

Published in final edited form as:

*Biochemistry*. 2020 January 21; 59(2): 171–174. doi:10.1021/acs.biochem.9b00678.

## Controlling Structure and Dimensions of a Disordered Protein via Mutations

Sneha Munshi, Divya Rajendran, Samyuktha Ramesh, Sandhya Subramanian, Kabita Bhattacharjee, Meagha Ramana Kumar, Athi N. Naganathan\*

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

### Abstract

The dimensions of intrinsically disordered proteins (IDPs) are sensitive to small energetic-entropic differences between intramolecular and protein–solvent interactions. This is commonly observed on modulating solvent composition and temperature. However, the inherently heterogeneous conformational landscape of IDPs is also expected to be influenced by mutations that can (de)stabilize pockets of local and even global structure, native and non-native, and hence the average dimensions. Here, we show experimental evidence for the remarkably tunable landscape of IDPs by employing the DNA-binding domain of CytR, a high-sequence-complexity IDP, as a model system. CytR exhibits a range of structure and compactness upon introducing specific mutations that modulate microscopic terms, including main-chain entropy, hydrophobicity, and electrostatics. The degree of secondary structure, as monitored by far-UV circular dichroism (CD), is strongly correlated to average ensemble dimensions for 14 different mutants of CytR and is consistent with the Uversky–Fink relation. Our experiments highlight how average ensemble dimensions can be controlled via mutations even in the disordered regime, the prevalence of non-native interactions and provide testable controls for molecular simulations.

Mutational perturbation of protein structures reveals position-specific and context-dependent information on protein structure network, folding, function, allostery, and epistasis.<sup>1–4</sup> A majority of studies have been performed on ordered proteins with a compact hydrophobic core and a well-defined three-dimensional structure.<sup>5–8</sup> However, elementary considerations of the intraprotein interaction or contact network highlight that mutations should not only modulate the folded versus unfolded state equilibrium but also tune the relative population of intermediate and excited states and thus the overall dimensions of the native ensemble.<sup>4,9</sup> Even small volume fluctuations are expected as the packing interaction can be weakened via specific mutations without changing the overall structure.<sup>12</sup> Testing this on folded proteins is

This is an open access article published under a Creative Commons Attribution (CC-BY) License, which permits unrestricted use, distribution and reproduction in any medium, provided the author and source are cited.

\*Corresponding Author: athi@iitm.ac.in.

#### Accession Codes

CytR: P0ACN7

#### Notes

The authors declare no competing financial interest.

a huge challenge as small changes in dimensions are difficult to extract from conventional experiments.

Intrinsically disordered proteins (IDPs), on the other hand, provide a wonderful testing bed for probing mutation-driven structural perturbations.<sup>13</sup> High-sequence-complexity IDPs, i.e., those protein sequences that are disordered despite exhibiting little compositional bias, are more interesting as they are expected to populate pockets of local structure, rich in native or non-native interactions (deduced from the folded conformation), apart from fully disordered states. For example, this can be observed in studies of denatured states of folded proteins.<sup>14–16</sup> It should, therefore, be possible to tune the dimensions of natural IDPs through small perturbations of basic thermodynamic factors including backbone conformational entropy, hydrophobicity, and charge–charge interactions, that (de)stabilize pockets of structure.

The CytR DNA-binding domain (referred to as CytR hereon) is a high-sequence-complexity IDP that acquires a compact three-helix bundle structure in the presence of DNA (Figure 1A).<sup>17</sup> The conformational ensemble of CytR is sensitive to temperature,<sup>18</sup> salt,<sup>19</sup> and DNA,<sup>17,19</sup> all of which arise from a combination of destabilizing electrostatics (Figure 1B) and weak packing in the folded conformation.<sup>20</sup> In this work, we control the dimensions of the disordered CytR via a combination of folded structure and sequence-alignment-based expectations, and rational engineering. Though we observe nonintuitive effects, a likely manifestation of non-native interactions within the disordered ensemble, we find a strong correlation between secondary structure and the apparent Stokes radius ( $R_s$ ), a first such observation in IDPs.

We have shown earlier that the P33A mutation in CytR reduces the secondary structure content by reducing the population of an excited folded-like conformation.<sup>21</sup> Main-chain entropic considerations<sup>22</sup> suggest that, if an amino acid with a higher flexibility is introduced at the same position, the secondary structure content is expected to be even lower, arising from the larger entropic destabilization of residual structure. True to this, the thermal dependence of the P33G mutant's CD signal is flat (blue in Figure 2A) unlike the weak structural transitions observed in the WT or the P33A mutant (red and green in Figure 2A, respectively). In addition to the secondary structure content, the mean dimensions as measured by  $R_s$  (in a calibrated size-exclusion chromatography column<sup>20</sup>), shows a trend where P33G is more expanded than the WT (Figure 2B). Another interesting position is residue K35 located in the electrostatically frustrated binding site of the folded conformation with a number of positive charges around it (Figure 1B). Eliminating this positive charge through the K35Q mutant not only enhances the secondary structure content (though only marginally) but also reduces the dimensions of the ensemble (cyan in Figure 2).

Similarly, we have shown recently that the double mutant A29V/A48M (DM; that introduces two large hydrophobic substitutions in the protein core) promotes collapse of the disordered ensemble into a compact folded ensemble while simultaneously increasing the secondary-structure content (mutations identified via sequence comparisons; magenta in Figure 2 and Supporting Figure S1).<sup>20</sup> Building on the folded-like conformation of the DM, we engineered two other mutants, Quad (A29V/A48M/R43A/K46A; black in Figure 2) and Pent (A29V/A48M/R43A/K46A/R28A; gray in Figure 2), that further eliminate the residual

electrostatic frustration. These two mutants again promote structure and compaction while differentially affecting the secondary structure content. However, our attempts to rationally engineer structure or loss of structure were not always successful. For example, the A29V mutant enhances the structure-compactness only in the presence of the R28Q substitution and not otherwise (dark cyan in Figure 3A). Thus, collapse and structure acquisition in polymer chains is an emergent property arising from long-range many-body effects that can be engineered not only via hydrophobicity (as shown before for the DM) but also via a combination of hydrophobicity and electrostatics (A29V/R28Q mutant).

Interestingly, a number of single-point mutations (R28Q, A29V, D34S, R43N, R43E, K46A, and A48M) modulate ensemble dimensions despite being disordered (Figure 3A,B). The average dimensions of CytR can be mutationally varied from  $\sim 22$  Å in the P33G mutant to  $\sim 16$  Å in the Pent mutant (Figure 3B; Figure S2 for the elution profiles). If one excludes DM, Quad, and Pent that induce large structural changes, the maximal change in Stokes radius is  $\sim 3.5$  Å, which is still significant as it corresponds to  $\sim 40\%$  reduction in the volume of the protein chain compared to the fully unfolded P33G mutant.

Mutations that eliminate unfavorable interactions (R43N, R43E, K46A) as expected from the folded conformation still result in an expansion of the ensemble, suggestive of non-native interactions in the disordered state (Figure 3B). The only exceptions to this are the R28Q and K35Q mutations that promote slight compaction and the D34S mutation that promotes slight expansion, upon elimination of unfavorable and favorable interactions, respectively (Figure 3B). Despite these varied outcomes on mutations, we find that it is possible to collapse all such variations into a single plot of ensemble dimensions versus secondary-structure content (Figure 3C;  $r = 0.96$  and  $p < 10^{-9}$ ). In fact, the dependence of relative CD signals on the relative molecular volumes agrees reasonably well with the Uversky-Fink relation<sup>23</sup> (Figure S3). We also observe a similar correlation between dimensions and secondary-structure content at different salt concentrations for the CytR WT combining SEC and analytical ultracentrifugation data (Figure S4). These observations are strong evidence that the magnitude of the far-UV CD signal at 222 nm (in absolute units) for a given protein can be an intimate indicator of the molecular dimension even in the disordered regime,<sup>24</sup> particularly for helical protein domains, and if appropriately calibrated.

Is the correlation between Stokes radius and secondary structure a manifestation of (de)population of only folded-like conformations upon mutation? To answer this, we modeled the conformational behavior of CytR and its mutants via the semiquantitative Wako–Saitō–Muñoz–Eaton (WSME) statistical mechanical model that considers an ensemble of  $2^{47}$  microstates (47 being the length of the folded region of CytR)<sup>25,26</sup> with contributions from packing interactions, electrostatics, and implicit solvation.<sup>27,28</sup> Since the model is Go-like,<sup>29</sup> it captures the effect of only those mutations that display native-structure-derived expectations (Figure 2 and Table S1). The predicted unfolding curves are very similar to experimental observations, both in terms of the order of stability and structure (Figure 4A), validating the model energetics and assumptions. All of the generated free-energy profiles display multiple minima whose relative populations are modulated on mutation (Figure 4B,C). This is also consistent with both explicit and implicit-solvent simulations of WT CytR that highlight a flat conformational landscape with numerous

competing sub-states.<sup>18,30</sup> In fact, even the folded-like DM displays broad, weakly cooperative, and probe-dependent unfolding.<sup>20</sup> Taken together, our observations suggest that the apparent Stokes radius represents an effective ensemble average of numerous partially structured states.

The solvent sensitivity of IDP dimensions arises from small imbalances between intrachain and chain–solvent interactions.<sup>15,16,31–34</sup> In this work, we tune this relative balance by introducing mutations in a high-sequence-complexity IDP, and thus indirectly affecting the solvent quality (as seen by the polymer chain) without modulating solvent composition. Our results establish that the average dimensions of CytR can be modulated via mutations in a continuous manner through both rational engineering and non-native effects. CytR and its mutants can thus serve as excellent model systems for simulations, and to particularly validate force-fields and water models without resorting to extreme solvent conditions or temperature.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Funding

This work was supported by the Wellcome Trust/DBT India Alliance Intermediate Fellowship IA/I/15/1/501837 awarded to A.N.N.

## Abbreviations

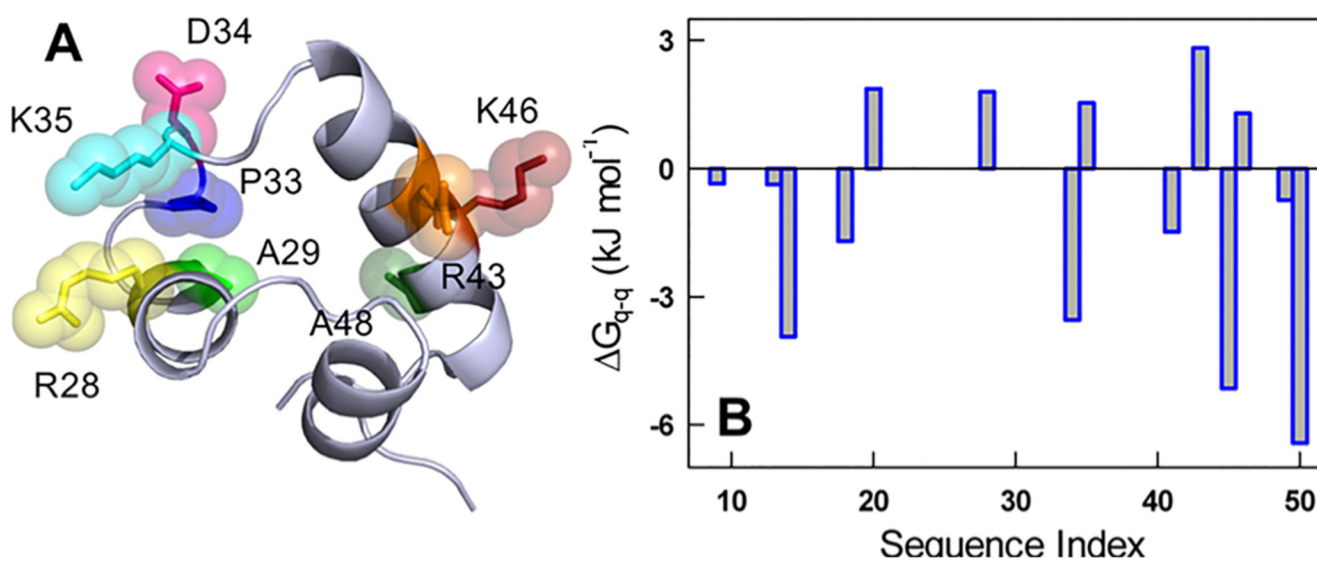
<b>CytR</b>	cytidine repressor
<b>WT</b>	wild-type
<b>CD</b>	circular dichroism
<b>MRE</b>	mean residue ellipticity
<b><math>R_s</math></b>	Stokes radius
<b>IDP</b>	intrinsically disordered protein
<b>WSME</b>	Wako–Saitô–Muñoz–Eaton

## References

- (1). Goldsmith M, Tawfik DS. Directed enzyme evolution: beyond the low-hanging fruit. *Curr Opin Struct Biol.* 2012; 22:406–412. [PubMed: 22579412]
- (2). Horovitz A, Fleisher RC, Mondal T. Double-mutant cycles: new directions and applications. *Curr Opin Struct Biol.* 2019; 58:10–17. [PubMed: 31029859]
- (3). Guarnera E, Berezovsky IN. On the perturbation nature of allostery: sites, mutations, and signal modulation. *Curr Opin Struct Biol.* 2019; 56:18–27. [PubMed: 30439587]
- (4). 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. [PubMed: 30268910]

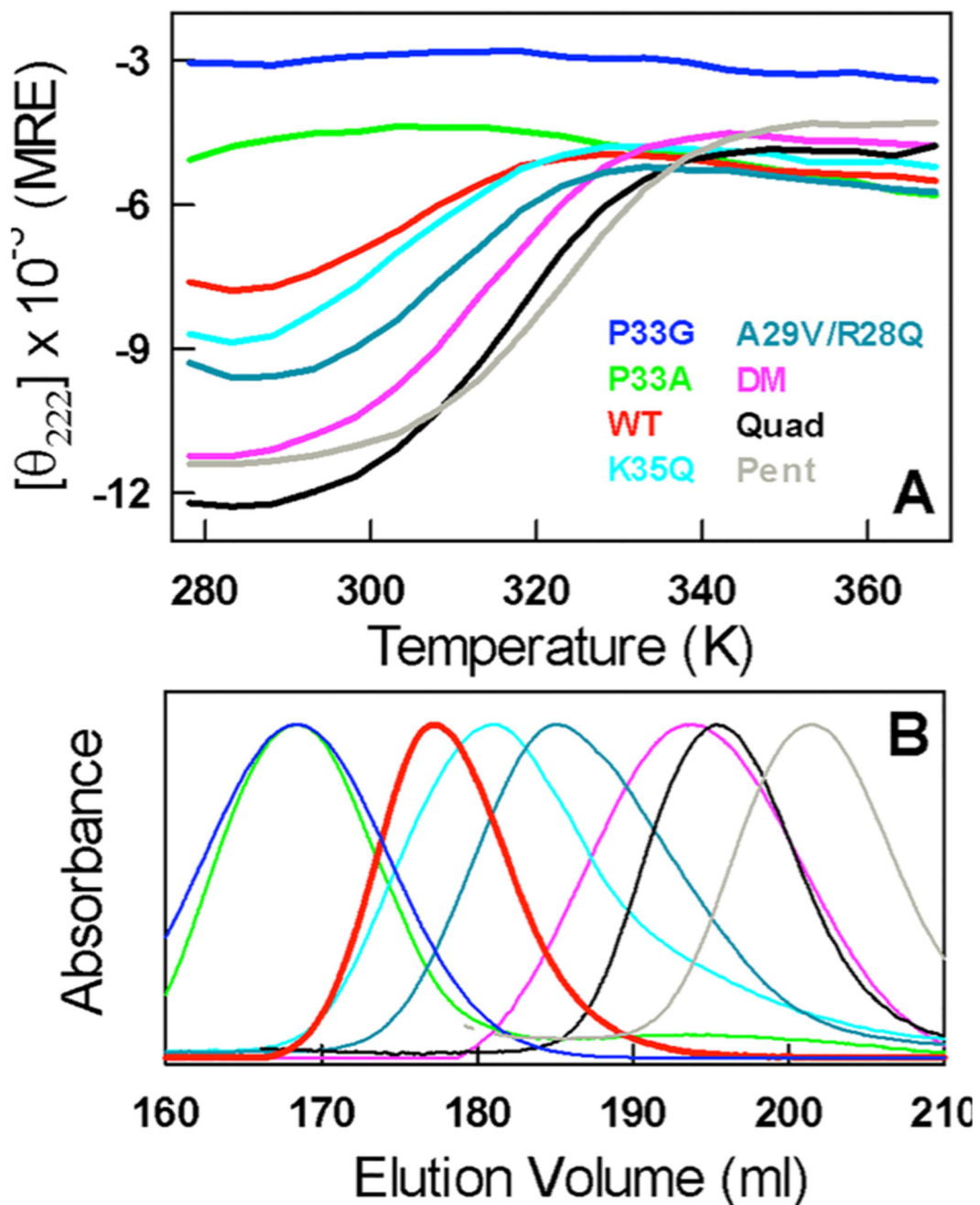
- (5). Xu J, Baase WA, Baldwin E, Matthews BW. The response of T4 lysozyme to large-to-small substitutions within the core and its relation to the hydrophobic effect. *Protein Sci.* 1998; 7:158–177. [PubMed: 9514271]
- (6). Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, Matthews BW. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science.* 1992; 255:178–183. [PubMed: 1553543]
- (7). 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 Struct Funct Genet.* 2013; 81:1069–1080. [PubMed: 23239146]
- (8). Daggett V, Fersht AR. Is there a unifying mechanism for protein folding? *Trends Biochem Sci.* 2003; 28:18–25. [PubMed: 12517448]
- (9). 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]
- (10). Tanford C, Kirkwood JG. Theory of Protein Titration Curves. I. General Equations for Impenetrable Spheres. *J Am Chem Soc.* 1957; 79:5333–5339.
- (11). Ibarra-Molero B, Loladze VV, Makhatadze GI, Sanchez-Ruiz JM. Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability. *Biochemistry.* 1999; 38:8138–8149. [PubMed: 10387059]
- (12). Law AB, Sapienza PJ, Zhang J, Zuo X, Petit CM. Native State Volume Fluctuations in Proteins as a Mechanism for Dynamic Allostery. *J Am Chem Soc.* 2017; 139:3599. [PubMed: 28094513]
- (13). Uversky VN. A decade and a half of protein intrinsic disorder: Biology still waits for physics. *Protein Sci.* 2013; 22:693–724. [PubMed: 23553817]
- (14). Bruun SW, Iesmantavicius V, Danielsson J, Poulsen FM. Cooperative formation of native-like tertiary contacts in the ensemble of unfolded states of a four-helix protein. *Proc Natl Acad Sci USA.* 2010; 107:13306–13311. [PubMed: 20624986]
- (15). 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]
- (16). Peran I, Holehouse AS, Carrico IS, Pappu RV, Bilsel O, Raleigh DP. Unfolded states under folding conditions accommodate sequence-specific conformational preferences with random coil-like dimensions. *Proc Natl Acad Sci USA.* 2019; 116:12301–12310. [PubMed: 31167941]
- (17). Moody CL, Tretyachenko-Ladokhina V, Laue TM, Senechal 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]
- (18). 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]
- (19). 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]
- (20). Munshi S, Subramanian S, Ramesh S, Golla H, Kalivarathan D, Kulkarni M, Campos LA, Sekhar A, Naganathan AN. Engineering Order and Cooperativity in a Disordered Protein. *Biochemistry.* 2019; 58:2389–2397. [PubMed: 31002232]
- (21). 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]
- (22). D'Aquino JA, Gomez J, Hilser VJ, Lee KH, Amzel LM, Freire E. The magnitude of the backbone conformational entropy change in protein folding. *Proteins Struct Funct Genet.* 1996; 25:143–156. [PubMed: 8811731]
- (23). Uversky VN, Fink AL. The chicken-egg scenario of protein folding revisited. *FEBS Lett.* 2002; 515:79–83. [PubMed: 11943199]

- (24). Narayan A, Bhattacharjee K, Naganathan AN. Thermally versus Chemically Denatured Protein States. *Biochemistry*. 2019; 58:2519–2523. [PubMed: 31083972]
- (25). Munoz V, Eaton WA. A simple model for calculating the kinetics of protein folding from three-dimensional structures. *Proc Natl Acad Sci USA*. 1999; 96:11311–11316. [PubMed: 10500173]
- (26). Wako H, Saito N. Statistical Mechanical Theory of Protein Conformation 0.2. Folding Pathway for Protein. *J Phys Soc Jpn*. 1978; 44:1939–1945.
- (27). Naganathan AN. Predictions from an Ising-like Statistical Mechanical Model on the Dynamic and Thermodynamic Effects of Protein Surface Electrostatics. *J Chem Theory Comput*. 2012; 8:4646–4656. [PubMed: 26605620]
- (28). Naganathan AN. A Rapid, Ensemble and Free Energy Based Method for Engineering Protein Stabilities. *J Phys Chem B*. 2013; 117:4956–4964. [PubMed: 23541220]
- (29). Taketomi H, Ueda Y, Go N. Studies on Protein Folding, Unfolding and Fluctuations by Computer-Simulation 0.1. Effect of Specific Amino-Acid Sequence Represented by Specific Inter-Unit Interactions. *Int J Pept Protein Res*. 1975; 7:445–459. [PubMed: 1201909]
- (30). 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]
- (31). 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]
- (32). Müller-Späh S, Soranno A, Hirschfeld V, Hofmann H, Ruegger S, Resssymond 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]
- (33). 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 USA*. 2014; 111:5213–5218. [PubMed: 24706910]
- (34). Holehouse AS, Pappu RV. Collapse Transitions of Proteins and the Interplay Among Backbone, Sidechain, and Solvent Interactions. *Annu Rev Biophys*. 2018; 47:19–39. [PubMed: 29345991]



**Figure 1.** (A) Structure of the folded conformation of the CytR DNA-binding domain highlighting residues that are mutated in the current study. (B) Residue-wise charge-charge interaction energy as calculated from the Tanford-Kirkwood algorithm.<sup>10,11</sup>

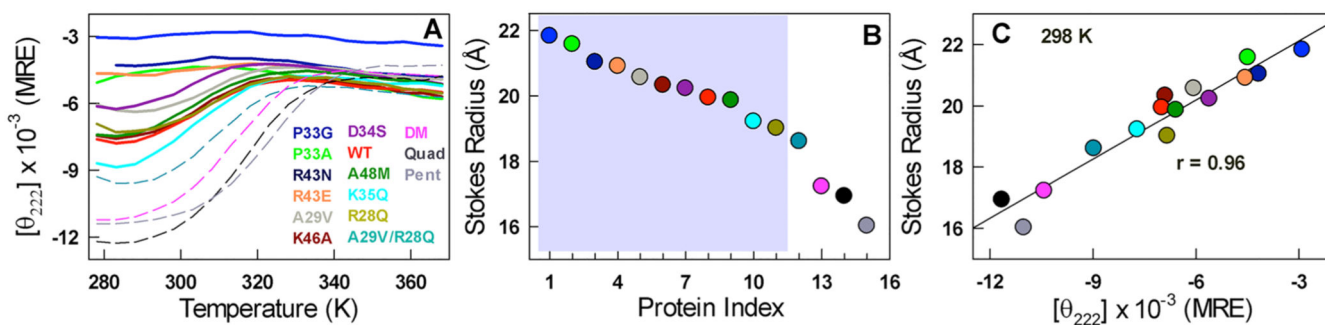




**Figure 2.**

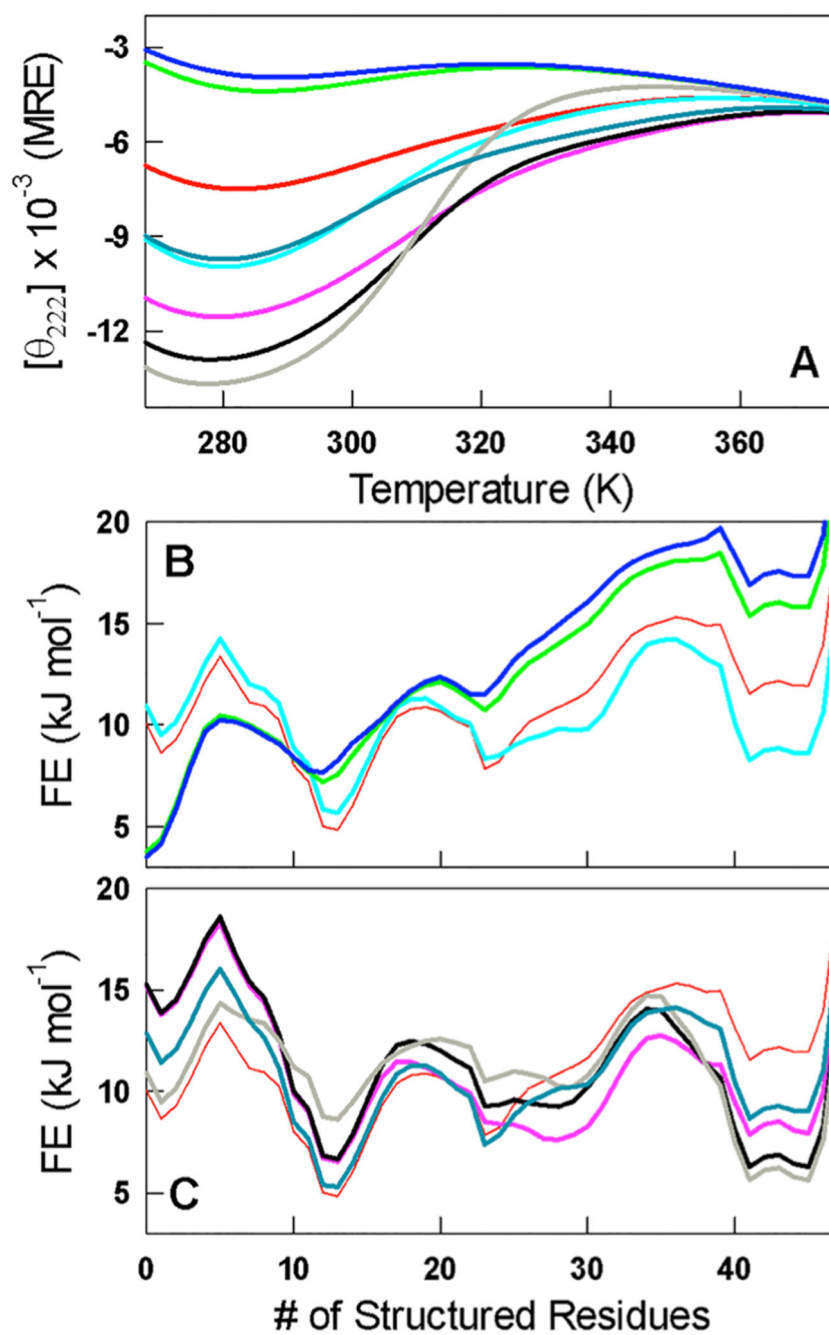
(A) Far-UV CD thermal unfolding curves of CytR and its mutants in 20 mM sodium phosphate buffer, pH 7.0, in mean residue ellipticity units of  $\text{deg. cm}^2 \text{dmol}^{-1}$ , as described in ref 20. (B) Size-exclusion chromatograms of the mutants in panel A at 298 K, 150 mM ionic strength ammonium acetate buffer at pH 8.0.<sup>20</sup> A high ionic strength is required to eliminate nonspecific interactions of proteins with the column matrix. The reported dimensions are thus lower estimates as CytR reduces its dimensions with salt.<sup>19</sup>





**Figure 3.**

(A) Far-UV CD thermal unfolding curves of WT CytR (red), single-point mutants (continuous curves), and multiple-point mutants (dashed curves). (B) Estimated apparent Stokes radius following the ordering and color code in panel A. The shaded region represents the disordered regime. (C) Correlation between apparent molecular dimensions (ordinate) and the secondary structure (abscissa).



**Figure 4.**

(A) Thermal unfolding curves predicted by the WSME model for the mutants of CytR following calibration of model parameters with the WT unfolding curve as described before.<sup>21,29</sup> The color coding is the same as in Figure 3A. (B and C) Predicted one-dimensional free-energy profiles of WT CytR (thin red), single-point mutants (panel B), and multiple-point mutants (panel C).