Biophysical Journal, Volume *112*

Supplemental Information

Opposing Intermolecular Tuning of Ca^{2+} Affinity for Calmodulin by Neurogranin and CaMKII Peptides

Pengzhi Zhang, Swarnendu Tripathi, Hoa Trinh, and Margaret S. Cheung

Supporting Text

I. Sample preparation

I.1 Intact Ng

Ng is a 78-residue intrinsically disordered protein with IQ motif (1, 2) composed of hydrophobic and basic amino acids. Available experimental measurement for residual structure of Ng is for the full-length protein Ng from mouse (3), therefore, the sequence is used in the modeling and parameterization of Ng, as shown below,

```
1               10              20              30              40              50
MDCCTESACSKPDDDILDIPLDDPGANAAAAKIQASFRGHMARKKIKSGE
                60              70
CGRKGPGPGGPGGAGGARGGAGGGPSGD 78
```

Following the work by Hoffman et al. (4), we divided the Ng protein into several units, including the acidic N terminal (residues 13-25), the IQ motif (residues 26-49) and the poly-Glycine C terminal (residues 50-78). The underscored sequence stands for the IQ motif which partially retains residual structure (residues 25-42) in the unbound state. To reproduce the residual structure, replica exchange method (REM) (5, 6) was employed for calculation of average helicity of the residual structure and the backbone model-free order parameter S^2.

I.2 $Ng_{13\text{-}49}$ peptide

Hoffman et al. showed from calorimetry and NMR experiments that the IQ motif alone does not reproduce similar Ng-mediated affinity of Ca^{2+} for Ca^{2+}-free CaM (apoCaM), or the pattern of intermolecular interaction between Ng and apoCaM (4). A combination of the acidic region in the N-terminal and the IQ motif yields the minimum composition of Ng for its function.

Therefore the Ng_{13-49} peptide is used to study the interaction with apoCaM (PDB ID: 1CFD) that allows the comparison of the results from computer simulations with several experimental measurements, including dissociation constant of apoCaM and Ng_{13-49}, the affinity of Ca^{2+} to CaM, and the changes in the chemical shifts of apoCaM upon Ng_{13-49} binding (4). The sequence of the Ng_{13-49} peptide is provided below,

```
                20                  30                  40
13 DDDILDIPLDDPGANAAAAKIQASFRGHMARKKIKSG 49
```

II. Coarse-grained protein or peptide models

II. 1 Hamiltonian and parametrization of the coarse-grained protein models

As described above, Ng is partly ordered in solution. The fragment G25-A42 has a fraction of 22% and 28% residual structure calculated from the C_α and the H_α chemical shifts in the nuclear magnetic resonance (NMR) experiment (3), respectively. The rest part of the peptide remains unstructured. In order to compare with results from several experiments (3, 4), peptide Ng_{13-49} comprising the IQ domain with adjacent acidic region as well as the full sequence Ng were used in the simulations (please find the sequences in session I from the *Supporting Information*.).

We adopted a side-chain-C_α coarse-grained model to represent protein/peptides developed by Cheung et al (7). In this model, each amino acid (except glycine, which is represented by a C_α bead) is represented by two beads: the C_α bead is located at the C_α position to represent the backbone atoms; the side-chain bead is located at the center of mass of the side-chain atoms to represent the side-chain atoms. The use of side-chain-C_α model enables us to capture large structural fluctuations of the proteins at a large timescale that warrants a wide search of the phase space.

The total potential energy E of Ng or $Ng_{13\text{-}49}$ is given by,

$$E^{Ng} = E_{bond} + E_{angle} + E_{dihedral} + E_{chirality} + E_{elec} + E_{HB} + E_{LJ} \qquad \text{Eq. (S1)}$$

The bonded interactions include the bond energy (E_{bond}), the bond-angle energy (E_{angle}), and the dihedral-angle energy ($E_{dihedral}$) the L-isomer restraint of chirality ($E_{chirality}$). E_{bond}, E_{angle}, and $E_{chirality}$ constrain the bond length, bond angle and side-chain chirality through harmonic potentials. The formula can be found in our previous work (8). The equilibrium bond length, bond angle and side-chain chirality parameters were taken from the crystal structure of apoCaM and IQ motif of Ng (PDB ID: 405E), and those of the missing segment were obtained from the structures predicted by the Sparks-X protein structure prediction server (9). Since these terms are mainly constrained by chemical rules, they do not vary significantly by the conformations.

The electrostatic energy (E_{elec}) between each two beads (C_α or side-chain bead) i and j with partial charges is described by Debye-Hückel potential (10) to include the screening effect of the electrolyte solution.

$$E_{elec}^{ij} = \frac{q_i q_j}{4\pi\epsilon_r\epsilon_0 r_{ij}} \exp\left(-r_{ij} \Big/ \sqrt{\frac{\epsilon_r\epsilon_0 k_B T}{2e^2 I}}\right) \qquad \text{Eq. (S2)}$$

where q_i (q_j) is the partial charges on bead i (j); e is the elementary charge (see the method of generating the partial charges in the next section II.2 from the *Supporting Information*); ϵ_r is the relative dielectric constant and is set to 80 for aqueous solution; ϵ_0 is the permittivity of the vacuum; r_{ij} is the distance between beads i and j; k_B is Boltzmann constant; T = 1.1 ε / k_B is the temperature; I = 0.15 M is the ionic strength.

For the backbone hydrogen bonding interactions, we adopted an angular-dependent function that captures directional properties of backbone hydrogen bonds (7),

$$E_{HB}^{ij} = \sum_{|j-i|>3} A(\rho)\varepsilon\left[\left(\frac{\rho_{ij}^0}{r_{ij}}\right)^6 - 2\left(\frac{\rho_{ij}^0}{r_{ij}}\right)^{12}\right] \quad \text{Eq. (S3)}$$

$$A(\rho) = \frac{1}{\left[1+(1-\cos^2\rho)(1-\frac{\cos\rho}{\cos\rho_\alpha})\right]^2} \quad \text{Eq. (S4)}$$

where the strength ε = 0.6 kcal/mol; r_{ij} is the distance between backbone beads *i* and *j*; and ρ_{ij}^0 = 4.6 Å is the typical length of a hydrogen bond; $A(\rho)$ measures the structural alignment of two interacting strands; ρ is the pseudo dihedral angle between backbone beads of the two interacting strands (7); ρ_α is the pseudo dihedral angle of a canonical helical turn, 0.466 rad. $A(\rho)$ preserves the tendency to form β-strands (ρ = 0 or π) or α-helices ($\rho = \rho_\alpha$) as it is maximized to 1 in either of the two cases.

For the non-bonded interactions (E_{LJ}), between C_α bead *i* from the backbone (bb) and side-chain (sc) bead *j*, a pure repulsive interaction was considered,

$$E_{LJ}^{sc-bb} = \varepsilon\left(\frac{\rho_{ij}^0}{r_{ij}}\right)^{12} \quad \text{Eq. (S5)}$$

where the strength ε = 0.6 kcal/mol; i $\rho_{ij}^0 = 0.9(\rho_i^0 + \rho_j^0)$, ρ_i^0 is the size of backbone bead *i*, 0.5 σ (σ = 3.8 Å) and ρ_j^0 is the size of side-chain bead *j*, which is the van der Waals radius of the side-chain. 0.9 is a scaling factor to remove clashes between bulky side-chains. For the non-bonded interactions between side-chain bead *i* and side-chain bead *j*, we applied the Lennard-Jones potential,

$$E_{LJ}^{sc-sc} = \varepsilon_{ij}\left[\left(\frac{\rho_{ij}^0}{r_{ij}}\right)^{12} - 2\left(\frac{\rho_{ij}^0}{r_{ij}}\right)^6\right] \quad \text{Eq. (S6)}$$

where ε_{ij} is solvent mediated interactions between the involved amino acids obtained from Betancourt-Thirumalai's study (11). $\rho_{ij}^0 = 0.9(\rho_i^0 + \rho_j^0)$, $\rho_{i(j)}^0$ is the van der Waals radius of side-chain bead *i* (*j*). r_{ij} is the distance between beads *i* and *j* in Eqs. S3, S5 and S6. Since the electrostatic interactions were explicitly included in the Hamiltonian, the Lennard-Jones potential coefficient ε_{ij} was rescaled as described in our previous work (8).

To best describe the intrinsic disorder nature of Ng or $Ng_{13\text{-}49}$, we employed a backbone dihedral angle potential that is sequence-specific but independent on protein topology. For this we used the model introduced by Karanicolas and Brooks (12), or the KB model. The dihedral angle between four adjacent α-carbons depends on the backbone dihedral angles of the two middle residues. Brooks' group produced a probability distribution for 400 possible ordered pairs of amino acid residues from a survey of the Protein Data Bank and thus related the probability distribution to potential energy ignoring the entropy contribution. The dihedral angle potential presents two minima corresponding to local α-helical and β-strand geometries. The statistical potential is modeled as a 4-term cosine series,

$$E_{dihedral}^{ijkl} = \varepsilon_{KB} \sum_{n=1}^{4} K_{ij,n} \left[1 + \cos(n\varphi_{ijkl} - \delta_{jk,n}) \right] \quad \text{(S7)}$$

where φ_{ijkl} is the dihedral angle formed by four consecutive C_α beads *i, j, k, l* with beads *j* and *k* in the middle, $K_{jk,n}$ and $\delta_{jk,n}$ are statistically determined constants. ε_{KB} is a factor to adjust the strength in relative to other interactions in the current model.

We tested a total of three types of dihedral potentials for a comparative study. The first dihedral angle potential is a sequence-based model as described above. The second dihedral angle potential is a structure-based potential, or SB model, where the structure of a specific

segment of Ng obtained from the crystal structure (PDB code: 4E50 for Ng) was used as the reference (please see Table S4 for the residues in this segment). The SB model is composed of two-term cosine-series,

$$E_{dihedral}^{ijkl} = \varepsilon_{SB} \sum_{n=1,3} k_{\varphi}^{n} \left[1 - \cos\left(n \times \left(\varphi_{ijkl} - \varphi_{ijkl}^{0}\right)\right)\right] \quad \text{Eq. (S8)}$$

where *i, j, k* and *l* are four consecutive C_α beads. φ_{ijkl} is the dihedral angle formed by those four beads. The equilibration position of the dihedral angle φ_{ijkl}^{0} for this specific segment G25-A42 was taken from the crystal structure (PDB code: 4E50 for Ng). $k_{\varphi}^{1} = 2k_{\varphi}^{3} = 2\varepsilon$. ε_{SB} is used to adjust the barrier of the dihedral angle potential. For the unstructured segment $\varepsilon_{SB} = 0$.

The third dihedral potential is the hybrid of the two models, or Hybrid model, as shown in Table S4. We replaced the dihedral angle potential of the unstructured segment (beyond G25-A42) from the SB model with a KB model.

Different barrier values of the structure-based dihedral potential ε_{SB} and the statistical dihedral potential ε_{KB} were tested. Fig. S7 shows the helicity of the fragment G25-A42 decreases with temperature in all the cases. At the temperature $T = 1.1\ \varepsilon/k_B$, for the SB model, the helicity fell within the experimental range when $\varepsilon_{SB} = 0.3$ (Fig. S7A); for the Hybrid model, the helical fraction fell within the experimental range with when $\varepsilon_{SB} = 0.3$, and ε_{KB} did not make any influence in the range from 1.8 to 2.2 (Fig. S7B); for the KB model, the helical fraction was best matched when ε_{KB} is in the range from 1.5 to 2.2 (Fig. S7C).

To determine which best represents the properties of Ng among all the three models, we further investigated them at the residue level. Therefore, for those models with parameters that matched the overall helicity (Fig. S7), we computed the nuclear magnetic resonance (NMR)

model free order parameter S^2 for the backbone beads from our coarse-grained simulations (please find the details of calculation in V.1 from the *Supporting Information*), and compared with data from the NMR experiments (3). The computed S^2 (Table S5) positively correlated with the experimental values in all cases. Interestingly, for the KB model, the correlation was overall higher than the SB and Hybrid model, especially for ε_{KB} = 1.8. Because Ng protein is intrinsically disordered (13), The KB dihedral angle potential of no bias to any specific structure enables sampling of a broad spectrum of Ng conformations, whereas the SB model and the Hybrid model of full or partial bias to a specific structure limit the flexibility to explore more conformations. Therefore, for modeling intrinsically disordered protein Ng, we adopted the sequence-based KB dihedral potential model. ε_{KB} = 1.8 was used in the following study of apoCaM-$Ng_{13\text{-}49}$ binding simulations.

II.2 Determination of partial charges for the coarse-grained models

II.2.a Neurogranin protein

We adopted a multi-scale method developed by Cheung group (14) to assign partial charges to the intact Ng protein. Firstly, we ran REMD simulations without electro-static interactions (please find the details in III.1 in the *Supporting Information*) and obtained the free energy surface $F(\Delta, \chi)$ as a function of asphericity Δ and overlap function χ. Δ measures the shape of the protein: it is like a rod or a sphere when Δ = 1 or 0, respectively (15). χ measures the similarity to the reference structure (16). Using the free energy, we thus selected about 400-600 frames from the REMD simulations of Ng protein through importance sampling (17). We reconstructed the coarse-grained structures into atomistic protein models (17), used H++ server (18) to predict the protonation states and computed the partial charges using the semi-empirical

quantum chemistry program MOPAC (19). We collected the partial charges on the backbone (side-chain) atoms as the charge for the C_α (sidechain) bead. We then obtained the averaged partial charges on the C_α and side-chain beads over all the structures. We repeated the same process for all the three models shown in Table S4 and the three sets of charges we obtained were highly similar to each other (the correlation coefficients were ~1.00). We therefore used only one set of charges for all the three models. The charges on Ng protein are provided in Table S9.

II.2.b Apo-calmodulin and neurogranin peptide system

We followed the same procedures as for the Ng protein. The free energy surface $F(d_{COM}, Z_{inter})$ was obtained from US simulations of apoCaM and $Ng_{13\text{-}49}$ without electrostatic interactions by using WHAM (20, 21) (please see details about the US simulations in III.2 in the *Supporting Information*). d_{COM} is the distance between center of mass of apoCaM and center of mass of $Ng_{13\text{-}49}$, and Z_{inter} is the total number of intermolecular contacts (please see the definition of a contact in V.3 from the *Supporting Information*). Two groups of partial charges were generated for apoCaM and $Ng_{13\text{-}49}$ according to the experimental conditions: the fluorescence experiments for measurement of the dissociation constant of apoCaM and $Ng_{13\text{-}49}$ were conducted at pH = 7.2 and at ionic strength I = 0.15 M; the NMR experiments for determining the change in chemical shifts of apoCaM upon binding $Ng_{13\text{-}49}$ were conducted at pH = 6.3 and at ionic strength I = 0.10 M. The protonation states of histidine residues were determined by the H++ server (18) before using MOPAC (19). The calculated partial charges of apoCaM and $Ng_{13\text{-}49}$ in the above two conditions were provided in Table S10-S13.

III. Coarse-grained Molecular Simulations

III.1 Replica Exchange Molecular Dynamics simulations

To study dynamics of unbound Ng protein, which has a rugged energy landscape because of its intrinsic disorder, we applied replica exchange molecular dynamics (REMD) (5, 6) to enhance the sampling. For the three dihedral angle models, a range of barriers of dihedral angel potential were investigated: for the SB (structure-based) model, ε_{SB} = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6 and 0.7; for the Hybrid model, ε_{SB} = 0.3, 0.4 for the part with residual structure (G25-A42), in combination with ε_{KB} = 1.8, 2.0 and 2.2 for the rest; for the KB model, ε_{KB} = 1.0, 1.5, 1.8, 2.0, 2.2, 2.5 and 3.0. For each of the models, 24 replicas were distributed at temperatures T = 0.80, 0.82, 0.83, 0.85, 0.87, 0.88, 0.90, 0.93, 0.97, 1.00, 1.03, 1.07, 1.10, 1.13, 1.17, 1.20, 1.23, 1.27, 1.30, 1.33, 1.40, 1.47, 1.57, 1.67 ε / k_B to produce ample exchanges between the replicas (ε is the reduced energy unit = 0.6 kcal/mol, and k_B is the Boltzmann constant). The average acceptance ratio of exchanging among the neighboring replicas ranges from 0.57 to 0.77. An exchange between neighboring replicas was attempted every 100,000 time steps. In order to explore more conformation space, in each REMD simulation, a total of 6 sets of distinct initial configurations were used, making up a total of 1200 exchange attempts for each model.

III.2 Umbrella sampling simulations

Umbrella sampling (US) method (22) was used to explore the thermodynamic properties of the apoCaM-Ng system. The distance between the center of mass of apoCaM and the center of mass of Ng_{13-49} d_{COM} was used as the reaction coordinate. d_{COM} was restrained by a harmonic force with spring constant 66.7 ε/σ^2 (σ = 3.8 Å is reduced length unit). The equilibrium positions of the harmonic force range from 0.2 σ to 20.0 σ with a bin size of 0.2 σ, making up a total of 100

windows. For each window, 10,000,000 time-steps of constrained molecular dynamic simulation were carried out.

To generate the initial structures at each window of d_{COM} for the US simulations, molecular dynamics simulations were carried out for the apoCaM-Ng_{13-49} complex from the bound form (PDB code: 4E50, in the coarse-grained model) at a high temperature $T = 1.33\ \varepsilon / k_B$ for dissociation (ε = 0.6 kcal/mol is reduced energy unit, and k_B is Boltzmann constant). A total of 5 sets of initial configurations for the following umbrella sampling simulations were generated from 5 dissociation trajectories.

US simulations were then performed without electrostatic interactions for generating ensemble of structures to generate ensemble of structures for partial charge determination (see 2.3 for the details). After we obtained the partial charges for apoCaM and Ng_{13-49}, US simulations were carried out with charge-charge interactions to study how the strength of non-electrostatic intermolecular interactions influences the binding affinity and to determine the optimal strength by comparing with the experiments (4). Using this strength of non-electrostatic intermolecular interactions, US simulations were carried up at appropriate experimental conditions to study the binding thermodynamic properties of apoCaM-Ng_{13-49} system. Each set of US simulations were performed from 5 different initial structures at the temperature $T = 1.1\ \varepsilon/k_B$. We analyzed the data using Weighted Histogram Analysis Method (WHAM) (20, 21).

IV. Calculation of difference in binding free energy using Jarzynski's equality from steered molecular dynamics simulations

IV.1 Selection of initial atomistic structures

holoCaM-Ng$_{13-49}$ complex: We employed coarse-grained molecular simulations to efficiently sample a broad ensemble of complex structures. We used the experimental measurements as a guide to strategically select several structures from the major cluster of coarse-grained complex structures. According to the crystal structure of apoCaM and IQ motif of Ng (NgIQ) (PDB ID: 4E50), the two EF-hand motifs from cCaM are open (EF-hand angles ≈ 101°) and the two EF-hand motifs from nCaM are closed (EF-hand angles ≈ 129°) (Table S3). The EF-hand angles were computed as the angle between the vectors that define the orientation of the two helices in the EF-hand motif; the vectors were defined from the center of mass of the C_α beads of the first 4 residues to the center of mass of the C_α beads of the last 4 residues in a helix of the EF-hand motif. Moreover, NgIQ binds to cCaM in the crystal structure and the NMR experiments (4) showed that Ng_{13-49} mainly interacts with cCaM. Therefore, we select the 4 structures from the simulations by following the criteria: (i) EF-hand motifs from cCaM must be open (EF-hand angles range from 85° to 105°) and EF-hand motifs from nCaM must be closed (EF-hand angles are greater than 125°); (ii) cCaM has more interactions with Ng_{13-49} than nCaM ($Z_c - Z_n > 0$). Using SCAAL method (17), we reconstructed all-atomistic structures of apoCaM-Ng_{13-49} from the selected sidechain-C_α configurations, as shown in Fig. S8. Four Ca^{2+} ions were added and their positions were estimated as the center of mass of sidechain beads of the corresponding Ca^{2+} co-ordination residues. The Ca^{2+} positions were further optimized by energy minimization on the all-atomistic structures of Ca^{2+}-CaM-Ng_{13-49} (holoCaM-Ng_{13-49}) using a gradient descent algorithm provided by the molecular dynamics software package GROMACS (version 5.0.6) (23) with AMBER99SB-ILDN force-field (24). The tolerance for the maximum force was set to 500 kJ / mol / nm to remove clashes between atoms.

holoCaM-CaMKII complex: We used the crystal structure of calcium-bound CaM-CaMKII complex (PDB ID: 1CDM) for Ca^{2+} binding free energy calculations.

holoCaM: We used the crystal structure of holoCaM (PDB ID: 1CLL) for assessing the Ca^{2+} binding free energy in the absence of the CaMBT as a reference.

IV.2 Protonation of the initial structures for pulling simulations

The charge distribution impacts accurate estimation of the binding free energy of the Ca^{2+} ions. We first took into consideration of the pH and ionic strength by using H++ server (18) to predict the protonation states of the all-atomistic structure. We input the following parameters to H++ as used in the experiment: pH = 7.4, ionic strength I = 0.15 M, the external dielectric constant is 78.4 and the internal dielectric constant is 4. After we obtained the protonated states of all the residues, we performed an energy minimization using steepest descent algorithm and the tolerance for the maximum force was set to 500 kJ / mol / nm. Then we assigned partial charges according to the geometry of the proteins (input is the protonated structure in PDB format) by using a semi-empirical quantum chemistry program MOPAC (19). We applied those protonated protein structures and partial charges for further all-atomistic calculations.

IV.3 Steered molecular dynamics simulations

Classical molecular dynamics simulations were carried out using GROMACS molecular dynamics package (version 5.0.6) (23). We used the AMBER99SB-ILDN force field (24) except the charge assignment. The partial charges were generated from MOPAC as explained in the

previous session. The rigid three-site TIP3P model (25) was used for water molecules. We neutralized the system with Na^+ and Cl^- ions and maintained an ionic strength of 150 mM.

The size of the box is about 10 nm x 10 nm x 10 nm. The proteins are placed at least 1 nm away from the edges of the cubic box. Periodic boundary conditions were employed to mimic the macroscopic settings for electrolytes. Electrostatic interactions between periodic images were treated using the particle mesh Ewald approach (26), with a grid size of 0.16 nm, fourth-order cubic interpolation and a tolerance of 10^{-5}. A cutoff of 10 Å was used for van der Waals interactions and real space Coulomb interactions as well as for updating neighbor lists.

We adopted a gradient descent algorithm for energy minimization. Then we gradually heated the system temperature by 50 K per 0.1 ns to 298.15 K using NVT ensemble. We carried out 1 ns NPT equilibration with heavy atoms of the proteins (including the four Ca^{2+} ions) fixed. The proteins as well as the Ca^{2+} ions were afterwards released and were further equilibrated for another 1 ns. All NPT simulations maintained a constant pressure of 1 bar and temperature of 298.15 K using the Parrinello-Rahman barostat (27). The bond lengths in proteins were constrained using the LINCS algorithm of Hess (28). The equation of motion was integrated using a 2-fs time steps. As discussed in the main text, the positions of the Ca^{2+} ions in the bound state change during the minimization and equilibration stage, which yields inaccurate estimation of the binding free energy, therefore, we froze the positions of the Ca^{2+} ions as well as the backbone heavy atoms of the protein (or protein complex) during these preparation stage. They were free to move afterwards in the pulling simulations.

For each of the initial structures including 1 structure for each of holoCaM and holoCaM-CaMKII and four selected structure of holoCaM-Ng_{13-49}, I pulled the Ca^{2+} from site III and IV independently to 2 nm away where the interaction between the Ca^{2+} and the corresponding Ca^{2+}

binding loop is negligible. During the pulling simulations, the positions of the proteins may shift and this would cause inaccurate estimation of the distance between Ca^{2+} and the corresponding Ca^{2+} binding loop. Therefore, we fixed the C_α atom of the 100th or 136th residue in CaM, which showed smallest RMSD in a separate equilibration simulation, while pulling Ca^{2+} from the binding site III or IV, respectively. The force constant *k* and speed of pulling *v* of the reference position are described in the next section. The direction of the pulling force was randomly assigned and pointed away from the center of mass of CaM to avoid clashes between CaM and the Ca^{2+}. The pulling direction was selected if the angle Ω between the pulling vector and the vector connecting center of mass of CaM and center of mass of the corresponding Ca^{2+} binding loop was within 90 degrees since Ω > 90 leads to a large work (as shown in Fig. S12). The displacement of the Ca^{2+} ion and the pulling forces were output every 20 fs for calculation of the work as shown in Eq. S9. The coordinates and velocities of the system were saved every 1 ps.

The setup of the pulling simulations is illustrated in Fig. S9. The Ca^{2+} is constrained to a reference position that is moving along $\vec{x}$ direction at a speed of $|\vec{v}|$ = 0.001 nm/ps. The force constant of the spring *k* is set to 10000 kJ / mol / nm^2 to guarantee that the Ca^{2+} strictly follows the reference position. Therefore, the pulling force is calculated as,

$$\vec{f}(t) = -k(\vec{x}(t) - \vec{v}t - \vec{x}_0) \qquad \text{Eq. (S9)}$$

where $\vec{x}$ ($\vec{x}_0$) is the instantaneous (initial) displacement of the Ca^{2+} from the center of mass of the Ca^{2+} binding loop.

IV.4 Justification of parameters used in steered molecular dynamics simulations

We set the pulling speed *v* and the spring constant of the external force *k* to effectively estimate Ca^{2+} binding free energy from steered molecular dynamics. We explored different

combinations of spring constant and pulling speed to rationalize the parameter setting. For example, to pull the Ca^{2+} from site III of holoCaM, we used k = 100, 1000, 10000, 100000 kJ/mol/nm^2 and v = 0.001, 0.01, 0.02 nm/ps. The simulation time was 4 ns, 0.4 ns and 0.2 ns to ensure that Ca^{2+} was pulled the same distance. In Fig. S10A, with fairly slow pulling speed, using k = 1000, 10000 and 10000 kJ/mol/nm^2, the Ca^{2+} follows the motion of the dummy bead. For k = 100 kJ/mol/nm^2, the pulling force is so weak to pull the Ca^{2+} out during the entire simulation. For k = 1000 kJ/mol/nm^2, there exists a lag in the beginning of the simulation, indicating a favored interaction between CaM and Ca^{2+} at site III. However, the tail that deviates from the straight line indicates that the thermal fluctuations in the unbound state of Ca^{2+} dominate its motion. By increasing v to 0.01 or 0.02 nm/ps, k = 1000, 10000 or 100000 kJ/mol/nm^2 meet the stiff spring approximation, shown in Figs. S9B and S9C. Comparing to v = 0.001 nm/ps, the displacement curves are much smoother.

During the exploration of factors that influence work values, we found that the pulling speed v is more sensitive. Jarzynski's group (29) showed that JE does not depend on the pulling speed. In their study, they showed for the same amount of simulation time, the estimation of the chemical potential of a simple Lennard-Jones fluid did not vary for different switching time (comparable to pulling speed v). However, in practice, the free energy calculation is more efficient with fewer slow pulling trajectories than more fast pulling trajectories given the same amount of simulation time (39). For a complex system of pulling Ca^{2+} form a holoCaM-CaMBT compound, the efficiency of convergence of distribution of work values depend on the pulling speed v. Moreover, it has been shown that slower pulling speed v reduces the perturbation from the pulling force compared to the level of the thermal fluctuations of the binding site (30). Therefore, in this study, we set a relatively slow pulling speed v = 0.001 nm/ps with a stiff spring

constant k = 10,000 kJ/mol/nm^2 to guarantee that the Ca^{2+} follows the moving reference bead that accounts for small work. To improve the efficiency of the free energy estimation, furthermore, we also employed cumulative integral (CI) extrapolation method developed by Zuckerman's group (31), which is discussed in the next session.

In Fig. S11, the typical force profiles lead to the same conclusion. At v = 0.001 nm/ps, k = 10,000 and 100,000 kJ/mol/nm^2 demonstrated the same trend in the force profiles except more thermal fluctuations in the latter, indicating a converged behavior of the dissociation of Ca^{2+} in this parameter range. Therefore, k = 10,000 kJ/mol/nm^2 were used for the pulling simulations.

IV.5 Estimation of binding free energy using Jarzynski's equality and Cumulative Integral extrapolation

After obtaining the work (w) distribution from the pulling simulations as explained in IV.3 in the *Supporting Information*, we estimated the binding free energy $\Delta G^{\text{holoCaM-CaMBT}}$ between a one-Ca^{2+}-missing complex of holoCaM-CaMBT (G_U) and a Ca^{2+}-loaded complex of holoCaM-CaMBT (G_B) according to JE (32),

$$\begin{aligned} \Delta G^{\text{holoCaM-CaMBT}} &= G_{\text{B}} - G_{\text{U}} \\ &= k_B T ln \langle e^{-\frac{W}{k_B T}} \rangle \\ &\approx k_B T ln \sum_{n=1}^{M} e^{-w_n/k_B T}/M \end{aligned} \qquad \text{Eq. (S10)}$$

where k_B is Boltzmann constant, T is the temperature, M is number of pulling simulations (M = 100 ~ 150), w_n is the work in nth pulling calculated as,

$$w_n(\text{B} \rightarrow \text{U}) = \int_0^t \overrightarrow{f_n}(\tau) \cdot \vec{v}_n d\tau \qquad \text{Eq. (S11)}$$

B and U stands for bound and unbound states of the Ca^{2+} ion, respectively.

However, accurate estimation of the binding free energy ΔG relies on the sufficient sampling of small work (32). Zuckerman's group (31) showed that using a cumulative integral (CI) to extrapolate the free energy estimation to $n \rightarrow \infty$ can help reduce the required simulation data by 5 to 40 fold (30, 31), where n is the number of work values. In Table S6, we showed the estimation of change in Ca^{2+} binding free energy (ΔΔG) by using JE or CI estimation. We found that JE or CI produced the same signs of the ΔΔG, however, the magnitudes were not the same. The discrepancy in the free energy calculation lies in the requirement of sufficient sampling of small work for direct JE calculations, which can be overcome by CI extrapolation (28).

V. Analyses

V.1 Calculation of model free order parameter S^2

A number of N (~ 500) representative coarse-grained structures were selected from the free energy surface of unbound Ng protein using importance sampling method (17). The NMR model-free order parameter S_i^2 (33) for the backbone N-H bond of residue *i* is calculated on the selected structures by relating to the Root Mean Square Fluctuations (RMSF) of the C_α bead through the following empirical relation (34),

$$S_i^2 = 1 - 0.5 * \ln(1 + \frac{RMSF_i}{23.6\text{Å}} * 10.0) \qquad \text{Eq. (S12)}$$

$$RMSF_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\boldsymbol{r}_i^j - \frac{1}{N}\sum_{k=1}^{N}\boldsymbol{r}_i^k)^2} \qquad \text{Eq. (S13)}$$

where $\boldsymbol{r}_i^j$ is the position of the C_α bead of residue *i* at frame *j* and *N* is the total number of frames. $RMSF_i$ was computed after structural alignment.

V.2 Definition of helicity

The fraction of helical structure, or helicity, of Ng (G25-A42) was calculated using the following formula (35),

$$f_H = \frac{1}{N-3}\sum_{i=1}^{N-3} < \Theta(\Delta\phi - |\phi_i - \phi_0|) > \quad \text{Eq. (S14)}$$

where $\Theta(x)$ is the Heaviside step function taking value 1 if $x \geq 0$ and value 0 otherwise. N is the total number of residues, i is the residue index, ϕ_i is the dihedral angle about the residues $i \sim i+3$ from the simulation, $\phi_0 = 49.50°$ is the dihedral angle in a perfect alpha helix and $\Delta\phi = 12.07°$. <•••> denotes ensemble average.

V.3 Definition of Z

We defined an order parameter Z to calculate the total number of intermolecular contacts between apoCaM and $Ng_{13\text{-}49}$. For each residue *i* from apoCaM and residue *j* from $Ng_{13\text{-}49}$, the sidechain-sidechain and backbone-backbone contacts (Z_{ij}^{ss} and Z_{ij}^{bb}, respectively) are determined as

$$Z_{ij}^{ss} = \Theta\left(c - \frac{d_{ij}^{ss}}{\rho_i + \rho_j}\right) \quad \text{Eq. (S15)}$$

$$Z_{ij}^{bb} = \Theta(c - \frac{d_{ij}^{bb}}{\rho_{HB}}) \quad \text{Eq. (S16)}$$

$d_{ij}^{ss}(d_{ij}^{bb})$ is the distance between side-chain (backbone) beads of residue *i* of apoCaM and residue *j* of $Ng_{13\text{-}49}$ in simulation, ρ_i (ρ_j) is van der Waals radius of residue *i(j)*, $\rho_{HB} = 4.66$ Å is the typical length of a hydrogen bond and cutoff c = 1.3. The total number of contacts Z is the summation of backbone-backbone and sidechain-sidechain contacts over all residues $Z = \sum_{ij}(Z_{ij}^{ss} + Z_{ij}^{bb})$.

V.4 Sampling protein configuration for structure analyses

The structures from US simulations are biased and not appropriate for structural analyses directly. We therefore used the Boltzmann reweighting method to sample an ensemble of unbiased structures. The probability of selecting a configuration *x* is

$$w(x) = \begin{cases} 1 & if\ p(x) \geq \rho \\ 0 & else \end{cases} \quad \text{where } p(x) = e^{-G(d_{COM}(x))/k_B T} \qquad \text{Eq. (S17)}$$

$p(x)$ is the probability of the configuration x in reweighted ensemble; $G(d_{\mathrm{COM}})$ is the reweighted/unbiased free energy obtained from the WHAM analyses along reaction coordinate d_{COM} (G_{min} is scaled to 0); ρ is a random number generated between 0 and 1; k_B is Boltzmann constant and T is the temperature. The ensemble generated after Boltzmann reweighting thus obeys the canonical distribution and is employed for subsequent analyses. Each structure from the biased ensemble was challenged by acceptance probability $w(x)$ and a total of 23,722 structures were sampled. This sample achieves a distribution of $P^{\mathrm{sample}}(d_{\mathrm{COM}})$. In order to assess the quality of the sampling, we computed the surprisal value compared with the original unbiased distribution $P^{\mathrm{ori}}(d_{\mathrm{COM}})$ defined by the following formula (36),

$$\text{surprisal} = \sum_r -P^{\mathrm{ori}}(r) \ln\left[P^{\mathrm{sample}}(r)/P^{\mathrm{ori}}(r)\right] \qquad \text{Eq. (S18)}$$

where the summation is over all the order parameter r (in this case d_{COM}). A surprisal value of 0.14 ensures that the sampled ensemble can well represent the original distribution.

V.5 Self-organized clustering algorithm

We applied the self-organized neural-net clustering algorithm (37-39) to determine the structures of the apoCaM-$\mathrm{Ng}_{13\text{-}49}$ complexes from the umbrella sampling simulations. In this clustering method a vector with M elements represents each conformation j, $x_j = [x_{1j},\ x_{2j},\ \dots,\ x_{Mj}]$, where $j = 1, 2, \dots N$ and N is the number of conformations selected for clustering analysis. The element x_{ij} ($i = 1, 2, \dots, M$) stands for the Euclidean distance between side-chains of the

polypeptide chain in the conformation j. To partition the various conformations into distinct clusters, the clusters are described by the cluster center and the size of the cluster is determined by a radius R_c. A given conformation is assigned to cluster k if the distance between the vector x_j and the center of the k^{th} cluster,

$$C_k = \frac{1}{M_k}\sum_{j=1}^{M_k} x_j, \qquad \text{Eq. (S19)}$$

where $C_k = [C_{1k}, C_{2k}, \ldots, C_{Mk}]$ and $M_k$is the total number of conformations in C_k.

Conformation j belongs to C_k if the Euclidean distance between conformation j and the cluster center k, $d_{jk} = \left|x_j - C_k\right| < R_c$, where R_c is a preassigned value. In the current study we used the native contact pairs from the unbound structure of apoCaM and $Ng_{13\text{-}49}$ as the M elements and a following cutoff distance R_c = 25 σ (σ = 3.8 Å) to categorizes the structures into seven distinct clusters.

V.6 Correlation between intra- and inter- molecular interactions in apoCaM-$Ng_{13\text{-}49}$ binding

In order to better understand the relation between interactions within the two proteins and intermolecular contacts, we built a correlation map between contacts (Fig. S4). The correlation between two contact pairs m and n is computed as in the following equation,

$$corr_{mn} = \frac{< q_m q_n > - < q_m >< q_n >}{\sqrt{< q_m^2 > - < q_m >^2}\sqrt{< q_n^2 > - < q_n >^2}} \qquad \text{Eq. (S20)}$$

where $q_{m(n)}$ is the contact state of the contact pair m (n): 1 if m (n) is a contact, 0 if not. The list of contacts are provided in Table S8. The correlation between those contacts falls into several categories.

a) Correlation: as shown in diagonal, the contact formation within apoCaM, between apoCaM and Ng_{13-49}, and within Ng_{13-49} are correlated.

b) Anti-correlation: contact formation between nCaM and Ng_{13-49} anti-correlates with contact formation within nCaM; contact formation between cCaM and Ng_{13-49} anti-correlates contact formation within cCaM. The anti-correlation tells that the contact formation within nCaM or that within cCaM competes with its interaction with Ng_{13-49}. We infer that binding of the target is responsible for the repacking of the nCaM or cCaM by direct competitive interaction with the Ca^{2+} binding loops.

c) No correlation: there is no apparent correlation among contact formation within nCaM and contact formation within cCaM, indicating the two domains of CaM are relative independent.

d) Mixed correlation: contact formation between nCaM and cCaM has a mixed correlation with contact formation between nCaM and Ng_{13-49} as well as contact formation between cCaM and Ng_{13-49}.

Supporting Figures

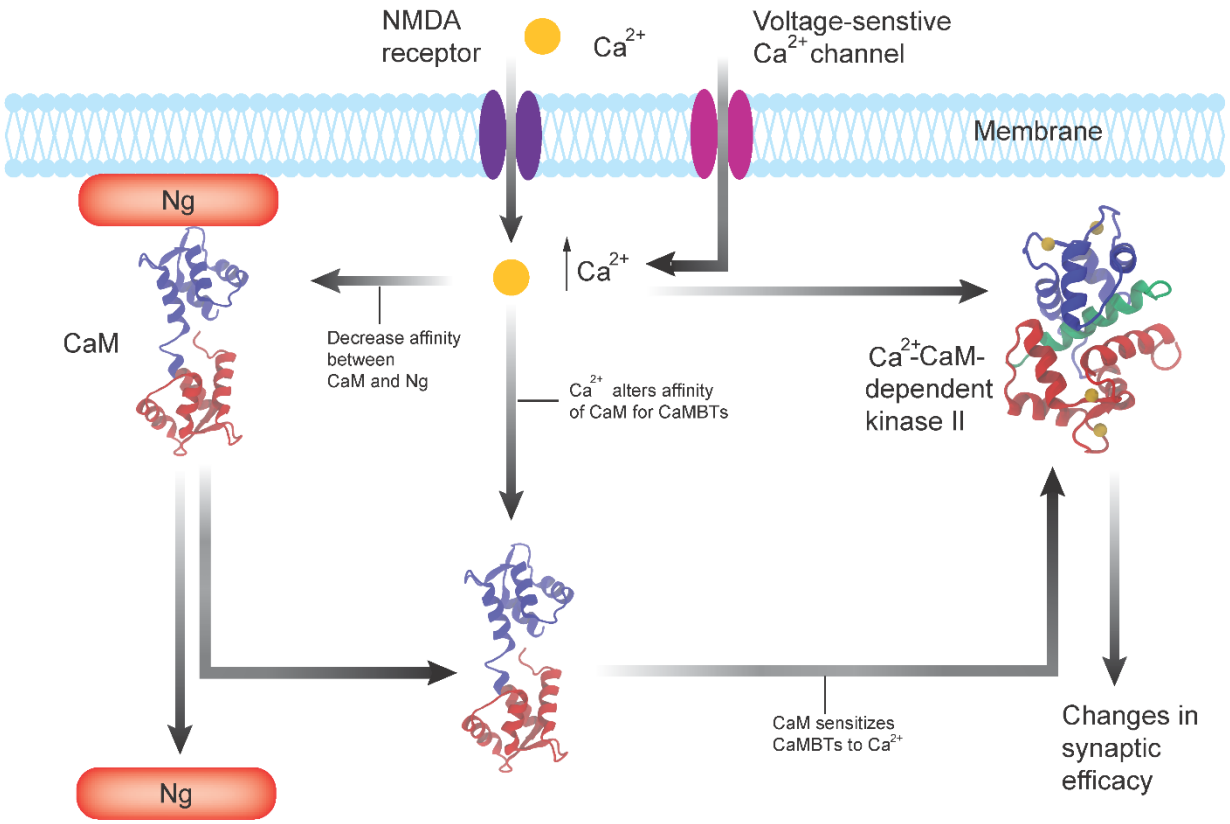

Fig. S1. Overview of the effects of CaM-dependent Ca^{2+} signaling and effects of CaM binding targets (CaMBTs) on changes in synaptic plasticity. Many of the effects of intracellular Ca^{2+} on synaptic plasticity are mediated through CaM-regulated proteins. Increase in intracellular Ca^{2+}, generated through the activity of NMDA (N-methyl-D-aspartate) receptors or voltage-sensitive Ca^{2+} channels, results in the release of CaM that is bound to Ng. CaM mediates the Ca^{2+} stimulation of CaMKII which is required for changes in synaptic plasticity. The structures of Ca^{2+}-free CaM (apoCaM, PDBID: 1CFD) and Ca^{2+}-CaM-CaMKII peptide (PDB ID: 1CDM) are provided. CaM is colored as follows, red → nCaM (residue 1 to 76), gray → central linker (residue 77 to 82), blue → cCaM (residue 83 to 148) and the CaMKII peptide is colored in green.

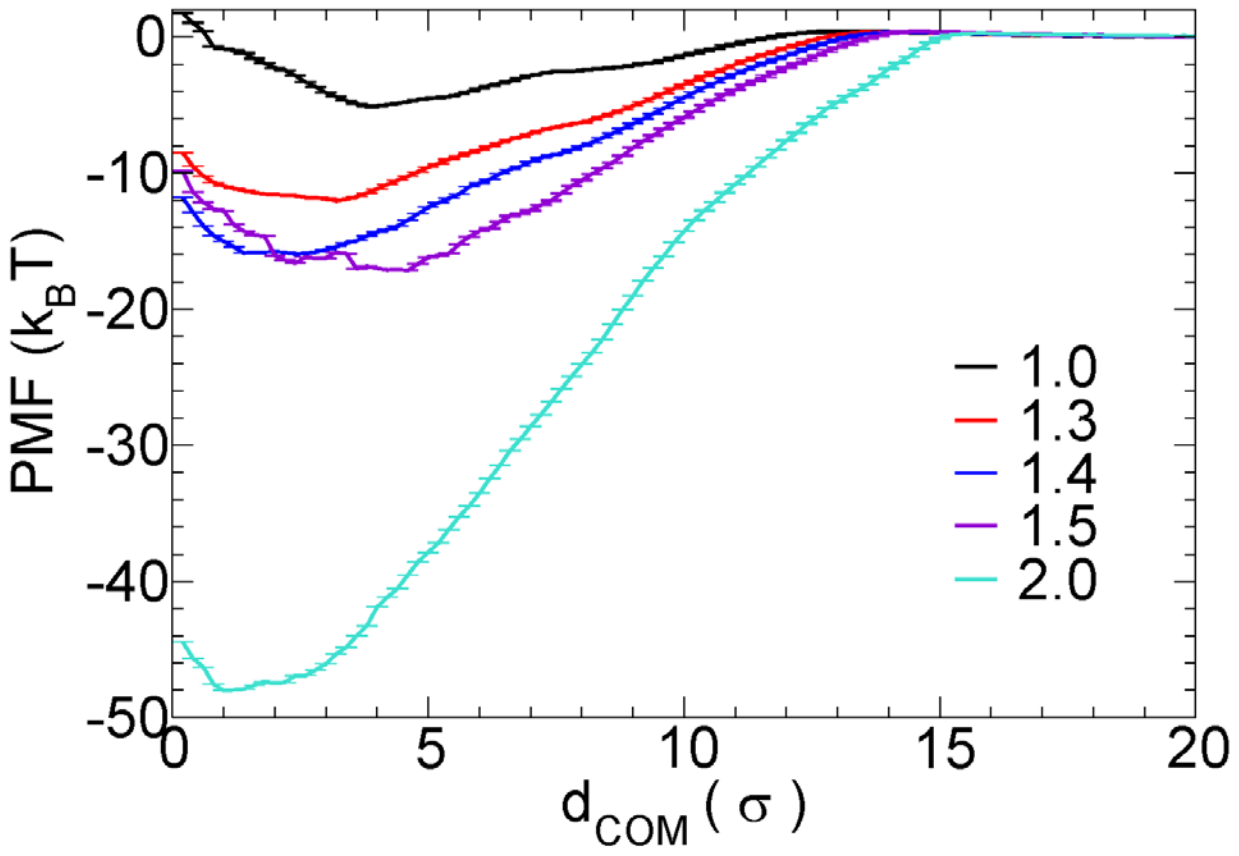

Fig. S2 Reweighted potential of mean force of apoCaM and $Ng_{13\text{-}49}$. The PMF was reweighted from umbrella sampling simulations using WHAM at varying scaling factors of the inter-molecular nonbonded interaction λ (excluding electrostatic interactions). KB statistical dihedral angle potential was employed. d_{COM} is the distance between center of mass of apoCaM and center of mass of $Ng_{13\text{-}49}$. σ equals 3.8 Å. T = 1.1 ε / k_B. ε = 0.6 kcal / mol.

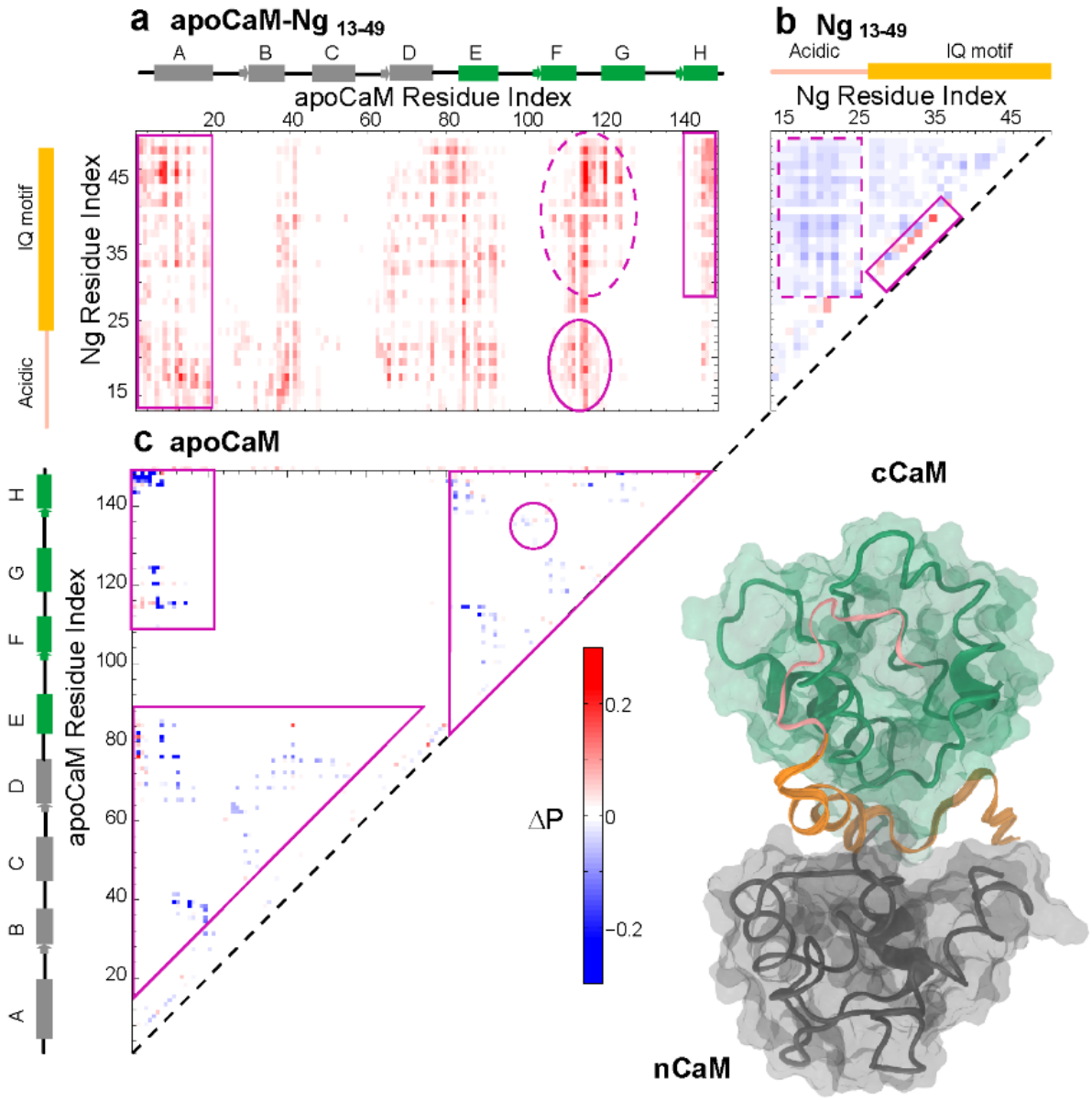

Fig. S3 Difference in probability of contact formation between the bound ensemble of the most dominant cluster (cluster 1) and the unbound ensemble. The difference maps for apoCaM intramolecular (a), $Ng_{13\text{-}49}$ intramolecular (b) and apoCaM-$Ng_{13\text{-}49}$ intermolecular (c) contacts are provided. d_{COM} = 20.0 s is used for the unbound state. The schematic representation of the helices of apoCaM and $Ng_{13\text{-}49}$ are provided along the axes. The representative structure, which has minimal root mean square deviation from the averaged structure, is in ribbon representation and colored according to the secondary structures: nCaM is in gray, cCaM is in green, acidic region of $Ng_{13\text{-}49}$ is in pink and IQ motif of $Ng_{13\text{-}49}$ is in orange. The regions of particular interest are encircled by ellipses, rectangles, circles and triangles. The color bar represents the difference in probability of contact formation ΔP between the bound and the

unbound conformations where positive ΔP indicates an increase of contact probabilities in the bound ensemble.

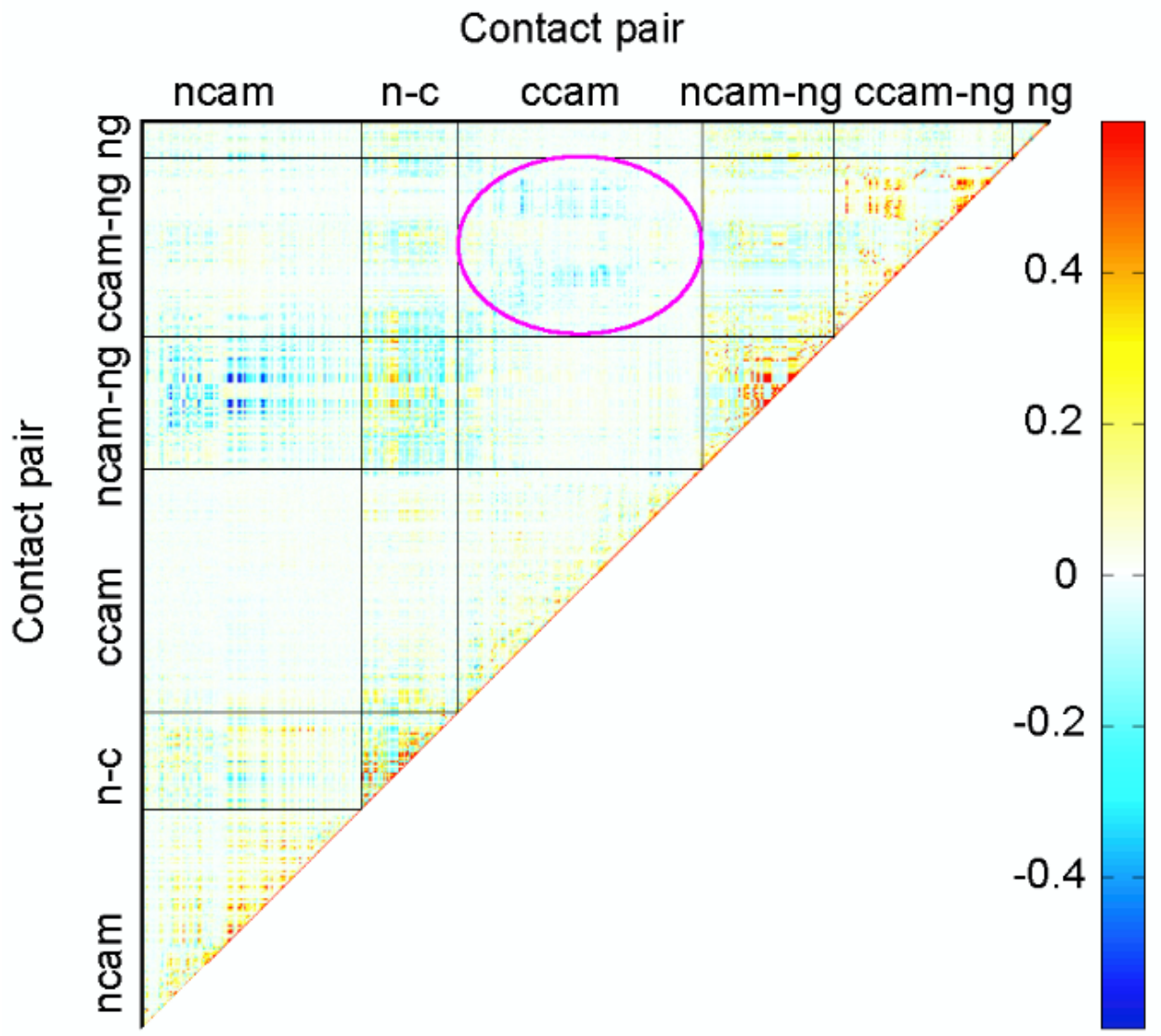

Fig. S4 Correlation map between contacts formation involving apoCaM and Ng_{13-49} in bound ensemble (cluster 1). The contact pair list is provided in Table S8. The blue color shows strong anti-correlation between involved contact pairs; the red color shows strong correlation; white means no correlation.

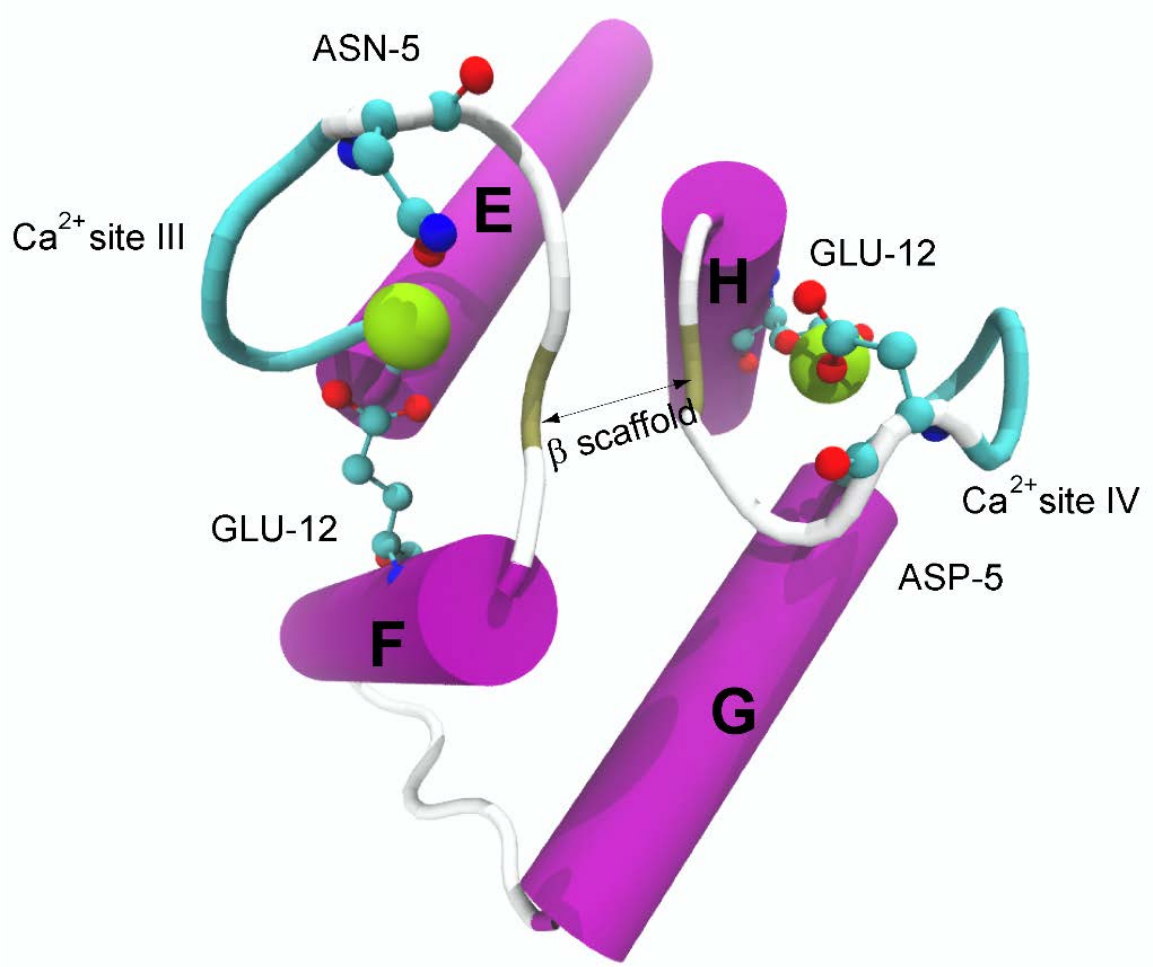

Fig. S5 Illustration of the EF-hand β-scaffold. The structure is from C-terminus of the crystal structure of holoCaM (PDB: 1CLL). The calcium ions are shown in yellow beads. The 5th (+Z coordinating ligand) and 12th (-Z coordinating ligand) residues in the Ca^{2+} binding loops are shown in ball-and-stick representation.

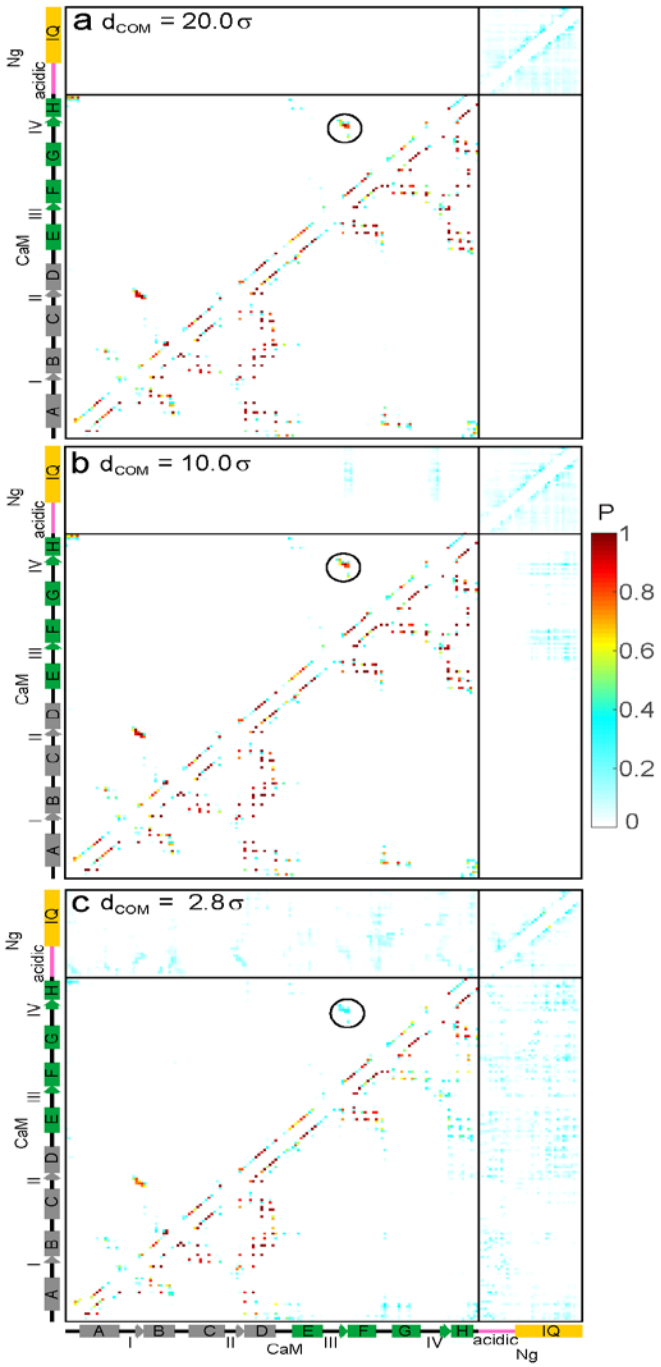

Fig. S6 Probability of contact formation in unbound, encounter and bound ensemble of apoCaM-Ng_{13-49}. In panel (a), the probability map of contact formation for the unbound state was calculated using the ensemble when apoCaM and Ng_{13-49} are well separated at d_{COM} = 20.0 σ. σ = 3.8 Å. (b) the probability map of contact formation for the encounter complex at d_{COM} = 10.0 σ and (c) in the bound state (d_{COM} = 2.8 σ) are provided. The upper and lower triangles show contacts maps for backbone-backbone conatcts and sidechain-sidechain contacts, respectively. The schematic representation of the helices of apoCaM and Ng_{13-49} are provided along the axes. The region encircled by a circle marks the interaction between the two β-strands in the EF-hand β-scaffold. The color bar represents the probability of contact formation P.

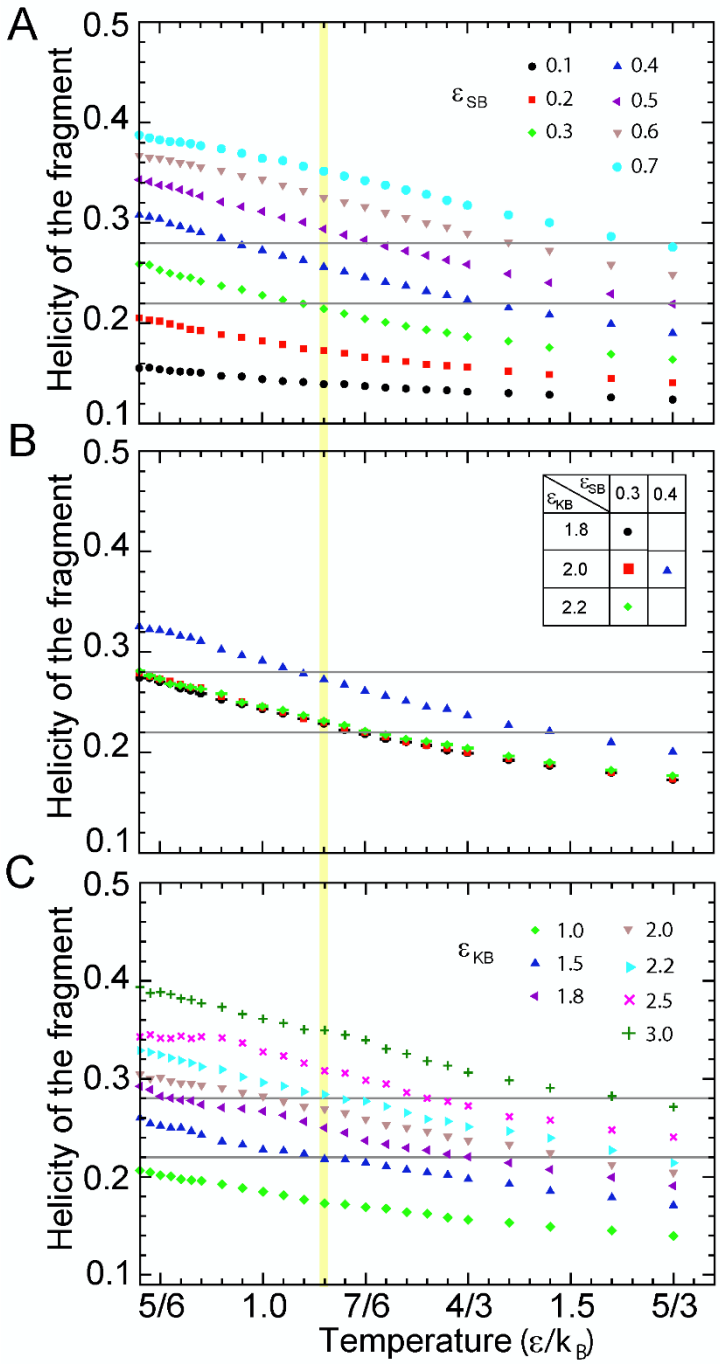

Fig. S7 Calculated helicity of the Ng protein fragment with residual structure matches with experiment using three dihedral angle potentials. The helicity of the fragment G25-A42 was computed from REMD simulations of the intact Ng protein with the coarse-grained side-chain C_α model. Three models of dihedral angle potential were used: (A) Structure-Based (SB) potential; (B) Hybrid potential; (C) Karanicolas-Brooks (KB) statistical potential. The two gray lines mark the upper and lower limit of the helicity of the segment from the experiment (3); the yellow shade marks the corresponding reduced temperature 1.1 ε / k_B. Helicity was estimated based on the dihedral angles between four consecutive C_α beads (see definition in V.2 in *Supporting Information*).

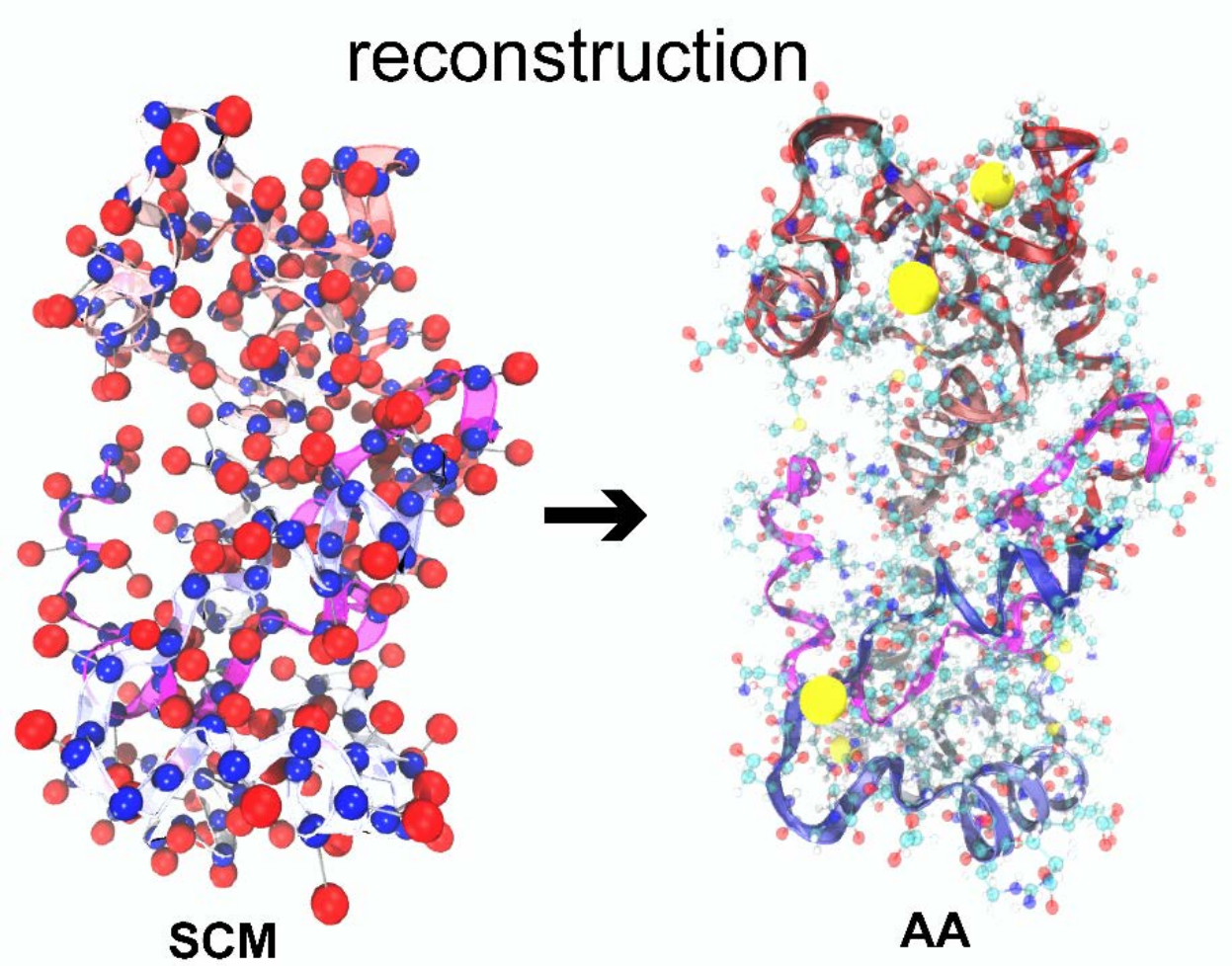

Fig. S8 Illustration of protein reconstruction from coarse-grained configuration to atomistic configuration for the holoCaM-Ng$_{13\text{-}49}$ complex. The backbone of CaM and Ng$_{13\text{-}49}$ is shown in ribbon representation. CaM is colored from red (N-Domain) to blue (C-Domain); Ng$_{13\text{-}49}$ is shown in magenta. (left) The Sidechain-C$_\alpha$ model (SCM), where the backbone atoms are represented by the C$_\alpha$ beads (blue balls) and sidechain atoms are represented by the sidechain beads (red balls). (right) All-atomistic (AA) configuration of the holoCaM-Ng$_{13\text{-}49}$ complex with Ca^{2+} (yellow balls) added to the Ca^{2+} binding loops. The other atoms are colored according to the atom names (oxygen atoms are in red, carbon in cyan, hydrogen in white, nitrogen in blue, etc.).

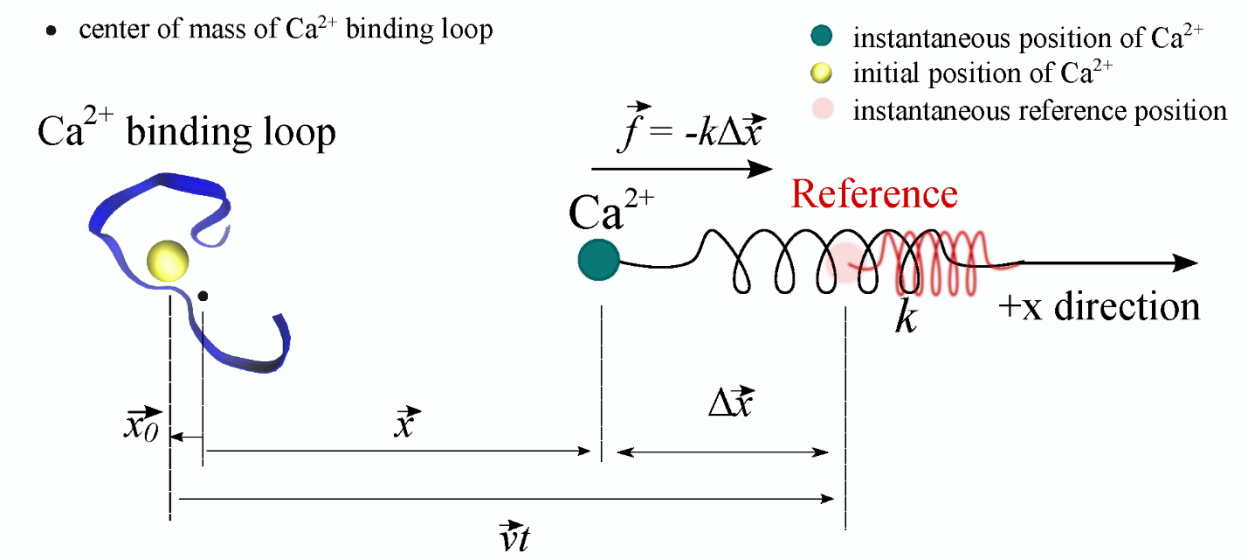

Fig. S9 Schematic illustration of pulling Ca^{2+} from one of the calcium-binding loop of CaM. The rest of the system including CaM and the solvent molecules is not shown in this illustration for better visualization.

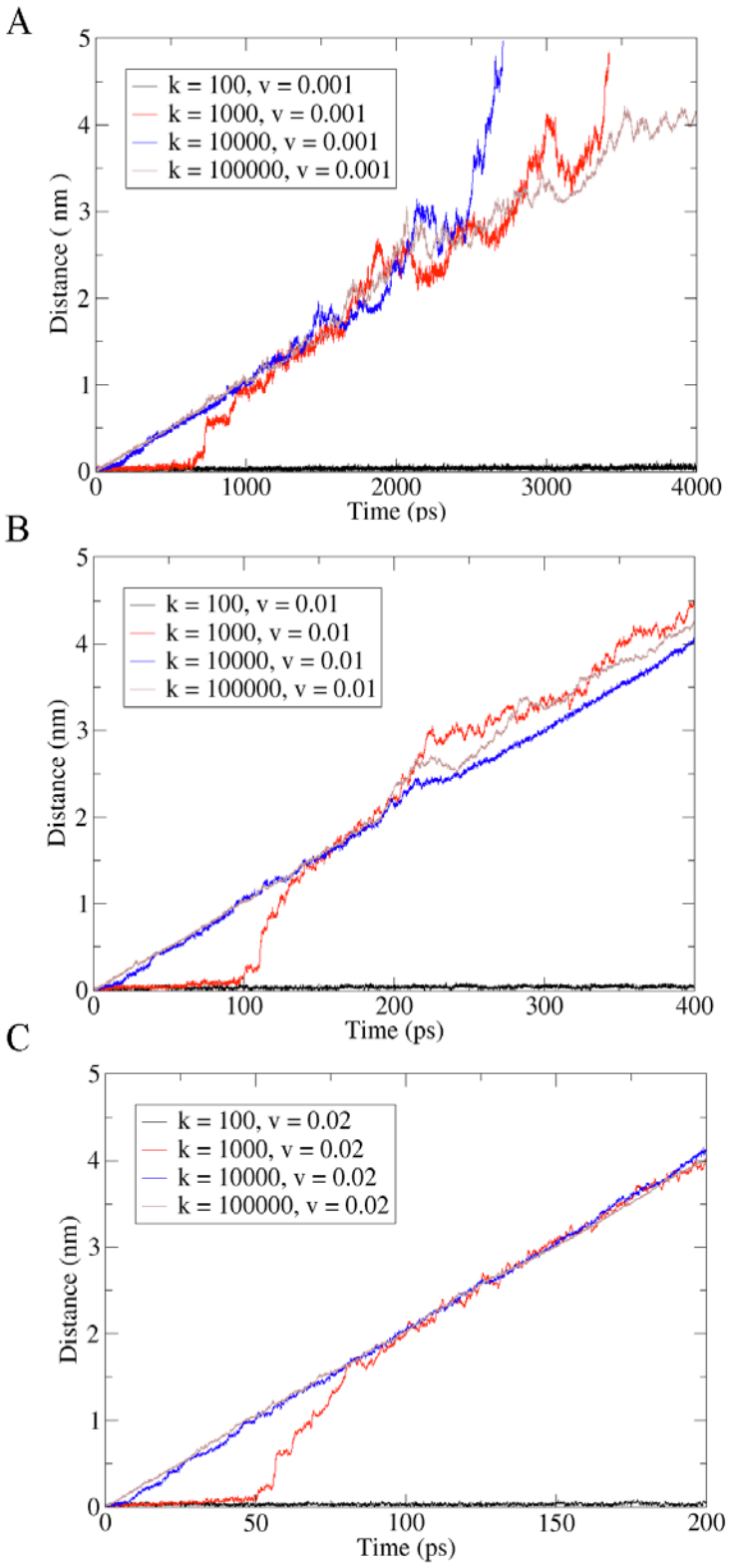

Fig. S10 The displacement of the Ca^{2+} from the Ca^{2+} binding loop during the pulling simulations of Ca^{2+} from the Ca^{2+} binding site III of holoCaM. The unit of k is kJ/mol/nm^2 and the unit of v is nm/ps. (A) (B) (C) show the case pulling speed v = 0.001, v = 0.01 and v = 0.02 nm/ps.

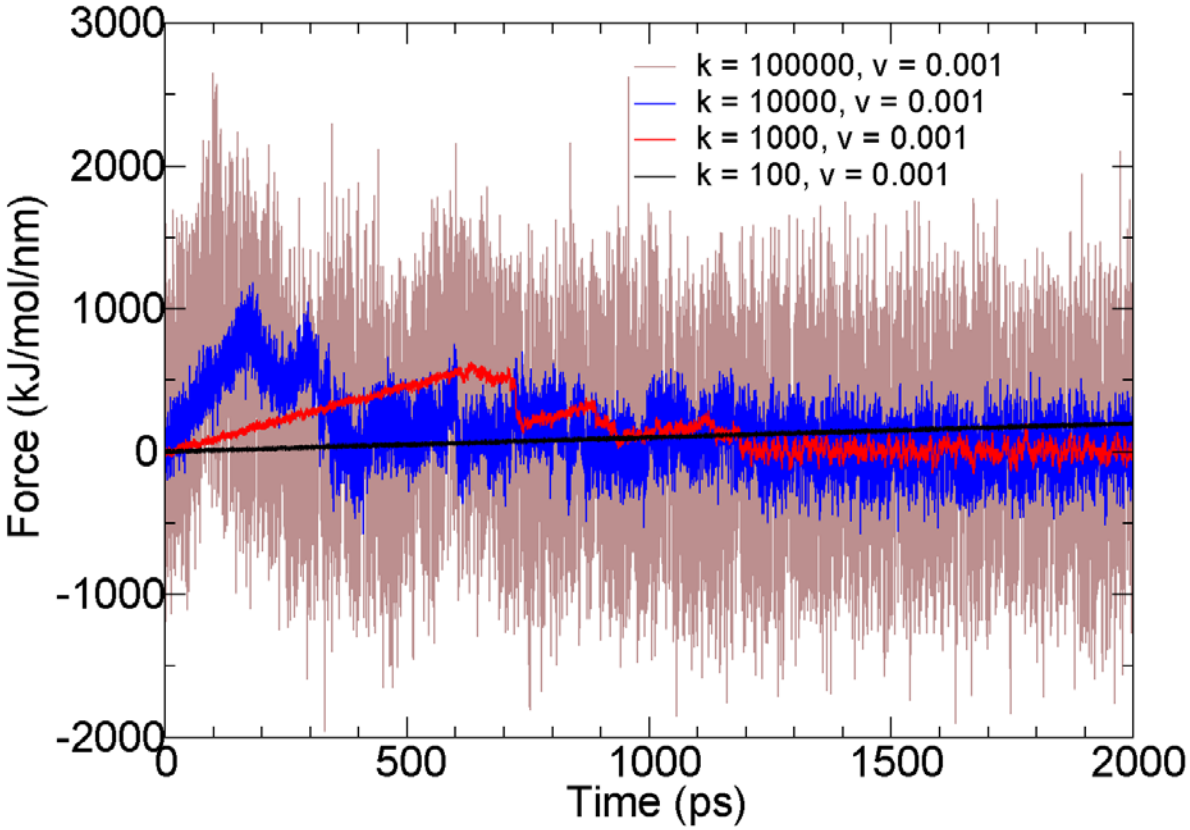

Fig. S11 The profile of external force in pulling simulations of Ca^{2+} from the Ca^{2+} binding site III of holoCaM. The pulling speed v = 0.001 nm / ps. The unit of spring constant k is kJ/mol/nm^2 and the unit of pulling speed v is nm/ps.

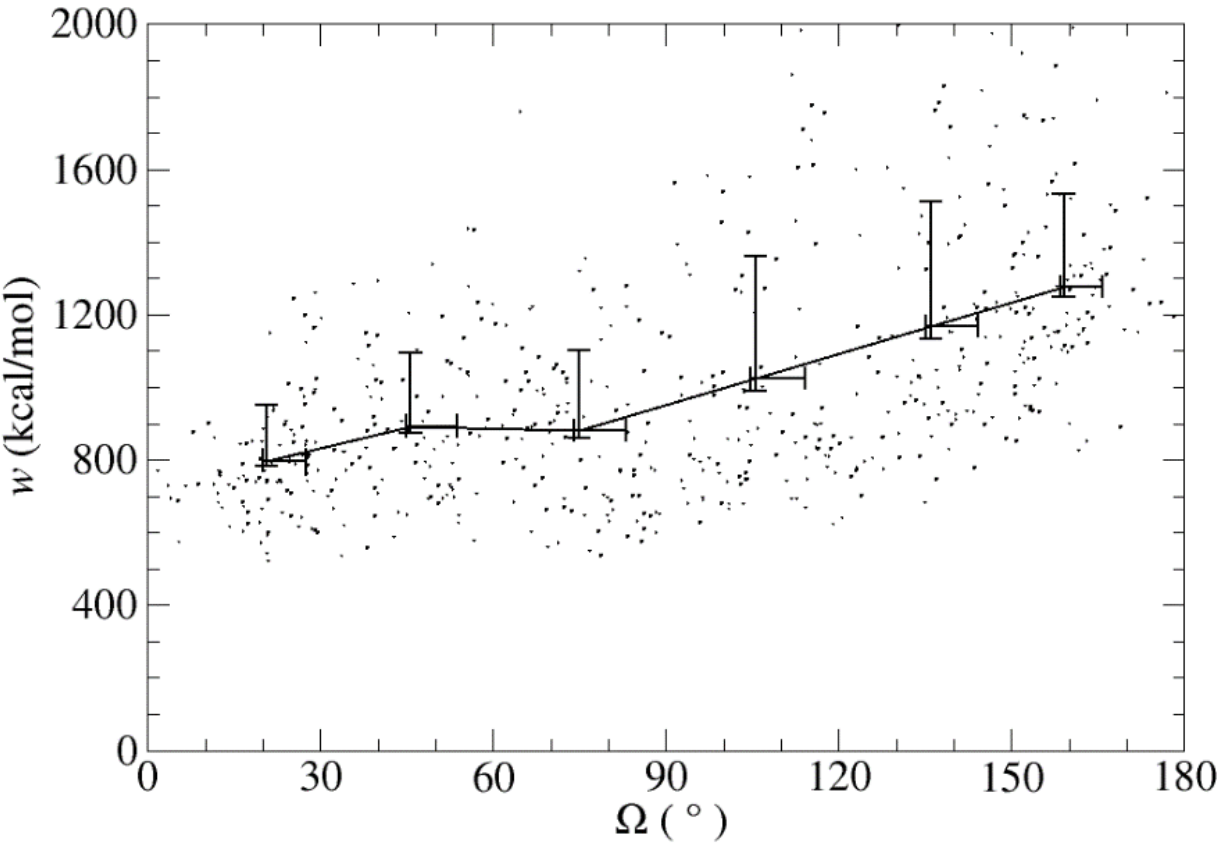

Fig. S12 Work done by external pulling force of pulling Ca^{2+} at different pulling angles. At each pulling angle zone Ω in [0, 30], [30, 60], [60, 90], [90, 120], [120, 150], [150, 180], the Ca^{2+} were pulled from site IV of holoCaM for 100 times. The average work, deviation (upper bar) as well as standard error (lower bar) are provided. The pulling speed v = 10 nm / ns and the spring constant of the pulling force k = 1000 kJ / mol / nm^2.

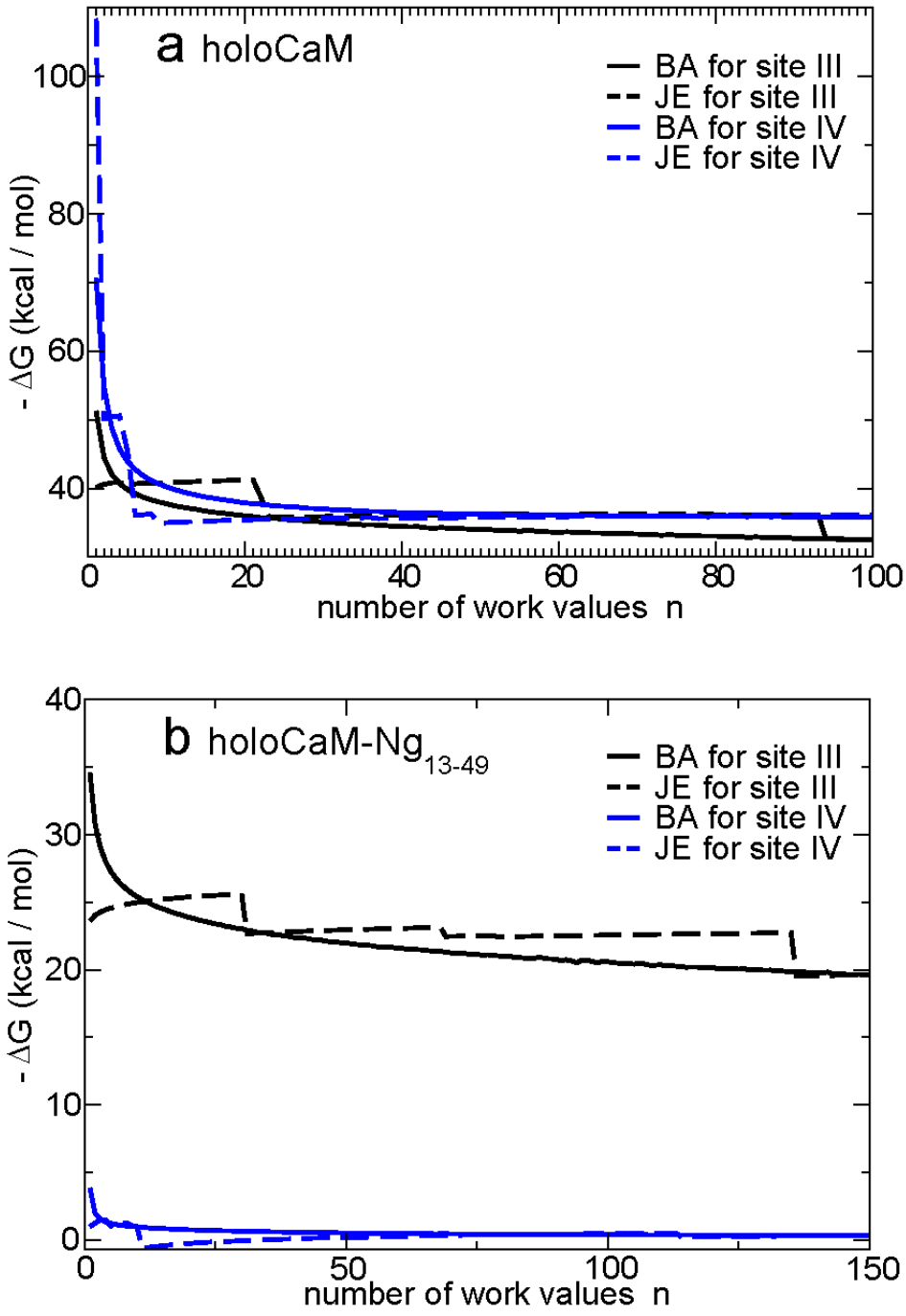

Fig. S13 The running Jarzynski equality (JE) estimate and subsampled block-averaged (BA) estimate are plotted as a function of the number of trajectories used in the estimate. The binding free energies of Ca^{2+} at site III and IV of holoCaM using the running JE method and the BA method are shown in panel (a) and the binding free energies of Ca^{2+} at site III and IV of holoCaM-Ng_{13-49} (using the conformation shown in Fig. 5) are shown in panel (b).

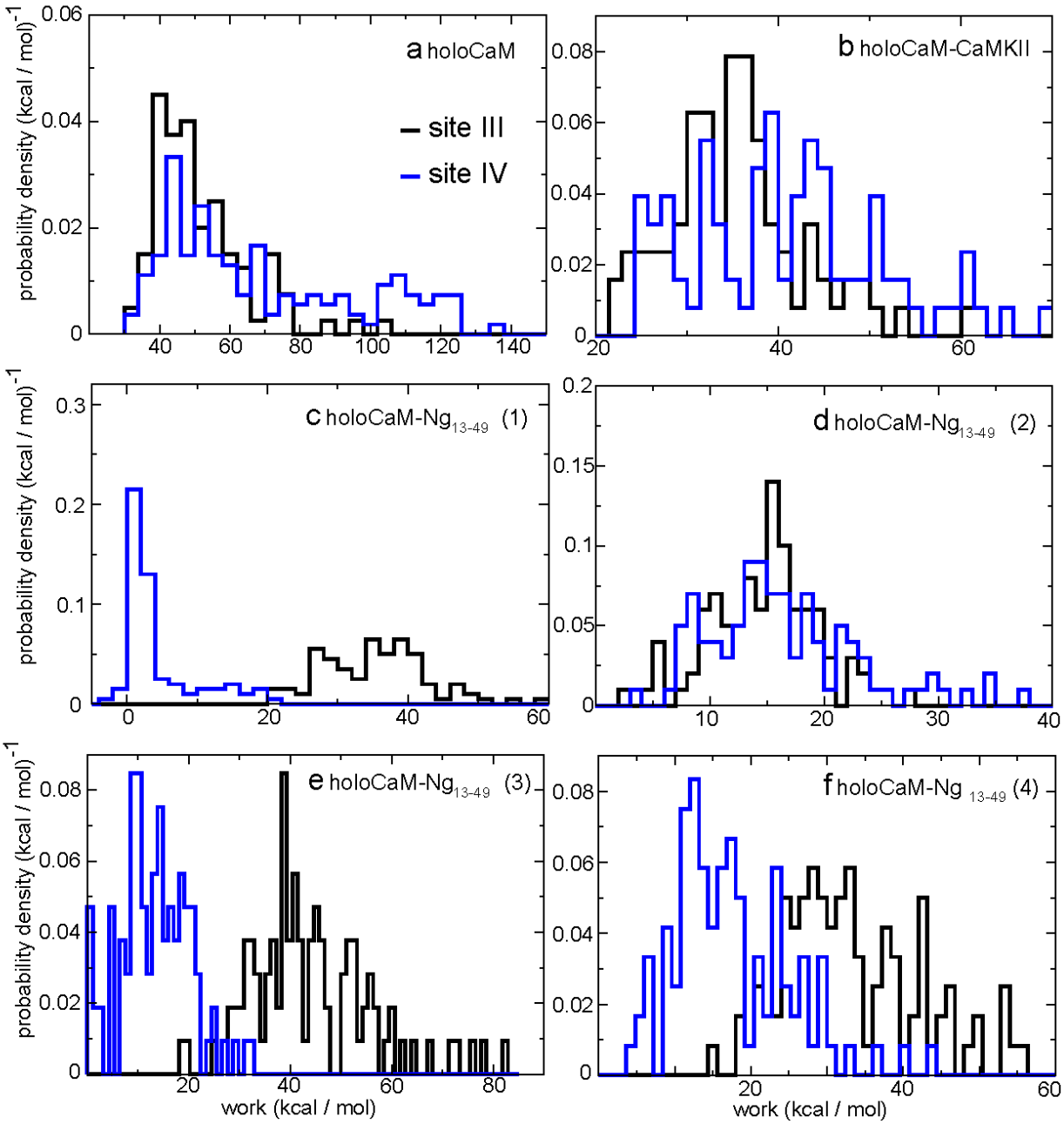

Figure S14 Distribution of work for dissociation of Ca^{2+}. (a) and (b) represent the holoCaM and holoCaM-Ng_{13-49}, respectively. (c-f) represent the four conformations of holoCaM-Ng_{13-49}.

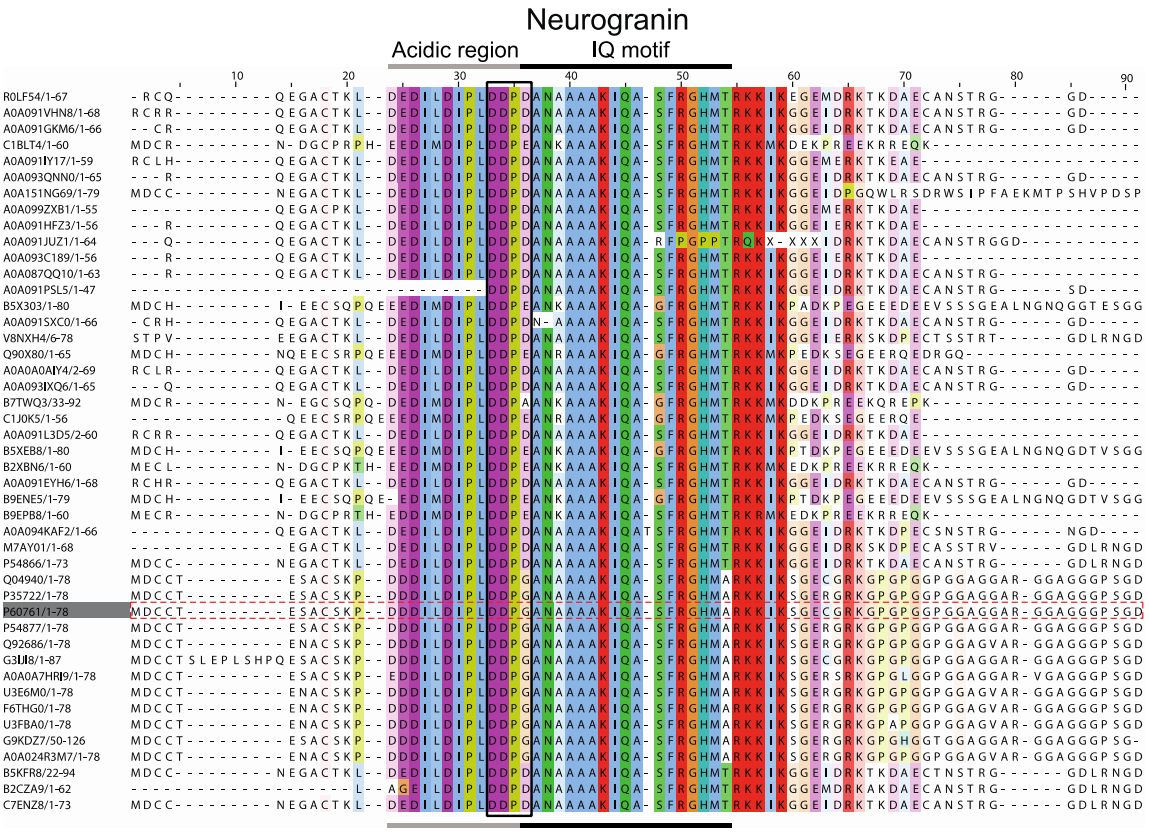

Fig. S15 Sequence alignment of neurogranin (Ng). 45 sequences of Ng (see Table S14 for the details of species associated with each sequence) were selected from UniProtKB (www.uniprot.org) by searching the keyword “neurogranin”. The sequences were aligned using the “Align” tool in UniProt and visualized in Jalview (40). The ClustalX color scheme in Jalview was used to highlight the residues and shading intensity of the color is based on the conservation over all the sequences in the alignment. The “acidic region” and the “IQ motif” of Ng are indicated by the grey and black line, respectively. The “DDPG (D/E/A)” motif is indicated inside the black area. The sequence of Ng from mouse (UniProt id P60761) that was used in this study is shown inside the red dotted area.

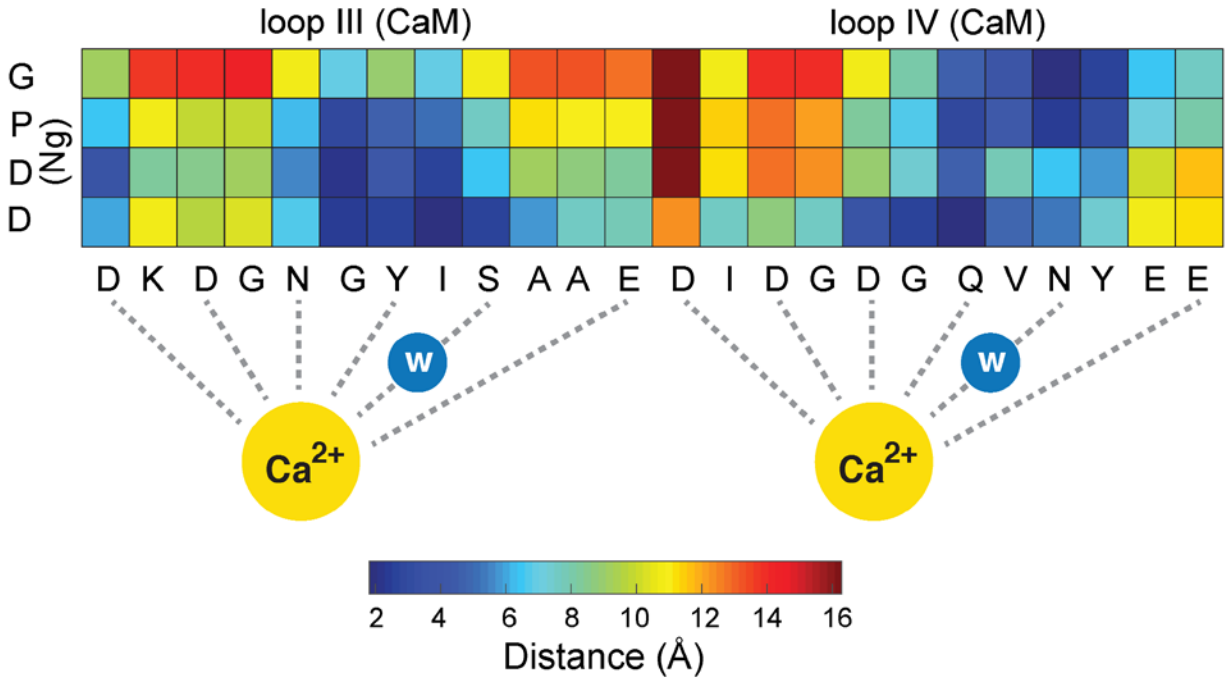

Fig. S16 Distance map between the Ca^{2+} binding loops (III and IV) of CaM and the "DDPG" motif of Ng. The distance here refers to the closest distance between the corresponding residues. The Ca^{2+} ions (shown as yellow spheres) coordinate (indicated by grey dotted lines) with the first, third, fifth, seventh, ninth (through water molecule (w), shown as a blue sphere) and twelfth residue of the loops. The distance to distance map is based on the holoCaM-Ng_{13-49} complex structure in Fig. 5a in the main text.

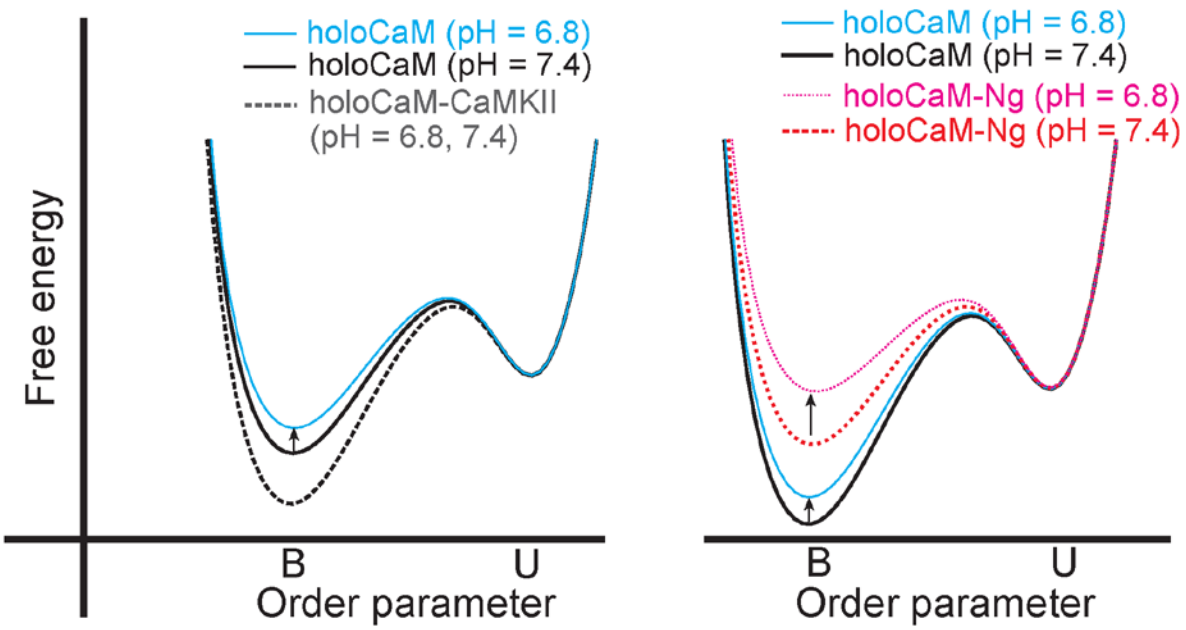

Fig. S17 Illustration of the changes in the binding free energy ΔG and ΔΔG from pH = 7.4 to pH = 6.8. $\Delta G = G_B - G_U$. $\Delta\Delta G = \Delta G^{holoCaM\text{-}CaMBT} - \Delta G^{holoCaM}$. B and U stand for bound and unbound states of the Ca^{2+}, respectively. The arrows show increase of ΔG for Ca^{2+} in all the three systems.

Supporting Tables

Table S1 Dissociation constant K_d of apoCaM and Ng_{13-49} with several values of strength of non-electrostatic intermolecular hydrogen bonding and van der Waals interactions λ. λ is shown in Eq 4. Experimentally measured K_d = 680 nM (4).

Strength λ	ΔPMF (k_BT)	K_d (nM)
1.0	5.12	6.0×10^{6}
1.3	12.04	5.9×10^{3}
1.4	15.94	119.0
1.5	17.16	35.4
2.0	48.01	1.4×10^{-11}

Table S2 The population of major clusters from computer simulations and the correlation coefficient of their computational "apparent chemical shifts" with the data from NMR experiments (4).

Cluster	1	2	3	4
Population	33%	26%	15%	12%
Correlation Coefficient	0.36	0.07	-0.1	0.07

Table S3 EF-hand angles in several forms of CaM. Definition of EF hand angles is provided in IV.1 from the *Supporting Information*.

	EF 1	EF 2	EF 3	EF 4
apoCaM (PDB: 1CFD)	131°	123°	127°	133°
holoCaM (PDB: 1CLL)	87°	84°	105°	96°
apoCaM-NgIQ (PDB: 4E50)	129°	129°	100°	101°
holoCaM-CaMKII (PDB: 1CDM)	84°	83°	98°	94°

Table S4 The composition of the three dihedral potentials for modeling Ng or Ng_{13-49}. SB stands structure-based and KB stands for Karanicolas-Brooks statistical potential (12).

fragment / models	from G25 to A42	Others
SB	structure-based	0
Hybrid	structure-based	sequence-based
KB	sequence-based	sequence-based

Table S5 Correlation between computed and experimentally-derived NMR model free order parameter S^2.

Models	SB		Hybrid	KB	
	ε_{SB} = 0.3	ε_{SB} = 0.4		ε_{KB} = 1.8	ε_{KB} = 2.0
Correlation Coefficient	0.40	0.73	0.29	0.78	0.64

Table S6 Difference in binding free energy of Ca^{2+} (ΔΔG) calculated from non-equilibrium molecular simulations and from the experiments at pH = 7.4. Direct use of Jarzynski's equality and a cumulants integral extrapolation was used in the calculation of binding free energy of Ca^{2+} from the simulations.

System		holoCaM-Ng_{13-49} (average)	holoCaM-CaMKII
ΔΔG (kcal/mol) Jarzynski's equality	Site III	8.5±2.7	-0.1
	Site IV	23.9±1.1	-3.4
ΔΔG (kcal/mol) CI extrapolation	Site III	9.2±2.2	-0.5
	Site IV	22.4±0.9	-7.0
ΔΔG (kcal/mol) experiment	Site III/IV	+2.5	-3.3

Table S7 Difference in binding free energy of Ca^{2+} (ΔΔG) for Ca^{2+} at site III calculated from non-equilibrium molecular simulations at pH = 6.8. Direct use of Jarzynski's equality and a cumulants integral extrapolation was used in the calculation of binding free energy of Ca^{2+} from the simulations.

System	holoCaM-Ng_{13-49} (average)	holoCaM-CaMKII
ΔΔG (kcal/mol) Jarzynski's equality	14.6±2.8	+0.7
ΔΔG (kcal/mol) CI extrapolation	10.5±2.1	-5.6

Table S8 Contact pair list used for correlation analysis. The non-specific contact pairs are determined from the difference contact map analysis. A contact pair is selected if the magnitude of the change in the probability of the contact pair in cluster 1 (Fig. S3) is greater than 0.2 from unbound state to the bound state. The contact pairs are listed in a sequence of categories: within nCaM, between nCaM and cCaM, within cCaM, between nCaM and Ng_{13-49}, between cCaM and Ng_{13-49} and within Ng_{13-49}. The sequence of apoCaM is from NMR structure (PDB: 1CFD) and sequence of Ng_{13-49} is provided in the *Materials and Methods* in the main text.

Index	Residue Index	Residue Index	Interaction Category	Index	Residue Index	Residue Index	Interaction Category
1	4	12	nCaM	2	4	65	nCaM
3	4	69	nCaM	4	4	8	nCaM
5	4	9	nCaM	6	5	9	nCaM
7	6	10	nCaM	8	7	11	nCaM

9	8	12	nCaM	10	9	13	nCaM
11	9	65	nCaM	12	9	69	nCaM
13	10	14	nCaM	14	11	15	nCaM
15	11	39	nCaM	16	12	16	nCaM
17	12	39	nCaM	18	12	41	nCaM
19	12	65	nCaM	20	12	68	nCaM
21	12	69	nCaM	22	12	72	nCaM
23	13	17	nCaM	24	13	20	nCaM
25	13	65	nCaM	26	14	18	nCaM
27	15	38	nCaM	28	15	39	nCaM
29	16	20	nCaM	30	16	27	nCaM
31	16	35	nCaM	32	16	38	nCaM
33	16	39	nCaM	34	16	65	nCaM
35	16	68	nCaM	36	18	38	nCaM
37	19	27	nCaM	38	19	31	nCaM
39	19	34	nCaM	40	19	35	nCaM
41	19	37	nCaM	42	19	38	nCaM
43	20	26	nCaM	44	20	27	nCaM
45	21	31	nCaM	46	22	26	nCaM
47	22	28	nCaM	48	22	31	nCaM
49	22	62	nCaM	50	24	60	nCaM
51	24	62	nCaM	52	26	62	nCaM
53	26	63	nCaM	54	26	64	nCaM
55	27	31	nCaM	56	27	32	nCaM
57	27	35	nCaM	58	27	62	nCaM
59	27	63	nCaM	60	27	68	nCaM
61	28	32	nCaM	62	28	62	nCaM
63	29	49	nCaM	64	29	52	nCaM
65	29	56	nCaM	66	29	62	nCaM
67	30	45	nCaM	68	30	49	nCaM
69	31	35	nCaM	70	32	36	nCaM
71	32	48	nCaM	72	32	52	nCaM
73	32	63	nCaM	74	32	68	nCaM
75	32	71	nCaM	76	34	38	nCaM
77	35	39	nCaM	78	35	68	nCaM
79	36	41	nCaM	80	36	42	nCaM
81	36	43	nCaM	82	36	48	nCaM
83	36	68	nCaM	84	36	71	nCaM
85	36	72	nCaM	86	36	75	nCaM

87	37	41	nCaM	88	37	42	nCaM
89	39	68	nCaM	90	39	72	nCaM
91	41	72	nCaM	92	41	75	nCaM
93	43	47	nCaM	94	43	48	nCaM
95	43	75	nCaM	96	44	48	nCaM
97	45	49	nCaM	98	46	50	nCaM
99	47	51	nCaM	100	47	75	nCaM
101	48	52	nCaM	102	48	71	nCaM
103	48	75	nCaM	104	49	53	nCaM
105	49	56	nCaM	106	50	54	nCaM
107	51	55	nCaM	108	51	71	nCaM
109	51	74	nCaM	110	51	75	nCaM
111	52	56	nCaM	112	52	63	nCaM
113	52	71	nCaM	114	53	57	nCaM
115	54	71	nCaM	116	54	74	nCaM
117	55	63	nCaM	118	55	67	nCaM
119	55	71	nCaM	120	56	60	nCaM
121	58	67	nCaM	122	63	67	nCaM
123	63	68	nCaM	124	63	71	nCaM
125	64	68	nCaM	126	65	69	nCaM
127	67	71	nCaM	128	67	74	nCaM
129	68	72	nCaM	130	69	73	nCaM
131	70	74	nCaM	132	71	75	nCaM
133	1	114	nCaM:cCaM	134	1	143	nCaM:cCaM
135	1	147	nCaM:cCaM	136	1	148	nCaM:cCaM
137	1	81	nCaM:cCaM	138	1	85	nCaM:cCaM
139	2	124	nCaM:cCaM	140	2	146	nCaM:cCaM
141	2	147	nCaM:cCaM	142	2	148	nCaM:cCaM
143	2	76	nCaM:cCaM	144	2	77	nCaM:cCaM
145	2	80	nCaM:cCaM	146	2	81	nCaM:cCaM
147	2	82	nCaM:cCaM	148	2	84	nCaM:cCaM
149	3	109	nCaM:cCaM	150	3	114	nCaM:cCaM
151	3	115	nCaM:cCaM	152	3	116	nCaM:cCaM
153	3	124	nCaM:cCaM	154	3	145	nCaM:cCaM
155	3	147	nCaM:cCaM	156	3	148	nCaM:cCaM
157	4	147	nCaM:cCaM	158	4	76	nCaM:cCaM
159	5	147	nCaM:cCaM	160	5	148	nCaM:cCaM
161	6	115	nCaM:cCaM	162	7	114	nCaM:cCaM
163	7	115	nCaM:cCaM	164	7	116	nCaM:cCaM

165	7	120	nCaM:cCaM	166	7	124	nCaM:cCaM
167	7	148	nCaM:cCaM	168	8	115	nCaM:cCaM
169	8	120	nCaM:cCaM	170	8	147	nCaM:cCaM
171	8	76	nCaM:cCaM	172	8	81	nCaM:cCaM
173	8	84	nCaM:cCaM	174	11	115	nCaM:cCaM
175	11	120	nCaM:cCaM	176	11	76	nCaM:cCaM
177	12	76	nCaM:cCaM	178	14	115	nCaM:cCaM
179	39	76	nCaM:cCaM	180	41	76	nCaM:cCaM
181	41	79	nCaM:cCaM	182	41	84	nCaM:cCaM
183	66	143	nCaM:cCaM	184	67	139	nCaM:cCaM
185	67	143	nCaM:cCaM	186	69	76	nCaM:cCaM
187	72	76	nCaM:cCaM	188	73	77	nCaM:cCaM
189	73	80	nCaM:cCaM	190	74	80	nCaM:cCaM
191	75	79	nCaM:cCaM	192	76	80	cCaM
193	76	81	cCaM	194	76	84	cCaM
195	77	138	cCaM	196	77	82	cCaM
197	78	83	cCaM	198	79	83	cCaM
199	79	84	cCaM	200	80	146	cCaM
201	81	145	cCaM	202	81	146	cCaM
203	81	85	cCaM	204	82	138	cCaM
205	82	139	cCaM	206	82	142	cCaM
207	82	143	cCaM	208	82	146	cCaM
209	82	86	cCaM	210	82	89	cCaM
211	83	87	cCaM	212	84	88	cCaM
213	85	112	cCaM	214	85	114	cCaM
215	85	142	cCaM	216	85	145	cCaM
217	85	146	cCaM	218	85	89	cCaM
219	86	138	cCaM	220	86	139	cCaM
221	86	142	cCaM	222	86	143	cCaM
223	86	90	cCaM	224	87	91	cCaM
225	88	112	cCaM	226	88	92	cCaM
227	89	105	cCaM	228	89	108	cCaM
229	89	109	cCaM	230	89	112	cCaM
231	89	114	cCaM	232	89	138	cCaM
233	89	141	cCaM	234	89	142	cCaM
235	89	145	cCaM	236	89	93	cCaM
237	89	100	cCaM	238	90	138	cCaM
239	90	94	cCaM	240	92	104	cCaM
241	92	107	cCaM	242	92	108	cCaM

243	92	111	cCaM	244	92	112	cCaM
245	92	100	cCaM	246	93	104	cCaM
247	93	138	cCaM	248	93	97	cCaM
249	93	99	cCaM	250	93	100	cCaM
251	94	104	cCaM	252	97	135	cCaM
253	99	135	cCaM	254	99	136	cCaM
255	99	137	cCaM	256	99	138	cCaM
257	100	104	cCaM	258	100	105	cCaM
259	100	108	cCaM	260	100	135	cCaM
261	100	136	cCaM	262	100	138	cCaM
263	101	105	cCaM	264	101	135	cCaM
265	102	106	cCaM	266	102	130	cCaM
267	102	131	cCaM	268	102	135	cCaM
269	102	136	cCaM	270	103	107	cCaM
271	104	108	cCaM	272	104	111	cCaM
273	105	109	cCaM	274	105	130	cCaM
275	105	136	cCaM	276	105	138	cCaM
277	105	141	cCaM	278	106	110	cCaM
279	106	118	cCaM	280	106	121	cCaM
281	106	122	cCaM	282	106	125	cCaM
283	106	126	cCaM	284	106	130	cCaM
285	106	131	cCaM	286	107	111	cCaM
287	108	112	cCaM	288	109	114	cCaM
289	109	116	cCaM	290	109	121	cCaM
291	109	124	cCaM	292	109	125	cCaM
293	109	130	cCaM	294	109	141	cCaM
295	109	145	cCaM	296	110	114	cCaM
297	110	118	cCaM	298	114	145	cCaM
299	116	120	cCaM	300	116	121	cCaM
301	116	124	cCaM	302	116	145	cCaM
303	117	121	cCaM	304	118	122	cCaM
305	120	124	cCaM	306	121	125	cCaM
307	122	126	cCaM	308	123	127	cCaM
309	123	144	cCaM	310	123	148	cCaM
311	124	128	cCaM	312	124	141	cCaM
313	124	144	cCaM	314	124	145	cCaM
315	124	148	cCaM	316	125	129	cCaM
317	125	130	cCaM	318	125	141	cCaM
319	126	131	cCaM	320	127	140	cCaM

321	127	144	cCaM	322	127	148	cCaM
323	128	140	cCaM	324	128	141	cCaM
325	128	144	cCaM	326	129	140	cCaM
327	130	135	cCaM	328	130	136	cCaM
329	130	141	cCaM	330	136	141	cCaM
331	137	141	cCaM	332	138	142	cCaM
333	139	143	cCaM	334	140	144	cCaM
335	141	145	cCaM	336	142	146	cCaM
337	143	147	cCaM	338	143	148	cCaM
339	144	148	cCaM	340	1	16	nCaM:Ng
341	1	18	nCaM:Ng	342	2	13	nCaM:Ng
343	2	18	nCaM:Ng	344	2	22	nCaM:Ng
345	2	40	nCaM:Ng	346	3	15	nCaM:Ng
347	3	16	nCaM:Ng	348	3	17	nCaM:Ng
349	3	18	nCaM:Ng	350	3	20	nCaM:Ng
351	3	41	nCaM:Ng	352	3	43	nCaM:Ng
353	3	45	nCaM:Ng	354	4	16	nCaM:Ng
355	4	21	nCaM:Ng	356	4	41	nCaM:Ng
357	6	15	nCaM:Ng	358	7	44	nCaM:Ng
359	8	32	nCaM:Ng	360	8	45	nCaM:Ng
361	9	17	nCaM:Ng	362	9	41	nCaM:Ng
363	11	18	nCaM:Ng	364	11	32	nCaM:Ng
365	11	44	nCaM:Ng	366	12	17	nCaM:Ng
367	12	18	nCaM:Ng	368	12	19	nCaM:Ng
369	12	21	nCaM:Ng	370	12	41	nCaM:Ng
371	13	17	nCaM:Ng	372	15	16	nCaM:Ng
373	15	17	nCaM:Ng	374	16	17	nCaM:Ng
375	16	19	nCaM:Ng	376	18	16	nCaM:Ng
377	19	16	nCaM:Ng	378	26	19	nCaM:Ng
379	26	20	nCaM:Ng	380	27	21	nCaM:Ng
381	29	22	nCaM:Ng	382	32	21	nCaM:Ng
383	35	17	nCaM:Ng	384	38	15	nCaM:Ng
385	38	16	nCaM:Ng	386	38	17	nCaM:Ng
387	39	15	nCaM:Ng	388	39	16	nCaM:Ng
389	39	17	nCaM:Ng	390	39	18	nCaM:Ng
391	39	19	nCaM:Ng	392	41	34	nCaM:Ng
393	62	20	nCaM:Ng	394	62	21	nCaM:Ng
395	62	22	nCaM:Ng	396	63	20	nCaM:Ng
397	63	21	nCaM:Ng	398	65	17	nCaM:Ng

399	65	19	nCaM:Ng	400	65	21	nCaM:Ng
401	65	41	nCaM:Ng	402	65	46	nCaM:Ng
403	67	38	nCaM:Ng	404	67	44	nCaM:Ng
405	67	47	nCaM:Ng	406	68	17	nCaM:Ng
407	68	19	nCaM:Ng	408	68	41	nCaM:Ng
409	69	19	nCaM:Ng	410	69	20	nCaM:Ng
411	69	21	nCaM:Ng	412	69	41	nCaM:Ng
413	70	38	nCaM:Ng	414	72	19	nCaM:Ng
415	72	41	nCaM:Ng	416	74	37	nCaM:Ng
417	74	38	nCaM:Ng	418	74	43	nCaM:Ng
419	74	47	nCaM:Ng	420	76	20	cCaM:Ng
421	76	41	cCaM:Ng	422	76	43	cCaM:Ng
423	77	43	cCaM:Ng	424	77	44	cCaM:Ng
425	80	43	cCaM:Ng	426	81	27	cCaM:Ng
427	81	43	cCaM:Ng	428	82	18	cCaM:Ng
429	82	32	cCaM:Ng	430	82	37	cCaM:Ng
431	82	38	cCaM:Ng	432	82	40	cCaM:Ng
433	82	41	cCaM:Ng	434	82	43	cCaM:Ng
435	82	44	cCaM:Ng	436	84	33	cCaM:Ng
437	84	34	cCaM:Ng	438	84	38	cCaM:Ng
439	85	17	cCaM:Ng	440	85	18	cCaM:Ng
441	85	33	cCaM:Ng	442	86	23	cCaM:Ng
443	86	38	cCaM:Ng	444	88	33	cCaM:Ng
445	89	17	cCaM:Ng	446	89	18	cCaM:Ng
447	89	19	cCaM:Ng	448	89	21	cCaM:Ng
449	89	23	cCaM:Ng	450	89	37	cCaM:Ng
451	99	21	cCaM:Ng	452	99	22	cCaM:Ng
453	99	23	cCaM:Ng	454	99	24	cCaM:Ng
455	99	34	cCaM:Ng	456	99	38	cCaM:Ng
457	100	21	cCaM:Ng	458	105	17	cCaM:Ng
459	105	19	cCaM:Ng	460	105	21	cCaM:Ng
461	106	21	cCaM:Ng	462	109	17	cCaM:Ng
463	109	19	cCaM:Ng	464	109	21	cCaM:Ng
465	112	17	cCaM:Ng	466	112	33	cCaM:Ng
467	112	37	cCaM:Ng	468	114	15	cCaM:Ng
469	114	17	cCaM:Ng	470	114	32	cCaM:Ng
471	114	33	cCaM:Ng	472	114	36	cCaM:Ng
473	114	41	cCaM:Ng	474	114	43	cCaM:Ng
475	114	44	cCaM:Ng	476	115	18	cCaM:Ng

477	115	23	cCaM:Ng	478	115	32	cCaM:Ng
479	115	36	cCaM:Ng	480	115	40	cCaM:Ng
481	115	42	cCaM:Ng	482	115	43	cCaM:Ng
483	115	44	cCaM:Ng	484	115	45	cCaM:Ng
485	116	32	cCaM:Ng	486	120	32	cCaM:Ng
487	120	43	cCaM:Ng	488	120	44	cCaM:Ng
489	120	45	cCaM:Ng	490	124	32	cCaM:Ng
491	125	21	cCaM:Ng	492	130	21	cCaM:Ng
493	135	20	cCaM:Ng	494	135	22	cCaM:Ng
495	135	36	cCaM:Ng	496	136	20	cCaM:Ng
497	137	20	cCaM:Ng	498	138	18	cCaM:Ng
499	138	19	cCaM:Ng	500	138	20	cCaM:Ng
501	138	22	cCaM:Ng	502	138	23	cCaM:Ng
503	138	27	cCaM:Ng	504	138	32	cCaM:Ng
505	138	33	cCaM:Ng	506	138	36	cCaM:Ng
507	138	40	cCaM:Ng	508	138	43	cCaM:Ng
509	138	44	cCaM:Ng	510	141	17	cCaM:Ng
511	141	19	cCaM:Ng	512	141	21	cCaM:Ng
513	142	16	cCaM:Ng	514	142	17	cCaM:Ng
515	142	18	cCaM:Ng	516	145	16	cCaM:Ng
517	145	17	cCaM:Ng	518	145	32	cCaM:Ng
519	145	33	cCaM:Ng	520	145	41	cCaM:Ng
521	145	43	cCaM:Ng	522	145	44	cCaM:Ng
523	146	16	cCaM:Ng	524	146	33	cCaM:Ng
525	146	35	cCaM:Ng	526	146	36	cCaM:Ng
527	146	42	cCaM:Ng	528	147	47	cCaM:Ng
529	18	32	Ng	530	20	26	Ng
531	20	27	Ng	532	20	28	Ng
533	22	27	Ng	534	23	27	Ng
535	27	31	Ng	536	27	32	Ng
537	28	32	Ng	538	29	33	Ng
539	30	34	Ng	540	30	36	Ng
541	32	36	Ng	542	32	37	Ng
543	33	37	Ng	544	34	38	Ng
545	36	40	Ng	546	37	41	Ng
547	38	43	Ng	548	40	44	Ng
549	41	45	Ng	550	41	46	Ng
551	43	47	Ng				

Table S9 Charge distribution on a coarse-grained side-chain C_α model of full length Ng.

Residue Index	Residue Name	Charge on C_α	Error of the C_α charge	Charge on side-chain	Error of the side-chain charge
1	MET	0.847	0.002	0.091	0.001
2	ASP	-0.028	0.002	-0.898	0.001
3	CYS	-0.037	0.002	0.035	0.001
4	CYS	-0.037	0.002	0.033	0.001
5	THR	-0.098	0.002	0.083	0.001
6	GLU	-0.052	0.002	-0.872	0.002
7	SER	-0.098	0.002	0.066	0.001
8	ALA	-0.022	0.002	0.049	0.001
9	CYS	-0.042	0.002	0.025	0.001
10	SER	-0.077	0.002	0.062	0.001
11	LYS	-0.061	0.003	1.000	0.001
12	PRO	-0.307	0.003	0.316	0.002
13	ASP	-0.010	0.002	-0.926	0.002
14	ASP	-0.037	0.002	-0.927	0.002
15	ASP	-0.040	0.002	-0.930	0.002
16	ILE	-0.070	0.002	0.078	0.001
17	LEU	-0.043	0.002	0.061	0.001
18	ASP	-0.053	0.002	-0.925	0.002
19	ILE	-0.073	0.003	0.080	0.001
20	PRO	-0.301	0.003	0.305	0.002
21	LEU	-0.066	0.002	0.063	0.001
22	ASP	-0.034	0.002	-0.930	0.002
23	ASP	-0.035	0.003	-0.929	0.002
24	PRO	-0.335	0.003	0.320	0.002
25	GLY	-0.182	0.002	0.188	0.001
26	ALA	-0.041	0.002	0.063	0.001
27	ASN	-0.047	0.003	0.011	0.001
28	ALA	-0.043	0.002	0.067	0.001
29	ALA	-0.064	0.002	0.064	0.001
30	ALA	-0.068	0.002	0.068	0.001
31	ALA	-0.057	0.002	0.069	0.001
32	LYS	-0.079	0.002	1.020	0.002
33	ILE	-0.089	0.002	0.103	0.001
34	GLN	-0.077	0.002	0.069	0.001
35	ALA	-0.046	0.002	0.064	0.001
36	SER	-0.091	0.002	0.079	0.001
37	PHE	-0.052	0.002	0.062	0.001
38	ARG	-0.079	0.002	1.027	0.002
39	GLY	-0.181	0.002	0.192	0.001
40	HIS	-0.049	0.002	0.049	0.001
41	MET	-0.069	0.002	0.069	0.001
42	ALA	-0.062	0.002	0.068	0.001
43	ARG	-0.062	0.002	1.020	0.002
44	LYS	-0.074	0.002	1.031	0.002

45	LYS	-0.077	0.002	1.030	0.002
46	ILE	-0.093	0.002	0.110	0.001
47	LYS	-0.092	0.002	1.028	0.002
48	SER	-0.089	0.002	0.069	0.001
49	GLY	-0.176	0.002	0.191	0.001
50	GLU	-0.060	0.002	-0.846	0.002
51	CYS	-0.056	0.002	0.032	0.001
52	GLY	-0.185	0.002	0.194	0.001
53	ARG	-0.066	0.002	1.017	0.002
54	LYS	-0.072	0.002	1.029	0.002
55	GLY	-0.193	0.003	0.192	0.001
56	PRO	-0.341	0.003	0.342	0.002
57	GLY	-0.197	0.002	0.188	0.001
58	PRO	-0.305	0.002	0.322	0.001
59	GLY	-0.192	0.002	0.189	0.000
60	GLY	-0.202	0.002	0.189	0.000
61	PRO	-0.303	0.002	0.325	0.001
62	GLY	-0.203	0.002	0.192	0.001
63	GLY	-0.186	0.002	0.191	0.001
64	ALA	-0.057	0.002	0.064	0.001
65	GLY	-0.197	0.002	0.195	0.001
66	GLY	-0.185	0.002	0.191	0.001
67	ALA	-0.044	0.002	0.061	0.001
68	ARG	-0.093	0.002	1.032	0.002
69	GLY	-0.186	0.002	0.193	0.001
70	GLY	-0.186	0.002	0.193	0.001
71	ALA	-0.058	0.002	0.066	0.001
72	GLY	-0.203	0.002	0.188	0.001
73	GLY	-0.197	0.002	0.191	0.001
74	GLY	-0.202	0.002	0.189	0.001
75	PRO	-0.295	0.002	0.319	0.001
76	SER	-0.092	0.002	0.057	0.001
77	GLY	-0.230	0.002	0.189	0.000
78	ASP	-0.928	0.002	-0.959	0.002

The unit of charge is $1.6*10^{-19}$ C.

Table S10 Charge distribution on a coarse-grained side-chain C_α model of apoCaM at pH = 7.2, I = 0.15 M.

Residue Index	Residue Name	Charge on C_α	Error of the C_α charge	Charge on side-chain	Error of the side-chain charge
1	ALA	1.564	0.004	0.293	0.002
2	ASP	-0.038	0.003	-0.889	0.002
3	GLN	-0.050	0.003	0.051	0.001
4	LEU	-0.068	0.002	0.081	0.001
5	THR	-0.093	0.002	0.073	0.001
6	GLU	-0.051	0.002	-0.915	0.002

7	GLU	-0.077	0.003	-0.872	0.002
8	GLN	-0.033	0.003	0.052	0.001
9	ILE	-0.104	0.002	0.088	0.001
10	ALA	-0.063	0.003	0.059	0.001
11	GLU	-0.059	0.003	-0.865	0.002
12	PHE	-0.049	0.003	0.049	0.002
13	LYS	-0.087	0.003	1.001	0.001
14	GLU	-0.057	0.002	-0.893	0.002
15	ALA	-0.030	0.002	0.059	0.001
16	PHE	-0.063	0.003	0.050	0.001
17	SER	-0.109	0.003	0.070	0.001
18	LEU	-0.062	0.002	0.078	0.001
19	PHE	-0.060	0.003	0.048	0.001
20	ASP	-0.043	0.003	-0.885	0.002
21	LYS	-0.083	0.002	0.986	0.001
22	ASP	-0.049	0.002	-0.914	0.002
23	GLY	-0.195	0.002	0.190	0.000
24	ASP	-0.062	0.002	-0.948	0.001
25	GLY	-0.152	0.002	0.186	0.001
26	THR	-0.064	0.003	0.064	0.001
27	ILE	-0.060	0.003	0.076	0.001
28	THR	-0.110	0.003	0.077	0.001
29	THR	-0.094	0.003	0.087	0.001
30	LYS	-0.070	0.002	0.998	0.002
31	GLU	-0.059	0.003	-0.866	0.001
32	LEU	-0.081	0.003	0.082	0.001
33	GLY	-0.181	0.003	0.191	0.001
34	THR	-0.108	0.003	0.083	0.001
35	VAL	-0.080	0.003	0.105	0.001
36	MET	-0.068	0.003	0.061	0.001
37	ARG	-0.074	0.003	1.003	0.002
38	SER	-0.090	0.003	0.075	0.001
39	LEU	-0.076	0.003	0.083	0.001
40	GLY	-0.194	0.003	0.188	0.001
41	GLN	-0.034	0.002	0.054	0.001
42	ASN	-0.789	0.003	0.744	0.002
43	PRO	-0.249	0.003	0.267	0.003
44	THR	-0.090	0.002	0.071	0.001
45	GLU	-0.057	0.002	-0.902	0.002
46	ALA	-0.051	0.002	0.047	0.001
47	GLU	-0.050	0.002	-0.880	0.002
48	LEU	-0.066	0.002	0.070	0.001
49	GLN	-0.067	0.002	0.043	0.001

50	ASP	-0.032	0.003	-0.947	0.001
51	MET	-0.036	0.003	0.053	0.001
52	ILE	-0.085	0.003	0.097	0.001
53	ASN	-0.038	0.002	-0.005	0.001
54	GLU	-0.058	0.002	-0.902	0.002
55	VAL	-0.073	0.003	0.094	0.001
56	ASP	-0.066	0.002	-0.913	0.001
57	ALA	-0.063	0.002	0.055	0.001
58	ASP	-0.056	0.002	-0.940	0.002
59	GLY	-0.188	0.002	0.192	0.000
60	ASN	-0.052	0.002	0.005	0.001
61	GLY	-0.160	0.003	0.188	0.001
62	THR	-0.083	0.003	0.074	0.001
63	ILE	-0.080	0.003	0.082	0.001
64	ASP	-0.024	0.002	-0.936	0.006
65	PHE	-0.742	0.004	0.744	0.008
66	PRO	-0.256	0.004	0.256	0.005
67	GLU	-0.051	0.003	-0.878	0.002
68	PHE	-0.037	0.003	0.054	0.008
69	LEU	-0.070	0.004	0.077	0.001
70	THR	-0.103	0.003	0.073	0.001
71	MET	-0.056	0.003	0.069	0.006
72	MET	-0.067	0.003	0.066	0.004
73	ALA	-0.034	0.003	0.055	0.001
74	ARG	-0.079	0.002	0.980	0.002
75	LYS	-0.056	0.003	0.994	0.001
76	MET	-0.051	0.003	0.071	0.006
77	LYS	-0.055	0.002	0.979	0.001
78	ASP	-0.046	0.002	-0.940	0.001
79	THR	-0.086	0.002	0.080	0.001
80	ASP	-0.012	0.003	-0.912	0.002
81	SER	-0.088	0.002	0.065	0.001
82	GLU	-0.064	0.002	-0.862	0.002
83	GLU	-0.068	0.002	-0.869	0.002
84	GLU	-0.061	0.002	-0.875	0.002
85	ILE	-0.093	0.003	0.094	0.001
86	ARG	-0.064	0.003	0.944	0.002
87	GLU	-0.073	0.003	-0.862	0.002
88	ALA	-0.043	0.003	0.061	0.001
89	PHE	-0.054	0.003	0.049	0.001
90	ARG	-0.072	0.002	0.974	0.002
91	VAL	-0.084	0.002	0.097	0.001
92	PHE	-0.075	0.003	0.054	0.001

93	ASP	-0.025	0.003	-0.892	0.001
94	LYS	-0.082	0.002	0.995	0.001
95	ASP	-0.055	0.002	-0.941	0.002
96	GLY	-0.193	0.002	0.195	0.001
97	ASN	-0.041	0.002	0.000	0.001
98	GLY	-0.165	0.002	0.189	0.001
99	TYR	-0.036	0.002	0.031	0.001
100	ILE	-0.074	0.002	0.094	0.001
101	SER	-0.097	0.003	0.060	0.001
102	ALA	-0.061	0.002	0.066	0.001
103	ALA	-0.044	0.002	0.052	0.001
104	GLU	-0.049	0.003	-0.874	0.001
105	LEU	-0.095	0.003	0.076	0.001
106	ARG	-0.056	0.003	0.995	0.002
107	HIS	-0.055	0.003	0.028	0.001
108	VAL	-0.080	0.003	0.094	0.001
109	MET	-0.053	0.003	0.060	0.001
110	THR	-0.088	0.003	0.081	0.001
111	ASN	-0.039	0.003	0.000	0.001
112	LEU	-0.073	0.003	0.081	0.001
113	GLY	-0.194	0.002	0.188	0.001
114	GLU	-0.061	0.002	-0.871	0.002
115	LYS	-0.080	0.002	0.988	0.001
116	LEU	-0.060	0.002	0.072	0.001
117	THR	-0.091	0.003	0.070	0.001
118	ASP	-0.025	0.002	-0.921	0.002
119	GLU	-0.067	0.002	-0.915	0.002
120	GLU	-0.060	0.002	-0.884	0.002
121	VAL	-0.084	0.002	0.086	0.001
122	ASP	-0.030	0.002	-0.903	0.001
123	GLU	-0.058	0.002	-0.888	0.002
124	MET	-0.049	0.002	0.053	0.001
125	ILE	-0.087	0.002	0.087	0.001
126	ARG	-0.060	0.002	0.949	0.002
127	GLU	-0.079	0.002	-0.913	0.002
128	ALA	-0.051	0.002	0.054	0.001
129	ASP	-0.019	0.002	-0.952	0.001
130	ILE	-0.077	0.002	0.077	0.001
131	ASP	-0.038	0.002	-0.950	0.002
132	GLY	-0.210	0.002	0.188	0.001
133	ASP	-0.037	0.002	-0.955	0.001
134	GLY	-0.174	0.002	0.190	0.001
135	GLN	-0.037	0.002	0.029	0.001

136	VAL	-0.078	0.002	0.097	0.001
137	ASN	-0.023	0.002	0.001	0.001
138	TYR	-0.051	0.003	0.033	0.001
139	GLU	-0.071	0.003	-0.893	0.002
140	GLU	-0.052	0.002	-0.893	0.002
141	PHE	-0.061	0.003	0.055	0.003
142	VAL	-0.091	0.003	0.093	0.001
143	GLN	-0.058	0.003	0.044	0.001
144	MET	-0.045	0.002	0.059	0.001
145	MET	-0.053	0.002	0.056	0.003
146	THR	-0.118	0.002	0.076	0.001
147	ALA	-0.050	0.003	0.042	0.001
148	LYS	-2.693	0.005	2.669	0.003

The unit of charge is $1.6*10^{-19}$ C.

Table S11 Charge distribution on a coarse-grained side chain C_α model of Ng_{13-49} at pH = 7.2, I = 0.15 M.

Residue Index	Residue Name	Charge on Cα	Error of the Cα charge	Charge on side-chain	Error of the side-chain charge
13	ASP	1.574	0.004	-1.659	0.004
14	ASP	-0.020	0.003	-0.921	0.002
15	ASP	-0.034	0.003	-0.911	0.002
16	ILE	-0.071	0.003	0.081	0.001
17	LEU	-0.060	0.002	0.061	0.001
18	ASP	-0.037	0.003	-0.916	0.002
19	ILE	-0.539	0.004	0.540	0.002
20	PRO	-0.265	0.003	0.258	0.003
21	LEU	-0.054	0.003	0.064	0.001
22	ASP	-0.042	0.003	-0.931	0.002
23	ASP	-0.818	0.004	-0.133	0.003
24	PRO	-0.282	0.003	0.261	0.002
25	GLY	-0.181	0.003	0.187	0.001
26	ALA	-0.032	0.002	0.054	0.001
27	ASN	-0.046	0.003	0.005	0.001
28	ALA	-0.040	0.003	0.064	0.001
29	ALA	-0.059	0.003	0.060	0.001
30	ALA	-0.066	0.003	0.062	0.001
31	ALA	-0.055	0.003	0.064	0.001
32	LYS	-0.079	0.003	0.996	0.001
33	ILE	-0.089	0.003	0.100	0.001
34	GLN	-0.072	0.002	0.059	0.001
35	ALA	-0.042	0.003	0.062	0.001
36	SER	-0.092	0.003	0.076	0.001
37	PHE	-0.048	0.003	0.057	0.001
38	ARG	-0.077	0.002	1.002	0.002
39	GLY	-0.175	0.002	0.187	0.001

40	HIS	-0.054	0.002	0.037	0.001
41	MET	-0.068	0.002	0.065	0.001
42	ALA	-0.057	0.003	0.064	0.001
43	ARG	-0.059	0.003	0.984	0.002
44	LYS	-0.061	0.003	0.997	0.002
45	LYS	-0.084	0.002	1.000	0.002
46	ILE	-0.089	0.003	0.106	0.001
47	LYS	-0.075	0.003	1.001	0.002
48	SER	-0.132	0.003	0.066	0.001
49	GLY	-1.086	0.002	0.174	0.001

The unit of charge is $1.6*10^{-19}$ C.

Table S12 Charge distribution on a coarse-grained side chain C_α model of apoCaM at pH = 6.3, I = 0.1 M.

Residue Index	Residue Name	Charge on C_α	Error on the C_α charge	Charge on the side-chain	Error on the side-chain charge
1	ALA	1.565	0.004	0.292	0.001
2	ASP	-0.041	0.003	-0.885	0.002
3	GLN	-0.052	0.002	0.053	0.001
4	LEU	-0.066	0.002	0.080	0.001
5	THR	-0.095	0.002	0.075	0.001
6	GLU	-0.051	0.002	-0.914	0.002
7	GLU	-0.072	0.003	-0.871	0.002
8	GLN	-0.040	0.003	0.053	0.001
9	ILE	-0.098	0.003	0.088	0.001
10	ALA	-0.061	0.003	0.058	0.001
11	GLU	-0.061	0.003	-0.868	0.002
12	PHE	-0.047	0.003	0.047	0.001
13	LYS	-0.090	0.003	0.998	0.001
14	GLU	-0.055	0.003	-0.893	0.002
15	ALA	-0.035	0.002	0.058	0.001
16	PHE	-0.055	0.003	0.053	0.001
17	SER	-0.118	0.003	0.070	0.001
18	LEU	-0.060	0.002	0.076	0.001
19	PHE	-0.054	0.003	0.047	0.001
20	ASP	-0.049	0.003	-0.885	0.002
21	LYS	-0.078	0.002	0.986	0.001
22	ASP	-0.049	0.003	-0.914	0.002
23	GLY	-0.196	0.002	0.191	0.001
24	ASP	-0.064	0.002	-0.945	0.001
25	GLY	-0.150	0.002	0.184	0.001
26	THR	-0.062	0.003	0.065	0.001
27	ILE	-0.066	0.003	0.077	0.001
28	THR	-0.107	0.003	0.079	0.001

29	THR	-0.095	0.003	0.089	0.001
30	LYS	-0.072	0.003	0.995	0.001
31	GLU	-0.058	0.003	-0.866	0.002
32	LEU	-0.079	0.003	0.082	0.001
33	GLY	-0.180	0.003	0.191	0.001
34	THR	-0.108	0.003	0.081	0.001
35	VAL	-0.082	0.003	0.108	0.001
36	MET	-0.066	0.003	0.061	0.001
37	ARG	-0.080	0.002	1.006	0.002
38	SER	-0.088	0.003	0.072	0.001
39	LEU	-0.077	0.003	0.085	0.001
40	GLY	-0.192	0.003	0.188	0.001
41	GLN	-0.033	0.002	0.054	0.001
42	ASN	-0.782	0.004	0.744	0.002
43	PRO	-0.254	0.004	0.265	0.002
44	THR	-0.092	0.002	0.070	0.001
45	GLU	-0.052	0.002	-0.904	0.002
46	ALA	-0.053	0.002	0.049	0.001
47	GLU	-0.052	0.002	-0.883	0.002
48	LEU	-0.064	0.002	0.069	0.001
49	GLN	-0.069	0.002	0.044	0.001
50	ASP	-0.031	0.002	-0.947	0.001
51	MET	-0.037	0.003	0.054	0.001
52	ILE	-0.084	0.003	0.096	0.001
53	ASN	-0.038	0.002	-0.006	0.001
54	GLU	-0.052	0.002	-0.902	0.002
55	VAL	-0.077	0.003	0.093	0.001
56	ASP	-0.062	0.002	-0.915	0.001
57	ALA	-0.063	0.002	0.053	0.001
58	ASP	-0.056	0.002	-0.940	0.001
59	GLY	-0.187	0.002	0.192	0.000
60	ASN	-0.055	0.002	0.004	0.001
61	GLY	-0.162	0.002	0.187	0.001
62	THR	-0.075	0.003	0.073	0.001
63	ILE	-0.083	0.003	0.079	0.001
64	ASP	-0.017	0.003	-0.945	0.001
65	PHE	-0.740	0.004	0.732	0.002
66	PRO	-0.260	0.003	0.251	0.002
67	GLU	-0.055	0.003	-0.880	0.002
68	PHE	-0.045	0.003	0.048	0.007
69	LEU	-0.061	0.002	0.075	0.001
70	THR	-0.103	0.002	0.072	0.001
71	MET	-0.053	0.003	0.061	0.001

72	MET	-0.066	0.003	0.066	0.006
73	ALA	-0.040	0.002	0.057	0.001
74	ARG	-0.073	0.002	0.978	0.002
75	LYS	-0.063	0.002	0.994	0.001
76	MET	-0.051	0.003	0.066	0.001
77	LYS	-0.056	0.003	0.980	0.001
78	ASP	-0.045	0.002	-0.939	0.001
79	THR	-0.083	0.002	0.079	0.001
80	ASP	-0.018	0.003	-0.912	0.001
81	SER	-0.088	0.002	0.067	0.001
82	GLU	-0.062	0.002	-0.865	0.002
83	GLU	-0.065	0.002	-0.869	0.002
84	GLU	-0.066	0.003	-0.875	0.002
85	ILE	-0.093	0.003	0.094	0.001
86	ARG	-0.067	0.002	0.945	0.002
87	GLU	-0.074	0.003	-0.859	0.002
88	ALA	-0.040	0.003	0.061	0.001
89	PHE	-0.055	0.003	0.049	0.001
90	ARG	-0.076	0.003	0.974	0.002
91	VAL	-0.081	0.003	0.097	0.001
92	PHE	-0.074	0.003	0.053	0.003
93	ASP	-0.025	0.003	-0.893	0.002
94	LYS	-0.078	0.002	0.999	0.001
95	ASP	-0.055	0.002	-0.940	0.001
96	GLY	-0.189	0.002	0.194	0.001
97	ASN	-0.044	0.002	0.000	0.001
98	GLY	-0.164	0.002	0.189	0.001
99	TYR	-0.031	0.002	0.032	0.001
100	ILE	-0.075	0.002	0.092	0.001
101	SER	-0.103	0.002	0.063	0.001
102	ALA	-0.062	0.002	0.070	0.001
103	ALA	-0.045	0.002	0.054	0.001
104	GLU	-0.050	0.003	-0.870	0.001
105	LEU	-0.096	0.003	0.080	0.001
106	ARG	-0.053	0.003	0.995	0.002
107	HIS	-0.048	0.003	0.973	0.002
108	VAL	-0.087	0.003	0.098	0.001
109	MET	-0.053	0.003	0.064	0.001
110	THR	-0.087	0.003	0.082	0.001
111	ASN	-0.043	0.003	0.002	0.001
112	LEU	-0.068	0.003	0.085	0.001
113	GLY	-0.193	0.002	0.187	0.001
114	GLU	-0.062	0.002	-0.870	0.002

115	LYS	-0.083	0.002	0.988	0.002
116	LEU	-0.061	0.003	0.074	0.001
117	THR	-0.090	0.003	0.068	0.001
118	ASP	-0.025	0.003	-0.921	0.002
119	GLU	-0.064	0.002	-0.913	0.002
120	GLU	-0.061	0.002	-0.883	0.002
121	VAL	-0.086	0.002	0.085	0.001
122	ASP	-0.029	0.002	-0.901	0.001
123	GLU	-0.061	0.002	-0.888	0.002
124	MET	-0.045	0.002	0.052	0.001
125	ILE	-0.084	0.003	0.086	0.001
126	ARG	-0.065	0.002	0.950	0.002
127	GLU	-0.075	0.002	-0.912	0.002
128	ALA	-0.051	0.002	0.053	0.001
129	ASP	-0.023	0.003	-0.955	0.001
130	ILE	-0.072	0.003	0.075	0.001
131	ASP	-0.040	0.002	-0.951	0.002
132	GLY	-0.211	0.002	0.190	0.001
133	ASP	-0.038	0.002	-0.955	0.001
134	GLY	-0.172	0.002	0.190	0.001
135	GLN	-0.042	0.002	0.030	0.001
136	VAL	-0.077	0.002	0.099	0.001
137	ASN	-0.025	0.002	0.003	0.001
138	TYR	-0.053	0.002	0.033	0.001
139	GLU	-0.068	0.002	-0.891	0.002
140	GLU	-0.050	0.002	-0.896	0.001
141	PHE	-0.059	0.003	0.048	0.001
142	VAL	-0.093	0.003	0.093	0.001
143	GLN	-0.056	0.002	0.045	0.001
144	MET	-0.047	0.002	0.060	0.001
145	MET	-0.052	0.002	0.058	0.001
146	THR	-0.118	0.002	0.077	0.001
147	ALA	-0.055	0.002	0.041	0.001
148	LYS	-2.680	0.005	2.666	0.003

The unit of charge is $1.6*10^{-19}$ C.

Table S13 Charge distribution on a coarse-grained side chain C_α model of Ng_{13-49}. pH = 6.3, I = 0.1 M.

Residue Index	Residue Name	Charge on C_α	Error of the C_α charge	Charge on the side-chain	Error of the side-chain charge
13	ASP	1.570	0.004	-1.657	0.003
14	ASP	-0.017	0.003	-0.918	0.002
15	ASP	-0.035	0.003	-0.909	0.002

16	ILE	-0.075	0.003	0.083	0.001
17	LEU	-0.059	0.003	0.063	0.001
18	ASP	-0.033	0.003	-0.918	0.002
19	ILE	-0.539	0.004	0.541	0.002
20	PRO	-0.269	0.003	0.257	0.002
21	LEU	-0.055	0.003	0.065	0.001
22	ASP	-0.039	0.002	-0.932	0.002
23	ASP	-0.813	0.004	-0.137	0.003
24	PRO	-0.286	0.003	0.265	0.002
25	GLY	-0.179	0.003	0.186	0.001
26	ALA	-0.031	0.002	0.056	0.001
27	ASN	-0.049	0.003	0.007	0.001
28	ALA	-0.045	0.003	0.063	0.001
29	ALA	-0.051	0.003	0.059	0.001
30	ALA	-0.070	0.003	0.063	0.001
31	ALA	-0.048	0.003	0.062	0.001
32	LYS	-0.084	0.003	0.998	0.001
33	ILE	-0.085	0.003	0.100	0.001
34	GLN	-0.074	0.002	0.060	0.001
35	ALA	-0.041	0.002	0.063	0.001
36	SER	-0.090	0.003	0.075	0.001
37	PHE	-0.050	0.003	0.057	0.001
38	ARG	-0.080	0.002	1.001	0.002
39	GLY	-0.174	0.003	0.189	0.001
40	HIS	-0.056	0.002	0.039	0.001
41	MET	-0.069	0.002	0.065	0.001
42	ALA	-0.062	0.002	0.067	0.001
43	ARG	-0.057	0.003	0.987	0.002
44	LYS	-0.064	0.003	0.996	0.002
45	LYS	-0.081	0.002	1.000	0.002
46	ILE	-0.091	0.003	0.105	0.001
47	LYS	-0.073	0.003	0.999	0.002
48	SER	-0.131	0.003	0.065	0.001
49	GLY	-1.086	0.002	0.176	0.001

The unit of charge is $1.6*10^{-19}$ C.

Table S14. Detail of the neurogranin (Ng) sequences obtained from UniprotKB (www.uniprot.org). 45 sequences of Ng with unique Uniprot ID were used for the sequence alignment (see Fig. S15).

Table S14(a)

Entry	Entry name	Status	Protein names
R0LF54	R0LF54_ANAPL	unreviewed	Neurogranin (Fragment)
A0A091VHN8	A0A091VHN8_NIPNI	unreviewed	Neurogranin (Fragment)
A0A091GKM6	A0A091GKM6_9AVES	unreviewed	Neurogranin (Fragment)
C1BLT4	C1BLT4_OSMMO	unreviewed	Neurogranin
A0A091IY17	A0A091IY17_CALAN	unreviewed	Neurogranin (Fragment)
A0A093QNN0	A0A093QNN0_9PASS	unreviewed	Neurogranin (Fragment)
A0A151NG69	A0A151NG69_ALLMI	unreviewed	Neurogranin
A0A099ZXB1	A0A099ZXB1_TINGU	unreviewed	Neurogranin (Fragment)
A0A091HFZ3	A0A091HFZ3_BUCRH	unreviewed	Neurogranin (Fragment)
A0A091JUZ1	A0A091JUZ1_COLST	unreviewed	Neurogranin (Fragment)
A0A093C189	A0A093C189_TAUER	unreviewed	Neurogranin (Fragment)
A0A087QQ10	A0A087QQ10_APTFO	unreviewed	Neurogranin (Fragment)
A0A091PSL5	A0A091PSL5_HALAL	unreviewed	Neurogranin (Fragment)
B5X303	B5X303_SALSA	unreviewed	Neurogranin
A0A091SXC0	A0A091SXC0_NESNO	unreviewed	Neurogranin (Fragment)
V8NXH4	V8NXH4_OPHHA	unreviewed	Neurogranin (Fragment)
Q90X80	Q90X80_GILMI	unreviewed	Neurogranin
A0A0A0AIY4	A0A0A0AIY4_CHAVO	unreviewed	Neurogranin (Fragment)
A0A093IXQ6	A0A093IXQ6_PICPB	unreviewed	Neurogranin (Fragment)
B7TWQ3	B7TWQ3_DANRE	unreviewed	Neurogranin (Fragment)
C1J0K5	C1J0K5_GILSE	unreviewed	Neurogranin (Fragment)
A0A091L3D5	A0A091L3D5_CATAU	unreviewed	Neurogranin (Fragment)
B5XEB8	B5XEB8_SALSA	unreviewed	Neurogranin
B2XBN6	B2XBN6_SALSA	unreviewed	Neurogranin
A0A091EYH6	A0A091EYH6_CORBR	unreviewed	Neurogranin (Fragment)
B9ENE5	B9ENE5_SALSA	unreviewed	Neurogranin
B9EPB8	B9EPB8_SALSA	unreviewed	Neurogranin
A0A094KAF2	A0A094KAF2_ANTCR	unreviewed	Neurogranin (Fragment)
M7AY01	M7AY01_CHEMY	unreviewed	Neurogranin (Fragment)
P54866	NEUG_SERCA	reviewed	Neurogranin (NG) (Canarigranin)

Q04940	NEUG_RAT	reviewed	Neurogranin (Ng) (Protein kinase C substrate 7.5 kDa protein) (RC3) [Cleaved into: NEUG(55-78)]
P35722	NEUG_BOVIN	reviewed	Neurogranin (NG) (B-50 immunoreactive C-kinase substrate) (BICKS) (p17) [Cleaved into: NEUG(55-78)]
P60761	NEUG_MOUSE	reviewed	Neurogranin (Ng) (RC3) [Cleaved into: NEUG(55-78)]
P54877	NEUG_CAPHI	reviewed	Neurogranin (NG) (Protein kinase C substrate 7.5 kDa protein) (RC3) [Cleaved into: NEUG(55-78)]
Q92686	NEUG_HUMAN	reviewed	Neurogranin (Ng) (RC3) [Cleaved into: NEUG(55-78)]
G3IJI8	G3IJI8_CRIGR	unreviewed	Neurogranin
A0A0A7HRI9	A0A0A7HRI9_HETGA	unreviewed	Neurogranin
U3E6M0	U3E6M0_CALJA	unreviewed	Neurogranin
F6THG0	F6THG0_MACMU	unreviewed	Neurogranin (Uncharacterized protein)
U3FBA0	U3FBA0_CALJA	unreviewed	Neurogranin
G9KDZ7	G9KDZ7_MUSPF	unreviewed	Neurogranin (Fragment)
A0A024R3M7	A0A024R3M7_HUMAN	unreviewed	Neurogranin (Protein kinase C substrate, RC3), isoform CRA_a
B5KFR8	B5KFR8_TAEGU	unreviewed	Putative neurogranin
B2CZA9	B2CZA9_COTJA	unreviewed	RC3/neurogranin (Fragment)
C7ENZ8	C7ENZ8_LARAR	unreviewed	RC3/neurogranin

Table S14(b)

Entry	Gene names	Organism	Length
R0LF54	Anapl_11925	Anas platyrhynchos (Mallard) (Anas boschas)	67
A0A091VHN8	Y956_07381	Nipponia nippon (Crested ibis) (Ibis nippon)	68
A0A091GKM6	N303_10815	Cuculus canorus (common cuckoo)	66
C1BLT4	NEUG	Osmerus mordax (Rainbow smelt) (Atherina mordax)	60
A0A091IY17	N300_07147	Calypte anna (Anna's hummingbird) (Archilochus anna)	59
A0A093QNN0	N305_02066	Manacus vitellinus (golden-collared manakin)	65
A0A151NG69	NRGN Y1Q_0010222	Alligator mississippiensis (American alligator)	109
A0A099ZXB1	N309_12965	Tinamus guttatus (White-throated tinamou)	55
A0A091HFZ3	N320_07902	Buceros rhinoceros silvestris	56
A0A091JUZ1	N325_03244	Colius striatus (Speckled mousebird)	64
A0A093C189	N340_06990	Tauraco erythrolophus (Red-crested turaco)	56
A0A087QQ10	AS27_01314	Aptenodytes forsteri (Emperor penguin)	63
A0A091PSL5	N329_06668	Haliaeetus albicilla (White-tailed sea-eagle)	47
B5X303	NEUG	Salmo salar (Atlantic salmon)	94
A0A091SXC0	N333_05461	Nestor notabilis (Kea)	66

V8NXH4	NRGN L345_07452	Ophiophagus hannah (King cobra) (Naja hannah)	78
Q90X80		Gillichthys mirabilis (Long-jawed mudsucker)	65
A0A0A0AIY4	N301_09391	Charadrius vociferus (Killdeer) (Aegialitis vocifera)	69
A0A093IXQ6	N307_15420	Picoides pubescens (Downy woodpecker) (Dryobates pubescens)	65
B7TWQ3	nrgna nrgn	Danio rerio (Zebrafish) (Brachydanio rerio)	92
C1J0K5		Gillichthys seta (Shortjaw mudsucker)	56
A0A091L3D5	N323_11621	Cathartes aura (Turkey vulture) (Vultur aura)	60
B5XEB8	NEUG	Salmo salar (Atlantic salmon)	94
B2XBN6		Salmo salar (Atlantic salmon)	60
A0A091EYH6	N302_11071	Corvus brachyrhynchos (American crow)	68
B9ENE5	NEUG	Salmo salar (Atlantic salmon)	93
B9EPB8	NEUG	Salmo salar (Atlantic salmon)	60
A0A094KAF2	N321_02863	Antrostomus carolinensis (Chuck-will's-widow) (Caprimulgus carolinensis)	66
M7AY01	UY3_13287	Chelonia mydas (Green sea-turtle) (Chelonia agassizi)	68
P54866	NRGN HAT14	Serinus canaria (Island canary) (Fringilla canaria)	73
Q04940	Nrgn	Rattus norvegicus (Rat)	78
P35722	NRGN	Bos taurus (Bovine)	78
P60761	Nrgn	Mus musculus (Mouse)	78
P54877	NRGN	Capra hircus (Goat)	78
Q92686	NRGN	Homo sapiens (Human)	78
G3IJI8	I79_024023	Cricetulus griseus (Chinese hamster) (Cricetulus barabensis griseus)	87
A0A0A7HRI9		Heterocephalus glaber (Naked mole rat)	78
U3E6M0	NRGN	Callithrix jacchus (White-tufted-ear marmoset)	78
F6THG0	NRGN	Macaca mulatta (Rhesus macaque)	78
U3FBA0	NRGN	Callithrix jacchus (White-tufted-ear marmoset)	78
G9KDZ7		Mustela putorius furo (European domestic ferret) (Mustela furo)	126
A0A024R3M7	NRGN hCG_1732008	Homo sapiens (Human)	78
B5KFR8		Taeniopygia guttata (Zebra finch) (Poephila guttata)	94
B2CZA9		Coturnix coturnix japonica (Japanese quail) (Coturnix japonica)	62
C7ENZ8		Larus argentatus (Herring gull)	73

References

1. Jurado, L. A., P. S. Chockalingam, and H. W. Jarrett. 1999. Apocalmodulin. Physiological reviews 79:661-682.
2. Bahler, M., and A. Rhoads. 2002. Calmodulin signaling via the IQ motif. FEBS Lett 513:107-113.
3. Ran, X., H. H. Miao, F. S. Sheu, and D. Yang. 2003. Structural and dynamic characterization of a neuron-specific protein kinase C substrate, neurogranin. Biochemistry 42:5143-5150.
4. Hoffman, L., A. Chandrasekar, X. Wang, J. A. Putkey, and M. N. Waxham. 2014. Neurogranin alters the structure and calcium binding properties of calmodulin. J Biol Chem 289:14644-14655.
5. Sugita, Y., and Y. Okamoto. 1999. Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 314:141-151.
6. Sanbonmatsu, K. Y., and A. E. Garcia. 2002. Structure of Met-enkephalin in explicit aqueous solution using replica exchange molecular dynamics. Proteins 46:225-234.
7. Cheung, M. S., J. M. Finke, B. Callahan, and J. N. Onuchic. 2003. Exploring the interplay between topology and secondary structural formation in the protein folding problem. J. Phys. Chem. B 107:11193-11200.
8. Wang, Q., P. Zhang, L. Hoffman, S. Tripathi, D. Homouz, Y. Liu, M. N. Waxham, and M. S. Cheung. 2013. Protein recognition and selection through conformational and mutually induced fit. Proceedings of the National Academy of Sciences of the United States of America 110:20545-20550.
9. Yang, Y., E. Faraggi, H. Zhao, and Y. Zhou. 2011. Improving protein fold recognition and template-based modeling by employing probabilistic-based matching between predicted one-dimensional structural properties of query and corresponding native properties of templates. Bioinformatics 27:2076-2082.
10. Debye, P., and E. Hückel. 1923. The theory of electrolytes. I. Lowering of freezing point and related phenomena. Physikalische Zeitschrift 24:185-206.
11. Betancourt, M. R., and D. Thirumalai. 1999. Pair potentials for protein folding: Choice of reference states and sensitivity of predicted native states to variations in the interaction schemes. Protein Sci. 8:361-369.
12. Karanicolas, J., and C. L. Brooks, 3rd. 2002. The origins of asymmetry in the folding transition states of protein L and protein G. Protein Sci 11:2351-2361.
13. Andreasen, T. J., C. W. Luetje, W. Heideman, and D. R. Storm. 1983. Purification of a novel calmodulin binding protein from bovine cerebral cortex membranes. Biochemistry 22:4615-4618.
14. Wang, Q., K. C. Liang, A. Czader, M. N. Waxham, and M. S. Cheung. 2011. The Effect of Macromolecular Crowding, Ionic Strength and Calcium Binding on Calmodulin Dynamics. PLoS Comput. Biol. 7:16.
15. Dima, R. I., and D. Thirumalai. 2004. Asymmetry in the shapes of folded and denatured states of proteins. J. Phys. Chem. B 108:6564-6570.
16. Camacho, C. J., and D. Thirumalai. 1993. Kinetics and Thermodynamics of Folding in Model Proteins. Proceedings of the National Academy of Sciences of the United States of America 90:6369-6372.

17. Samiotakis, A., D. Homouz, and M. S. Cheung. 2010. Multiscale investigation of chemical interference in proteins. J Chem Phys 132:175101.
18. Anandakrishnan, R., B. Aguilar, and A. V. Onufriev. 2012. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. Nucleic Acids Res 40:W537-541.
19. Stewart, J. J. 2007. Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. J Mol Model 13:1173-1213.
20. Kumar, S., D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg. 1992. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules .1. The Method. J Comput Chem 13:1011-1021.
21. Grossfield, A. 2013. WHAM: an implementation of the weighted histogram analysis method. http://membrane.urmc.rochester.edu/content/wham Version 2.0.4.
22. Roux, B. 1995. The Calculation of the Potential of Mean Force Using Computer-Simulations. Comput Phys Commun 91:275-282.
23. Pall, S., M. J. Abraham, C. Kutzner, B. Hess, and E. Lindahl. 2015. Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS. Lect Notes Comput Sc 8759:3-27.
24. Lindorff-Larsen, K., S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw. 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins-Structure Function and Bioinformatics 78:1950-1958.
25. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of Simple Potential Functions for Simulating Liquid Water. Journal of Chemical Physics 79:926-935.
26. Darden, T., D. York, and L. Pedersen. 1993. Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. Journal of Chemical Physics 98:10089-10092.
27. Parrinello, M., and A. Rahman. 1981. Polymorphic Transitions in Single-Crystals - a New Molecular-Dynamics Method. J Appl Phys 52:7182-7190.
28. Hess, B., H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije. 1997. LINCS: A linear constraint solver for molecular simulations. J Comput Chem 18:1463-1472.
29. Hendrix, D. A., and C. Jarzynski. 2001. A "fast growth" method of computing free energy differences. Journal of Chemical Physics 114:5974-5981.
30. Zhang, D., J. Gullingsrud, and J. A. McCammon. 2006. Potentials of mean force for acetylcholine unbinding from the alpha7 nicotinic acetylcholine receptor ligand-binding domain. J Am Chem Soc 128:3019-3026.
31. Ytreberg, F. M., and D. M. Zuckerman. 2004. Efficient use of nonequilibrium measurement to estimate free energy differences for molecular systems. J Comput Chem 25:1749-1759.
32. Jarzynski, C. 1997. Nonequilibrium equality for free energy differences. Phys Rev Lett 78:2690-2693.
33. Lipari, G., and A. Szabo. 1982. Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .1. Theory and Range of Validity. J Am Chem Soc 104:4546-4559.
34. Berjanskii, M., and D. S. Wishart. 2006. NMR: prediction of protein flexibility. Nat Protoc 1:683-688.

35. Kudlay, A., M. S. Cheung, and D. Thirumalai. 2009. Crowding effects on the structural transitions in a flexible helical homopolymer. Phys Rev Lett 102:118101.
36. Weinkam, P., E. V. Pletneva, H. B. Gray, J. R. Winkler, and P. G. Wolynes. 2009. Electrostatic effects on funneled landscapes and structural diversity in denatured protein ensembles. Proc Natl Acad Sci U S A 106:1796-1801.
37. Carpenter, G. A., and S. Grossberg. 1987. ART 2: self-organization of stable category recognition codes for analog input patterns. Appl Opt 26:4919-4930.
38. Cheung, M. S., and D. Thirumalai. 2007. Effects of crowding and confinement on the structures of the transition state ensemble in proteins. J. Phys. Chem. B 111:8250-8257.
39. Guo, Z., and D. Thirumalai. 1997. The nucleation-collapse mechanism in protein folding: evidence for the non-uniqueness of the folding nucleus. Folding and Design 2:377-391.
40. Waterhouse, A. M., J. B. Procter, D. M. Martin, M. Clamp, and G. J. Barton. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189-1191.