Electrostatic control of calcineurin's intrinsically-disordered regulatory domain binding to calmodulin

Bin Sun[a], Erik C. Cook[b], Trevor P. Creamer[b], Peter M. Kekenes-Huskey[a,*]

[a] Department of Chemistry, University of Kentucky, 505 Rose St., Chemistry-Physics Building, Lexington, KY, USA 40506
[b] Department of Molecular and Cellular Biochemistry, University of Kentucky, 741 South Limestone, St. Lexington, KY, USA 40536

ABSTRACT

Calcineurin (CaN) is a serine/threonine phosphatase that regulates a variety of physiological and pathophysiological processes in mammalian tissue. The calcineurin (CaN) regulatory domain (RD) is responsible for regulating the enzyme's phosphatase activity, and is believed to be highly-disordered when inhibiting CaN, but undergoes a disorder-to-order transition upon diffusion-limited binding with the regulatory protein calmodulin (CaM). The prevalence of polar and charged amino acids in the regulatory domain (RD) suggests electrostatic interactions are involved in mediating calmodulin (CaM) binding, yet the lack of atomistic-resolution data for the bound complex has stymied efforts to probe how the RD sequence controls its conformational ensemble and long-range attractions contribute to target protein binding. In the present study, we investigated via computational modeling the extent to which electrostatics and structural disorder facilitate CaM/CaN association kinetics. Specifically, we examined several RD constructs that contain the CaM binding region (CAMBR) to characterize the roles of electrostatics versus conformational diversity in controlling diffusion-limited association rates, via microsecond-scale molecular dynamics (MD) and Brownian dynamic (BD) simulations. Our results indicate that the RD amino acid composition and sequence length influence both the dynamic availability of conformations amenable to CaM binding, as well as long-range electrostatic interactions to steer association. These findings provide intriguing insight into the interplay between conformational diversity and electrostatically-driven protein-protein association involving CaN, which are likely to extend to wide-ranging diffusion-limited processes regulated by intrinsically-disordered proteins.

1. Introduction

Calcineurin (CaN) is a ubiquitously expressed protein that controls myriad developmental and signaling processes [1, 2]. It is chiefly regulated by calmodulin (CaM), one of the most prolific proteins in terms of its role in shaping intracellular signal transduction cascades. Despite the fundamental importance of CaM-regulated CaN phosphatase activity in organism physiology, the molecular mechanisms governing this process are incompletely understood. CaM/CaN is a prototypical example of a protein/protein complex involving a globular protein (CaM) and an intrinsically disordered binding domain (CaN) [3, 4], for which structural details of the protein/protein complex are restricted to intact CaM bound to a small fragment of the CaN regulatory peptide. In this regard, the CaM/CaN complex is similar to the tens of CaM/protein target complexes [5] that have resisted structure determination methods beyond the binding of short peptides. Remarkably, despite the CaN regulatory domain presenting little stabilized secondary structure, the CaM/CaN complex binds with picomolar affinity [6], afforded in part by rapid, diffusion-limited binding.

CaN is heterodimeric protein consisting of two domains: chain A (57–61 kDa) and chain B (19 kDa) [2, 7], while CaM (17 kDa) is comprised of two alpha-helix rich domains capable of binding Ca^{2+}. At

Ca^{2+} concentrations typical of resting cells (50 to 100 nM) [8], CaN phosphatase activity is negligible and CaM is believed to be in Ca^{2+}-free state [9]. Under these conditions, the CaN catalytic domain is inhibited by the protein's auto-inhibitory domain (AID). At rising Ca^{2+} concentrations, the CaN AID dissociates from the catalytic domain. CaM binding to the AID-containing CaN regulatory domain (RD) (Ser373 to Thr468) is a critical determinant of this process [4, 10]. Interestingly, like many intrinsically disordered peptide (IDP)-containing complexes, well-defined secondary structure is observed only upon binding a protein target [11–14].

In absence of hydrophobic residues [15] that would otherwise promote collapse of protein into a molten apolar core, many IDPs such as the CaN RD are polyampholytic [16]. Metrics like the net charge per residue (NCPR) have been proposed to relate charge density in IDPs to ensemble properties including compactness and shape [17, 18]. Formally, NCPR $= |f_+ - f_-|$ where f_+ and f_- are fractions of positively and negatively charged residues, respectively, and fraction of charged residues (FCR) is calculated as $\frac{N_c}{N_t}$ where N_c is the number of charged residues and N_t is the total number of residues. Generally, IDPs with large NCPR values (> 0.25) tend to adopt more extended conformations due to repulsive intra-molecular electrostatic interactions [17, 19]. However, the distribution of charged residues in the CaN RD is
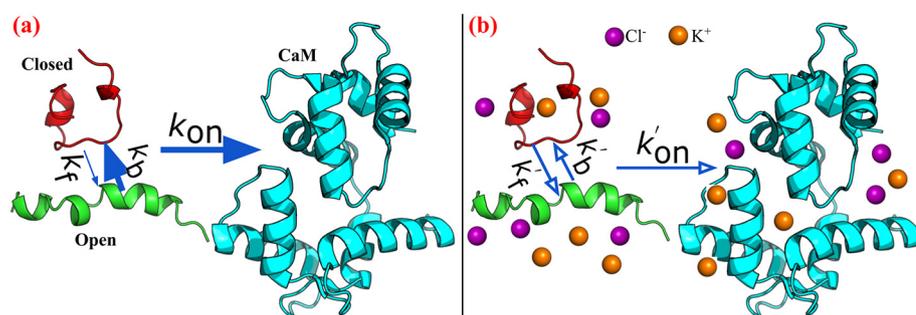
Fig. 1. Schematic of CaN peptide binding to CaM. The green and red peptides represent CaM conformations which are capable and incapable of binding to CaM (colored in cyan). The association rates between intrinsically-disordered CaN peptides and CaM are controlled by open/closed state gating kinetics (depicted by k_b and k_f) and the CaM/CaN diffusional encounter rate (depicted by k_{on}). Our study demonstrates that low (a) and high (b) ionic strengths influence both contributions to afford diffusion limited encounters

heterogeneous, thus such sequence-dependent metrics have limited utility in determining localized properties, such as the availability of binding motifs to target proteins.

We hypothesized therefore that RD sequence charge composition (as measured by NCPR) and ionic strength influence the dynamic availability of conformations amenable to CaM binding, while long-range electrostatic interactions drive diffusion-limited association (see Fig. 1). To investigate this hypothesis, we utilized long-timescale MD simulations to probe the highly dynamic conformational ensembles comprising the RD constructs, toward delineating the extent to which conformational gating kinetics and long-range electrostatic interactions govern IDP/protein association. A chief outcome of this work is our demonstration that charge-sensitive 'local' and long range factors, namely CN RD conformational dynamics and RD/CaM electrostatic interactions, can jointly facilitate diffusion-limited target association.

2. Methods

Three CaN RD constructs were predicted by Rosetta [20] (see Sect. 2.1) and were subject to microsecond scaled molecular dynamics (MD) via Amber14 [21] (see Sect. 2.2). 2D replica exchange umbrella sampling (REUS) potential of mean force (PMF) calculations of the CaN RD constructs were performed by NAMD2.11 [22] (see Sect. 2.3). The MD trajectories were used to characterize the conformational dynamics of the CaN RD constructs through Markov state modeling (MSM) via Aqualab [23] (see Sect. 2.4). The diffusional encounters of CaN RD constructs with CaM were simulated using BrownDye [24] (see Sect. 2.5), from which effective association rates that account for conformational dynamics were estimated as described in Sect. 2.6.

2.1. Structure preparation

The N-domain (residue numbers 3 to 75) and C-domain (residue numbers 76 to 147) of CaN were extracted from the crystal structure (PDB ID: 4Q5U (25)). For CaN peptides, three different peptides with varying lengths and charge distributions were considered: 1) pCaN: native binding region for CaM. 2) lpCaN: elongated pCaN with five additional residues added to two ends of pCaN, respectively. 3) lpcCaN: charge mutated lpCaN having EESE to KKSK mutations at the C-terminal end. Since diffusion limited binding between CaM and CaN are suggested for both intact CaN and its regulatory domain (arXiv:1611.04080v1), we postulated that these comparatively shorter constructs could capture key factors governing diffusion limited association in a computationally-tractable manner. Rosetta [20] was used to model initial conformations for the CaN peptides. The *ab initio* structural prediction was conducted by running the "AbinitioRe- lax.li- nuxgccrelease" installed on our local computing resources. The parameters used in present study are similar to that listed in [26]. The example flags set (parameters) of pCaN structural prediction are provided in the supplement (see Sect. S3.1). The fragment libraries (e.g., frag3 and frag9) were generated via the online server (http://robetta. bakerlab.org/fragmentsubmit.jsp). The number of output conformations was set to ten. According to the energy score, for each CaN

peptide, the conformation with the lowest energy was selected for further MD sampling as described below.

2.2. Molecular dynamic simulation

We performed MD simulations to extensively explore the conformational space of the CaN peptides. The Amber ff99SB-ILDN [27] force field was chosen, given its improved recapitulation of experimentally-observed IDP ensembles, in contrast to common forcefields that tend to predict overly collapsed states for IDPs [28]. MD was performed with Amber14 [21]. The implicit solvent model (igb = 2 with salt concentration = 0.15 M) was used. The rationale for choosing implicit solvent model was based on a recent study [29] that indicated a combination of ff99SB-ILDN with implicit solvent model achieves reasonably accurate sampling for IDPs. The cutoff value for non-bond interactions was set as 999 Å. The starting structure was first subjected to 50,000 steps of energy minimization. The minimized structure was slowly heated from 1 to 298.15 K by using the Berendsen Thermostat within 800 ps. During the MD process, the time interval was set to 2 fs and the SHAKE [30] constraints were applied on bonds involving hydrogen atoms. The initial temperature in the heating stage was set equal to 1 instead of 0 and ig was set to −1, which randomized initial velocity distributions as a means of ensuring the simulations are independent. For each peptide, three MD simulations were performed (total 15 μs production run for each peptide). To study the effect of ionic strength on sampling, we ran analogous simulations with salt concentration = 1.5 M, resulting in a total of 30 μs production run for each peptide (15 μs at 0.15 M and 15 μs at 1.5 M ionic strength). Although 1.5 M ionic strength is non-physiological it provides reasonable screening of electrostatic interactions compared to 0.15 M. We analyzed the root mean square fluctuations (RMSF) for each peptide, based on aligning all trajectory frames to the first frame Contacts between heavy atoms were analyzed via CPPTRAJ in Amber with a distance cutoff of 7 Å for residue pairs which are at least 5 residues apart (i and $i + 5$) in sequence are considered.

2.3. Two-dimensional replica-exchange umbrella sampling (REUS) potential of mean force calculations

Two-dimensional potential of mean force (PMF) calculations were performed to characterize the free energy surface associated with each peptide ensemble. Two reaction coordinates (RCs) were defined: 1) α which describes the α-helical content of the peptide (ranging from 0.1 to 0.9) and 2) radius of gyration (R_g) of the peptide (ranging from 5 to 32 Å). Each RC range was divided into nine bins resulting in total 81 windows (with interval being 0.1 and 3 Å for α and R_g, respectively). The two force constants of the harmonic potentials imposed on these two RCs were 1.000×10^3 kcal mol^{-1} U^{-2} for α and 2.5 kcal mol^{-1} Å^{-2} for R_g. For each peptide, the representative structure from the most populated cluster was chosen as the starting structure. NAMD2.11 [22] was chosen to perform the 2D REUS calculations due to it's *colvar* module which supports various user-defined collective variables. The CHARMM36 [31, 32] force field was used in the 2D

REUS calculations. For each window, the simulation length was set to 20 ns from which the last 15 ns were used to calculate free energy by WHAM [33].

2.4. Markov state model (MSM) analysis via Aqualab

For each peptide, a 1D kinetic trajectory was created from the 15 μs MD trajectory describing the state change along simulation time. Open states were defined based on examining Browndye-predicted association rates as a function of root mean squared deviations (RMSD); the RMSD value below which the association rate rapidly increased above negligible values was utilized as the open/closed state cutoff criterion. We note the 'closed states' were not necessarily precluded from binding, but we assumed that the timescale for induced fit was slow relative to conformational changes in dependent of the target. These criteria differed among the three peptides, based on our Figs. S10 and S9. For pCaN, the open state was defined as RMSD < 7 Å while for lpCaN and lpcCaN the open state are defined as RMSD < 5 Å. Using these cutoffs, Markovian networks were created based on the 1D trajectory (rates of open to closed and vice versa) via Aqualab [23], for which conditions such as detailed balance [34] were imposed to define P, the equilibrium probability matrix and T, the transition probability matrix.

2.5. Brownian dynamic (BD) simulations

The binding of CaN peptide and N/C-terminal domains of CaM were treated as two independent events and were thus simulated separately by using the BrownDye package [24]. For each peptide, ten conformations for each RMSD cluster were randomly selected to perform BD simulations with the N- and C-terminal domains of CaM. PDB2PQR [35] was used to generate the pqr files for CaM N/C domains and the selected conformations of CaN peptides from MD trajectory with radii and point charge parameters adapted from the AMBER99 [36] force field. The generated pqr files were then passed into APBS [37] to evaluate the electrostatic potential of these structures. APBS was used to numerically solve the linearized Poisson-Boltzmann equation assuming an ionic strength of 0.15 M and 1.5 M NaCl:

$$- \epsilon \nabla^2 \psi = \sum \rho_i q_i - \kappa^2 \psi \qquad [1]$$

where ψ is the electrostatic potential, $\rho_i q_i$ is the charge distribution of fixed charge i, and κ is the inverse of Debye length. The Debye length reflects the scale over which mobile charges could screen out electric potential fields. All other APBS parameters were left at their default values.

For the BD simulation, the reaction criterion was chosen to be six pairs of contacts with distances 10 Å. This reaction criterion is more restrictive than the four contact (< 4.5 Å) criterion used in arXiv:1611.04080v1, which we found necessary to resolve binding kinetics for diverse peptide conformations. The contact list was created via the "make_rxn_pairs" routine in Browndye package based on the pCaN-CaM complex crystal structure (PDB ID: 4Q5U (25)) with distance cutoff being 5 Å. 10,000 single trajectory simulations for each system were conducted on 10 parallel processors using nam-simulation. Thus for each peptide, the total number of BD trajectories was about 1 million. The reaction rate constants were calculated using "compute-rate-constant" from the BrownDye package.

To estimate the association rate and its sensitivity to ionic strength, we computed association rates for the terminal domains separately, assuming that both components bind independently,

$$\frac{1}{k_{on}} = \frac{1}{k_{Cterm}} + \frac{1}{k_{Nterm}} \qquad (2)$$

where the rates in the right hand side correspond to the association rates for the C and N terminal domains, respectively. We anticipate that this expression may under estimate the rate of complex formation, given that tethered binding partners generally exhibit higher

efficiencies for forming intact complexes [38, 39].

2.6. Effective association rate combined with gating kinetics

The effective association rate constant after taking conformational dynamics into account was given by Szabo et al. [40]:

$$k_{\text{eff}} = \frac{K_D K_{eq} k_b Z [k_f + k_b]}{k_f (K_{eq} + K_D Z [k_f + k_b]) + k_b Z [k_f + k_b](K_D + K_{eq})} \qquad (3)$$

where

$$Z [k_f + k_b] = 1 + ((k_f + k_b)R^2/D)^{0.5} \qquad (4)$$

where K_D is the association rate when the peptide remains open and in the present study, K_D is the BD simulated association rate constant of the open state CaN peptides with CaM (e.g., $K_D = k_{on,\ open}$). K_{eq} is a characteristic constant indicating the extent to which the association is diffusion-controlled (see [40] for more details). In this study, we set $K_{eq} = 1 \times 10^{20}\,\mathrm{M^{-1}\,s^{-1}}$ and justified by showing in Fig. S8 the sensitivity of k_{eff} to K_{eq}, for which k_{eff} values for all peptides were effectively constant for K_{eq} is larger than $1 \times 10^6\,\mathrm{M^{-1}\,s^{-1}}$ (see also Sect. S.1). k_f and k_b are the conversion rates between the open and closed state determined from the MSM analysis. R is the distance at which interactions between peptide and target are centrosymmetric. Following [41], typical R and diffusion coefficient, D, values for moderately-sized proteins are 40 Å and 20 Å^2/ns. D may also be obtained from BD by setting R equal to the average b-radius values from BD simulations and evaluating [24, 41].

$$D = \frac{K_D}{4\pi R f_\infty} \qquad (5)$$

where f_∞ is the reaction probability, which estimated as approximately 1×10^{-4} by the BD simulations. All code written in support of this publication are publicly available at https://bitbucket.org/pkhlab/pkh-lab-analyses (2018_CaM-CaN). Simulation input files and generate data are available upon request.

3. Results and discussion

3.1. Molecular simulations confirm the intrinsically-disordered structure of the CaN regulatory domain

Several studies using circular dichroism spectroscopy, hydrogen-deuterium exchange mass spectrometry, Fourier transform infrared spectroscopy and X-ray crystallography indicate that the nearly one-hundred amino acids of the CaN RD domain (Ser373 to Thr468 [3, 4]) form an intrinsically disordered ensemble [42–44, 4]. Of these, approximately twenty amino acids (Ala391-Arg414) comprising the CAMBR adopt an alpha-helix in the presence of CaM [25]. Here we examine three RD constructs (pCaN, lpCaN and lpcCaN, see Fig. 2) that present diffusion-limited association with CaM (arXiv:1611.04080v1). pCaN spans residues A391 to R414 and has been co-crystallized with CaM (PDB ID: 4Q5U [25]). lpCaN includes three additional polar residues (Ser3, Ser32 and Ser34) and four additional charged residues (Asp1, Glu30, Glu31 and Glu33) while lpcCaN is a peptide of the same length with substitutions of three glutamic acids with lysines at the C-termini of lpCaN. The set of constructs considered here span a range of charge densities that we later demonstrate tune CaM/CaN association kinetics. pCaN and lpcCaN have similar NCPR values of 0.291 and 0.264, which are considerably larger than the value for lpCaN (0.088); previous works [17, 19], suggest NCPR scores above 0.25 reflect extended IDP conformations given the propensity for repulsive intramolecular interactions, whereas those below this threshold are comparatively compact. We expected therefore that 1) the CaN peptides lack well-resolved secondary structure characteristic of a folded protein and 2) the ensemble of lpCaN should be modestly more compact than

Fig. 2. Amino acid sequences of three CaN peptide constructs examined here, including their respective FCR and NCPR scores. pCaN: native CAMBR of CaN (Ala391- Arg414). lpCaN: five predominantly negatively charged residues affixed to the pCaN termini. lpcCaN: lpCaN construct with three positively-charged substitutions at the C-terminus. In present study, the residue numbers of lpCaN and lpcCaN are counted from 1 to 34 while pCaN has the residue numbers from 6 to 29.

that of lpcCaN, given that latter has higher charge density.

To investigate the hypothesis, we performed 5 μs MD simulations in triplicate (total 15 μs) at 0.15 M and 1.5 M ionic strength, respectively. The choice of physiological (0.15 M) and high ionic strength was intended to probe the contribution of intra-peptide electrostatic interactions to ensemble properties, as such interactions would be screened at

high ionic strength. While simulations of IDPs of up to 100 residues have been reported elsewhere [19], the breadth of simulations used in this study restricted our construct sizes to 24 to 34 a.a. Our MD simulations indicate that the heavy atom root mean squared fluctuations (RMSF) for each residue in Fig. S1(e-f) are shown to be larger than 5 Å for all three peptides at both ionic strengths, which is consistent with the high mobility loop scores reported in Fig. S1(b). These data suggest that the peptides do not form stable folded structures in solution regardless of ionic strength.

The MD-generated structures present a multitude of conformations, ranging from loosely-formed, hairpin-like configurations to extended structures. lpCaN presents perhaps the most folded, 'hairpin-like' character, as corroborated by intramolecular contacts reported through contact map analyses in Fig. S2. Among these contacts are prominent interactions between Arg12-Glu30, Arg23-Glu30 and Arg13-Glu32, which we attribute to transient salt-bridge formation. For lpcCaN, the mutation of negative residues (Glu30, Glu31 and Glu33) to the positively charged residues (Lys30, Lys31 and Lys33) appears to disrupt these intramolecular contacts, thereby yielding a more extended conformational ensemble relative to lpCaN. Given the similar NCPR values of pCaN and lpcCaN, we expected pCaN would similarly present fewer intramolecular contacts than lpCaN. Surprisingly, pCaN exhibited contacts similar to lpCaN, that is, both peptides have comparable intra-contacts. Later we will demonstrate that it is in fact the interconversion
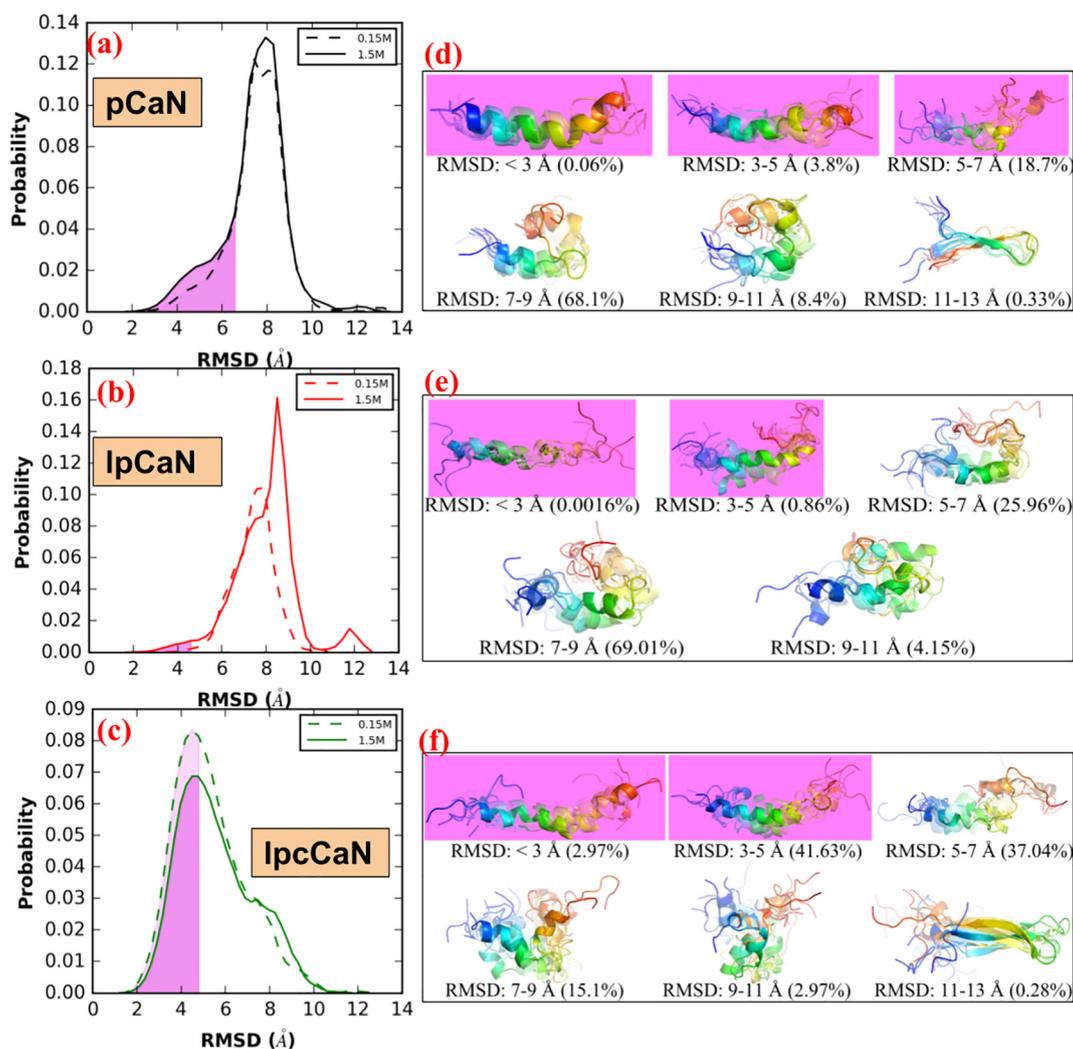

Fig. 3. Distribution of RMSD (with respect to bound-pCaN crystal structure in PDB 4Q5U) in the MD of each CaN peptide at 0.15 M ionic strength and 1.5 M ionic strength, respectively(a-c). The shaded area colored in violet denotes the open state-like conformation. The representative structures (colored in rainbow with N-termini as blue and C-termini as red) for each RMSD range and percentage of conformations within this RMSD range were also shown(d-f).
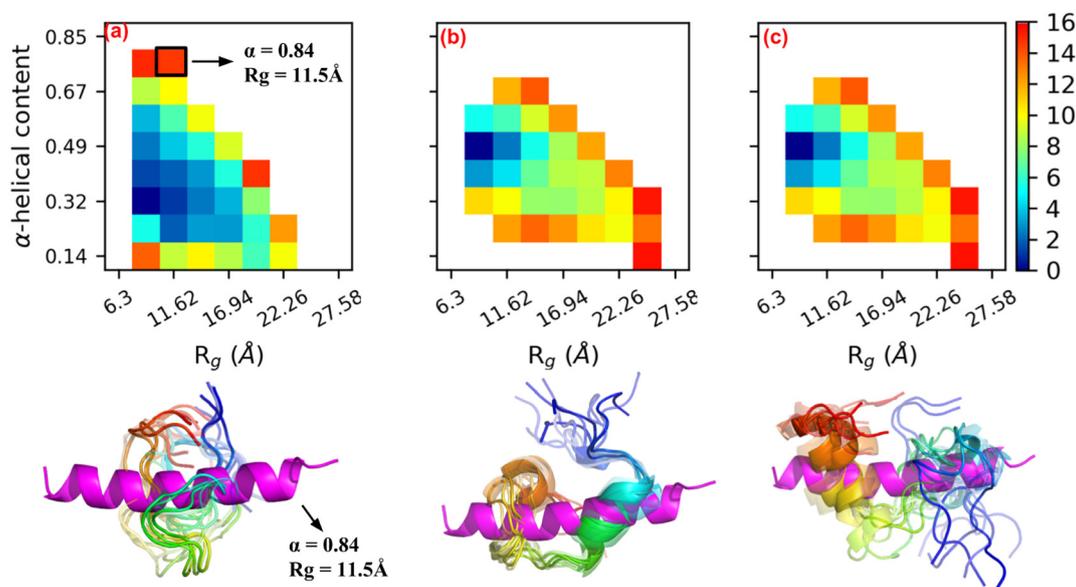
Fig. 4. Two dimensional PMFs for pCaN (a), lpCaN (b) and lpcCaN (c) at 0.15 M ionic strength. The x and y axis depict α-helical and Rg reaction coordinates, respectively. The y axis of (b) and (c) are hidden for clarity. For each peptide, ten randomly selected structures (colored in rainbow with N-termini as blue and C-termini as red) from lowest energy area are compared against bound state pCaN conformation (colored in magenta) from PDB 4Q5U (α = 0.844, Rg = 11.54 Å). The unit of color bar is kbT where kb is Boltzmann constant and $T = 298$ K is temperature.
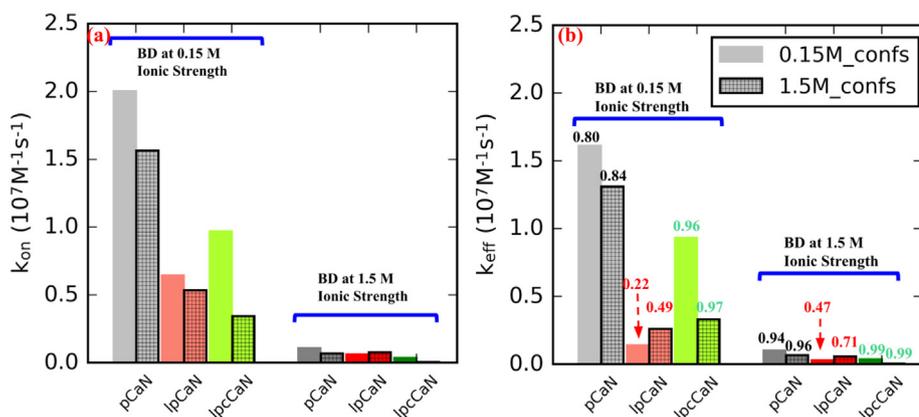

Fig. 5. Association rate constants between CaN peptide and CaM before (a) and after (b) taking CaN peptide's conformational dynamic (from MSM modeling) into consideration using Eq. (3). The bars without grids and with grids depict results in which CaN peptide conformations were sampled at 0.15 M and 1.5 M ionic strength, respectively. In (a), k_{on} was calculated via Eq. (2) where k_{Cterm} and k_{Nterm} are the average values of ten randomly selected conformations from each peptide's open state. In (b) the numbers above each bar represent the ratios of k_{eff} to k_{on}.

kinetics, not the average structural properties, of lpCaN and pCaN that dictate binding kinetics. Additionally, we found that increasing ionic strength to 1.5 M screens the electrostatic interactions between residues comprising the reported salt bridges. As a result, we observe for pCaN and lpcCaN that the structures become modestly more extended on average.

It is important to acknowledge that while implicit solvent simulations permit greater degrees of conformation sampling compared to their all-atom counterparts [45], the utilization of an implicit solvent comes with certain limitations. Among these include the propensity to overestimate alpha-helical character and to alter the interaction strengths of salt-bridging amino acid pairs [46]. Similarly, in the event that charged amino acids might chelate counterions in solution, as exemplified by acidic EF-hands in diverse Ca^{2+}-binding proteins [47], implicit solvent may underestimate the strength of such interactions. It would therefore be of interest in future studies to assess the significance of these limitations in IDPs using explicit all-atom simulations of comparable length to implicit solvent trajectories.

Nevertheless, to support the formation of the CaN/CaM protein-protein interaction (PPI), the CaN CAMBR must be revealed to the solvent-exposed CaM surface. The exposure of the CAMBR could occur spontaneously, which would promote binding by presenting mutually compatible conformations independent of the complementary species,

or via an induced-fit mechanism when the binding partner, CaM, is present. In the previous section, we indicated that the peptides have considerable structural variability, therefore here we determine whether this variability confers greater access to the CAMBR *independent* of CaM.

In Fig. 3(a–c), we report the RMSD of the CAMBR binding region for each configuration from the MD simulations, relative to the extended, alpha-helical pCaN conformation that is compatible with the CaM binding surface. From these simulations, we identify conformations that are amenable to CaM binding ("open" state) and those unsuitable for CaM binding ("closed" state), using a cutoff of pCaN: 7 Å, lpCaN and lpcCaN: 5 Å. We utilized a more restrictive criterion for the longer constructs, as the 7 Å cutoff assumed for pCaN yielded structures that were incompatible with CaM. RMSD values below the cutoff more closely resemble the fully-extended reference structure, whereas values above this cutoff are more compact. As shown in Fig. 3(a–c), all three peptides adopt a small percentage of CaM-compatible configurations as measured by RMSD and the percentages appear to be insensitive to ionic strength. These data additionally indicate that lpCaN (NCPR = 0.088) has the smallest percentage of CaM-campatible structures as assessed by RMSD compared with the bound CaN complex, relative to lpcCaN (NCPR = 0.264) and pCaN (NCPR = 0.291).

In Fig. 3(d–f), we present the structures of the most probable

conformations based on RMSD clustering analysis. Significantly, each peptide was observed to partially fold into an α-helix, indicating that bound-like 'residual' structures can spontaneously form in the absence of the binding partner, as has been reported for other IDPs [48–51]. The open state probabilities we determined represented a small, but significant fraction (pCaN: ~20%; lpCaN: ~1%; lpcCaN: ~45%) of the conformations sampled. Importantly, these data indicate extended/CaM-compatible conformations form spontaneously and in a charge density (NCPR)-dependent manner (see Fig. 3(a-c)). We speculate that the tendency for a percentage of the conformational ensemble to assume an extended pose relative to a hairpin fold suggests that intramolecular repulsion may partially destabilize the formation of loose hairpins. This effect would be exacerbated with charge densities of increasing magnitude, such as those reflected in the NCPR values for pCaN and lpcCaN, and relatively diminished for low NCPR peptides like lpCaN.

To establish a thermodynamic basis for the trends of greater conformational diversity for the high NCPR cases (pCaN and lpcCaN) relative to the low NCPR case (lpCaN), we report potential of mean force (PMF) calculations for these peptides as a function of α-helical character, a measure of secondary structure formation, and R_g, a measure of compactness (see Fig. 4). Such PMFs have been used to characterize the propensity for IDPs to assume specific ensemble characteristics, including IDP compactness [52, 53]. Each construct preferentially adopted smaller α-helical character than the 84% reflected in pCaN when bound to CaM. *Lacking CaM, unfolded CaN RD states dominate the conformational distribution, thus CaM is apparently necessary to induce the formation of native alpha helical character in the RD ensemble.*

Interestingly, we observe that the range of R_g and α-helical values within a few k_bT of the energy minima (0 k_bT) are larger for the high NCPR cases compared to lpCaN. These data mirror our findings for the histogram of RMSD distributions in Fig. 3, with the low NCPR case presenting a narrower distribution relative to the high NCPR cases. Further, the PMF data support the observation for the lpCaN and lpcCaN peptides that the former structure assumes a more compact, hairpin-like configuration relative to the latter, as we observed in Fig. 3(d–f). This indicates that the high NCPR cases access a larger range of conformations in their IDP ensembles that overlap with the CaM-bound structures, in contrast to the more narrowly peaked distributions presented for the low NCPR (lpCaN) configuration. Our results are consistent with the work done by Mao et al. [17] for protamine IDPs, which demonstrated that globule-to-coil transitions were more favored with increasing of NCPR values.

3.2. CaN regulatory domain ensemble conformational dynamics are rapid and have modest ionic-strength dependence

Our unconstrained MD and PMF calculations both indicate that the CaN RD peptides do not assume an "openstate" compatible with CaM on average, but rather there exist infrequent, CaM-compatible configurations. In this regard, one can view the accessibility of the pre-folded CAMBR domain to CaM as a 'gating' event, which in principle could control the apparent binding rate for this process [54, 55]. Given that our previous work (arXiv:1611.04080v1) in which CaN peptides were assumed to have fully CaM-compatible and rigid conformations demonstrated that all three peptides are capable of binding CaM, albeit at substantially different rates, we hypothesized that the appearance of bound-like structures before binding serves to nucleate loosely-associated CaM-compatible transient encounter states with low alpha-helical character. This nucleation could then permit 'induced folding' in the presence of CaM to access alpha-helix rich bound-states.

As a first step toward probing this hypothesis, we first estimated the transition kinetics between open and closed states identified in Sect. 3.1.2 using Markov state model (MSM) analysis. Intuitively, we would expect that higher rates of accessing CaM-compatible open states would maximize the CaM/CaN association rate. We note here that we defined

the open state as consisting of conformations below the 7 Å (for pCaN) and 5 Å (for lpCaN and lpcCaN) RMSD cutoff used in Fig. S10 (the red dash lines depict the RMSD of open and closed state for each peptide), while all conformations with dissimilar RMSDs were lumped into a single closed state. We verify that the states are essentially Markovian as the correlation times were negligible beyond roughly tens of nanoseconds (see Fig. S4) which is faster than the diffusion encounter time. Overall, based on this partition of MD data, the transition rates between closed and open states are rapid (the slowest rate is at $1 \times 10^7 \, s^{-1}$, see Table S1) and lead to the short-lived open states shown in Fig. S5 (average life times of open state for all three peptides under both ionic strengths are around 0.2 ns).

Ionic strength was shown to have negligible impact on the RMSD of our predicted peptide structures relative to CaM-bound conformation. However, given the pronounced role of electrostatics in facilitating protein/protein association rates and protein folding [56, 57], we sought to determine whether transition kinetics between conformations were influenced by ionic strength. Hence, we compared MSM rate predictions for MD generated structures at low (0.15 M) and high (1.5 M) ionic strength. Here we found that for pCaN and lpcCaN, increasing ionic strength does not affect the gating rates between open and closed states. However, for lpCaN, increasing ionic strength increased k_f from $1 \times 10^7 \, s^{-1}$ to $1 \times 10^8 \, s^{-1}$. As a result the lifetime of its closed states decreased from 12.83 to 4.42 ns, as shown in Fig. S5. Hence, the open and closing kinetics of peptides with high NCPR appear to be less sensitive to ionic strength, compared to structures with low NCPR. These results concur with findings from Liu et al. [58], for which they demonstrated that the fast-phase structural fluctuations as measured by Fluorescence correlation spectroscopy (FCR) for the IDP Sic1 disappeared with decreasing ionic strength. At first glance, it is surprising that the ionic strength did not appreciably alter the average properties of the conformation ensemble. However, we anticipate that the change in ionic strength, while significant, was insufficient to strongly disfavor the desolvation of the charged groups in order to drive hydrophobic collapse.

3.3. Long-range interactions promote rapid CaM/CaN association

Our results thus far indicate that the CaN RD peptides adopt CaM-compatible conformations in the absence of CaM rapidly, albeit transiently. Here we determine the compatibility of these transient states with the CaM/CaN binding interface using Brownian dynamic (BD) simulations. Specifically, we sought to evaluate two hypotheses: 1) that frequent presentation of CaN open states promote near diffusion-limited association rates and 2) that long-range electrostatic interactions are exploited in PPIs involving IDPs. Motivating our first hypothesis are recent indications that target-compatible residual structures of the isolated p53 up-regulated modulator of apoptosis (PUMA) IDP form spontaneously as a function of ionic strength and electrolyte composition [59]. For the latter hypothesis, we adopt the paradigm of electrostatically-driven association of globular proteins [60, 61, 62, 41], which depends critically on the notion of a transient encounter complex [63, 64]. The encounter complex serves as the rate determining step in PPI formation, whereby a protein loosely binds to its protein target, before adopting the fully-formed bound configuration. However, unlike PPIs involving globular partners that typically feature regions of complementarily-charged hydrophilic patches [65, 66], such patches may only be transiently presented in IDPs. Though these results offer specific insights into the ubiquitous Ca^{2+} signaling partners CaM and CaN, these trends may vary in importance depending on the IDP sequence, particularly those with vastly different amino acid charge densities and distributions.

We tested these hypotheses by assuming each peptide must achieve a minimal number of 'native contacts' with the CaM N-terminal and C-terminal domains to constitute a transient encounter complex. The native contacts were obtained by analyzing the crystal structure of

CaM-pCaN complex, in which key interactions between CaM and pCaN were extracted to guide the BD simulations. From this standpoint, the lenient conditions for association are tantamount to the notion of a transient encounter complex [67, 68], which is formed upon association of two binding partners prior to forming the fully-bound complex. Because we test the first hypothesis using conformations generated from the MD simulations *without* CaM, this test bears similarity to the conformational selection paradigm [69], though we emphasize CaM is likely required to completely form the bound complex from the transient encounter state. However, the abundance of IDP-target protein complexes that exhibit diffusion-limited association suggests that the folding rate from the encounter state to the bound complex is rapid, as elaborated in Sect. S.1. In Fig. S10 and Fig. S9 we show that the MD-generated open states presented in each of the peptide configurations are compatible with both the N- and C-terminal CaM domains to varying degrees, as the open state of each peptide gives BD-simulated k_{on}s in the diffusion-limited regime ($> 1 \times 10^7\,\mathrm{M}^{-1}\,\mathrm{s}^{-1}$). Notably, all peptides considered here are capable of forming the transient encounter complex with CaM at rates that decrease with increasing ionic strength (see also Sect. S.4.1). While not explored in this study, binding affinities via methods including Molecular Mechanics-Generalized Born Surface area (MM-GBSA) techniques [70, 71] could help rank configurations most compatible with the bound complex.

3.4. Rapid IDP conformational ensemble dynamics promote rapid association

Lastly, we investigated the role of conformational gating rates on the effective association rates based on the stochastic gating model postulated by Szabo et al. [40]. In this gating model, k_{on} is the expected 'ideal' association rate between CaN peptide and CaM, assuming that the CaN peptide remains fixed in the 'CaM-compatible open state. Second, the characteristic diffusion time, defined as $t_D = R^2/D$ by Szabo et al. [40], is compared to k_b and k_f which reflect the conversion time scales between open/closed states of CaN peptide as determined from the MSM analysis. In the gating model, there are two limits that bound the effective rates: 1) given gating rates that are significantly faster than the characteristic diffusion time ($(k_f + k_b)^{-1} \ll t_D$), the effective association rate k_{eff} is equivalent to the rate associated with the open state, that is, $k_{eff} = k_{on}$. 2) given gating rates significantly slower than the diffusional encounter rate ($(k_f + k_b)^{-1} \gg t_D$), k_{eff} is weighted by the equilibrium probability that the open state is available, that is $k_{eff} = k_b(k_f + k_b)^{-1}k_{on}$.

Rates associated with intermediate regimes are obtained by evaluating Eq. (3) using the MSM-estimated gating rates. Based on the data in Table S2 we show in Fig. 5 for pCaN and lpcCaN that k_{eff} and k_{on} are comparable (e.g. $k_{eff}/k_{on} \to 1$), indicating a marginal effect of conformational gating on the association rate. This arises because the conformational transition rates are of the order $1 \times 10^9\,\mathrm{s}^{-1}$, roughly 100 times faster than the characteristic diffusion time ($t_D \sim 80\,\mathrm{ns}$, assuming R = 40 A and D = 20 A^2/ns for typical proteins) (see Table S2). Moreover, the rates are strongly attenuated at $1.5\,\mathrm{M}$ relative to low ionic strength conditions of $0.15\,\mathrm{M}$, which suggest the strong role of long-range electrostatic interactions in promoting association. These data indicate that diffusion-limited association kinetics are realized in the CaN IDP constructs, though the effective rate depends both on ensemble gating kinetics and long-range electrostatic interactions.

An important factor to consider in this model is that the inability of the closed state to bind as rapid as the open state does not render the former 'unsuitable' for binding. Rather, we anticipate its binding rate will be much slower relative to the open states, as the latter of which requires lesser structural reorganization to assume the correct binding pose. We support this speculation based on Gō model predictions summarized in Sect. S4, for which we demonstrated a greater degree of frustration of transition from closed-like conformations to open-like. Namely, for closed state pCaN conformations, the BD-simulated

complex presents an average fraction of native contact as $Q_n = 0.24$ and $Q_c = 0.22$, for the percent native contacts of the N and C terminal CaM domains, respectively while the open state pCaN conformations yield, $Q_n = 0.56$ and $Q_c = 0.59$ (see Figs. S12 and S11). Since lower native contact values will likely encounter greater frustration in folding landscapes compared to near native states, and reduced frustration results in increased rates of folding kinetics [72], we anticipate that induced-fit binding occurs more slowly than the binding of open-state conformations.

We acknowledge that experimental validation would have nicely complemented the model predictions. However, given the availability of experimental kinetics data from stopped-flow measurements of pCaN/CaN binding to CaM (arXiv:1611.04080v1), a computational model consistent with these data offers experimentally-verifiable insights into molecular mechanisms governing their binding. Of these, the couplings between peptide charge distribution and ionic strength on the dynamics of an IDP conformational ensemble is one that could be examined through spectroscopic techniques including time-resolved Forster resonance energy transfer [25] or spin-relaxation times from nuclear magnetic resondance spectroscopy [73]. As with many techniques used to resolve structural characteristics of IDP conformational ensembles, the measurements present ensemble-averaged information, therefore atomistic-scale, computational models will continue to be a strong complement to such probes for characterization of IDP behavior.

4. Conclusions

Our studies of CaN conformational dynamics and CaM/CaN association reveal several interesting features. While the role of charge distribution in IDPs has been shown to be a strong predictor of ensemble structure including compactness [17, 19], our simulations reveal that measures such as NCPR may offer predictive estimates for the ionic strength sensitivity of conformation transition kinetics. Namely, higher NCPR structures are more likely to adopt conformations that complement their binding target, and are less sensitive to changes in ionic strength that may influence gating kinetics. However, it is important to note that this trend may not generalize to necessarily all IDPs, given the wide range of protein/protein association rates ($< 1 \times 10^3$ to $> 1 \times 10^9\,\mathrm{M}^{-1}\,\mathrm{s}^{-1}$ [41]) reported in the literature, which hints at the possibility of different assembly mechanisms. Second, we demonstrate that long-range electrostatic interactions can play a paramount role in determining the kinetics of forming PPIs involving intrinsically-disordered partners, while protein-solvent and protein-protein electrostatic interactions govern the kinetics of presenting target-compatible binding motifs. Together, these factors suggest that IDPs can achieve diffusion-limited association by controlling conformational gating, so long as a conformation amenable to association is rapidly sampled. Overall our findings build upon the growing understanding of the roles of conformation selection and induced fit in dictating PPIs, both identifying how conformational selection can accelerate association, despite potential requirements for induced fitting in order to adopt the final binding pose.

Our study focused on CaN's binding interaction with CaM, of which the latter regulates a staggering array of eukaryotic signaling cascades through forming PPIs with target protein [5]. What sets CaM apart from other such hubs is the surprisingly diverse variety of targets it regulates, despite presenting a single isoform across all mammalian species [74]. In part, its ability to regulate this diversity is attributed to the conformational heterogeneity of the CaM binding interface [75] it is capable of forming. These findings provide intriguing insight into the interplay between conformational diversity and electrostatically-driven protein-protein association involving CaN, which are likely to extend to wide-ranging processes regulated by intrinsically-disordered proteins. As such, exploiting IDP composition to tune PPI kinetics could offer new tools to probe and modulate important biochemical signal transduction pathways.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbagen.2018.07.027.

References

[1] J. Jain, P.G. McCafffrey, Z. Miner, T.K. Kerppola, J.N. Lambert, G.L. Verdine, T. Curran, A. Rao, The T-cell transcription factor NFATp is a substrate for calcineurin and interacts with Fos and Jun, Nature 365 (1993) 353–355.

[2] F. Rusnak, P. Mertz, Calcineurin: form and function, Physiol. Rev. 80 (4) (2000) 1483–1521.

[3] S.-A. Yang, C.B. Klee, Low affinity Ca^{2+}-binding sites of calcineurin B mediate conformational changes in calcineurin A, Biochemistry 39 (51) (2000) 16147–16154, https://doi.org/10.1021/bi001321q.

[4] J. Rumi-Masante, F.I. Rusinga, T.E. Lester, T.B. Dunlap, T.D. Williams, A.K. Dunker, D.D. Weis, T.P. Creamer, Structural basis for activation of calcineurin by calmodulin, J. Mol. Biol. 415 (2) (2012) 307–317.

[5] P. Kursula, The many structural faces of calmodulin: a multitasking molecular jackknife, Amino Acids 46 (10) (2014) 2295–2304.

[6] A.R. Quintana, D. Wang, J.E. Forbes, M.N. Waxham, Kinetics of calmodulin binding to calcineurin, Biochem. Biophys. Res. Commun. 334 (2) (2005) 674–680.

[7] C. B. Klee, H. Ren, X. Wang, Regulation of the calmodulin-stimulated protein phosphatase, calcineurin, J. Biol. Chem. 273 (22) (1998) 13367–13370. arXiv:http://www.jbc.org/content/273/22/13367.full.pdf+html, doi:https://doi.org/10.1074/jbc.273.22.13367. URL http://www.jbc.org/content/273/22/13367.short

[8] R.S. Lewis, Calcium signaling mechanisms in t lymphocytes, Annu. Rev. Immunol. 19 (1) (2001) 497–521, https://doi.org/10.1146/annurev.immunol.19.1.497 11244045.

[9] A. Persechini, B. Cronk, The relationship between the free concentrations of Ca2 + and Ca2 + -calmodulin in intact cells, J. Biol. Chem. 274 (11) (1999) 6827–6830.

[10] T.B. Dunlap, E.C. Cook, J. Rumi-Masante, H.G. Arvin, T.E. Lester, T.P. Creamer, The distal helix in the regulatory domain of calcineurin is important for domain stability and enzyme function, Biochemistry 52 (48) (2013) 8643–8651.

[11] P.E. Wright, H. Dyson, Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm, J. Mol. Biol. 293 (2) (1999) 321–331.

[12] H. Dyson, P.E. Wright, Coupling of folding and binding for unstructured proteins, Curr. Opin. Struct. Biol. 12 (1) (2002) 54–60.

[13] H.J. Dyson, P.E. Wright, Intrinsically Unstructured Proteins and Their Functions, (2005), https://doi.org/10.1038/nrm1589.

[14] K. Sugase, H.J. Dyson, P.E. Wright, Mechanism of coupled folding and binding of an intrinsically disordered protein, Nature 447 (7147) (2007) 1021–1025.

[15] J.A. Marsh, J.D. Forman-Kay, Sequence determinants of compaction in intrinsically disordered proteins, Biophys. J. 98 (10) (2010) 2383–2390, https://doi.org/10.1016/j.bpj.2010.02.006 URL http://www.sciencedirect.com/science/article/pii/S0006349510002341.

[16] A.V. Dobrynin, R.H. Colby, M. Rubinstein, Polyampholytes, J. Polym. Sci. B Polym. Phys. 42 (19) (2004) 3513–3538.

[17] A. H. Mao, S. L. Crick, A. Vitalis, C. L. Chicoine, R. V. Pappu, Net charge per residue modulates conformational ensembles of intrinsically disordered proteins, Proc. Natl. Acad. Sci. 107 (18) (2010) 8183–8188. arXiv:http://www.pnas.org/content/107/18/8183.full.pdf, doi:https://doi.org/10.1073/pnas.0911107107. URL http://www.pnas.org/content/107/18/8183.abstract

[18] S. Müller-Späth, A. Soranno, V. Hirschfeld, H. Hofmann, S. Rüegger, L. Reymond, D. Nettels, B. Schuler, Charge interactions can dominate the dimensions of intrinsically disordered proteins, Proc. Natl. Acad. Sci. 107 (33) (2010) 14609–14614. arXiv:http://www.pnas.org/content/107/33/14609.full.pdf, doi:https://doi.org/10.1073/pnas.1001743107. URL http://www.pnas.org/content/107/33/14609.abstract

[19] K.P. Sherry, R.K. Das, R.V. Pappu, D. Barrick, Control of transcriptional activity by design of charge patterning in the intrinsically disordered ram region of the notch receptor, Proc. Natl. Acad. Sci. 114 (44) (2017) E9243–E9252.

[20] K.T. Simons, R. Bonneau, I. Ruczinski, D. Baker, Ab initio protein structure prediction of casp iii targets using rosetta, Proteins: Struct., Funct., Bioinf. 37 (S3) (1999) 171–176.

[21] D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko,

[22] R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, J. Wang Walker, R.M. Wolf, X. Wu, P.A. Kollman, Amber 14, University of California, San Francisco, 2014.

[22] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with namd, J. Comput. Chem. 26 (16) (2005) 1781–1802, https://doi.org/10.1002/jcc.20289. URL https://doi.org/10.1002/jcc.20289.

[23] G. Berezovska, D. Prada-Gracia, S. Mostarda, F. Rao, Accounting for the kinetics in order parameter analysis: lessons from theoretical models and a disordered peptide, J. Chem. Phys. 137 (19) (2012) 194101, https://doi.org/10.1063/1. 4764868.

[24] G.A. Huber, J.A. McCammon, Browndye: A software package for Brownian dynamics, Comput. Phys. Commun. 181 (11) (2010) 1896–1905, https://doi.org/10.1016/j.cpc.2010.07.022 URL http://www.sciencedirect.com/science/article/pii/S0010465510002559.

[25] T. B. Dunlap, H.-F. Guo, E. C. Cook, E. Holbrook, J. Rumi-Masante, T. E. Lester, C. L. Colbert, C. W. Vander Kooi, T. P. Creamer, Stoichiometry of the calcineurin regulatory domain–calmodulin complex, Biochemistry 53 (36) (2014) 5779–5790, pMID: 25144868. arXiv:http://dx.doi.org/10.1021/bi5004734, doi:https://doi.org/10.1021/bi5004734. URL https://doi.org/10.1021/bi5004734

[26] P. Xiong, M. Wang, X. Zhou, T. Zhang, J. Zhang, Q. Chen, H. Liu, Protein design with a comprehensive statistical energy function and boosted by experimental selection for foldability, Nat. Commun. 5 (2014) 5330.

[27] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror, D.E. Shaw, Improved side-chain torsion potentials for the amber ff99sb protein force field, Proteins: Struct., Funct., Bioinf. 78 (8) (2010) 1950–1958, https://doi.org/10.1002/prot.22711.

[28] J. Henriques, C. Cragnell, M. Skepö, Molecular dynamics simulations of intrinsically disordered proteins: Force field evaluation and comparison with experiment, J. Chem. Theory Comput. 11 (7) (2015) 3420–3431, pMID: 26575776. arXiv:https://doi.org/10.1021/ct501178z, doi:https://doi.org/10.1021/ct501178z. URL https://doi.org/10.1021/ct501178z

[29] I. Maffucci, A. Contini, An updated test of amber force fields and implicit solvent models in predicting the secondary structure of helical, β-hairpin, and intrinsically disordered peptides, J. Chem. Theory Comput. 12 (2) (2016) 714–727.

[30] J.-P. Ryckaert, G. Ciccotti, H.J. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, J. Comput. Phys. 23 (3) (1977) 327–341, https://doi.org/10.1016/0021-9991(77)90098-5 URL http://www.sciencedirect.com/science/article/pii/0021999177900985.

[31] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (18) (1998) 3586–3616.

[32] A.D. MacKerell, M. Feig, C.L. Brooks, Improved treatment of the protein backbone in empirical force fields, J. Am. Chem. Soc. 126 (3) (2004) 698–699.

[33] A. Grossfield, Wham: The Weighted Histogram Analysis Method, URL http://membrane.urmc.rochester.edu/content/wham.

[34] G.R. Bowman, K.A. Beauchamp, G. Boxer, V.S. Pande, Progress and challenges in the automated construction of Markov state models for full protein systems, J. Chem. Phys. 131 (12) (2009) 124101.

[35] T. J. Dolinsky, P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe, N. A. Baker, Pdb2pqr: expanding and upgrading automated preparation of biomolecular structures for molecular simulations, Nucleic Acids Res. 35.

[36] J. Wang, P. Cieplak, P.A. Kollman, How well does a restrained electrostatic potential (resp) model perform in calculating conformational energies of organic and biological molecules? J. Comput. Chem. 21 (12) (2000) 1049–1074.

[37] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, Electrostatics of nanosystems: Application to microtubules and the ribosome, Proc. Natl. Acad. Sci. 98 (18) (2001) 10037–10041. arXiv:http://www.pnas.org/content/98/18/10037.full.pdf, doi:https://doi.org/10.1073/pnas.181342398. URL http://www.pnas.org/content/98/18/10037.abstract

[38] D. Van Valen, M. Haataja, R. Phillips, Biochemistry on a leash: the roles of tether length and geometry in signal integration proteins, Biophys. J. 96 (4) (2009) 1275–1292.

[39] B.A. Shoemaker, J.J. Portman, P.G. Wolynes, Speeding molecular recognition by using the folding funnel: the fly-casting mechanism, Proceedings of the …, Vol. 97 National Acad Sciences, 2000(p. 8868).

[40] A. Szabo, D. Shoup, S. H. Northrup, J. A. McCammon, Stochastically gated diffusion influenced reactions, J. Chem. Phys. 77 (9) (1982) 4484–4493. arXiv:https://doi.org/10.1063/1.444397, doi:https://doi.org/10.1063/1.444397. URL https://doi.org/10.1063/1.444397

[41] G. Schreiber, G. Haran, H.-X. Zhou, Fundamental aspects of protein-protein association kinetics, Chem. Rev. 109 (3) (2009) 839–860, https://doi.org/10.1021/cr800373w 19196002.

[42] C.R. Kissinger, H.E. Parge, D.R. Knighton, C.T. Lewis, L.A. Pelletier, A. Tempczyk, V.J. Kalish, K.D. Tucker, R.E. Showalter, E.W. Moomaw, L.N. Gastinel, N. Habuka, X. Chen, F. Maldonado, J.E. Barker, R. Bacquet, J.E. Villafranca, Crystal structures of human calcineurin and the human FKBP12–FK506–calcineurin complex, Nature 378 (6557) (1995) 641–644.

[43] A. S. Manalan, C. B. Klee, Activation of calcineurin by limited proteolysis, Proc. Natl. Acad. Sci. 80 (14) (1983) 4291–4295. arXiv:http://www.pnas.org/content/80/14/4291.full.pdf. URL http://www.pnas.org/content/80/14/4291

[44] X. Shen, H. Li, Y. Ou, W. Tao, A. Dong, J. Kong, C. Ji, S. Yu, The secondary structure of calcineurin regulatory region and conformational change induced by calcium/calmodulin binding, J. Biol. Chem. 283 (17) (2008) 11407–11413. arXiv:http://www.jbc.org/content/283/17/11407.full.pdf + html, doi:https://doi.org/10.1074/jbc.M708513200. URL http://www.jbc.org/content/283/17/11407.abstract

[45] A. Onufriev, Chapter 7 - implicit solvent models in molecular dynamics simulations: a brief overview, Annual Reports in Computational Chemistry, Vol. 4 Elsevier, 2008, pp. 125–137, , https://doi.org/10.1016/S1574-1400(08)00007-8 URL http://www.sciencedirect.com/science/article/pii/S1574140008000078.

[46] Z. Ruhong, Free energy landscape of protein folding in water: explicit vs. implicit solvent, Proteins: Struct., Funct., Bioinf. 53 (2) 148–161. doi:https://doi.org/10.1002/prot.10483.

[47] Z. Grabarek, Structural basis for diversity of the EF-hand calcium-binding proteins, J. Mol. Biol. 359 (3) (2006) 509–525.

[48] M. Fuxreiter, I. Simon, P. Friedrich, P. Tompa, Preformed structural elements feature in partner recognition by intrinsically unstructured proteins, J. Mol. Biol. 338 (5) (2004) 1015–1026, https://doi.org/10.1016/j.jmb.2004.03.017 http://www.sciencedirect.com/science/article/pii/S0022283604003079.

[49] S.G. Sivakolundu, D. Bashford, R.W. Kriwacki, Disordered p27kip1 exhibits intrinsic structure resembling the cdk2/cyclin a-bound conformation, J. Mol. Biol. 353 (5) (2005) 1118–1128, https://doi.org/10.1016/j.jmb.2005.08.074 URL http://www.sciencedirect.com/science/article/pii/S0022283605010545.

[50] M.-K. Yoon, V. Venkatachalam, A. Huang, B.-S. Choi, C.M. Stultz, J.J. Chou, Residual structure within the disordered c-terminal segment of p21waf1/cip1/sdi1 and its implications for molecular recognition, Protein Sci. 18 (2) (2009) 337–347, https://doi.org/10.1002/pro.34.

[51] M. Kjaergaard, K. Teilum, F.M. Poulsen, Conformational selection in the molten globule state of the nuclear coactivator binding domain of cbp, Proc. Natl. Acad. Sci. 107 (28) (2010) 12535–12540, https://doi.org/10.1073/pnas.1001693107.

[52] D. Song, W. Wang, W. Ye, D. Ji, R. Luo, H.-F. Chen, ff14idps force field improving the conformation sampling of intrinsically disordered proteins, Chem. Biol. Drug Des. 89 (1).

[53] X. Guo, J. Han, R. Luo, H.-F. Chen, Conformation dynamics of the intrinsically disordered protein c-myb with the ff99idps force field, RSC Adv. 7 (2017) 29713–29721, https://doi.org/10.1039/C7RA04133K.

[54] H.-X. Zhou, S.T. Wlodek, J.A. McCammon, Conformation gating as a mechanism for enzyme specificity, Proc. Natl. Acad. Sci. 95 (16) (1998) 9280–9283.

[55] H.-X. Zhou, From induced fit to conformational selection: a continuum of binding mechanism controlled by the timescale of conformational transitions, Biophys. J. 98 (6) (2010) L15–L17.

[56] M. Roca, B. Messer, A. Warshel, Electrostatic contributions to protein stability and folding energy, FEBS Lett. 581 (10) (2007) 2065–2071, https://doi.org/10.1016/j.febslet.2007.04.025 URL http://www.sciencedirect.com/science/article/pii/S0014579307004103.

[57] H.-X. Zhou, X. Pang, Electrostatic interactions in protein structure, folding, binding, and condensation, Chem. Rev. 0 (0) (2018).

[58] B. Liu, D. Chia, V. Csizmok, P. Farber, J.D. Forman-Kay, C.C. Gradinaru, The effect of intrachain electrostatic repulsion on conformational disorder and dynamics of the sic1 protein, J. Phys. Chem. B 118 (15) (2014) 4088–4097.

[59] B.I.M. Wicky, S.L. Shammas, J. Clarke, Affinity of idps to their targets is modulated by ion-specific changes in kinetics and residual structure, Proc. Natl. Acad. Sci. 114 (37) (2017) 9882–9887, https://doi.org/10.1073/pnas.1705105114.

[60] G. Schreiber, A. R. Fersht, Rapid, electrostatically assisted association of proteins, Nat. Struct. Biol. 3 (427).

[61] F.B. Sheinerman, R. Norel, B. Honig, Electrostatic aspects of protein–protein interactions, Curr. Opin. Struct. Biol. 10 (2) (2000) 153–159, https://doi.org/10.1016/S0959-440X(00)00065-8 URL http://www.sciencedirect.com/science/article/pii/S0959440X00000658.

[62] M. Vijayakumar, K.-Y. Wong, G. Schreiber, A.R. Fersht, A. Szabo, H.-X. Zhou, Electrostatic enhancement of diffusion-controlled protein-protein association: comparison of theory and experiment on barnase and barstar11edited by b. honig, J. Mol. Biol. 278 (5) (1998) 1015–1024, https://doi.org/10.1006/jmbi.1998.1747 URL http://www.sciencedirect.com/science/article/pii/S0022283698917478.

[63] R. Gabdoulline, R. Wade, Simulation of the diffusional association of barnase and barstar, Biophys. J. 72 (5) (1997) 1917–1929.

[64] R. Alsallaq, H.-X. Zhou, Electrostatic rate enhancement and transient complex of protein–protein association, Proteins: Struct., Funct., Bioinf. 71 (1) (2008) 320–335, https://doi.org/10.1002/prot.21679.

[65] L.-P. Lee, B. Tidor, Optimization of binding electrostatics: charge complementarity in the barnase-barstar protein complex, Protein Sci. 10 (2) (2001) 362–377, https://doi.org/10.1110/ps.40001.

[66] D.F. Green, B. Tidor, Design of improved protein inhibitors of hiv-1 cell entry: optimization of electrostatic interactions at the binding interface, Proteins: Struct., Funct., Bioinf. 60 (4) (2005) 644–657, https://doi.org/10.1002/prot.20540.

[67] R.R. Gabdoulline, R.C. Wade, On the protein–protein diffusional encounter complex, J. Mol. Recognit. 12 (4) (1999) 226–234, https://doi.org/10.1002/(SICI)1099-1352(199907/08)12:4 < 226:: AID-JMR462 > 3.0.CO;2-P.

[68] C. Tang, J. Iwahara, M.G. Clore, Visualization of transient encounter complexes in protein–protein association, Nature 444 (2006) 383–386.

[69] D.D. Boehr, R. Nussinov, P.E. Wright, The role of dynamic conformational ensembles in biomolecular recognition, Nat. Chem. Biol. 5 (11) (2009) 789–796, https://doi.org/10.1038/nchembio.232.

[70] A. Onufriev, D.A. Case, D. Bashford, Effective born radii in the generalized born approximation: the importance of being perfect, J. Comput. Chem. 23 (2002) 1297–1304.

[71] A. Onufriev, D. Bashford, D.A. Case, Exploring protein native states and large-scale conformational changes with a modified generalized born model, Proteins: Struct., Funct., Bioinf. 55 (2) (2004) 383–394.

[72] F.O. Tzul, K.L. Schweiker, G.I. Makhatadze, Modulation of folding energy landscape by charge–charge interactions: linking experiments with computational modeling, Proc. Natl. Acad. Sci. 112 (3) (2015) 201410424.

[73] P.R.L. Markwick, T. Malliavin, M. Nilges, Structural biology by NMR: structure, dynamics, and interactions, PLoS Comput. Biol. 4 (9) (2008) e1000168.

[74] D. Chin, A.R. Means, Calmodulin: a prototypical calcium sensor, Trends Cell Biol. 10 (8) (2000) 322–328.

[75] D. Shukla, A. Peck, V. S. Pande, Conformational heterogeneity of the calmodulin binding interface, Nat. Commun. 7.

[76] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G.D. Peterson, R. Roskies, J.R. Scott, N. Wilkins-Diehr, Xsede: accelerating scientific discovery, Comput. Sci. Eng. 16 (5) (2014) 62–74.