Published in final edited form as:
Biochemistry. 2019 May 14; 58(19): 2389–2397. doi:10.1021/acs.biochem.9b00182.

Engineering Order and Cooperativity in a Disordered Protein

Sneha Munshi[#1], **Sandhyaa Subramanian**[#1], **Samyuktha Ramesh**[1], **Hemashree Golla**[1], **Kalivarathan Divakar**[2], **Madhurima Kulkarni**[3], **Luis A. Campos**[4], **Ashok Sekhar**[3], and **Athi N. Naganathan**[1,*]

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

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

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

[4]National Biotechnology Center, Consejo Superior de Investigaciones Científicas, Darwin 3, Campus de Cantoblanco, 28049 Madrid, Spain

[#] These authors contributed equally to this work.

Abstract

Structural disorder in proteins arises from a complex interplay between weak hydrophobicity and unfavorable electrostatic interactions. The extent to which 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

[*]**Corresponding Author**: athi@iitm.ac.in, Phone: +91-44-2257 4140.

Supporting Information accompanies this paper.
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 apparent heat capacity of the CytR DM and a two-state fit to the corresponding DSC curve, representative refolding kinetic traces of the CytR DM at select temperatures, two-state fits to chemical denaturation curves of the CytR DM and PurR I40A mutant, kinetic amplitudes from stopped-flow refolding experiments as a function of temperature.

Accession Codes
CytR (P0ACN7) and PurR (P0ACP7).

Competing Financial Interests
The authors declare no competing financial interests.

Europe PMC Funders Author Manuscripts

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 only a marginally higher affinity than the wild-type, though non-trivial differences are observed in the binding to non-cognate DNA. Our work provides direct experimental evidence for 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 in protein design and functional manipulation.

Keywords

Electrostatics; Hydrophobicity; Molten-globule; Rugged landscape; Frustration

Introduction

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 for an evolutionary selection for disorder, the advantages of which can range from efficient regulation, large 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 on the role of small sequence modifications on the residual structure, dynamics, populations of partially structured states and their functional role, aspects that were earlier studied on denatured states of folded proteins.6–9

Accordingly, the basic thermodynamic forces that drive disorder in proteins have received considerable interest over the last decade10–14 starting with the seminal work of Uversky and coworkers.15 IDP sequences are enriched in 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/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, Gln) cannot populate a well-defined structure due to the energetically degenerate nature of intra-chain hydrogen-bond interactions. In contrast to low complexity sequences, disorder tendency in high complexity sequences is determined by the combined effect of several factors, that are yet to be fully understood, including weak hydrophobicity, strategic positioning of Gly/Pro residues, destabilizing electrostatics and correlated sequence effects.18

Europe PMC Funders Author Manuscripts

In the majority of protein engineering studies on IDPs, mutations modulate only the local structure or the secondary structural propensity while the protein remains predominantly unstructured.19–22 These works lead to interesting questions: can an IDP be engineered to fold to a compact structure via minimal hydrophobic substitutions? What is the consequence of structural ordering on the conformational landscape and hence the binding affinity to its partner ligand? Particularly, the role of protein sequence hydrophobicity and its non-additive effect on protein stability has been predicted and tested via multiple approaches in well-folded domains.23–28 The non-additivity has its origins in complex solvation effects and the precise and directional nature of packing interactions that are almost always interlinked when studying folded proteins.29–37 Disordered domains, on the other hand, can serve as excellent model systems to understand how higher order packing effects emerge upon mutations in the sequence, as packing interactions are minimal or non-existent to start with. Moreover, if the design of a folded variant is successful, the resulting changes in ensemble dimensions and stability will be more dramatic, providing a cleaner signal. These two terms should also provide a direct measure of the extent to which solvation (and hence 'hydrophobic effect') contributes to polymer compaction to a folded structure (and not to a non-specific globule) while establishing multiple weak non-local interactions that determine cooperativity.38–43

We answer these questions in the current work employing the disordered DNA-binding domain (DBD) of CytR as a model system (termed CytR from hereon)44, a member of the LacI DBD family. CytR exhibits weak helicity and dimensions of a disordered protein, populates an excited folded-like conformation in its ensemble45, 46 but undergoes a second-order-like non-specific collapse transition with increase in temperature.47 It folds on approaching DNA, driven by electrostatic screening and a continuous conformational selection mechanism, a feature that can also be mimicked by modulating salt concentration in the buffer.48 CytR thus serves as a highly tunable system whose baseline properties (in the presence and absence of DNA) are well established. In this work, we identify two positions along the sequence of CytR (A29 and A48) that are critical for promoting disorder through sequence and structural analysis. Single-point hydrophobic substitutions at these positions have little effect on the structure, while the double-mutation – A29V/A48M – promotes collapse and significant structural ordering. The conformational landscape of this double mutant is shown to be complex with non-trivial effects on function. Importantly, we provide direct experimental evidence for the non-additive nature of hydrophobicity and experimentally demonstrate the emergent nature of order and cooperativity expected of polymeric systems.

Methods

CytR, PurR and their mutants were expressed and purified as described before.47, 48 All experiments were recorded in pH 7.0, 20 mM phosphate buffer (43 mM ionic strength) with or without urea. Buffers were freshly prepared in MilliQ water, filtered and degassed prior to experiments. The protein samples were filtered using 0.22 μm syringe filter (Millipore) and the absorbance was measured using a UV-Vis spectrophotometer (Jasco). An extinction coefficient of 1490 M^{-1} cm^{-1} and 6990 M^{-1} cm^{-1}, was used for calculating the concentrations of CytR and PurR, respectively. Experimental conditions for CytR, PurR and their variants

are identical unless otherwise mentioned. The details of dynamic light scattering experiments and electrostatic potential calculations can be found in references 47 and 48, respectively.

Far- and Near-UV Circular Dichroism

Far-UV CD spectra, as a function of temperature or urea, were recorded in a Jasco J-815 spectrophotometer connected to a Peltier at protein concentrations of ~25 μM in a 1 mm pathlength quartz cuvette. Near-UV CD experiments were performed at protein concentrations of ~100 μM and ~70 μM for CytR and PurR, respectively, in a 10 mm quartz cuvette.

Quantum Yield Measurements

Fluorescence experiments were performed in a Chirascan-Plus qCD instrument (Applied Photophysics Ltd., UK). Emission spectra were accumulated 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 the 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 of 43 mM ionic strength. NMR samples of $^{15}N/^{13}C$ CytR WT and ^{15}N CytR DM had concentrations in the 600-700 μM range. NMR data was acquired in a 14.1 T (600 MHz) Agilent spectrometer equipped with a cryogenically cooled triple resonance probe. Spectra were processed using NMRPipe49 and overlays were obtained using Sparky.50

Differential Scanning Calorimetry

Scanning calorimetry measurements were performed in a MicroCal VP-Capillary DSC (Malvern Ltd., UK) 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 coworkers.51

Stopped-Flow Kinetics

Refolding and unfolding kinetic traces as a function of urea and temperature were recorded in a Chirascan SF.3 Stopped Flow instrument (Applied Photophysics Ltd., UK) coupled to a thermostated water bath. A PMT detector recorded the total fluorescence on excitation by a 280 nm LED while a 295 nm cut-off filter before the detector filtered out the scattered photons. For refolding experiments, a sample of unfolded CytR in 4M of urea (mutants in 6M urea) and unfolded PurR in 8M urea were rapidly mixed with an excess of 43 mM ionic

strength buffer (1:10 mixing) resulting in final protein concentrations of ~18 μM and ~ 10 μM for CytR and PurR, respectively. For unfolding experiments, folded PurR and its mutants in denaturant-free buffer was rapidly mixed with excess urea and the time-dependent unfolding signal 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 one minute.

Binding Anisotropy Profiles

300 nM of double stranded udpO (5'-ATTTATGCAACGCA-3') labeled with Alexa-532 at the 5`-end (IBA Lifesciences) was titrated with increasing protein concentrations ranging from ~1 nm - 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 collected at 580 nm in a Chirascan-Plus qCD instrument (Applied Photophysics Ltd.). Binding studies were also performed with 300 nM of 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 anisotropy of Alexa-532 labeled udpO on 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 chloride and 1 mM EDTA, pH 6.0. The final DNA concentration was 300 nM. Six kinetic traces were recorded at an interval of one minute 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 near-continuous manner on increasing the concentration of salt48 signaling large electrostatic frustration in the native state (Figure 1A, 1B). 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 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 increase on increasing 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 at all ionic strength conditions explored. This observation is first evidence that differences in packing interactions determine the disordered tendency of CytR compared to 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 the residue index (Figure 1C). Coupling distances are a

measure of the strength of the intra-protein interaction network derived from both 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 the residues when compared to PurR and LacR (Figure 1C and S2). Sequence alignment indicates similar hydrophobic tendency across all the apolar positions in CytR (stars in Figure 1D), except for the 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 and 2D). Taking a cue from sequence-structure analysis, we mutated the residue positions A29 and A48 to bulkier hydrophobic side-chains to enhance hydrophobicity and potentially the 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 can be observed in the far-UV CD spectrum and the thermal unfolding curve (Figure 2B, 2C). However, introducing both the mutations in the sequence, as in the double mutant A29V/A48M (DM), significantly enhanced the helical structure and stability highlighting the non-additive nature of hydrophobic effect. The far-UV CD spectrum at the lowest temperature shows the characteristic features of a helical protein with minima at ~208/222 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 the CytR DM exhibits a folding transition with an apparent T_m of 318 ± 0.4 K (red in Figure 2C).

The observations above suggest that the 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, 2F) thus validating the critical role of these two positions for promoting order in the LacI family. The V21A mutant of PurR is in fact very similar to the WT CytR in terms of the overall unfolding curve (compare blue in Figure 2C with 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 propensity55, 56). The fact that we observe these mutations to promote order in CytR and disorder in PurR, respectively, is strong evidence for 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 the CytR DM well-folded or does it exhibit properties of a molten-globule? To answer this question, we acquired HSQC NMR spectra on the ^{15}N labeled sample of CytR WT and DM. The spectrum of the WT at 288 K is typical of a disordered protein with poor chemical shift dispersion in the ^{1}H 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 the CytR DM is significantly more intense than the WT with the unfolding curve displaying features similar to that from far-UV CD (Figure 3B, 3C) but with a lower apparent T_m of 312 ± 0.7 K. Accordingly, the 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 SEC, the hydrodynamic radii (Rh) 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 the consequence of the longer tails at both the 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 towards the folded state (i.e. with a favorable gradient towards the folded state) compared to the WT that shows the opposite behavior (favorable gradient towards 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 have 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 probe the average signal. In this regard, the absolute heat capacity profile of the CytR DM displays no pre-transition 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 the CytR DM involves cooperative long-range effects that are not observable in the WT disordered domain.

^{1}H-^{15}N HSQC spectra acquired at different temperatures reveal that the intensities of DM CytR amide resonances decrease with increasing temperature (Figure 4B). Such a decrease is usually a fingerprint of conformational exchange in the μs-ms timescale 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 intensities 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 ^{1}H 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 DM CytR consistent with the expectation from scanning calorimetry experiments.

The 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 pre-transition regions are observed on 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 increasing 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 folding;63 Figure 4E). This is strong evidence for a complex unfolding mechanism in the CytR DM. In contrast, the I40A mutant of PurR that displays a similar chemical denaturation midpoint as the CytR DM shows weak chevron-like behavior at 285 K (black in Figure 4E) despite a lower cooperativity: a two-state analysis of the equilibrium melt results in an apparent *m*-value of 4.4±0.7 and 2.2±1.1 kJ/(mol.M) for the CytR DM and PurR I40A, respectively (Figure S7).

We have earlier 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 towards the disordered state. The experiments on CytR WT are characterized by a single kinetic phase whose amplitude decreases with temperature signaling a decrease in population of this excited state.46 Since folded conformations dominate the CytR DM landscape, the kinetic phases exhibit proportionately larger amplitudes that again decrease with increasing temperatures signaling a decrease in the population of the folded-like conformation (Figure S8). The observed rate constants, on the other hand, increase only by a factor of two compared to the WT despite the large favorable gradient towards the folded state (Figure 4F). The relaxation rate constants increase linearly with temperature (in the log-scale) ranging from 60 s^{-1} to 300 s^{-1} between 283 and 303 K. The slow rate constants are suggestive of a slower folding diffusion coefficient in the CytR DM that is consistent with the observations on CytR WT.46

Non-Trivial Effects of Structural Order on Binding Properties

The CytR WT 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 can be accounted

for by Hill coefficients less than 1 ($n_H < 1$). We explained these observations 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 the 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 the WT for the cognate *udp* half-site (Figure 5A-C). The final anisotropy value is also marginally higher than the WT in accordance with the extracted stronger binding affinity (Figure 5D). The 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 $>10^9$ M^{-1} s^{-1}, similar to the WT48 (Figure S9).

The binding profile of the 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, CytR WT shows two overlapping titration profiles on PurR DNA ($K_{1/2,1}$ ~0.4 μM, $K_{1/2,2}$ >20 μM; the fitting errors are of the same order as shown in Figure 5B, 5C) while the CytR DM displays a single binding profile with an apparent $K_{1/2}$ ~40 μM (Figure 5E). Moreover, the final anisotropy value is significantly higher in the DM compared to the WT. Similarly, the CytR DM binds the poly-AC control with a marginally higher affinity and larger heterogeneity ($K_{1/2}$~5 μM, n_H=0.62) compared to the WT ($K_{1/2}$ ~8 μM, 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 for 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 on 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 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 a gain in solvent entropy - the true definition of ‘hydrophobic effect’) is challenging to extract through studies on folded proteins, as small compaction or swelling of the native ensembles are challenging to deduce from conventional experiments; moreover, the native state frequently redistributes the dynamics and strength of packing to accommodate core mutations37, 54 thus confounding a direct interpretation. However, the engineered A29V/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.

Since 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 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 non-local interactions in the folded structure of the CytR DM induce a weakly cooperative and heterogeneous unfolding transition as evidenced from differences in apparent melting temperatures (~312 K, 318 K 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, the 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 ~300 s^{-1} at 303 K) hinting at a slow folding diffusion coefficient (as in the WT46) because of a rugged landscape with multiple conformational sub-states slowing down the access to the folded state.

We demonstrate that the binding affinity of the CytR DM to the natural *udp* half-site is stronger (though only marginally) when compared to the disordered WT. Though increasing intrinsic helical propensity increases the binding affinity of IDPs to their partners, the mere marginal affinity increase observed is unexpected given the large structural compaction observed in the CytR DM. Therefore, disorder does not seem to be a requirement for CytR binding to its natural half-site, though we observe non-trivial differences in the anisotropy profiles with non-cognate 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 on full-length CytR. Higher ionic strength under *in vivo* conditions (~150 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 to understand 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 employing the equilibrium data on ensemble dimensions (determined by solvent quality) and helicity (determined by the relative hydrogen bond strength) of CytR WT 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. The CytR DM can thus be thought of 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.

Acknowledgement

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

Abbreviations

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

References

1. Uversky VN. A decade and a half of protein intrinsic disorder: Biology still waits for physics. Protein Sci. 2013; 22:693–724. [PubMed: 23553817]
2. van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer J, Jones DT, Kim PM, et al. Classification of intrinsically disordered regions and proteins. Chem Rev. 2014; 114:6589–6631. [PubMed: 24773235]
3. Fuxreiter M, Tompa P. Fuzzy complexes: a more stochastic view of protein function. Adv Exp Med Biol. 2012; 725:1–14. [PubMed: 22399315]
4. Cumberworth A, Lamour G, Babu MM, Gsponer J. Promiscuity as a functional trait: intrinsically disordered regions as central players of interactomes. Biochem J. 2013; 454:361–369. [PubMed: 23988124]
5. Flock T, Weatheritt RJ, Latysheva NS, Babu MM. Controlling entropy to tune the functions of intrinsically disordered regions. Curr Opin Struct Biol. 2014; 26C:62–72.
6. Meng WL, Lyle N, Luan BW, Raleigh DP, Pappu RV. Experiments and simulations show how long-range contacts can form in expanded unfolded proteins with negligible secondary structure. Proc Natl Acad Sci USA. 2013; 110:2123–2128. [PubMed: 23341588]
7. Meng W, Raleigh DP. Analysis of electrostatic interactions in the denatured state ensemble of the N-terminal domain of L9 under native conditions. Proteins. 2011; 79:3500–3510. [PubMed: 21915914]
8. Ozenne V, Noel JK, Heidarsson PO, Brander S, Poulsen FM, Jensen MR, Kragelund BB, Blackledge M, Danielsson J. Exploring the Minimally Frustrated Energy Landscape of Unfolded ACBP. J Mol Biol. 2014; 426:722–734. [PubMed: 24211721]
9. Waldauer SA, Bakajin O, Lapidus LJ. Extremely slow intramolecular diffusion in unfolded protein L. Proc Natl Acad Sci U S A. 2010; 107:13713–13717. [PubMed: 20643973]
10. Das RK, Pappu RV. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. Proc Natl Acad Sci USA. 2013; 110:13392–13397. [PubMed: 23901099]
11. Wuttke R, Hofmann H, Nettels D, Borgia MB, Mittal J, Best RB, Schuler B. Temperature-dependent solvation modulates the dimensions of disordered proteins. Proc Natl Acad Sci U S A. 2014; 111:5213–5218. [PubMed: 24706910]
12. Merlino A, Pontillo N, Graziano G. A driving force for polypeptide and protein collapse. Phys Chem Chem Phys. 2016; 19:751–756. [PubMed: 27929162]
13. Holehouse AS, Pappu RV. Collapse Transitions of Proteins and the Interplay Among Backbone, Sidechain, and Solvent Interactions. Annu Rev Biophys. 2018; 47:19–39.

14. Müller-Späth S, Soranno A, Hirschfeld V, Hofmann H, Ruegger S, Reymond L, Nettels D, Schuler B. Charge interactions can dominate the dimensions of intrinsically disordered proteins. Proc Natl Acad Sci USA. 2010; 107:14609–14614. [PubMed: 20639465]
15. Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins: Struct Funct Genet. 2000; 41:415–427. [PubMed: 11025552]
16. Das RK, Ruff KM, Pappu RV. Relating sequence encoded information to form and function of intrinsically disordered proteins. Curr Opin Struct Biol. 2015; 32:102–112. [PubMed: 25863585]
17. Sherry KP, Das RK, Pappu RV, Barrick D. 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. 2017; 114:E9243–E9252. [PubMed: 29078291]
18. Rajasekaran N, Gopi S, Narayan A, Naganathan AN. Quantifying Protein Disorder through Measures of Excess Conformational Entropy. J Phys Chem B. 2016; 120:4341–4350. [PubMed: 27111521]
19. Iesmantavicius V, Dogan J, Jemth P, Teilum K, Kjaergaard M. Helical Propensity in an Intrinsically Disordered Protein Accelerates Ligand Binding. Angew Chem Int Ed. 2014; 53:1548–1551.
20. Kjaergaard M, Norholm AB, Hendus-Altenburger R, Pedersen SF, Poulsen FM, Kragelund BB. Temperature-dependent structural changes in intrinsically disordered proteins: formation of alpha-helices or loss of polyproline II? Protein Sci. 2010; 19:1555–1564. [PubMed: 20556825]
21. Dahal L, Kwan TOC, Hollins JJ, Clarke J. Promiscuous and Selective: How Intrinsically Disordered BH3 Proteins Interact with Their Pro-survival Partner MCL-1. J Mol Biol. 2018; 430:2468–2477. [PubMed: 29654795]
22. Karlsson OA, Chi CN, Engstrom A, Jemth P. The transition state of coupled folding and binding for a flexible beta-finger. J Mol Biol. 2012; 417:253–261. [PubMed: 22310047]
23. Dill KA. Additivity principles in biochemistry. J Biol Chem. 1997; 272:701–704. [PubMed: 8995351]
24. Xu J, Baase WA, Baldwin E, Matthews BW. The response of T4 lysozyme to large-to-small substitutions within the core and its relation to the hydrophobic effect. Protein Sci. 1998; 7:158–177. [PubMed: 9514271]
25. Chan HS. Modeling protein density of states: additive hydrophobic effects are insufficient for calorimetric two-state cooperativity. Proteins. 2000; 40:543–571. [PubMed: 10899781]
26. Chen J, Stites WE. Energetics of side chain packing in staphylococcal nuclease assessed by systematic double mutant cycles. Biochemistry. 2001; 40:14004–14011. [PubMed: 11705392]
27. Shimizu S, Chan HS. Anti-cooperativity and cooperativity in hydrophobic interactions: Three-body free energy landscapes and comparison with implicit-solvent potential functions for proteins. Proteins. 2002; 48:15–30. [PubMed: 12012334]
28. Southall NT, Dill KA, Haymet ADJ. A view of the hydrophobic effect. J Phys Chem B. 2002; 106:521–533.
29. Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, Matthews BW. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. Science. 1992; 255:178–183. [PubMed: 1553543]
30. Hurley JH, Baase WA, Matthews BW. Design and structural analysis of alternative hydrophobic core packing arrangements in bacteriophage T4 lysozyme. J Mol Biol. 1992; 224:1143–1159. [PubMed: 1569571]
31. Baase WA, Eriksson AE, Zhang XJ, Heinz DW, Sauer U, Blaber M, Baldwin EP, Wozniak JA, Matthews BW. Dissection of protein structure and folding by directed mutagenesis. Faraday Discuss. 1992:173–181. [PubMed: 1290931]
32. Ratnaparkhi GS, Varadarajan R. 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. 2000; 39:12365–12374. [PubMed: 11015216]
33. Kaya H, Chan HS. Solvation effects and driving forces for protein thermodynamic and kinetic cooperativity: How adequate is native-centric topological modeling? J Mol Biol. 2003; 326:911–931. [PubMed: 12581650]
34. Dill KA, Truskett TM, Vlachy V, Hribar-Lee B. Modeling water, the hydrophobic effect, and ion solvation. Annu Rev Biophys Biomol Struct. 2005; 34:173–199. [PubMed: 15869376]

35. Roche J, Caro JA, Dellarole M, Guca E, Royer CA, Garcia-Moreno BE, Garcia AE, Roumestand C. Structural, energetic, and dynamic responses of the native state ensemble of staphylococcal nuclease to cavity-creating mutations. Proteins. 2013; 81:1069–1080. [PubMed: 23239146]
36. Achoch M, Dorantes-Gilardi R, Wymant C, Feverati G, Salamatian K, Vuillon L, Lesieur C. Protein structural robustness to mutations: an in silico investigation. Phys Chem Chem Phys. 2016; 18:13770–13780. [PubMed: 26688116]
37. Naganathan AN. Modulation of allosteric coupling by mutations: from protein dynamics and packing to altered native ensembles and function. Curr Opin Struct Biol. 2019; 54:1–9.
38. Shakhnovich EI. Theoretical studies of protein-folding thermodynamics and kinetics. Curr Opin Struct Biol. 1997; 7:29–40. [PubMed: 9032061]
39. Shoemaker BA, Wang J, Wolynes PG. Structural correlations in protein folding funnels. Proc Natl Acad Sci USA. 1997; 94:777–782. [PubMed: 9023333]
40. Klimov DK, Thirumalai D. Cooperativity in protein folding: from lattice models with sidechains to real proteins. Fold Des. 1998; 3:127–139. [PubMed: 9565757]
41. Chan HS, Zhang Z, Wallin S, Liu Z. Cooperativity, Local-Nonlocal Coupling, and Nonnative Interactions: Principles of Protein Folding from Coarse-Grained Models. Leone SR, Cremer PS, Groves JT, Johnson MA. Ann Rev Phys Chem. 2011:301–326. [PubMed: 21453060]
42. Muñoz V, Campos LA, Sadqi M. Limited cooperativity in protein folding. Curr Opin Struct Biol. 2016; 36:58–66. [PubMed: 26845039]
43. Malhotra P, Udgaonkar JB. How cooperative are protein folding and unfolding transitions? Protein Sci. 2016; 25:1924–1941. [PubMed: 27522064]
44. Moody CL, Tretyachenko-Ladokhina V, Laue TM, Senear DF, Cocco MJ. Multiple Conformations of the Cytidine Repressor DNA-Binding Domain Coalesce to One upon Recognition of a Specific DNA Surface. Biochemistry. 2011; 50:6622–6632. [PubMed: 21688840]
45. Naganathan AN, Orozco M. The conformational landscape of an intrinsically disordered DNA-binding domain of a transcription regulator. J Phys Chem B. 2013; 117:13842–13850. [PubMed: 24127726]
46. Munshi S, Rajendran D, Naganathan AN. Entropic Control of an Excited Folded-Like Conformation in a Disordered Protein Ensemble. J Mol Biol. 2018; 430:2688–2694. [PubMed: 29885328]
47. Munshi S, Gopi S, Subramanian S, Campos LA, Naganathan AN. Protein plasticity driven by disorder and collapse governs the heterogeneous binding of CytR to DNA. Nucleic Acids Res. 2018; 46:4044–4053. [PubMed: 29538715]
48. Munshi S, Gopi S, Asampille G, Subramanian S, Campos LA, Atreya HS, Naganathan AN. Tunable order-disorder continuum in protein-DNA interactions. Nucleic Acids Res. 2018; 46:8700–8709. [PubMed: 30107436]
49. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR. 1995; 6:277–293. [PubMed: 8520220]
50. Lee W, Tonelli M, Markley JL. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics. 2015; 31:1325–1327. [PubMed: 25505092]
51. Guzman-Casado M, Parody-Morreale A, Robic S, Marqusee S, Sanchez-Ruiz JM. Energetic evidence for formation of a pH-dependent hydrophobic cluster in the denatured state of Thermus thermophilus ribonuclease H. J Mol Biol. 2003; 329:731–743. [PubMed: 12787674]
52. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res. 2004; 32:W665–667. [PubMed: 15215472]
53. Rajasekaran N, Naganathan AN. A self-consistent structural perturbation approach for determining the magnitude and extent of allosteric coupling in proteins. Biochem J. 2017; 474:2379–2388. [PubMed: 28522638]
54. Rajasekaran N, Suresh S, Gopi S, Raman K, Naganathan AN. A general mechanism for the propagation of mutational effects in proteins. Biochemistry. 2017; 56:294–305. [PubMed: 27958720]

55. Chakrabartty A, Baldwin RL. Stability of alpha-helices. Adv Protein Chem. 1995; 46:141–176. [PubMed: 7771317]

56. Pace CN, Scholtz JM. A helix propensity scale based on experimental studies of peptides and proteins. Biophys J. 1998; 75:422–427. [PubMed: 9649402]

57. Uversky VN. Size-exclusion chromatography in structural analysis of intrinsically disordered proteins. Methods Mol Biol. 2012; 896:179–194. [PubMed: 22821524]

58. Lumry R, Biltonen RL, Brandts JF. Validity of the “Two-State” Hypothesis for Conformational Transitions of Proteins. Biopolymers. 1966; 4:917–944. [PubMed: 5975643]

59. Muñoz V, Sanchez-Ruiz JM. Exploring protein folding ensembles: a variable barrier model for the analysis of equilibrium unfolding experiments. Proc Natl Acad Sci USA. 2004; 101:17646–17651. [PubMed: 15591110]

60. Ibarra-Molero B, Naganathan AN, Sanchez-Ruiz JM, Muñoz V. Modern Analysis of Protein Folding by Differential Scanning Calorimetry. Methods Enzymol. 2016; 567:281–318. [PubMed: 26794359]

61. Gomez J, Hilser VJ, Xie D, Freire E. The Heat-Capacity of Proteins. Proteins. 1995; 22:404–412. [PubMed: 7479713]

62. Makhatadze GI, Privalov PL. Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. J Mol Biol. 1990; 213:375–384. [PubMed: 2342113]

63. Naganathan AN, Doshi U, Muñoz V. Protein folding kinetics: Barrier effects in chemical and thermal denaturation experiments. J Am Chem Soc. 2007; 129:5673–5682. [PubMed: 17419630]

Europe PMC Funders Author Manuscripts

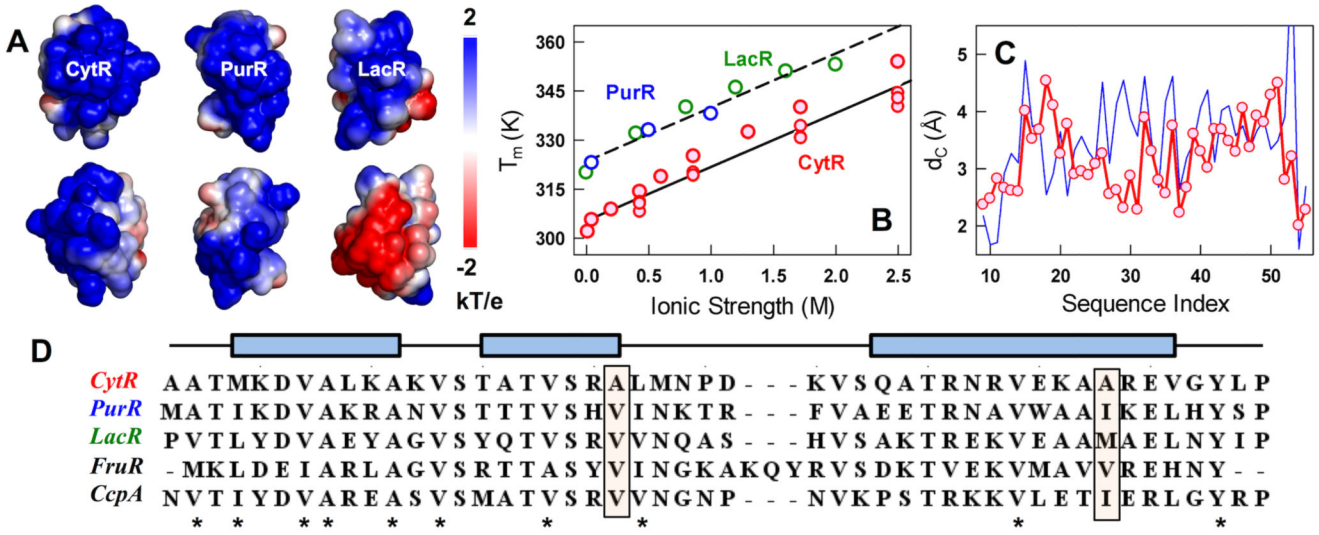

Figure 1.

Origins of disorder in the LacI DBD family. (A) Electrostatic potential maps of CytR, PurR and LacR (298 K, 43 mM ionic strength, pH 7.0) highlighting the electrostatic frustration at the DNA-binding face (top row) and the opposite face (bottom row). Note that the DNA-bound conformation of CytR is used for the calculation. (B) Observed changes in melting temperatures from far-UV CD experiments as a function of ionic strength for CytR (red), PurR (blue) and LacR (green). (C) Predicted differences in coupling distances as a function of the residue index for the DNA-bound CytR (red) and apo-PurR (blue). The PDB numbering of CytR starting from 9 is used as a reference (PDB id: 2L8N/2LCV). (D) Multiple sequence alignment of LacI family members with known structures. Stars represent the highly conserved apolar positions across all sequences. Rectangular filled boxes highlight positions 29 and 48 in CytR with small apolar side-chains contrasting with the rest of the sequences that display bulkier side-chains. Blue rectangular boxes above the sequences show the helical boundaries.

Europe PMC Funders Author Manuscripts

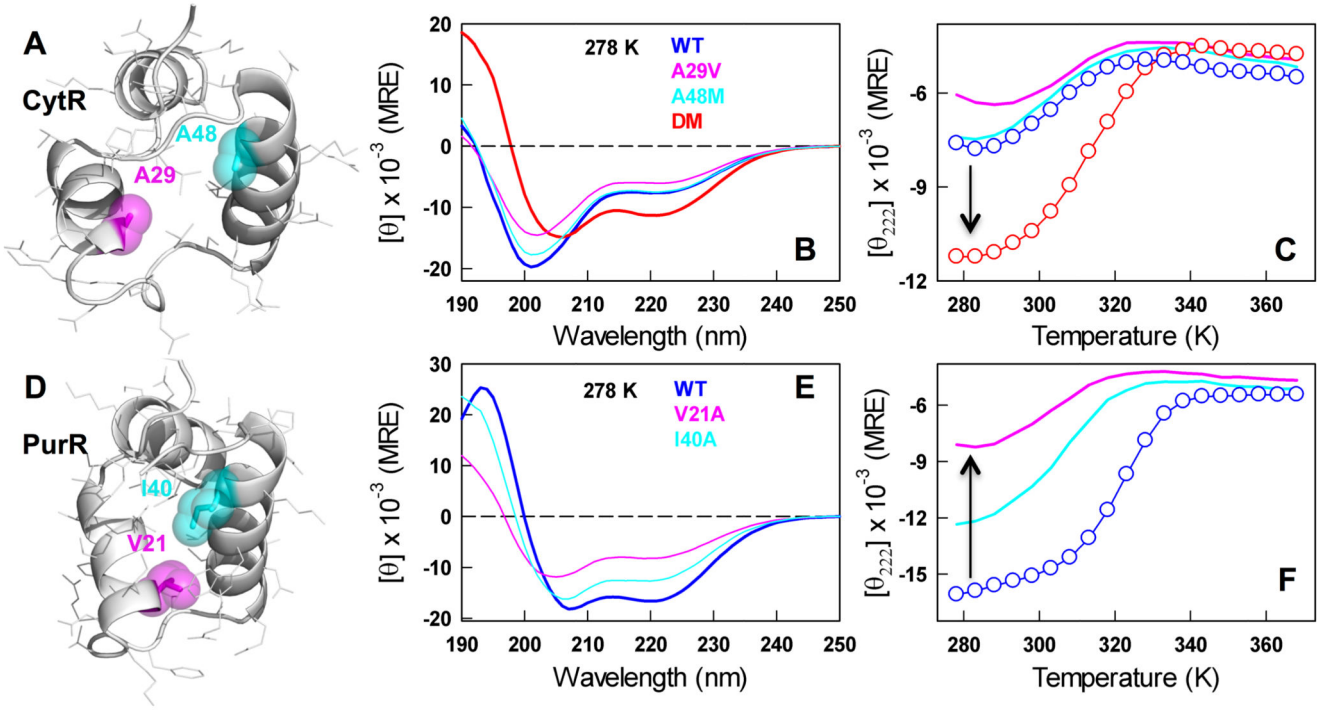

Figure 2.

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

Europe PMC Funders Author Manuscripts

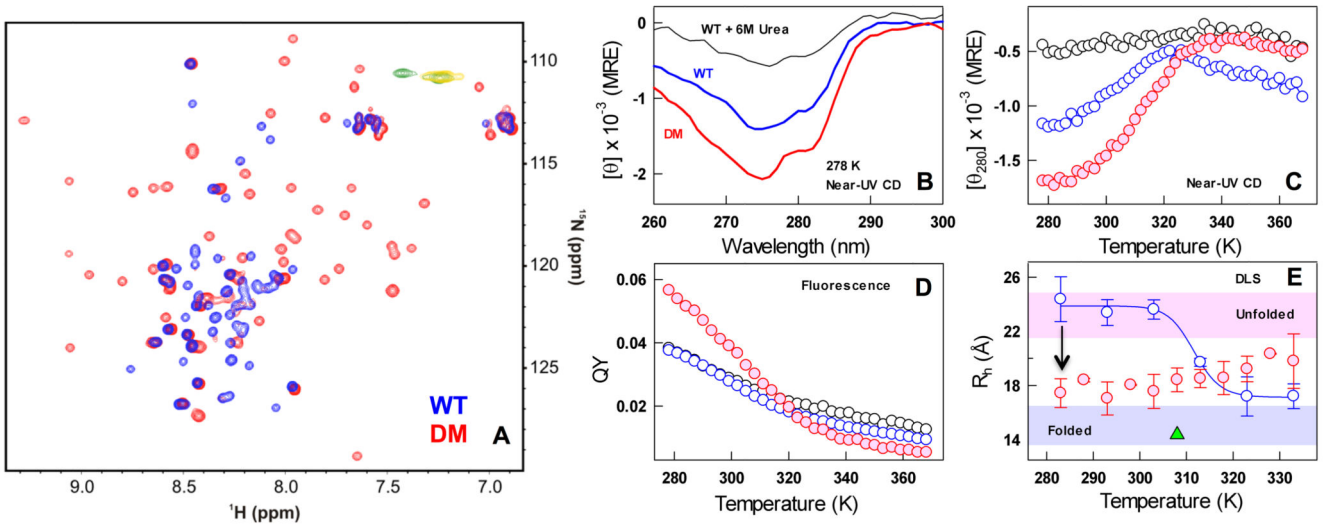

Figure 3.

The native ensemble of CytR DM is compact and structured. (A) Overlay of CytR WT (blue) and DM (red) HSQC spectra acquired at 288 K. Yellow and green contours in the overlay are the peaks with negative intensity for CytR WT and the DM, respectively. (B-C) Near-UV CD spectra (panel B) and the unfolding curves (panel C). (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 the CytR WT (blue) and 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 Rh of CytR from the PDB structure (residues 9-55).

Europe PMC Funders Author Manuscripts

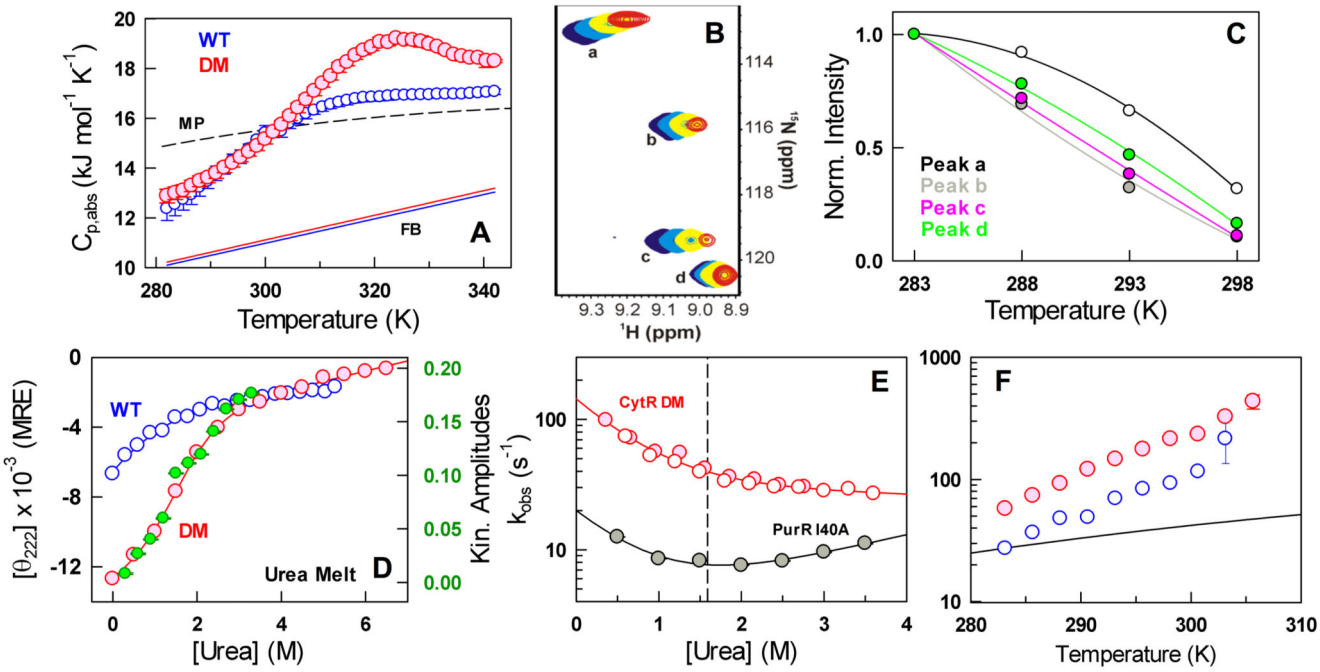

Figure 4.

Complex unfolding mechanism of the CytR DM. (A) Absolute heat capacity profiles of the WT and the double mutant (DM) together with the Freire folded baseline61 (FB) and Makhatadze-Privalov unfolded baseline (MP).62 Note the excess heat capacity peak for the DM around 325 K. (B) ^{1}H-^{15}N HSQC spectra of the CytR DM at temperatures of 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) The observed relaxation rate constants as a function of urea for the CytR DM (red) and PurR I40A (filled gray circles) at 285 K. The open red and filled red circles represent 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).

Europe PMC Funders Author Manuscripts

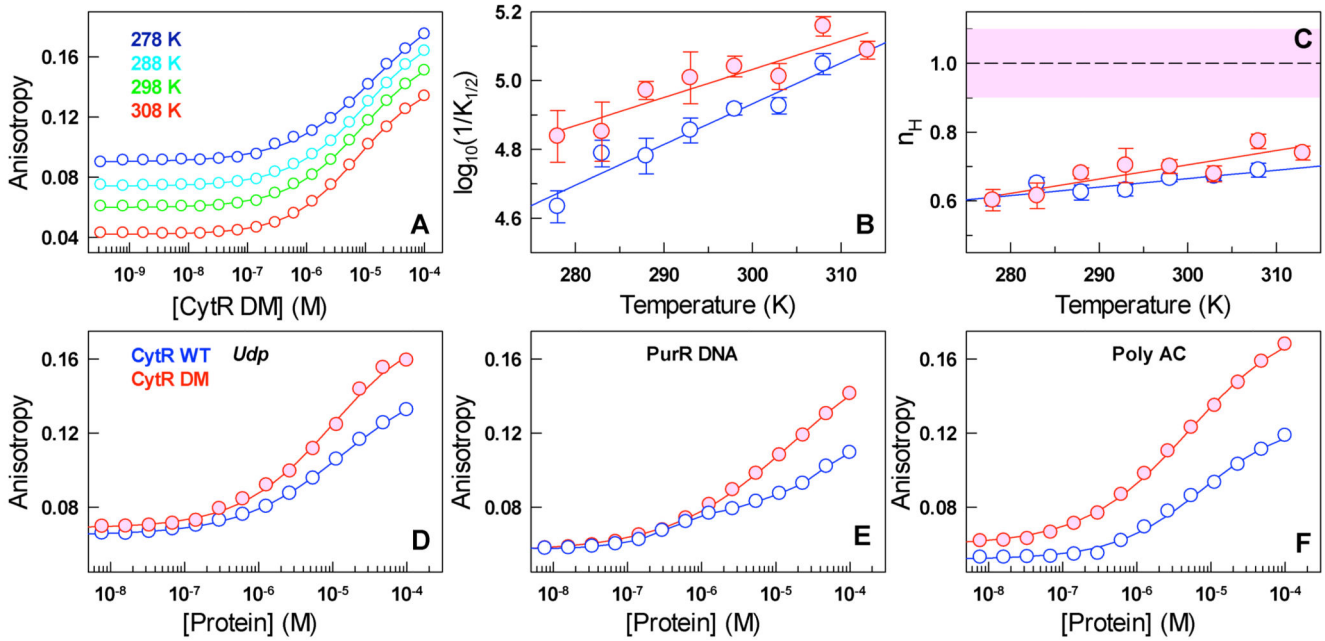

Figure 5.

(A) Anisotropy titration profiles of the CytR DM on Alexa-532 labeled *udp* half-site. The curves are fits to the Hill equation. (B, C) The apparent binding affinities and Hill coefficients as a function of temperature. Blue and red represent CytR WT and the 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 CytR WT (blue) and the DM (red) on to different DNA sequences at 293 K.

Europe PMC Funders Author Manuscripts