

Engineering Order and Cooperativity in a Disordered Protein

Sneha Munshi,^{†,⊥} Sandhya Subramanian,^{†,⊥} Samyuktha Ramesh,[†] Hemashree Golla,[†] Divakar Kalivarathan,[‡] Madhurima Kulkarni,[§] Luis A. Campos,^{||} Ashok Sekhar,[§] and Athi N. Naganathan^{*,†,⊥}

[†]Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai 600036, India

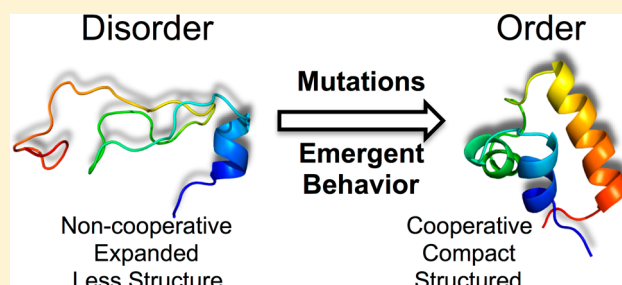
[‡]Department of Biotechnology, National Institute of Technology Warangal, Warangal 506004, India

[§]Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India

^{||}National Biotechnology Center, Consejo Superior de Investigaciones Científicas, Darwin 3, Campus de Cantoblanco, 28049 Madrid, Spain

S Supporting Information

ABSTRACT: Structural disorder in proteins arises from a complex interplay between weak hydrophobicity and unfavorable electrostatic interactions. The extent to which the hydrophobic effect contributes to the unique and compact native state of proteins is, however, confounded by large compensation between multiple entropic and energetic terms. Here we show that protein structural order and cooperativity arise as emergent properties upon hydrophobic substitutions in a disordered system with non-intuitive effects on folding and function. Aided by sequence–structure analysis, equilibrium, and kinetic spectroscopic studies, we engineer two hydrophobic mutations in the disordered DNA-binding domain of CytR that act synergistically, but not in isolation, to promote structure, compactness, and stability. The double mutant, with properties of a fully ordered domain, exhibits weak cooperativity with a complex and rugged conformational landscape. The mutant, however, binds cognate DNA with an affinity only marginally higher than that of the wild type, though nontrivial differences are observed in the binding to noncognate DNA. Our work provides direct experimental evidence of the dominant role of non-additive hydrophobic effects in shaping the molecular evolution of order in disordered proteins and *vice versa*, which could be generalized to even folded proteins with implications for protein design and functional manipulation.



Intrinsically disordered proteins (IDPs), many of which act as hubs in protein–protein interaction networks, play a vital role in cellular signaling, transcription, and regulation in eukaryotes, bacteria, and viruses. Their abundance is evidence of an evolutionary selection for disorder, the advantages of which can range from efficient regulation, high flexibility enabling simultaneous binding to multiple partners, switching functions upon post-translational modifications, and intrinsic sensitivity to solvent conditions to a large capture radius.^{1–5} Many IDPs fold upon binding to their partner proteins and acquire a structure that is determined by both its intrinsic sequence properties and the details of the molecular surface presented by the partner. Work on IDPs has therefore overturned the conventional structure–function paradigm and has opened up a healthy debate about the role of small sequence modifications on the residual structure, dynamics, and populations of partially structured states and their functional role, aspects that have been studied previously for denatured states of folded proteins.^{6–9}

Accordingly, the basic thermodynamic forces that drive disorder in proteins have received considerable interest over

the past decade,^{10–14} starting with the seminal work of Uversky and co-workers.¹⁵ IDP sequences are enriched with charged residues and display a low hydrophobicity index, the combination of which promotes disorder due to a weak driving force for compaction and the large desolvation penalty associated with burying charged residues. In fact, most low-complexity sequences (specifically those rich in polar and/or charged residues) are disordered with the relative composition and patterning of charged residues determining the ensemble dimensions¹⁶ and even function.¹⁷ Sequences rich in glycine residues are disordered primarily due to the excess backbone flexibility, while those rich in uncharged polar residues (Ser, Thr, Asn, and Gln) cannot populate a well-defined structure due to the energetically degenerate nature of intrachain hydrogen-bond interactions. In contrast to low-complexity sequences, the disorder tendency in high-complexity sequences

Received: March 4, 2019

Revised: April 18, 2019

Published: April 19, 2019

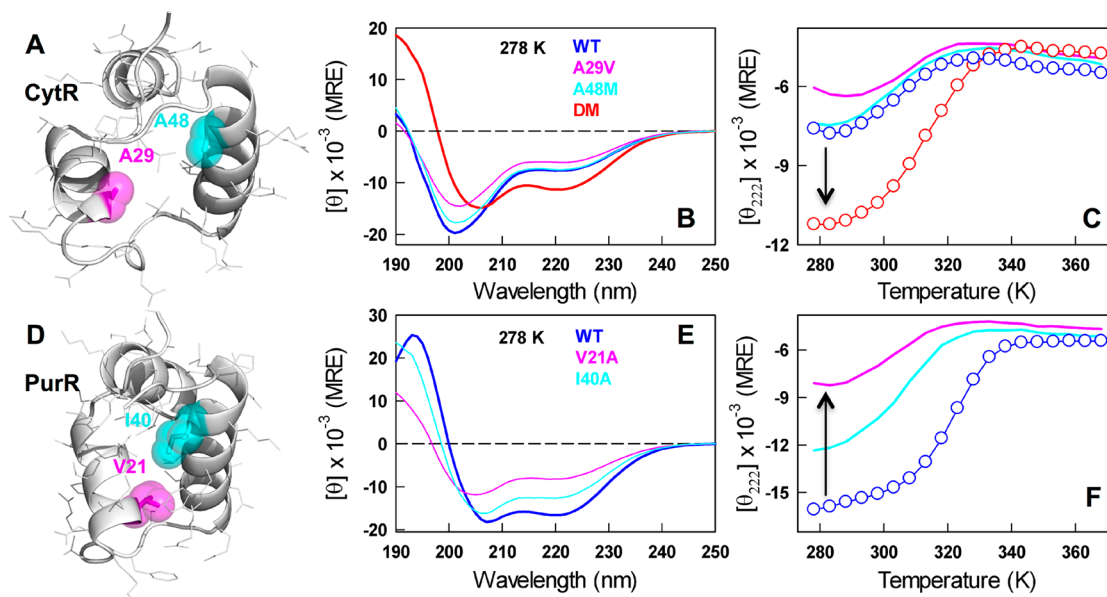


Figure 2. Engineering order in a disordered protein. The top and bottom row represent CytR and PurR, respectively. All experimental data were recorded at pH 7.0 in 43 mM ionic strength phosphate buffer. (A) Structure of the DNA-bound conformation of CytR (PDB entries 2L8N and 2LCV) highlighting the loosely packed hydrophobic core. (B and C) Far-UV CD spectra of WT CytR and single- and double-point mutants and their unfolding curves. The arrow signals the increase in helical structure on mutations. (D) The structure of apo-PurR (PDB entry 1PRU) shows a well-packed hydrophobic core and slightly different helical orientations compared to those of folded CytR. (E and F) Far-UV CD spectra of WT PurR and single-point mutants and their unfolding curves. The arrow signals the decrease in helical structure on mutations (panel F).

performed at protein concentrations of ~ 100 and ~ 70 μM for CytR and PurR, respectively, in a 10 mm path length quartz cuvette.

Quantum Yield Measurements. Fluorescence experiments were performed in a Chirscan-Plus qCD instrument (Applied Photophysics Ltd.). Emission spectra were recorded in the wavelength range of 280–500 nm upon excitation of CytR (~ 26 μM) and PurR (~ 10 μM) at 274 and 295 nm, respectively. NATA, dissolved in water (~ 12 μM), was used as a reference to calculate the quantum yield.

NMR Experiments. Samples of both wild-type (WT) and A29V/A48M CytR (CytR DM) were prepared in 22 mM phosphate buffer (pH 7), and 10% D_2O was added to give a final solution with an ionic strength of 43 mM. NMR samples of WT [$^{15}\text{N}/^{13}\text{C}$]CytR and [^{15}N]CytR DM had concentrations in the range of 600–700 μM . NMR data were acquired in a 14.1 T (600 MHz) Agilent spectrometer equipped with a cryogenically cooled triple-resonance probe. Spectra were processed using NMRPipe,⁴⁹ and overlays were obtained using Sparky.⁵⁰

Differential Scanning Calorimetry. Scanning calorimetry measurements were performed in a MicroCal VP-Capillary differential scanning calorimeter (Malvern Ltd.) coupled to an automated sample injector. The instrument was equilibrated by extensively recording water and buffer scans before each experiment. The apparent heat capacity was measured by recording scans for five different concentrations ranging from ~ 52 to ~ 110 μM at a scan rate of 1 K/min. Three scans were recorded for every concentration to check for reversibility. Buffer scans were recorded before and after every protein scan to check for the reproducibility of buffer–buffer baselines. The data are finally reported in absolute heat capacity units following the method of Sanchez-Ruiz and co-workers.⁵¹

Stopped-Flow Kinetics. Refolding and unfolding kinetic traces as a function of urea and temperature were recorded in a

Chirscan SF.3 stopped-flow instrument (Applied Photophysics Ltd.) coupled to a thermostated water bath. A PMT detector recorded the total fluorescence on excitation by a 280 nm light-emitting diode, while a 295 nm cutoff filter before the detector filtered out the scattered photons. For refolding experiments, unfolded CytR in 4 M urea (mutants in 6 M urea) and unfolded PurR in 8 M urea were rapidly mixed with an excess of 43 mM ionic strength buffer (1:10 mixing), resulting in final protein concentrations of ~ 18 and ~ 10 μM for CytR and PurR, respectively. For unfolding experiments, folded PurR and its mutants in denaturant-free buffer were rapidly mixed with excess urea and the time-dependent unfolding signal was monitored. For temperature-dependent refolding kinetics, traces were recorded from 278 to 318 K at an interval of 2.5 K. At each temperature and urea concentration, six scans were recorded (with 1000 data points each) at an interval of 1 min.

Binding Anisotropy Profiles. Double-stranded udpO (5'-ATTTATGCAACGCA-3', at a concentration of 300 nM) labeled with Alexa-532 at the 5'-end (IBA Lifesciences) was titrated with increasing protein concentrations ranging from ~ 1 nM to 100 μM . Both DNA and protein were dissolved in 50 mM sodium phosphate buffer, 30 mM sodium chloride, and 1 mM EDTA (pH 6.0). The samples were excited at 530 nm, and the emission was collected at 580 nm in a Chirscan-Plus qCD instrument (Applied Photophysics Ltd.). Binding studies were also performed with 300 nM 14-mer poly-AC (5'-ACACACACACACAC-3') and the PurR cognate site (5'-TACGCAAACGTTTGCGT-3').

Association Kinetics. Binding kinetics were monitored by recording time-dependent changes in the anisotropy of Alexa-532-labeled udpO upon 1:1 protein:DNA mixing under pseudo-first-order conditions (i.e., at protein concentrations >10 times that of DNA) at 278 K. Both DNA and protein were dissolved in 50 mM sodium phosphate buffer, 30 mM sodium

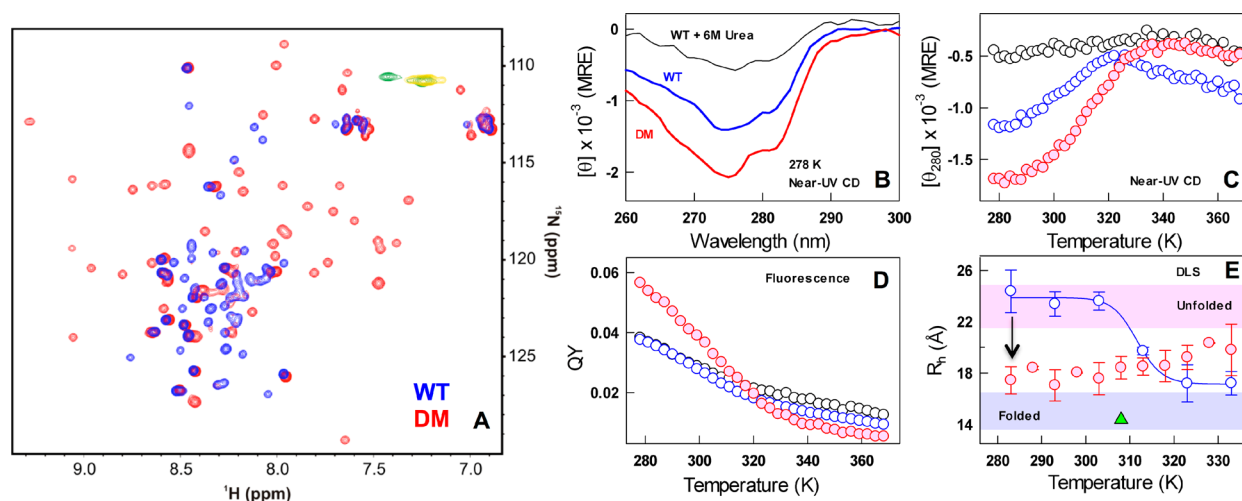


Figure 3. The native ensemble of CytR DM is compact and structured. (A) Overlay of WT CytR (blue) and CytR DM (red) HSQC spectra acquired at 288 K. Yellow and green contours in the overlay are the peaks with negative intensity for WT CytR and CytR DM, respectively. (B) Near-UV CD spectra and (C) unfolding curves. (D) Fluorescence QY as a function of temperature following the color code in panel B. Note the enhanced QY of the DM at low temperatures. (E) Estimates of hydrodynamic radii (ordinate) from DLS experiments for WT CytR (blue) and CytR DM (red). The DM is more compact at the lowest temperatures compared to the WT (arrow). The shaded regions represent the dimensions of unfolded and folded proteins from size-scaling estimates.⁵⁷ The green triangle represents the estimated R_h of CytR from the PDB structure (residues 9–55).

chloride, and 1 mM EDTA (pH 6.0). The final DNA concentration was 300 nM. Six kinetic traces were recorded at an interval of 1 min for every protein concentration (see Figure 5) and averaged.

RESULTS AND DISCUSSION

Origins of Disorder in CytR. Disordered CytR folds to a stable structure in a nearly continuous manner with an increase in the concentration of salt,⁴⁸ signaling strong electrostatic frustration in the native state (Figure 1A,B). Electrostatic potential maps⁵² of known folded models of the LacI DBD family point to similar, unfavorable positive potential on the DNA-binding face (Figure 1A and Figure S1). This is not surprising, as these proteins have been evolutionarily selected to bind DNA, which exhibits a large negative electrostatic potential. This pattern, on the other hand, hints that electrostatics need not necessarily be the primary reason for disorder in CytR as both PurR and LacR are well-folded in the absence of DNA. True to this expectation, the stability of even the ordered domains increases with an increase in the salt concentration (PurR and LacR in Figure 1B). The slopes of the plots of melting temperatures (T_m) versus ionic strength are parallel, with the ordered domains more stable by ~ 18 K under all ionic strength conditions explored. This observation is the first evidence that differences in packing interactions determine the disorder tendency of CytR compared to those of its folded counterparts.

The differences in packing density can be more vividly observed in the plot of coupling distances (d_C) as a function of residue index (Figure 1C). Coupling distances are a measure of the strength of the intraprotein interaction network derived from the first and second shells of interactions starting from a folded structure.^{53,54} A larger d_C reflects stronger long-range energetic coupling and vice versa, thus allowing a precise measure of packing across the entire structure.^{53,54} The folded domain of CytR displays a smaller d_C across nearly all of the residues when compared to those of PurR and LacR (Figure 1C and Figure S2). Sequence alignment indicates similar

hydrophobic tendencies across all of the apolar positions in CytR (stars in Figure 1D), except for sites 29 and 48 (rectangular boxes in Figure 1D). The latter positions are less bulky in CytR (alanine in both positions) compared to valine/methionine/isoleucine in the folded domains.

Order from Disorder via Mutations. The folded structures reveal strikingly stronger packing in PurR in contrast with a loosely packed hydrophobic core in CytR (Figure 2A,D). Taking a cue from sequence–structure analysis, we mutated A29 and A48 to bulkier hydrophobic side chains to enhance hydrophobicity and potentially long-range coupling to promote folding (Figure 2A). Neither of the single-point mutations, A29V or A48M, induced additional structure or increased the stability of CytR. In fact, the A29V mutation reduced the residual helicity at the lowest temperatures as one can see in the far-UV CD spectrum and the thermal unfolding curve (Figure 2B,C). However, introducing both of the mutations into the sequence, as in the double mutant A29V/A48M (DM), significantly enhanced the helical structure and stability, highlighting the non-additive nature of the hydrophobic effect. The far-UV CD spectrum at the lowest temperature shows the characteristic features of a helical protein with minima at ~ 208 and 222 nm and a maximum at ~ 190 nm that is distinct from any of the other variants (red in Figure 2B). The thermal unfolding curve of CytR DM exhibits a folding transition with an apparent T_m of 318 ± 0.4 K (red in Figure 2C).

The observations described above suggest that positions 29 and 48 play a crucial role in promoting structure and long-range order in CytR. If this is indeed the case, the reverse mutations (V to A and I to A) at the same positions in the folded counterparts should promote disorder (Figure 2D). True to this expectation, the single-point mutations V21A and I40A in PurR dramatically decrease the helical content and stability (Figure 2E,F), thus validating the critical role of these two positions in promoting order in the LacI family. The V21A mutant of PurR is in fact very similar to WT CytR in terms of the overall unfolding curve (compare blue in Figure 2C with

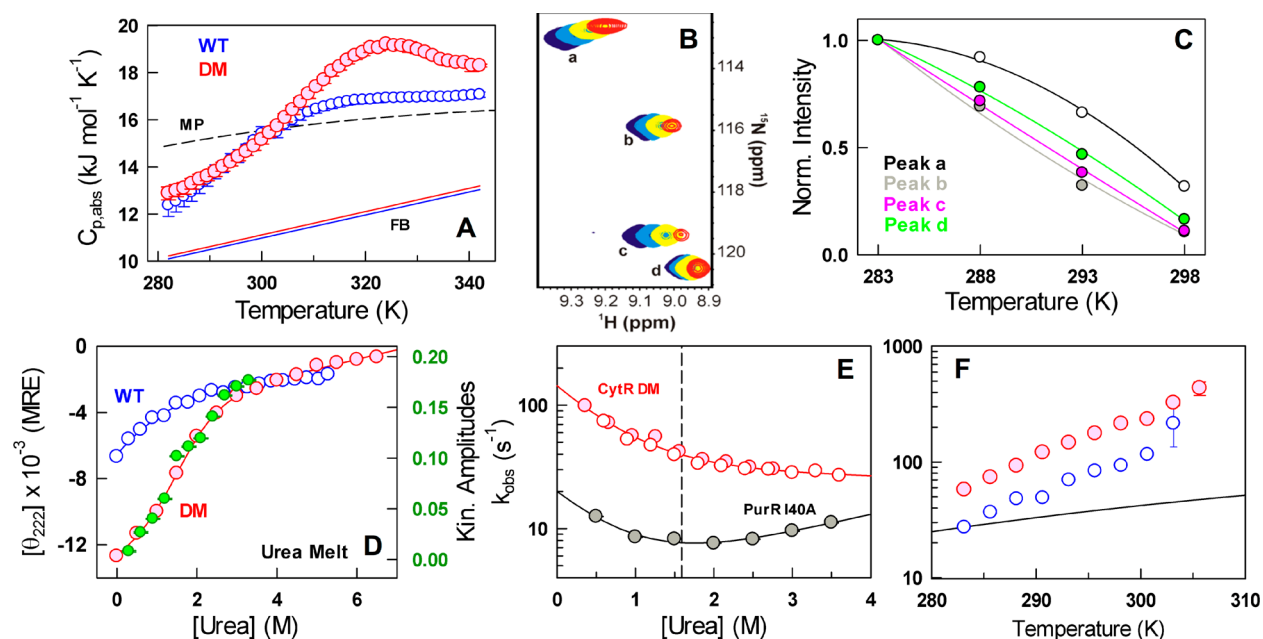


Figure 4. Complex unfolding mechanism of CytR DM. (A) Absolute heat capacity profiles of the WT and the double mutant (DM) together with the Freire folded baseline⁶¹ (FB) and Makhatazde–Privalov unfolded baseline (MP).⁶² Note the excess heat capacity peak for the DM around 325 K. (B) ^1H – ^{15}N HSQC spectra of CytR DM at 283 K (blue), 288 K (cyan), 293 K (yellow), and 298 K (red) for specific resonances labeled from a to d. (C) Normalized intensities of the resonances shown in panel B. (D) Urea unfolding curve of the DM at 285 K together with the amplitudes from kinetic experiments (green and right axis). The urea unfolding curve of the WT at 298 K is shown for reference (blue). (E) Observed relaxation rate constants as a function of urea for CytR DM (red) and PurR I40A (filled gray circles) at 285 K. The empty red and filled red circles represent data from unfolding and refolding experiments, respectively. The vertical dashed line signals the chemical denaturation midpoint of 1.6 M. (F) Temperature-dependent relaxation rate constants for the WT (blue) and the DM (red) together with the expectation from changes in solvent viscosity alone (black).

magenta in Figure 2F), consistent with the larger coupling distance of this position with the rest of the structure (Figure 1C). It is important to note that the engineered mutations to valine and methionine in CytR should reduce the helical propensity while the reverse mutations in PurR should enhance the helical propensity (as alanine has the highest helical propensity^{55,56}). The fact that we observe these mutations to promote order in CytR and disorder in PurR is strong evidence of the role of long-range packing interactions driven by hydrophobicity in determining the helicity. Moreover, the A29V/A48I mutation in CytR exhibits a similar structural enhancement (Figure S3), highlighting the critical role of sequence hydrophobicity at these positions.

A Compact Native Ensemble in the CytR Double Mutant. Is the native ensemble of CytR DM well-folded, or does it exhibit properties of a molten globule? To answer this question, we recorded HSQC NMR spectra on the ^{15}N -labeled sample of WT CytR and CytR DM. The spectrum of the WT at 288 K is typical of a disordered protein with poor chemical shift dispersion in the ^1H dimension (blue in Figure 3A). In contrast, the spectrum of CytR DM shows significant chemical shift dispersion, demonstrating that it folds into an ordered conformation (red in Figure 3A). CytR has a sole tyrosine in its sequence at position 53 that brings together helices 1 and 3, serving as a unique probe for three-dimensional structure. The near-UV CD spectrum of CytR DM is significantly more intense than that of the WT with the unfolding curve displaying features similar to those of far-UV CD (Figure 3B,C) but with a lower apparent T_m of 312 ± 0.7 K. Accordingly, the quantum yield (QY) of the sole tyrosine is

also enhanced in the double mutant relative to the WT (Figure 3D).

The two single-point mutants of CytR (A29V and A48M) show similar retention times in size-exclusion chromatography (SEC), indicating that they are as unstructured as the WT (Figure S4). On the contrary, the double mutant exhibits a longer retention time characteristic of a more folded protein (Figure S3). Consistent with the SEC results, the hydrodynamic radii (R_h) are estimated to be ~ 24 and ~ 17.4 Å from DLS experiments for the WT and DM, respectively, at 283 K [assuming a spherical particle (Figure 3E)]. The larger dimension of CytR DM (residues 1–66) compared to the NMR model (residues 9–55, green triangle in Figure 3E) is a consequence of the longer tails at both N- and C-termini that are not accounted for in the latter. Taken together, the different spectroscopic probes point to a compact, well-packed, and highly helical native ensemble for the CytR double mutant at the lowest temperatures. The double mutation therefore tilts the folding free energy profile toward the folded state (i.e., with a favorable gradient toward the folded state) compared to the WT that shows the opposite behavior (favorable gradient toward the unfolded state).

Rugged Non-Two-State Landscape. Does a structurally well-defined native ensemble translate to strongly cooperative or two-state-like unfolding? Thermal unfolding monitored by scanning calorimetry experiments has been at the forefront of determining the nature of thermodynamic transitions in proteins.^{58,59} Specifically, heat capacity is a measure of the second moment of enthalpy, i.e., enthalpic fluctuations or the variance,⁶⁰ which is more sensitive to underlying conformational distributions than conventional measurements that

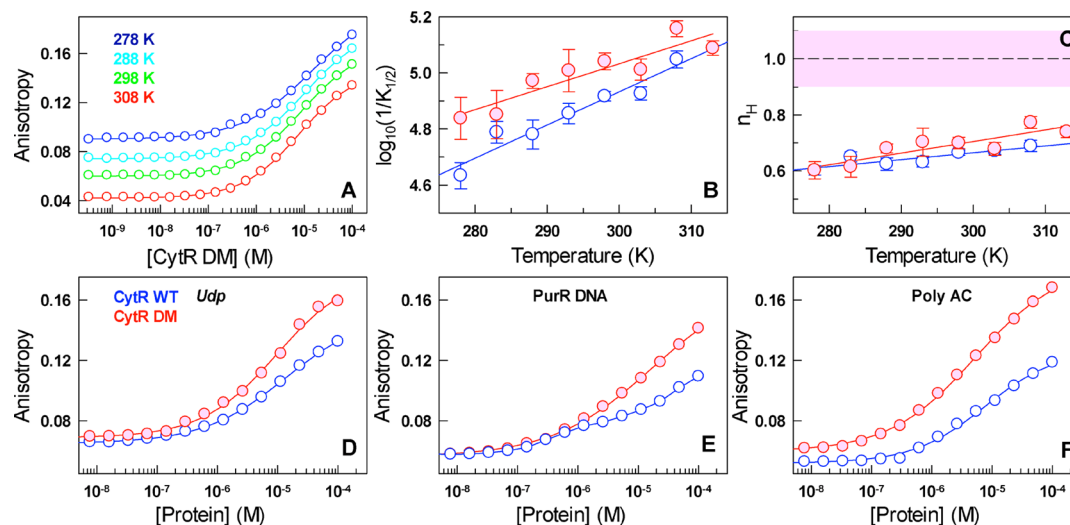


Figure 5. (A) Anisotropy titration profiles of CytR DM on the Alexa-532-labeled *udp* half-site. The curves are fits to the Hill equation. (B and C) Apparent binding affinities and Hill coefficients, respectively, as a function of temperature. Blue and red represent data for WT CytR and CytR DM, respectively. The dashed line and the corresponding shaded area in panel C highlight the expectation for a binding event with a single well-defined affinity. (D–F) Titration profiles of WT CytR (blue) and CytR DM (red) for different DNA sequences at 293 K.

probe the average signal. In this regard, the absolute heat capacity profile of CytR DM displays no pretransition and increases linearly at low temperatures (Figure 4A and Figure S5A), suggestive of large conformational heterogeneity with increasing temperatures well before the apparent denaturation midpoint. The fit to a two-state equilibrium model results in crossing baselines, a folded baseline with a large slope and deviation from the Freire baseline, indicating a non-two-state unfolding transition (Figure S5B). The weak excess heat capacity peak (T_m of 322 ± 0.9 K) is, however, evidence that the unfolding of CytR DM involves cooperative long-range effects that are observed in the WT disordered domain.

^1H - ^{15}N HSQC spectra acquired at different temperatures reveal that the intensities of DM CytR amide resonances decrease with an increase in temperature (Figure 4B). Such a decrease is usually a fingerprint of conformational exchange on the micro- to millisecond time scale and arises from exchange broadening at higher temperatures, as intensities otherwise increase at elevated temperatures because of faster rotational diffusion. A plot of the normalized variation in intensity with temperature shows that the decrease observed here is residue-specific, again consistent with what is expected for conformational exchange, with each residue having its own unique ^1H and ^{15}N chemical shift differences between the exchanging conformations (Figure 4C). Taken together, the temperature-dependent NMR data point to the presence of prevalent conformational exchange in CytR DM consistent with the expectation from scanning calorimetry experiments.

CytR DM displays a chemical denaturation midpoint of just 1.6 ± 0.6 M at 285 K (Figure 4E), indicating that the long-range interactions, though present, are weak. No pretransition regions are observed upon chemical denaturation, consistent with DSC measurements. Stopped-flow kinetic experiments are well-defined by a single kinetic phase (Figure S6) whose amplitudes match with equilibrium (green circles in Figure 4D). The observed relaxation rate constants decrease with an increase in urea concentration but surprisingly do not increase beyond the chemical denaturation midpoint as expected of a two-state system [a necessary but not a sufficient condition for two-state folding⁶³ (Figure 4E)]. This is strong evidence of a

complex unfolding mechanism in CytR DM. In contrast, the I40A mutant of PurR that displays a chemical denaturation midpoint similar to that of CytR DM shows weak chevronlike behavior at 285 K (black in Figure 4E) despite a lower cooperativity: a two-state analysis of the equilibrium melt results in apparent m -values of 4.4 ± 0.7 and 2.2 ± 1.1 $\text{kJ mol}^{-1} \text{M}^{-1}$ for CytR DM and PurR I40A, respectively (Figure S7).

We have previously shown that the disordered CytR samples an excited folded-like conformation in the native ensemble from temperature-dependent fluorescence kinetics.⁴⁶ The fluorescence signal arises from Y53 that packs against the first helix upon populating a folded-like conformation even when the landscape is tilted toward the disordered state. The experiments with WT CytR are characterized by a single kinetic phase whose amplitude decreases with temperature, signaling a decrease in the population of this excited state.⁴⁶ Because folded conformations dominate the CytR DM landscape, the kinetic phases exhibit proportionately larger amplitudes that again decrease with an increase in temperature, signaling a decrease in the population of the folded-like conformation (Figure S8). The observed rate constants, on the other hand, increase by a factor of only 2 compared to that of the WT despite the large favorable gradient toward the folded state (Figure 4F). The relaxation rate constants increase linearly with temperature (on a log scale) ranging from 60 to 300 s^{-1} between 283 and 303 K. The slow rate constants are suggestive of a slower folding diffusion coefficient in CytR DM that is consistent with the observations on WT CytR.⁴⁶

Nontrivial Effects of Structural Order on Binding Properties. WT CytR binds its cognate *udp* half-site with a weak apparent binding affinity ($1/K_{1/2}$) that increases with temperature.⁴⁷ The binding isotherms are broad and are well described by Hill coefficients of <1 ($n_H < 1$). We explained these observations by invoking binding heterogeneity at the level of single molecules that translate to broad binding isotherms and attributing these to the disordered nature of CytR.⁴⁷ The expectation is that CytR DM with a dramatically altered structure, stability, and folding landscape should display equivalent deviations in its binding properties. To our surprise,

we find that CytR DM exhibits similarly broad binding isotherms with a binding affinity and Hill coefficients only marginally higher than those of the WT for the cognate *udp* half-site (Figure 5A–C). The final anisotropy value is also marginally higher than that of the WT in accordance with the extracted higher binding affinity (Figure 5D). CytR DM further binds to the *udp* half-site faster than the dead time of the stopped-flow instrument under pseudo-first-order conditions (i.e., excess protein concentrations), setting the lower bound on k_{on} for binding at $>10^9 \text{ M}^{-1} \text{ s}^{-1}$, similar to that of the WT⁴⁸ (Figure S9).

The binding profile for binding of CytR DM to PurR DNA (binding site of the PurR DBD) and the poly-AC control show deviations from that of the WT, the origins of which are not obvious. Specifically, WT CytR shows two overlapping titration profiles on PurR DNA ($K_{1/2,1} \sim 0.4 \mu\text{M}$, and $K_{1/2,2} > 20 \mu\text{M}$; the fitting errors are of the same order of magnitude as those shown in panels B and C of Figure 5), while the CytR DM displays a single binding profile with an apparent $K_{1/2}$ of $\sim 40 \mu\text{M}$ (Figure 5E). Moreover, the final anisotropy value is significantly higher in the DM than in the WT. Similarly, the CytR DM binds the poly-AC control with a marginally higher affinity and greater heterogeneity ($K_{1/2} \sim 5 \mu\text{M}$; $n_H = 0.62$) compared to the WT ($K_{1/2} \sim 8 \mu\text{M}$; $n_H = 0.73$), with the former exhibiting a large final anisotropy value even upon accounting for the baseline differences at low protein concentrations (Figure 5F). The differences in the final anisotropy values are evidence of a more complex binding behavior potentially arising from changes in DNA structure or even in the bound conformation, orientation, or position of the CytR DM that cannot be readily extracted from the current experiments.

CONCLUSIONS

In this work, we have successfully engineered order in a disordered domain and explored the consequences for folding and function. We find that a mere two mutations separate the disordered CytR from its folded counterpart. CytR attains a compact structure only in the presence of both of the bulkier mutations, while the individual single-point substitutions have no measurable effect. In this regard, the precise role of apolar residues in compacting polymeric chains (i.e., a decrease in solvent-excluded volume that is driven by an increase in solvent entropy, the true definition of “hydrophobic effect”) is challenging to extract through studies of folded proteins, as small compaction or swelling of the native ensembles is challenging to deduce from conventional experiments; moreover, the native state frequently redistributes the dynamics and strength of packing to accommodate core mutations,^{37,54} thus confounding a direct interpretation. However, the engineered A29V and A48M mutations in CytR demonstrate the non-additive and non-intuitive nature of hydrophobic forces that drive compaction of protein chains in a model-independent manner.

Because mutations in CytR have been introduced through knowledge-based approaches at specific sequence positions (and not randomly), the protein ensemble not only reduces its dimensions but also orders itself into a folded-like conformation, despite the lower helical propensity of both of the mutations. This suggests that chain compaction and acquisition of secondary structure are interconnected, especially when the helices themselves have little intrinsic propensity: effective hydrophobicity clearly determines the

compaction (Figure 2), while backbone hydrogen bonds and long-range van der Waals packing stabilize the structure. The nonlocal interactions in the folded structure of CytR DM induce a weakly cooperative and heterogeneous unfolding transition as evidenced from differences in apparent melting temperatures (~ 312 , 318 , and 322 K from near-UV, far-UV CD, and DSC, respectively), broad and atypical heat capacity profiles, and non-chevron kinetics. In other words, CytR DM seems to unfold via heterogeneous molten globule-like conformations despite exhibiting a well-folded structure due to the combination of weak packing and destabilizing electrostatics. Accordingly, the observed relaxation rate constants are slow (just $\sim 300 \text{ s}^{-1}$ at 303 K), hinting at a slow folding diffusion coefficient (as in the WT⁴⁶) because of a rugged landscape with multiple conformational substates slowing the access to the folded state.

We demonstrate that the binding affinity of the CytR DM for the natural *udp* half-site is stronger (though only marginally) when compared to that of the disordered WT. Though increasing intrinsic helical propensity increases the binding affinity of IDPs for their partners, the mere marginal affinity increase observed is unexpected given the large structural compaction observed in CytR DM. Therefore, disorder does not seem to be a requirement for CytR binding to its natural half-site, though we observe nontrivial differences in the anisotropy profiles with noncognate DNA. It is possible that either disorder in CytR has been evolutionarily selected for regulatory purposes or the large negative electrostatic potential of DNA is strong enough to promote binding despite the extent of residual disorder, but a clearer picture will emerge only after studies of full length CytR. A higher ionic strength under *in vivo* conditions ($\sim 150 \text{ mM}$) could also promote partial folding of CytR given that the conformational landscape of this protein is remarkably sensitive to salt concentrations.⁴⁸ From the viewpoint of simulations, CytR and its mutational variants can serve as excellent model systems for understanding the interplay between disorder–compaction and folding–binding given the small size of the system. Finally, water models, implicit solvent energy functions, and the ability of force fields to capture the non-additive effects can be tested directly by employing the equilibrium data on ensemble dimensions (determined by solvent quality) and helicity (determined by the relative hydrogen-bond strength) of WT CytR and its variants.

Taken together, our work highlights how secondary and tertiary structure (compaction) and cooperativity arise as emergent properties in a disordered system driven purely by sequence hydrophobicity. CytR DM can thus be considered as an “intermediate” between highly cooperative two-state proteins and disordered systems, a less-explored sequence space with atypical unfolding and functional behaviors. Additional mutations at strategic positions should enhance the residual long-distance energetic coupling, shedding light on the precise contributions from hydrophobicity and electrostatics in determining stability and cooperativity in CytR. It should also be possible to engineer order in disordered domains with compact binding poses (and not those with extended binding conformations) via similar minimal sequence modifications involving hydrophobic residues, a route potentially taken by Nature to tune dimensions and functional–allosteric responses of protein ensembles.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.9b00182.

Electrostatic potential maps of the LacI family members, perturbation analysis of LacR structure, thermal unfolding curve of the CytR double mutant A29V/A48I, size-exclusion chromatography profiles of CytR mutants, concentration dependence of the apparent heat capacity of CytR DM and a two-state fit to the corresponding DSC curve, representative refolding kinetic traces of CytR DM at select temperatures, two-state fits to chemical denaturation curves of CytR DM and PurR I40A, and kinetic amplitudes from stopped-flow refolding experiments as a function of temperature (PDF)

Accession Codes

CytR, P0ACN7; PurR, P0ACP7.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: athi@iitm.ac.in. Telephone: +91-44-2257 4140.

ORCID

Athi N. Naganathan: 0000-0002-1655-7802

Author Contributions

[†]S.M. and S.S. contributed equally to this work.

Funding

This work was supported by Wellcome Trust/DBT India Alliance Intermediate Fellowship IA/I/15/1/501837 to A.N.N. A.S. acknowledges the Department of Science and Technology, India, for the Ramanujan Fellowship.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge the FIST facility sponsored by the Department of Science and Technology (DST), India, at the Department of Biotechnology, IITM, for the instrumentation.

■ ABBREVIATIONS

CD, circular dichroism; DSC, differential scanning calorimetry; NMR, nuclear magnetic resonance; DLS, dynamic light scattering.

■ REFERENCES

- Uversky, V. N. (2013) A decade and a half of protein intrinsic disorder: Biology still waits for physics. *Protein Sci.* 22, 693–724.
- van der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D. T., Kim, P. M., Kriwacki, R. W., Oldfield, C. J., Pappu, R. V., Tompa, P., Uversky, V. N., Wright, P. E., and Babu, M. M. (2014) Classification of intrinsically disordered regions and proteins. *Chem. Rev.* 114, 6589–6631.
- Fuxreiter, M., and Tompa, P. (2012) Fuzzy complexes: a more stochastic view of protein function. *Adv. Exp. Med. Biol.* 725, 1–14.
- Cumberworth, A., Lamour, G., Babu, M. M., and Gsponer, J. (2013) Promiscuity as a functional trait: intrinsically disordered regions as central players of interactomes. *Biochem. J.* 454, 361–369.
- Flock, T., Weatheritt, R. J., Latysheva, N. S., and Babu, M. M. (2014) Controlling entropy to tune the functions of intrinsically disordered regions. *Curr. Opin. Struct. Biol.* 26C, 62–72.

(6) Meng, W. L., Lyle, N., Luan, B. W., Raleigh, D. P., and Pappu, R. V. (2013) Experiments and simulations show how long-range contacts can form in expanded unfolded proteins with negligible secondary structure. *Proc. Natl. Acad. Sci. U. S. A.* 110, 2123–2128.

(7) Meng, W., and Raleigh, D. P. (2011) Analysis of electrostatic interactions in the denatured state ensemble of the N-terminal domain of L9 under native conditions. *Proteins: Struct., Funct., Genet.* 79, 3500–3510.

(8) Ozenne, V., Noel, J. K., Heidarsson, P. O., Brander, S., Poulsen, F. M., Jensen, M. R., Kragelund, B. B., Blackledge, M., and Danielsson, J. (2014) Exploring the Minimally Frustrated Energy Landscape of Unfolded ACBP. *J. Mol. Biol.* 426, 722–734.

(9) Waldauer, S. A., Bakajin, O., and Lapidus, L. J. (2010) Extremely slow intramolecular diffusion in unfolded protein L. *Proc. Natl. Acad. Sci. U. S. A.* 107, 13713–13717.

(10) Das, R. K., and Pappu, R. V. (2013) Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U. S. A.* 110, 13392–13397.

(11) Wuttke, R., Hofmann, H., Nettels, D., Borgia, M. B., Mittal, J., Best, R. B., and Schuler, B. (2014) Temperature-dependent solvation modulates the dimensions of disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.* 111, 5213–5218.

(12) Merlino, A., Pontillo, N., and Graziano, G. (2017) A driving force for polypeptide and protein collapse. *Phys. Chem. Chem. Phys.* 19, 751–756.

(13) Holehouse, A. S., and Pappu, R. V. (2018) Collapse Transitions of Proteins and the Interplay Among Backbone, Sidechain, and Solvent Interactions. *Annu. Rev. Biophys.* 47, 19–39.

(14) Müller-Späh, S., Soranno, A., Hirschfeld, V., Hofmann, H., Ruegger, S., Reymond, L., Nettels, D., and Schuler, B. (2010) Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14609–14614.

(15) Uversky, V. N., Gillespie, J. R., and Fink, A. L. (2000) Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins: Struct., Funct., Genet.* 41, 415–427.

(16) Das, R. K., Ruff, K. M., and Pappu, R. V. (2015) Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* 32, 102–112.

(17) Sherry, K. P., Das, R. K., Pappu, R. V., and Barrick, D. (2017) Control of transcriptional activity by design of charge patterning in the intrinsically disordered RAM region of the Notch receptor. *Proc. Natl. Acad. Sci. U. S. A.* 114, E9243–E9252.

(18) Rajasekaran, N., Gopi, S., Narayan, A., and Naganathan, A. N. (2016) Quantifying Protein Disorder through Measures of Excess Conformational Entropy. *J. Phys. Chem. B* 120, 4341–4350.

(19) Iesmantavicius, V., Dogan, J., Jemth, P., Teilum, K., and Kjaergaard, M. (2014) Helical Propensity in an Intrinsically Disordered Protein Accelerates Ligand Binding. *Angew. Chem., Int. Ed.* 53, 1548–1551.

(20) Kjaergaard, M., Norholm, A. B., Hendus-Altenburger, R., Pedersen, S. F., Poulsen, F. M., and Kragelund, B. B. (2010) Temperature-dependent structural changes in intrinsically disordered proteins: formation of alpha-helices or loss of polyproline II? *Protein Sci.* 19, 1555–1564.

(21) Dahal, L., Kwan, T. O. C., Hollins, J. J., and Clarke, J. (2018) Promiscuous and Selective: How Intrinsically Disordered BH3 Proteins Interact with Their Pro-survival Partner MCL-1. *J. Mol. Biol.* 430, 2468–2477.

(22) Karlsson, O. A., Chi, C. N., Engstrom, A., and Jemth, P. (2012) The transition state of coupled folding and binding for a flexible beta-finger. *J. Mol. Biol.* 417, 253–261.

(23) Dill, K. A. (1997) Additivity principles in biochemistry. *J. Biol. Chem.* 272, 701–704.

(24) Xu, J., Baase, W. A., Baldwin, E., and Matthews, B. W. (1998) The response of T4 lysozyme to large-to-small substitutions within the core and its relation to the hydrophobic effect. *Protein Sci.* 7, 158–177.

- (25) Chan, H. S. (2000) Modeling protein density of states: additive hydrophobic effects are insufficient for calorimetric two-state cooperativity. *Proteins: Struct., Funct., Genet.* 40, 543–571.
- (26) Chen, J., and Stites, W. E. (2001) Energetics of side chain packing in staphylococcal nuclease assessed by systematic double mutant cycles. *Biochemistry* 40, 14004–14011.
- (27) Shimizu, S., and Chan, H. S. (2002) Anti-cooperativity and cooperativity in hydrophobic interactions: Three-body free energy landscapes and comparison with implicit-solvent potential functions for proteins. *Proteins: Struct., Funct., Genet.* 48, 15–30.
- (28) Southall, N. T., Dill, K. A., and Haymet, A. D. J. (2002) A view of the hydrophobic effect. *J. Phys. Chem. B* 106, 521–533.
- (29) Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* 255, 178–183.
- (30) Hurlley, J. H., Baase, W. A., and Matthews, B. W. (1992) Design and structural analysis of alternative hydrophobic core packing arrangements in bacteriophage T4 lysozyme. *J. Mol. Biol.* 224, 1143–1159.
- (31) Baase, W. A., Eriksson, A. E., Zhang, X. J., Heinz, D. W., Sauer, U., Blaber, M., Baldwin, E. P., Wozniak, J. A., and Matthews, B. W. (1992) Dissection of protein structure and folding by directed mutagenesis. *Faraday Discuss.* 93, 173–181.
- (32) Ratnaparkhi, G. S., and Varadarajan, R. (2000) Thermodynamic and structural studies of cavity formation in proteins suggest that loss of packing interactions rather than the hydrophobic effect dominates the observed energetics. *Biochemistry* 39, 12365–12374.
- (33) Kaya, H., and Chan, H. S. (2003) Solvation effects and driving forces for protein thermodynamic and kinetic cooperativity: How adequate is native-centric topological modeling? *J. Mol. Biol.* 326, 911–931.
- (34) Dill, K. A., Truskett, T. M., Vlachy, V., and Hribar-Lee, B. (2005) Modeling water, the hydrophobic effect, and ion solvation. *Annu. Rev. Biophys. Biomol. Struct.* 34, 173–199.
- (35) Roche, J., Caro, J. A., Dellarole, M., Guca, E., Royer, C. A., Garcia-Moreno E., B., Garcia, A. E., and Roumestand, C. (2013) Structural, energetic, and dynamic responses of the native state ensemble of staphylococcal nuclease to cavity-creating mutations. *Proteins: Struct., Funct., Genet.* 81, 1069–1080.
- (36) Achoch, M., Dorantes-Gilardi, R., Wymant, C., Feverati, G., Salamatián, K., Vuillon, L., and Lesieur, C. (2016) Protein structural robustness to mutations: an in silico investigation. *Phys. Chem. Chem. Phys.* 18, 13770–13780.
- (37) Naganathan, A. N. (2019) Modulation of allosteric coupling by mutations: from protein dynamics and packing to altered native ensembles and function. *Curr. Opin. Struct. Biol.* 54, 1–9.
- (38) Shakhnovich, E. I. (1997) Theoretical studies of protein-folding thermodynamics and kinetics. *Curr. Opin. Struct. Biol.* 7, 29–40.
- (39) Shoemaker, B. A., Wang, J., and Wolynes, P. G. (1997) Structural correlations in protein folding funnels. *Proc. Natl. Acad. Sci. U. S. A.* 94, 777–782.
- (40) Klimov, D. K., and Thirumalai, D. (1998) Cooperativity in protein folding: from lattice models with sidechains to real proteins. *Folding Des.* 3, 127–139.
- (41) Chan, H. S., Zhang, Z., Wallin, S., and Liu, Z. (2011) Cooperativity, Local-Nonlocal Coupling, and Nonnative Interactions: Principles of Protein Folding from Coarse-Grained Models. *Annu. Rev. Phys. Chem.* 62, 301–326.
- (42) Muñoz, V., Campos, L. A., and Sadqi, M. (2016) Limited cooperativity in protein folding. *Curr. Opin. Struct. Biol.* 36, 58–66.
- (43) Malhotra, P., and Udgaonkar, J. B. (2016) How cooperative are protein folding and unfolding transitions? *Protein Sci.* 25, 1924–1941.
- (44) Moody, C. L., Tretyachenko-Ladokhina, V., Laue, T. M., Senear, D. F., and Cocco, M. J. (2011) Multiple Conformations of the Cytidine Repressor DNA-Binding Domain Coalesce to One upon Recognition of a Specific DNA Surface. *Biochemistry* 50, 6622–6632.
- (45) Naganathan, A. N., and Orozco, M. (2013) The conformational landscape of an intrinsically disordered DNA-binding domain of a transcription regulator. *J. Phys. Chem. B* 117, 13842–13850.
- (46) Munshi, S., Rajendran, D., and Naganathan, A. N. (2018) Entropic Control of an Excited Folded-Like Conformation in a Disordered Protein Ensemble. *J. Mol. Biol.* 430, 2688–2694.
- (47) Munshi, S., Gopi, S., Subramanian, S., Campos, L. A., and Naganathan, A. N. (2018) Protein plasticity driven by disorder and collapse governs the heterogeneous binding of CytR to DNA. *Nucleic Acids Res.* 46, 4044–4053.
- (48) Munshi, S., Gopi, S., Asampille, G., Subramanian, S., Campos, L. A., Atreya, H. S., and Naganathan, A. N. (2018) Tunable order-disorder continuum in protein-DNA interactions. *Nucleic Acids Res.* 46, 8700–8709.
- (49) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277–293.
- (50) Lee, W., Tonelli, M., and Markley, J. L. (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* 31, 1325–1327.
- (51) Guzman-Casado, M., Parody-Morreale, A., Robic, S., Marqusee, S., and Sanchez-Ruiz, J. M. (2003) Energetic evidence for formation of a pH-dependent hydrophobic cluster in the denatured state of *Thermus thermophilus* ribonuclease H. *J. Mol. Biol.* 329, 731–743.
- (52) Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* 32, W665–667.
- (53) Rajasekaran, N., and Naganathan, A. N. (2017) A self-consistent structural perturbation approach for determining the magnitude and extent of allosteric coupling in proteins. *Biochem. J.* 474, 2379–2388.
- (54) Rajasekaran, N., Suresh, S., Gopi, S., Raman, K., and Naganathan, A. N. (2017) A general mechanism for the propagation of mutational effects in proteins. *Biochemistry* 56, 294–305.
- (55) Chakrabarty, A., and Baldwin, R. L. (1995) Stability of alpha-helices. *Adv. Protein Chem.* 46, 141–176.
- (56) Pace, C. N., and Scholtz, J. M. (1998) A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J.* 75, 422–427.
- (57) Uversky, V. N. (2012) Size-exclusion chromatography in structural analysis of intrinsically disordered proteins. *Methods Mol. Biol.* 896, 179–194.
- (58) Lumry, R., Biltonen, R. L., and Brandts, J. F. (1966) Validity of the “Two-State” Hypothesis for Conformational Transitions of Proteins. *Biopolymers* 4, 917–944.
- (59) Muñoz, V., and Sanchez-Ruiz, J. M. (2004) Exploring protein folding ensembles: a variable barrier model for the analysis of equilibrium unfolding experiments. *Proc. Natl. Acad. Sci. U. S. A.* 101, 17646–17651.
- (60) Ibarra-Molero, B., Naganathan, A. N., Sanchez-Ruiz, J. M., and Muñoz, V. (2016) Modern Analysis of Protein Folding by Differential Scanning Calorimetry. *Methods Enzymol.* 567, 281–318.
- (61) Gomez, J., Hilser, V. J., Xie, D., and Freire, E. (1995) The Heat-Capacity of Proteins. *Proteins: Struct., Funct., Genet.* 22, 404–412.
- (62) Makhatazde, G. I., and Privalov, P. L. (1990) Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. *J. Mol. Biol.* 213, 375–384.
- (63) Naganathan, A. N., Doshi, U., and Muñoz, V. (2007) Protein folding kinetics: Barrier effects in chemical and thermal denaturation experiments. *J. Am. Chem. Soc.* 129, 5673–5682.