calcium/Calmodulin Binding. J. Biol. Chem 283, 11407–11413. [PubMed: 18296442]

(16). Fruman DA, Klee CB, Bierer BE, and Burakoff SJ, (May 1992). Calcineurin phosphatase activity in T lymphocytes is inhibited by FK 506 and cyclosporin A. Proc. Natl. Acad. Sci 89, 3686–3690. [PubMed: 1373887]

(17). Parsons JN, Wiederrecht GJ, Salowe S, Burbaum JJ, Rokosz LL, Kincaid RL, and O'Keefe SJ, (July 1994). Regulation of calcineurin phosphatase activity and interaction with the FK-506?? FK-506 binding protein complex. J. Biol. Chem 269, 19610–19616. [PubMed: 7518461]

(18). Li J, Jia ZG, Zhou WC, and Wei Q, (Nov. 2009). Calcineurin regulatory subunit B is a unique calcium sensor that regulates calcineurin in both calcium-dependent and calcium-independent manner. Proteins: Struct., Funct., Bioinf 77, 612–623.

(19). Harish BM, Saraswathi R, Vinod D, and Devaraju KS, (May 2016). Discovery of a latent calcineurin inhibitory peptide from its autoinhibitory domain by docking, dynamic simulation, and in vitro methods. J. Biomol. Struct. Dyn 34, 983–992. [PubMed: 26111023]

(20). Rodríguez A, Roy J, Martínez-Martínez S, López-Maderuelo MD, Niño-Moreno P, Ortí L, Pantoja-Uceda D, Pineda-Lucena A, Cyert MS, and Redondo JM, (Mar. 2009). A Conserved Docking Surface on Calcineurin Mediates Interaction with Substrates and Immunosuppressants. Mol. Cell 33, 616–626. [PubMed: 19285944]

(21). Grigoriu S, Bond R, Cossio P, Chen JA, Ly N, Hummer G, Page R, Cyert MS, and Peti W, (Feb. 2013). The Molecular Mechanism of Substrate Engagement and Immunosuppressant Inhibition of Calcineurin. PLoS Biol 11, ed. by Petsko, G. A., e1001492. [PubMed: 23468591]

(22). Zhao Y, Zhang J, Shi X, Li J, Wang R, Song R, Wei Q, Cai H, and Luo J, (Aug. 2016). Quercetin targets the interaction of calcineurin with LxVP-type motifs in immunosuppression. Biochimie 127, 50–58. [PubMed: 27109380]

(23). Song R, Li J, Zhang J, Wang L, Tong L, Wang P, Yang H, Wei Q, Cai H, and Luo J, (Nov. 2017). Peptides derived from transcription factor EB bind to calcineurin at a similar region as the NFAT-type motif. Biochimie 142, 158–167. [PubMed: 28890387]

(24). Nagulapalli M, Parigi G, Yuan J, Gsponer J, Deraos G, Bamm VV, Harauz G, Matsoukas J, De Planque MR, Gerothanassis IP, Babu MM, Luchinat C, and Tzakos AG, (Mar. 2012). Recognition pliability is coupled to structural heterogeneity: A calmodulin intrinsically disordered binding region complex. Structure 20, 522–533. [PubMed: 22405011]

(25). Gsponer J, Christodoulou J, Cavalli A, Bui JM, Richter B, Dobson CM, and Vendruscolo M, (May 2008). A Coupled Equilibrium Shift Mechanism in Calmodulin-Mediated Signal Transduction. Structure 16, 736–746. [PubMed: 18462678]

(26). Liu F, Chu X, Lu HP, and Wang J, (May 2017). Molecular mechanism of multispecific recognition of Calmodulin through conformational changes. Proc. Natl. Acad. Sci 114, E3927–E3934. [PubMed: 28461506]

(27). Sun B, Cook EC, Creamer TP, and Kekenes-Huskey PM, (2018). Electrostatic control of calcineurin's intrinsically-disordered regulatory domain binding to calmodulin. Biochim. Biophys. Acta, Gen. Subj 1862, 2651–2659. [PubMed: 30071273]

(28). Dunlap TB, Guo H-F, Cook EC, Holbrook E, Rumi-Masante J, Lester TE, Colbert CL, Vander Kooi CW, and Creamer TP, (2014). Stoichiometry of the Calcineurin Regulatory Domain Calmodulin Complex. Biochemistry 53, 5779–5790. [PubMed: 25144868]

(29). Nevola L, and Giralt E, (2015). Modulating protein protein interactions: the potential of peptides. Chem. Commun 51, 3302–3315.

(30). Jochim AL, and Arora PS, (Aug. 2009). Assessment of helical interfaces in protein-protein interactions. Mol. Biosyst 5, 924–926. [PubMed: 19668855]

(31). Bullock BN, Jochim AL, and Arora PS, (2011). Assessing helical protein interfaces for inhibitor design. J. Am. Chem. Soc 133, 14220–14223. [PubMed: 21846146]

(32). Pierce BG, Wiehe K, Hwang H, Kim B-H, Vreven T, and Weng Z, (2014). ZDOCK server: interactive docking prediction of protein protein complexes and symmetric multimers. Bioinformatics 30, 1771–1773. [PubMed: 24532726]

(33). Chaudhury S, Berrondo M, Weitzner BD, Muthu P, Bergman H, and Gray JJ, (2011). Benchmarking and analysis of protein docking performance in Rosetta v3.2. PLoS One 6, 10.1371/journal.pone.0022477.

(34). Hu X, Lee MS, and Wallqvist A, (Dec. 2009). Interaction of the disordered Yersinia effector protein YopE with its cognate chaperone SycE. Biochemistry 48, 11158–11160. [PubMed: 19877667]

(35). Schiffer JM, Malmstrom RD, Parnell J, Ramirez-Sarmiento C, Reyes J, Amaro RE, and Komives EA, (Aug. 2016). Model of the Ankyrin and SOCS Box Protein, ASB9, E3 Ligase Reveals a Mechanism for Dynamic Ubiquitin Transfer. Structure 24, 1248–1256. [PubMed: 27396830]

(36). Bui JM, and Gsponer J, (Aug. 2014). Phosphorylation of an intrinsically disordered segment in Ets1 shifts conformational sampling toward binding-competent substates. Structure 22, 1196–1203. [PubMed: 25017730]

(37). Van Lierop JE, Wilson DP, Davis JP, Tikunova S, Sutherland C, Walsh MP, and David Johnson J, (2002). Activation of smooth muscle myosin light chain kinase by calmodulin. Role of LYS30 and GLY40. J. Biol. Chem 277, 6550–6558. [PubMed: 11748245]

(38). Zhi G, Abdullah SM, and Stull JT, (Apr. 1998). Regulatory segments of Ca2+/calmodulin-dependent protein kinases. J. Biol. Chem 273, 8951–8957. [PubMed: 9535879]

(39). Soderling TR, and Stull JT, (2001). Structure and regulation of calcium/calmodulin-dependent protein kinases. Chem. Rev 101, 2341–2351. [PubMed: 11749376]

(40). Case D, et al. Amber 16, University of California, San Francisco., Apr. 2016.

(41). Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, and Shaw DE, (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins: Struct., Funct., Bioinf 78, 1950–1958.

(42). Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, and Woods RJ, The Amber biomolecular simulation programs. J. Comput. Chem 26.

(43). Hawkins GD, Cramer CJ, and Truhlar DG, (1996). Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. J. Phys. Chem 100, 19824–19839.

(44). Patriksson A, and van der Spoel D, (2008). A temperature predictor for parallel tempering simulations. Phys. Chem. Chem. Phys 10, 2073–2077. [PubMed: 18688361]

(45). Zhang W, Wu C, and Duan Y, (2005). Convergence of replica exchange molecular dynamics. The Journal of Chemical Physics 123, 154105. [PubMed: 16252940]

(46). Patel S, Vierling E, and Tama F, (June 2014). Replica exchange molecular dynamics simulations provide insight into substrate recognition by small heat shock proteins. Biophys. J 106, 2644–2655. [PubMed: 24940782]

(47). Ryckaert J-P, Ciccotti G, and Berendsen HJ, (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys 23, 327–341.

(48). Mintseris J, Pierce B, Wiehe K, Anderson R, Chen R, and Weng Z, (2007). Integrating statistical pair potentials into protein complex prediction. Proteins: Struct., Funct., Bioinf 69, 511–520.

(49). Chen R, and Weng Z, (May 2002). Docking unbound proteins using shape complementarity, desolvation, and electrostatics. Proteins: Structure, Function and Genetics 47, 281–294.

(50). Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, and Klein ML, (1983). Comparison of simple potential functions for simulating liquid water. J. Chem. Phys 79, 926–935.

(51). Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, and Simmerling C, (Aug. 2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J. Chem. Theory Comput 11, 3696–3713. [PubMed: 26574453]

(52). Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, and Haak JR, (1984). Molecular dynamics with coupling to an external bath. J. Chem. Phys 81, 3684–3690.

(53). Malmstrom RD, Lee CT, Van Wart AT, and Amaro RE, (2014). Application of Molecular-Dynamics Based Markov State Models to Functional Proteins. J. Chem. Theory Comput 10, 2648–2657. [PubMed: 25473382]

(54). Genheden S, and Ryde U, (May 2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin. Drug Discovery 10, 449–461.

(55). Immadisetty K, and D Madura J, (2013). A review of monoamine transporter-ligand interactions. Curr. Comput. Aided. Drug. Des 9, 556–568. [PubMed: 24138394]

(56). Roe DR, and Cheatham TE, (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J. Chem. Theory Comput 9, 3084–3095. [PubMed: 26583988]

(57). Kabsch W, and Sander C, (Dec. 1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen bonded and geometrical features. Biopolymers 22, 2577–2637. [PubMed: 6667333]

(58). Fiorin G, Klein ML, and Hénin J, (Dec. 2013). Using collective variables to drive molecular dynamics simulations. Mol. Phys 111, 3345–3362.

(59). Cho MJ, Vaghy PL, Kondo R, Lee SH, Davis JP, Rehl R, Heo WD, and David Johnson J, (1998). Reciprocal regulation of mammalian nitric oxide synthase and calcineurin by plant calmodulin isoforms. Biochemistry 37, 15593–15597. [PubMed: 9843363]

(60). Kursula P, (Oct. 2014). The many structural faces of calmodulin: A multitasking molecular jackknife. Amino Acids 46, 2295–2304. [PubMed: 25005783]

(61). Das RK, and Pappu RV, (2013). Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. Proc. Natl. Acad. Sci 110, 13392–13397. [PubMed: 23901099]

(62). Das RK, Ruff KM, and Pappu RV, (June 2015). Relating sequence encoded information to form and function of intrinsically disordered proteins. Curr. Opin. Struct. Biol 32, 102–112. [PubMed: 25863585]

(63). Guharoy M, and Chakrabarti P, (Aug. 2007). Secondary structure based analysis and classification of biological interfaces: Identification of binding motifs in protein-protein interactions. Bioinformatics 23, 1909–1918. [PubMed: 17510165]

(64). de A Marques M, Parvatiyar MS, Yang W, de Oliveira G, and Pinto JR, (Dec. 2018). The missing links within troponin. Arch. Biochem. Biophys 663, 95–100. [PubMed: 30584890]

(65). Dunlap TB, Kirk JM, Pena EA, Yoder MS, and Creamer TP, (Apr. 2013). Thermodynamics of binding by calmodulin correlates with target peptide α-helical propensity. Proteins: Struct., Funct., Bioinf 81, 607–612.

(66). Zamoon J, Nitu F, Karim C, Thomas DD, and Veglia G, (2005). Mapping the interaction surface of a membrane protein: unveiling the conformational switch of phospholamban in calcium pump regulation. Proc. Natl. Acad. Sci 102, 4747. [PubMed: 15781867]

(67). Zhou G, Pantelopulos GA, Mukherjee S, and Voelz VA, (Aug. 2017). Bridging Microscopic and Macroscopic Mechanisms of p53-MDM2 Binding with Kinetic Network Models. Biophys. J 113, 785–793. [PubMed: 28834715]

(68). Paul F, Noé F, and Weikl TR, (May 2018). Identifying Conformational-Selection and Induced-Fit Aspects in the Binding-Induced Folding of PMI from Markov State Modeling of Atomistic Simulations. J. Phys. Chem. B 122, 5649–5656. [PubMed: 29522679]

(69). Grishaev A, Anthis NJ, and Clore GM, (2012). Contrast-matched small-angle x-ray scattering from a heavy-atom-labeled protein in structure determination: Application to a lead-substituted calmodulin-peptide complex. J. Am. Chem. Soc 134, 14686–14689. [PubMed: 22908850]

(70). Dogan J, Gianni S, and Jemth P, (Mar. 2014). The binding mechanisms of intrinsically disordered proteins. Phys. Chem. Chem. Phys 16, 6323–6331. [PubMed: 24317797]

(71). Hoeflich KP, and Ikura M, (Mar. 2002). Calmodulin in action: Diversity in target recognition and activation mechanisms 108, 739–742.

(72). Blumenthal DK, Takio K, Edelman AM, Charbonneau H, Titani K, Walsh KA, and Krebs EG, (1985). Identification of the calmodulin-binding domain of skeletal muscle myosin light chain kinase. Proc. Natl. Acad. Sci 82, 3187–3191. [PubMed: 3858814]

(73). Fitzsimons DP, Herring BP, Stull JT, and Gallagher PJ, (Nov. 1992). Identification of basic residues involved in activation and calmodulin binding of rabbit smooth muscle myosin light chain kinase. J. Biol. Chem 267, 23903–23909. [PubMed: 1429728]

(74). Siddiqui JK, Tikunova SB, Walton SD, Liu B, Meyer M, de Tombe PP, Neilson N, Kekenes-Huskey PM, Salhi HE, Janssen PM, Biesiadecki BJ, and Davis JP, (Dec. 2016). Myofilament Calcium Sensitivity: Consequences of the effective concentration of troponin I. Front. Physiol 7, 632. [PubMed: 28066265]

(75). Van Valen D, Haataja M, and Phillips R, (2009). Biochemistry on a Leash: The Roles of Tether Length and Geometry in Signal Integration Proteins. Biophys. J 96, 1275–1292. [PubMed: 19217847]

(76). Shoemaker BA, Portman JJ, and Wolynes PG, (2000). Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proc. Natl. Acad. Sci 97, 8868. [PubMed: 10908673]

(77). Fruman DA, Pai SY, Burakoff SJ, and Bierer BE, (July 1995). Characterization of a mutant calcineurin A alpha gene expressed by EL4 lymphoma cells. Mol. Cell. Biol 15, 3857–3863. [PubMed: 7791792]

(78). Sagoo JK, Fruman DA, Wesselborg S, Walsh CT, and Bierer BE, (Dec. 1996). Competitive inhibition of calcineurin phosphatase activity by its autoinhibitory domain. Biochem. J 320 (Pt 3, 879–84. [PubMed: 9003375]

(79). Borcherds W, Becker A, Chen L, Chen J, Chemes LB, and Daughdrill GW, (May 2017). Optimal Affinity Enhancement by a Conserved Flexible Linker Controls p53 Mimicry in MdmX. Biophys. J 112, 2038–2042. [PubMed: 28487147]

(80). Stemmer PM, and Klee CB, (June 1994). Dual Calcium Ion Regulation of Calcineurin by Calmodulin and Calcineurin B. Biochemistry 33, 6859–6866. [PubMed: 8204620]

(81). Peng K, Radivojac P, Vucetic S, Dunker AK, and Obradovic Z, (Apr. 2006). Length-dependent prediction of protein in intrinsic disorder. BMC Bioinf 7, 208.

(82). Lin YS, Bowman GR, Beauchamp KA, and Pande VS, (Jan. 2012). Investigating how peptide length and a pathogenic mutation modify the structural ensemble of amyloid beta monomer. Biophys. J 102, 315–324. [PubMed: 22339868]

(83). Fluitt AM, and De Pablo JJ, (Sept. 2015). An Analysis of Biomolecular Force Fields for Simulations of Polyglutamine in Solution. Biophys. J 109, 1009–1018. [PubMed: 26331258]

(84). Henriques J, Cragnell C, and Skep M, (2015). Molecular Dynamics Simulations of Intrinsically Disordered Proteins: Force Field Evaluation and Comparison with Experiment. J. Chem. Theory Comput 11, 3420–3431. [PubMed: 26575776]

(85). Guo Z, Mohanty U, Noehre J, Sawyer TK, Sherman W, and Krilov G, (Apr. 2010). Probing the α-helical structural stability of stapled p53 peptides: Molecular dynamics simulations and analysis: Research article. Chem. Biol. Drug Des 75, 348–359. [PubMed: 20331649]

(86). Miller Y, Ma B, and Nussinov R, (May 2010). Zinc ions promote Alzheimer A aggregation via population shift of polymorphic states. Proc. Natl. Acad. Sci 107, 9490–9495. [PubMed: 20448202]

(87). Best RB, and Mittal J, (Apr. 2011). Free-energy landscape of the GB1 hairpin in all-atom explicit solvent simulations with different force fields: Similarities and differences. Proteins: Struct., Funct., Bioinf 79, 1318–1328.

(88). Qin Y, Liu J, Li X, and Wei Q, (Mar. 2005). Preparation and characterization of a single-chain calcineurin-calmodulin complex. Biochim. Biophys. Acta, Proteins Proteomics 1747, 171–178.

(89). Swulius MT, and Waxham MN, (Sept. 2008). Ca2+/calmodulin-dependent protein kinases. Cell. Mol. Life Sci 65, 2637–2657. [PubMed: 18463790]

(90). Myers JB, Zaegel V, Coultrap SJ, Miller AP, Bayer KU, and Reichow SL, (Aug. 2017). The CaMKII holoenzyme structure in activation-competent conformations. Nat. Commun 8, 15742. [PubMed: 28589927]

(91). Zhang M, Pascal JM, and Zhang J-F, (Mar. 2013). Unstructured to structured transition of an intrinsically disordered protein peptide in coupling Ca2+-sensing and SK channel activation. Proc. Natl. Acad. Sci 110, 4828–4833. [PubMed: 23487779]

(92). Towns J, Cockerill T, Dahan M, Foster I, Gaither K, Grimshaw A, Hazlewood V, Lathrop S, Lifka D, Peterson GD, Roskies R, Scott JR, and Wilkins-Diehr N, (Sept. 2014). XSEDE: Accelerating Scientific Discovery. Comput. Sci. Eng 16, 62–74.
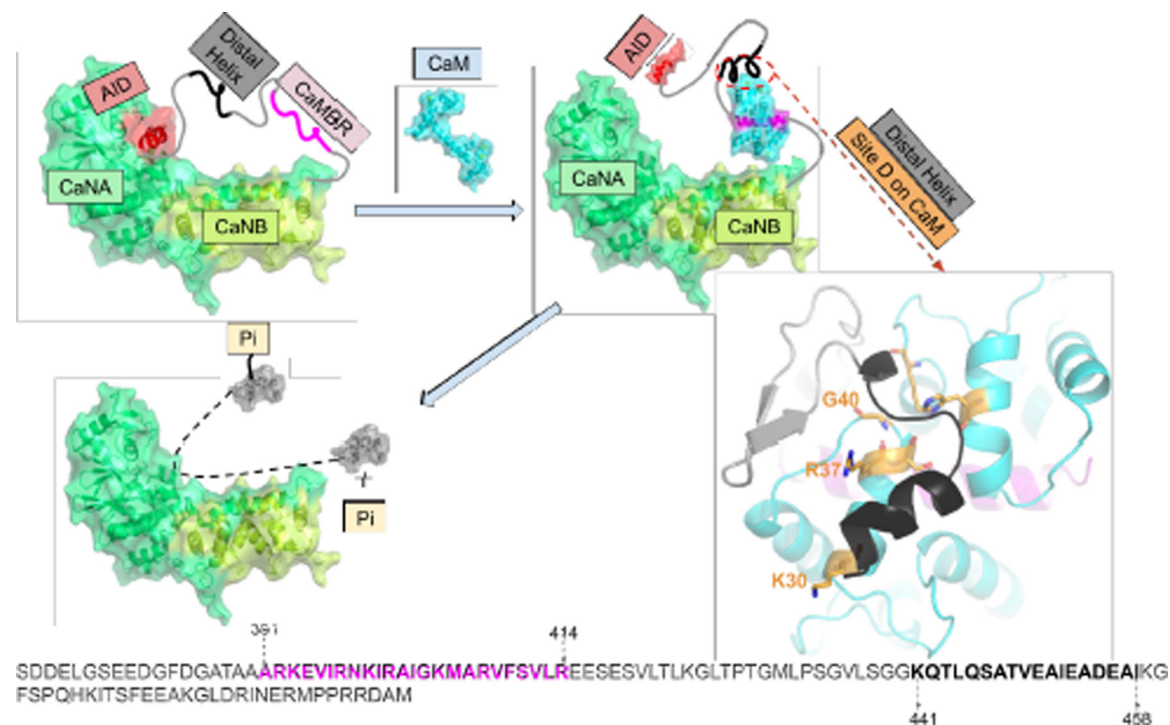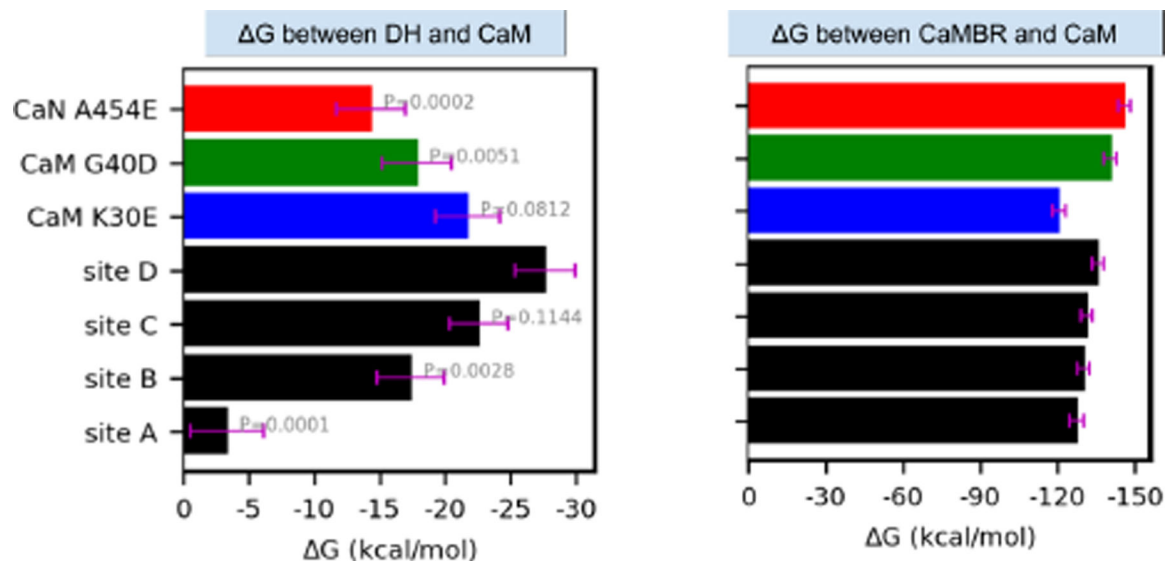
Figure 1:
Refined model of calcineurin (CaN) activation by calmodulin (CaM) through direct binding of the 'distal helix' to CaM, based on the mechanism initially proposed in (1). The two chains of CaN (CaNA and CaNB) are colored in limegreen and lime, respectively. AID is colored in red. CaM is colored in cyan, CaMBR is colored in magenta. The amino acid sequence of CaN RD is shown at the bottom of the panel with CaMBR and the distal helix region colored in magenta and black, respectively. In the absence of CaM, CaN is inhibited by its auto-inhibitory domain (AID). After CaM binds the CaM binding region (CaMBR) in CaN's regulatory domain, a secondary interaction between CaM and a 'distal helix' ultimately removes the AID from the CaN catalytic site. The activated CaN catalyzes the dephosphorylation of target proteins essential to myriad physiological functions.
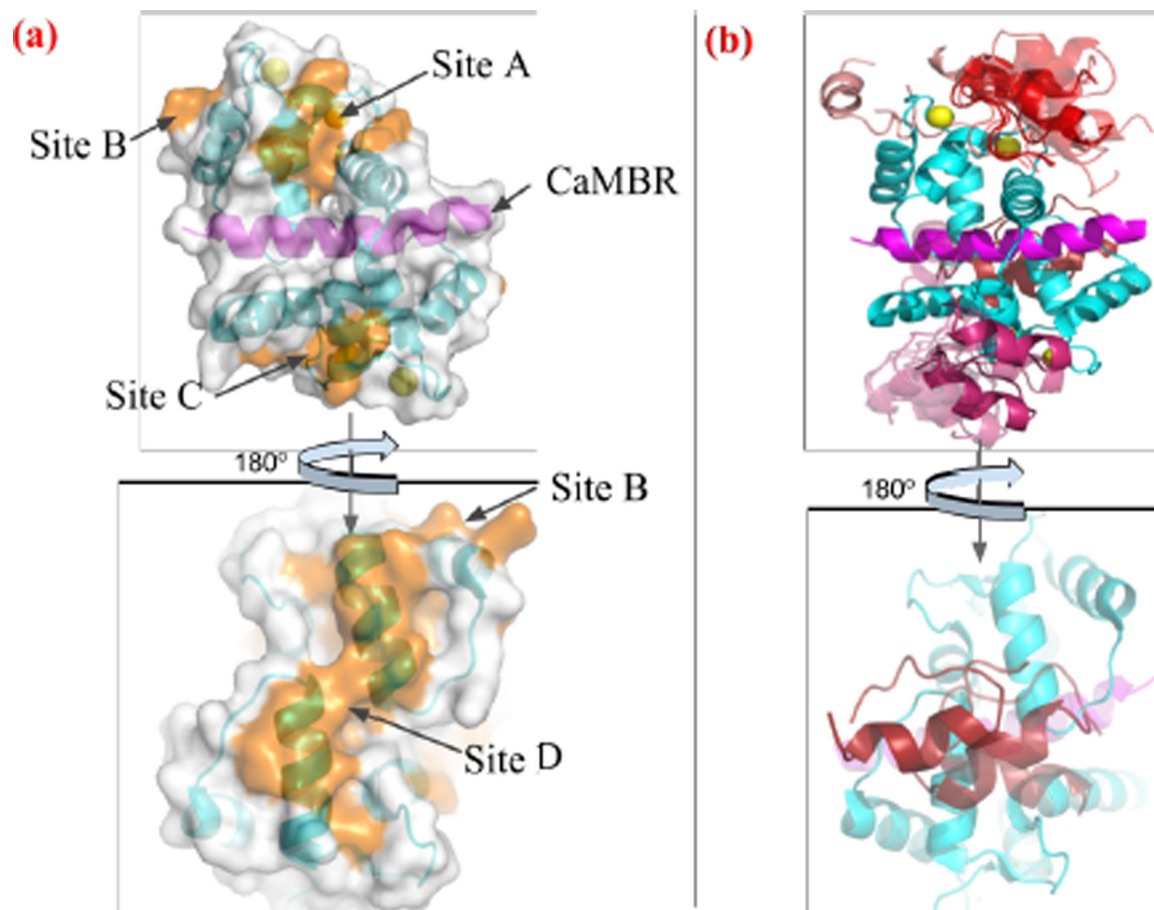
Figure 2:
(a) Four tentative binding sites (orange) on the surface of CaM-CaMBR complex. CaM is colored in cyan, CaMBR is colored in magenta and Ca^{2+} ions are colored in yellow. (b) ZDOCK predicted conformations of the distal helix interacting with CaM/CaMBR complex at each site. Predicted distal helix conformations from site A to D are colored as red, salmon, warmpink and firebrick, respectively.
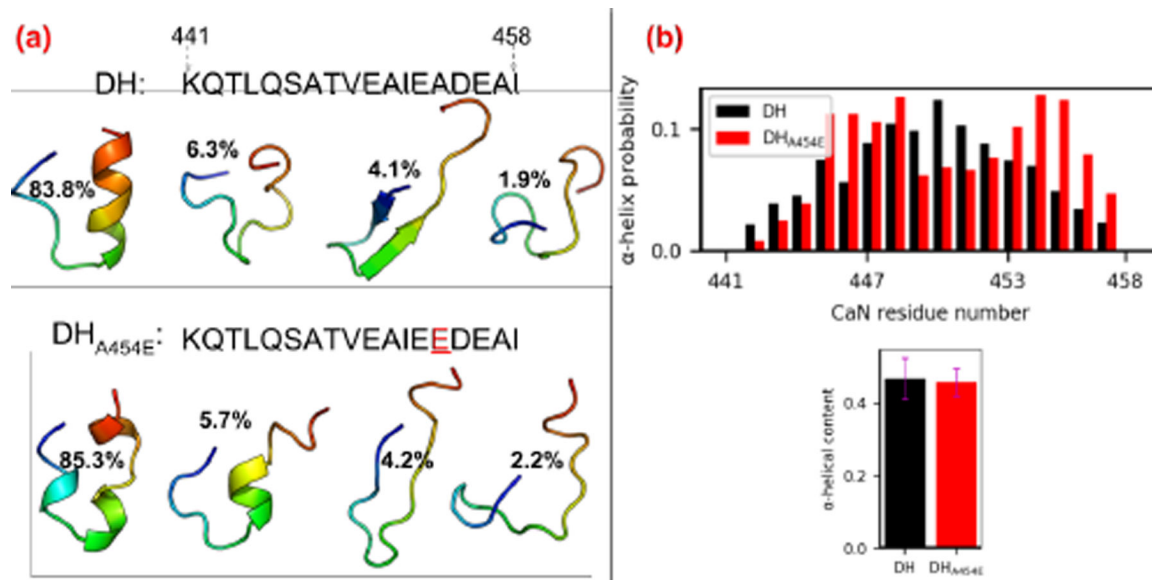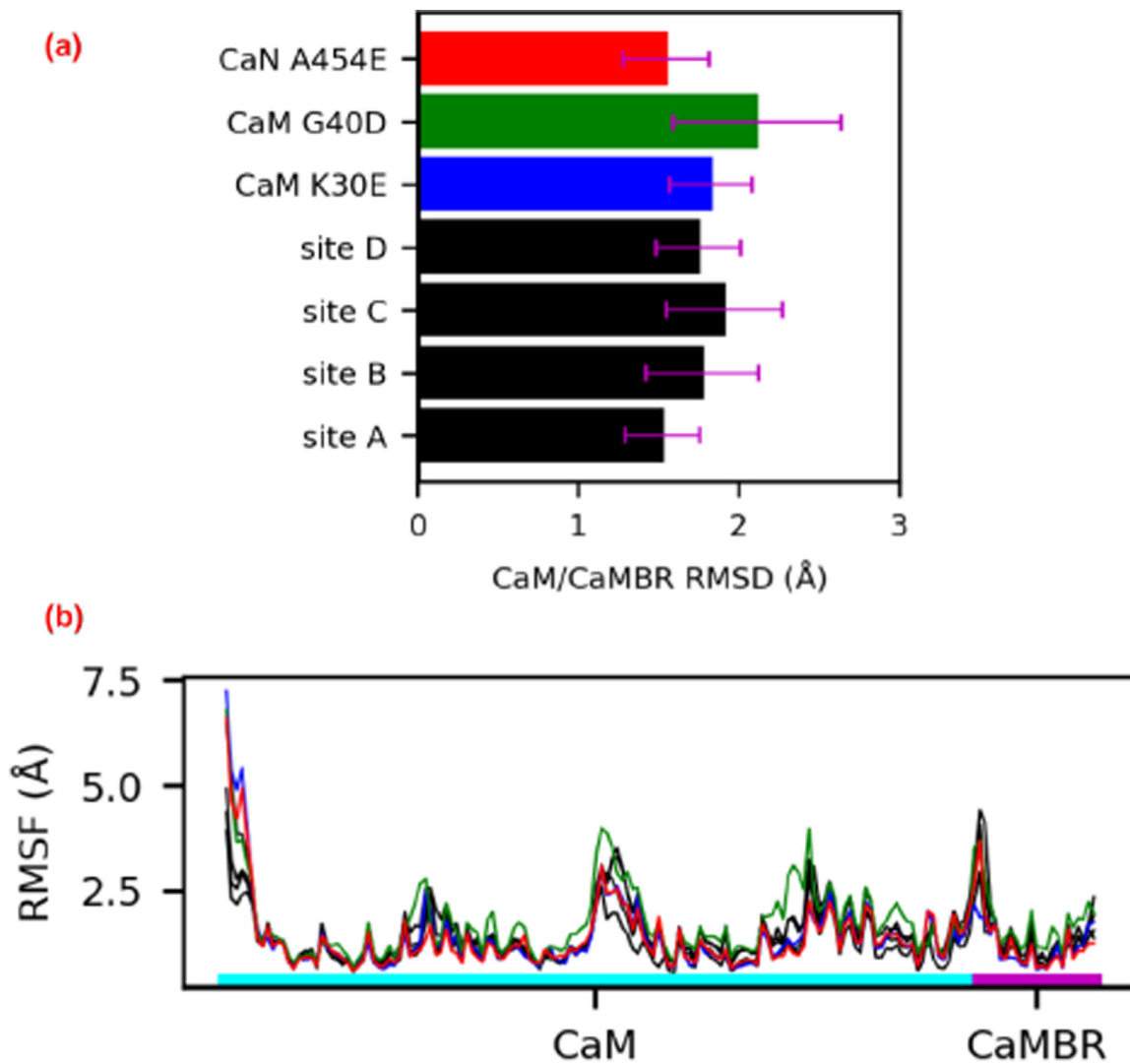
Figure 3:
(a) Sequence of distal helix/DH$_{A454E}$ and representative structures of the four most populated clusters from 100 ns REMD simulations. The structures are colored in rainbow with N-terminus as blue and C-terminus as red. (b) Secondary structure probability of each residue calculated from REMD trajectory via CPPTRAJ with DSSP algorithm. The lower panel shows the total α-helix content of two fragments calculated via the COLVAR module of VMD.
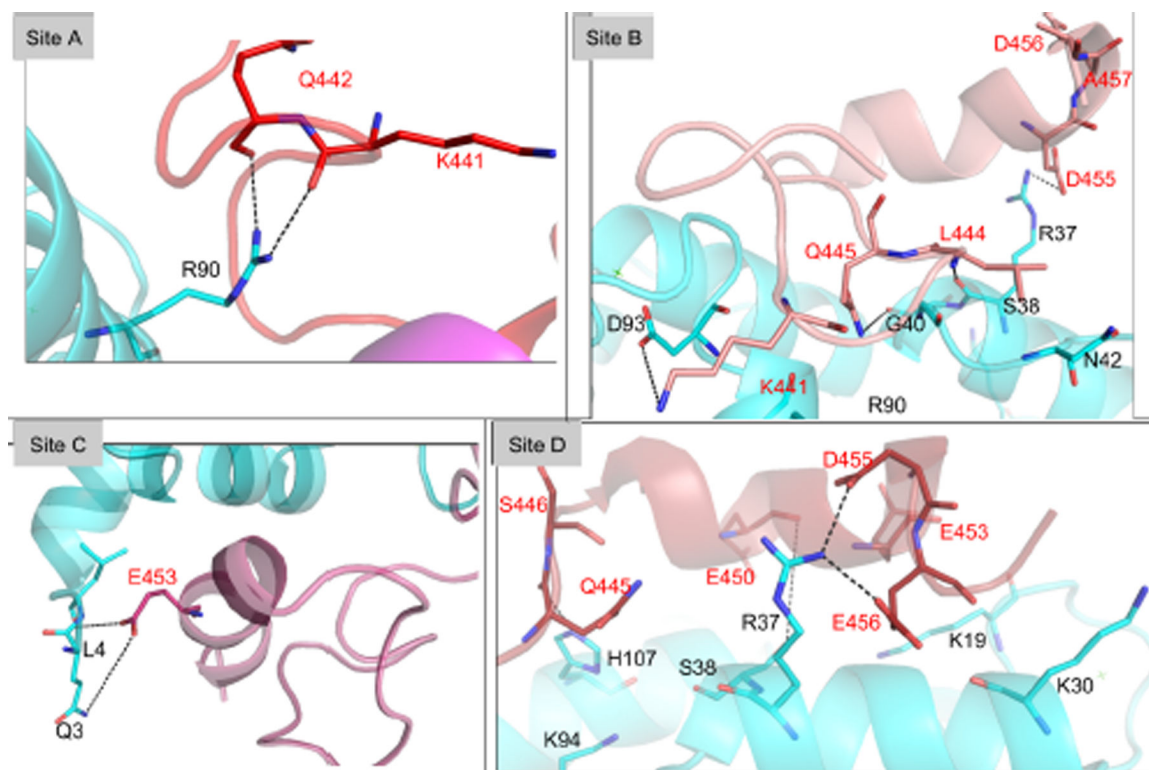
Figure 4:

Approximate binding free energies between CaM and the distal helix (left) or CaMBR regions (right) via Molecular Mechanics-Generalized Born and Surface Area continuum solvation (MM-GBSA). Black bars correspond to wild-type CaN, whereas colored bars utilize the A454E CaN and CaM variants. The calculation was conducted on frames extracted every 2 ns from MD trajectories. The error bar represents standard error of the mean. The values above bars in the left panel are P values of each case with null hypothesis that their mean values are equal to site D.
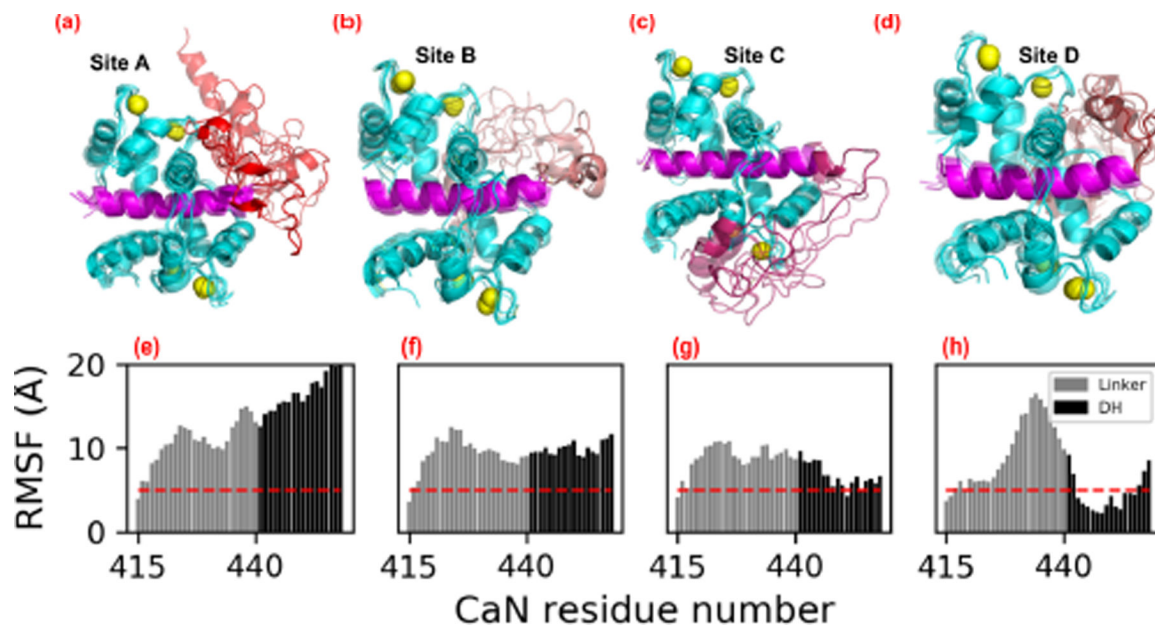
Figure 5:
(a) Root mean squared deviations (RMSD) of the peptide backbone atoms of CaM and CaMBR from μs-length MD simulations. The reference structure for the RMSD calculation was the CaM/CaMBR crystal structure (PDB ID: 4q5u). (b) Root mean squared fluctuations (RMSF) of non-hydrogen atoms in CaM and CaMBR.
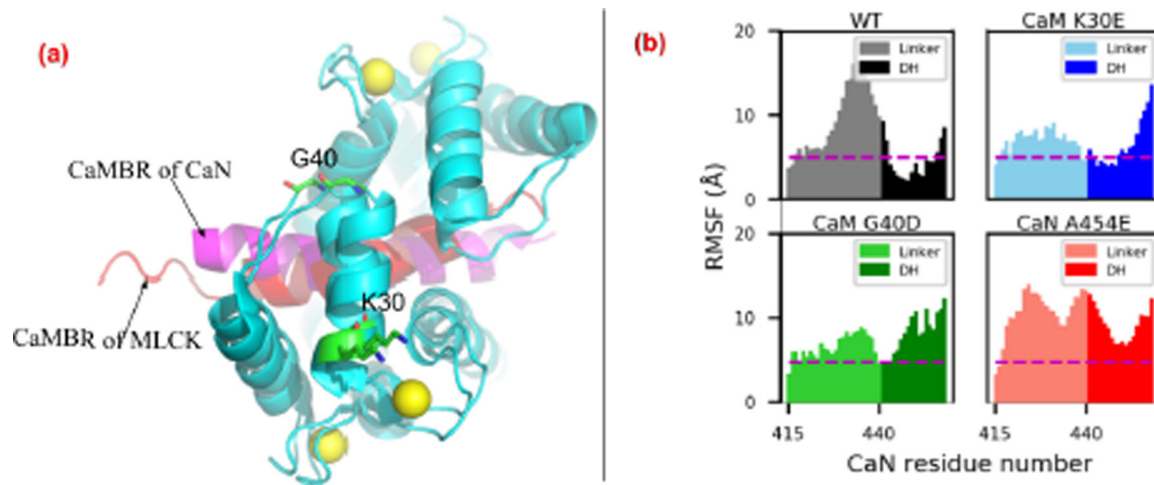
Figure 6:
Interaction between the linker/distal helix of CaN and CaM at sites A-D. Key residues at the interaction surface are shown in sticks with black labels for CaM residues and red labels for distal helix residues. See Table S4 for duration of these interactions.
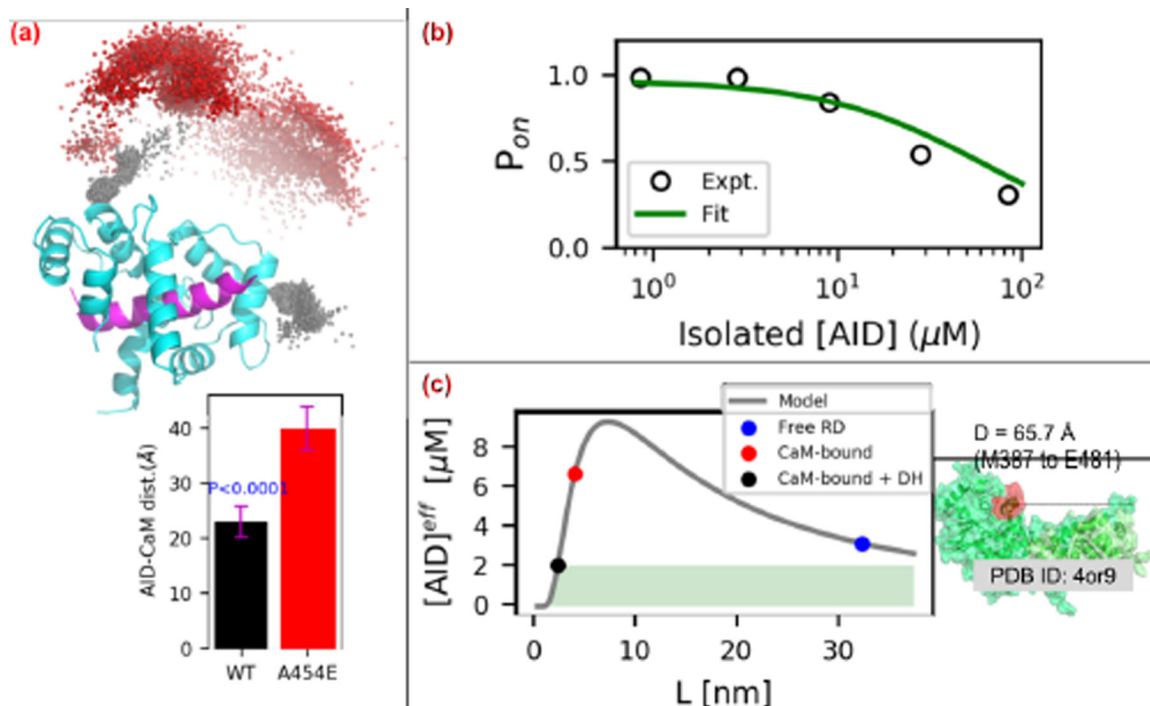
Figure 7:

(a–d) Representative structures of distal helix/CaM from microsecond length MD simulations initialized from ZDOCK-predicted distal helix poses. CaM is colored in cyan, CaMBR is colored in magenta and Ca^{2+} ions are depicted as yellow spheres. The linker and distal helix regions in site A–D are colored as red, salmon, warmpink and firebrick, respectively. (e–h) Non-hydrogen atom RMSFs of linker and distal helix residue calculated from MD simulations of each site, as an indicator of binding stability. The red dash line depicts RMSF values of 5 Å. During the MD simulations, distal helix structures initiated at site B migrated toward site D (Fig. S3).

Figure 8:
(a) Comparison of CaM-petide complex structure from CaN and MLCK (PDB ID: 2lv6 (69)). K30 and G40 are labeled (shown as sticks) based on their implication in the activation of the CaM target MLCK (37) and proximity to site D determined by our simulations. (b) Non-hydrogen RMSF of linker and distal helix in WT and mutants. The dash line depicts RMSF value as 5 Å. The hydrogen bonds formed between the distal helix and CaM of these variants are shown in Fig. S9.

Figure 9:
(a) Distribution of AID center of mass (COM) relative to the CaM/CaMBR complex. The black and red spheres represent the COMs of AID in WT and A454E cases, respectively. The lower panel depicts distance between COMs of AID and CaM. The P-value above the WT bar is for the null hypothesis that the WT center of mass is the same as that of A454E. (b) Fitting of the competitive-inhibitor model (Eq. 5) to experimental data from (77). (c) Effective AID concentrations calculated via Eq. 4. The shaded green area represents effective [AID] that leads to CaN's activation. Right panel illustrates the assumed distance between CaMBR and catalytic site. The value is set as 66 Å in this study.

Table 1:

Kinetic parameters of pNPP dephosphorylation with WT CaM and two site D variants. P-values given by Welch's t-test for difference of means with unequal variance.

CaM	K_M (mM)	SD	p-value
WT	27.6	1.3	-
K30E	46.0	2.8	0.002
G40D	35.5	2.2	0.008