Denatured State Ensembles with the Same Radii of Gyration Can Form Significantly Different Long-Range Contacts

ARTICLE *in* BIOCHEMISTRY · NOVEMBER 2013

Impact Factor: 3.02 · DOI: 10.1021/bi4008337 · Source: PubMed

4 AUTHORS, INCLUDING:

Bowu Luan

8 PUBLICATIONS **55** CITATIONS

SEE PROFILE

Nicholas James Lyle

Washington University in St. Louis

19 PUBLICATIONS **216** CITATIONS

SEE PROFILE

Rohit V Pappu

Washington University in St. Louis

108 PUBLICATIONS **2,856** CITATIONS

SEE PROFILE

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Denatured State Ensembles with the Same Radii of Gyration Can Form Significantly Different Long-Range Contacts

Bowu Luan,[†] Nicholas Lyle,[‡] Rohit V. Pappu,[‡] and Daniel P. Raleigh*[,†,§]

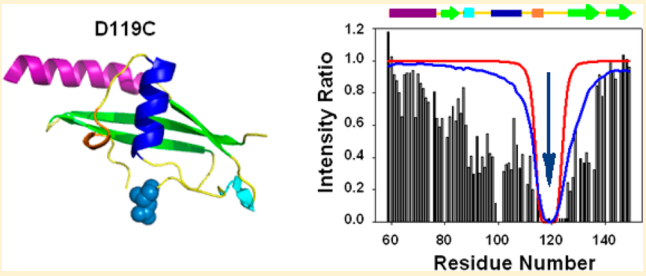
[†]Department of Chemistry, Stony Brook University, Stony Brook, New York 11794-3400, United States

[‡]Department of Biomedical Engineering, Washington University in St. Louis, One Brookings Drive, Campus Box 1097, St. Louis, Missouri 63130-4899, United States

[§]Graduate Program in Biochemistry and Structural Biology and Graduate Program in Biophysics, Stony Brook University, Stony Brook, New York 11794, United States

Supporting Information

ABSTRACT: Defining the structural, dynamic, and energetic properties of the unfolded state of proteins is critical for an in-depth understanding of protein folding, protein thermodynamics, and protein aggregation. Here we analyze long-range contacts and compactness in two apparently fully unfolded ensembles of the same protein: the acid unfolded state of the C-terminal domain of ribosomal protein L9 in the absence of high concentrations of urea as well as the urea unfolded state at low pH. Small angle X-ray scattering reveals that the two states are expanded with values of R_g differing by <7%. Paramagnetic relaxation enhancement (PRE) nuclear magnetic resonance studies, however, reveal that the acid unfolded state samples conformations that facilitate contacts between residues that are distant in sequence while the urea unfolded state ensemble does not. The experimental PRE profiles for the acid unfolded state differ significantly from these predicted using an excluded volume limit ensemble, but these long-range contacts are largely eliminated by the addition of 8 M urea. The work shows that expanded unfolded states can sample very different distributions of long-range contacts yet still have similar radii of gyration. The implications for protein folding and for the characterization of unfolded states are discussed.



Quantitative characterization of denatured state ensembles (DSEs) of proteins, also termed the unfolded or denatured state, is important for understanding the mechanism of protein folding. The DSE is the starting point of protein folding, the thermodynamic reference state for protein stability, and it can be targeted by rational protein design.[1−7] Studies of DSEs can also reveal factors that impact protein misfolding and modulate the tendency for protein aggregation *in vitro* and *in vivo* and amyloid formation.[8−12] The exploration of the mechanisms and biological function of intrinsically disordered proteins (IDPs) largely depends on the characterization of the properties of unstructured and partially structured states and therefore has much in common with studies of the DSE.[13,14]

The properties of the DSE can vary considerably depending upon the conditions used to populate it. Under near-native conditions, the DSE can be compact with significant residual structure, while more expanded and less structured DSEs are usually populated under strongly denatured conditions. Small angle X-ray scattering (SAXS) is frequently used to study the overall compactness of the DSEs and provides the radius of gyration (R_g) and in favorable cases more information.[15−20] DSEs that have the same value of R_g are often assumed to be similarly unfolded.[16,17,21]

Under strongly denaturing conditions, the DSE expands to make favorable interactions with the solvent, and the R_g of proteins without disulfide cross-links follows a nontrivial power law relationship, which scales with the number of amino acids in the peptide chain, N, as $N^{0.59}$.[16,17,22] Similar scaling behavior is observed for polymers modeled as self-avoiding random walks.[23] Observation of an R_g value consistent with this scaling is often taken to mean a protein is fully unfolded; however, this scaling does not preclude the possibility of detectable, low-likelihood native and non-native contacts within expanded DSEs, even under strongly denaturing conditions.[2−6,8,13,24−36] However, it is unclear if different DSEs generated for the same protein under different conditions, all with similar R_g values, will exhibit similar patterns of and likelihoods of native and non-native contacts. This issue is important because such contacts might contribute directly to folding and might influence the tendency to aggregate. It is also important because it potentially highlights the need to go beyond measurements of R_g alone as an adjudicator of the degree of unfoldedness and as a descriptor of unfolded states.

Here we examine the 92-residue C-terminal domain of ribosomal protein L9 in the acid-induced DSE and in the low-pH urea-induced DSE to determine if the conformational

Received: June 27, 2013
Revised: November 15, 2013
Published: November 26, 2013

ensembles are different between different DSEs that have similar R_g values. The domain, denoted as CTL9, has a mixed parallel, antiparallel β-sheet, two loops with partial α-helix 3_{10} turns, and two α-helices (Figure 1). CTL9 can be denatured by

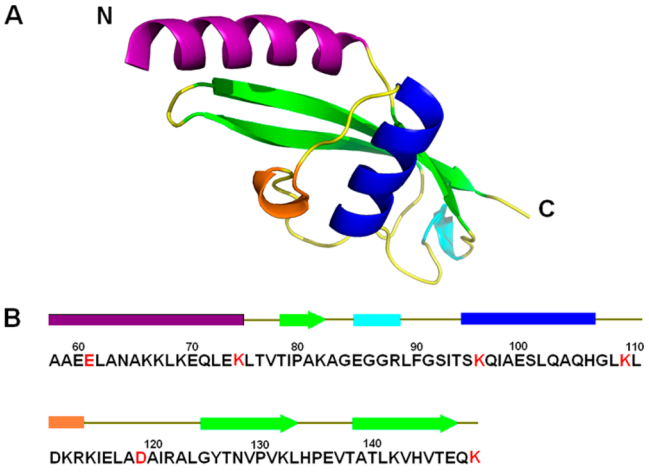


Figure 1. (A) Structure of the C-terminal domain of ribosomal protein L9 (CTL9), corresponding to residues 58−149 of L9, shown as a ribbon diagram (Protein Data Bank entry 1DIV). The N- and C-termini are labeled. (B) Primary sequence of CTL9. The sites used to attach the spin-labels are colored red. A schematic representation of the secondary structure elements of CTL9, using the same color coding that was employed for the ribbon diagram, is shown above the sequence. Helices are represented as rectangles and β-sheets as arrows.

lowering the pH, or by adding denaturant, and folds in a two-state fashion.[37−41] Previous studies have used nuclear magnetic resonance (NMR) chemical shift analysis to show that the two ensembles differ in their residual secondary structure but have similar hydrodynamic radii, R_h.[40,41] In particular, more helical structure was detected in the acid unfolded state.[40] R_h is a useful parameter for probing compactness, but it is determined indirectly from NMR diffusion measurements, involves comparison to an added internal standard, and requires assumptions about the hydrodynamic properties of the ensemble. R_g measurements offer an alternative probe of the compactness of the DSE, but involve assumptions different from those of the diffusion measurements and are, in some sense, a more direct measure of the ensemble. However, the NMR studies and hydrodynamic measurements were unable to define the nature of any long-range contacts. This information is required for a full description of the DSE, and one might expect that states that are equally expanded should have a similar level of long-range contacts. We show that this is not the case, and we also directly demonstrate that states that have very similar values of R_g, which furthermore are consistent with the scaling expected for expanded highly unfolded states, can differ dramatically in the extent of detectable long-range contacts.

NMR paramagnetic relaxation enhancement (PRE) measurements were used to detect long-range contacts, and the R_g was measured using SAXS for two different DSEs of CTL9. The two ensembles have very similar R_g distributions, but exhibit significantly different patterns of long-range contacts. The urea unfolded DSE contains few long-range contacts and is described well by an excluded volume model, while the acid unfolded DSE displays extensive long-range interactions.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification. Wild-type ^{15}N-labeled CTL9 and its cysteine mutants were overexpressed in *Escherichia coli* strain BL21 cells and purified as described previously.[37] The identity of the protein was confirmed by matrix-assisted laser desorption ionization time-of-flight mass spectroscopy and analytical high-performance liquid chromatography (HPLC).

Small Angle X-ray Scattering (SAXS) Experiments and Data Analysis. Experiments were performed on the acid-induced DSE and on the urea-induced DSE. Samples of wild-type CTL9 were prepared using the same buffers used for the NMR PRE experiments. Scattering experiments were performed at beamline X9 at National Synchrotron Light Source I (Brookhaven National Laboratory, Upton, NY). Samples were injected continuously into a 1 mm capillary during the measurement at a rate of 0.67 μL/s to prevent radiation damage. The exposure time for each measurement was 30 s. Each sample was measured three times and then averaged before data analysis. pyXS (http://www.bnl.gov/ps/x9/software/pyXS.asp) was used for buffer subtraction, and the data were then processed using standard procedures via PRIMUS.[42] The scattering patterns were analyzed using the ensemble optimization method (EOM) to determine the size distribution for the DSE in urea and in acid. EOM creates a pool of structures based on the protein primary sequence, and then an ensemble of protein structures that best fits the experimental data is selected.[43]

Circular Dichroism (CD) Spectroscopy. Thermal unfolding experiments with wild-type CTL9 and the six cysteine mutants were performed with an Applied Photophysics Chirascan CD instrument, over the temperature range of 4−92 °C with an interval of 2 °C. The absorbance of the spin-label prevents CD-monitored unfolding studies of the derivatized mutants. The protein concentration was 16−19 μM, in 10 mM MOPS [3-(*N*-morpholino)propanesulfonic acid] and 150 mM NaCl (pH 7.5). The signal at 222 nm was monitored. All thermal unfolding curves were analyzed by nonlinear least-squares fitting using SigmaPlot.

A pH titration curve was measured for wild-type CTL9 at 25 °C, in 10 mM sodium acetate and 150 mM NaCl, over the pH range of 1.8−12.0. Urea unfolding data were collected at 25 °C on wild-type CTL9, using an AVIV Instruments model 202SF CD instrument. The native protein sample was titrated with urea in 0.25 M increments until the final urea concentration reached 10 M. The buffer was the same as that used for the thermal unfolding experiments, and the urea concentration was determined by measuring the refractive index.

Paramagnetic Relaxation Enhancement (PRE) Experiments. Single-cysteine mutants of CTL9 were dissolved in 600 μL of NMR buffer, containing 10% D_2O. An MTSL [(1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-yl)methyl methanesulfonate] solution was added, and then the pH was adjusted to 2.0. After being incubated at room temperature for 12 h, the sample was loaded onto a Sephadex G25 column to remove the excess MTSL. The completeness of labeling was assessed using HPLC and by the intensity of the Cys NMR peaks. The Cys resonance is bleached in the labeled sample because of the attachment of MTSL, but is not bleached in the unlabeled sample. Diamagnetic samples were prepared by adding TCEP [tris(2-carboxyethyl)phosphine]. Urea-denatured samples were prepared by adding urea to the samples immediately after the

dx.doi.org/10.1021/bi4008337 | *Biochemistry* 2014, 53, 39−47

desalting process. The final urea concentration was 8 M, determined by measuring the refractive index. The final protein concentration was 250 μM.

$^{1}H-^{15}N$ correlated heteronuclear single-coherence (HSQC) experiments were performed on both samples for paramagnetic and diamagnetic forms. The spectra were recorded at 25 °C, using 1024 × 256 complex points with eight scans per increment. The spectral widths were 6009.6 and 1517.8 Hz for the ^{1}H and ^{15}N dimensions, respectively. Three-dimensional TOCSY-HSQC spectra were recorded on the ^{15}N-labeled diamagnetic samples to confirm peak assignments. The mixing time was 75 ms, and the data matrix was 1024 (direct ^{1}H dimension) × 128 (indirect ^{1}H dimension) × 128 (^{15}N dimension). Spectral widths for the direct ^{1}H dimension, the indirect ^{1}H dimension, and the ^{15}N dimension were 6009.6, 6009.6, and 1500.0 Hz, respectively.

NMR Data Processing and Determination of PREs. All spectra were processed using NMRPipe[44] and visulized via NMRView.[45] The ratio of the intensity for a particular residue was calculated as I_{para}/I_{dia}, where I_{para} is the intensity of the paramagnetic form, with spin-label MTSL; and I_{dia} is the peak intensity for the diamagnetic sample, with the spin-label MTSL reduced. Lower intensity ratios correspond to larger PRE effects and indicate an interaction with the spin-label.

Two different models were used to calculate baseline PRE effects for a highly unfolded chain: a Gaussian distribution model and an excluded volume (EV) model.

Generation of the Gaussian Distribution Model. In this model, a Gaussian distribution of root-mean-square end-to-end distances is used to describe the distances between each residue and the site of the spin-label:[33,35]

$$\langle r^2 \rangle = nl^2 \left[\frac{1+\alpha}{1-\alpha} - \frac{2\alpha(1-\alpha^n)}{n(1-\alpha)^2} \right] \quad (1)$$

where n is the number of residues between residue i and the site of the spin-label, r is the end-to-end distance between a residue and the site of the spin-label, l (3.8 Å) is the link length of the chain, and α is the cosine of the bond angle supplements for the freely rotating chain model, which was taken to be 0.8, based on experimentally determined estimates of the statistical segment lengths in poly-L-alanine. R_{2P} is the contribution of the paramagnetic relaxation enhancement to the transverse relaxation rate and was calculated using

$$R_{2P} = \frac{K}{r^6}\left(4\tau_c + \frac{3\tau_c}{1+\omega_H{}^2\tau_c{}^2}\right) \quad (2)$$

where r is the distance between a given residue and the site of the spin-label, K is a constant equal to 1.23×10^{-32} cm^6 s^{-2}, τ_c is the effective correlation time (3.8 ns), and ω_H is the proton Larmor frequency. The peak intensity ratios between the paramagnetic and diamagnetic forms were calculated using

$$\frac{I_P}{I_D} = \frac{R_{2D}\exp(-R_{2P}t)}{R_{2D}+R_{2P}} \quad (3)$$

where R_{2D} is the transverse relaxation rate of the backbone amide protons in the diamagnetic form. The average value was determined to be 13.5 s^{-1} using a one-dimensional NMR, and t is set to 12 ms, equal to the total duration of the INEPT delays in the HSQC pulse sequence.

Genration of the EV Ensemble. We used the CAMPARI software package for Metropolis Monte Carlo (MC) simulations based on the ABSINTH implicit solvation model and underlying force field paradigm,[46] and parameters from the OPLSS-AA/L molecular mechanics force field, specifically parameters in abs3.2_opls.prm.[47] The internal degrees of freedom included the backbone φ, ψ, and ω and side chain χ dihedral angles. Details regarding the MC moveset were detailed by Meng et al.[48] Results were averaged over 10 independent simulations. Each simulation underwent 4×10^7 MC moves, not including an equilibration phase of 1×10^5 moves. The radius of gyration over CTL9 was accumulated every 1×10^4 MC moves.

RESULTS

The folding of CTL9 has been characterized previously in terms of its kinetics and thermodynamics, and the cold denaturation of a destabilized mutant has been probed.[37,39,49−52] The stability and folding rate of CTL9 are strongly dependent on pH, due in part to the three His residues in the protein, and the domain can be unfolded by a decrease in pH, as well as by addition of a denaturant. CD-monitored titration curves indicate the transformation from the native folded state to the DSE is well fit by a two-state model (Figure S1 of the Supporting Information). The acid unfolding transition is complete by pH 2.8 in the absence of urea, and the urea unfolding transition is complete by 5.5 M urea at pH 5.6 as judged by CD. The degree of secondary structure in the two DSEs is difficult to determine from CD because the strong absorbance of urea limits the accessible wavelength range. In addition, the CD signal from short α-helices can differ from standard spectra.[53] However, previous studies used NMR to probe residual secondary structure in both of these states. $C_\alpha{}^1H$, $^{13}C_\omega$ and $^{13}C_\beta$ chemical shifts were analyzed and revealed that in the acid unfolded state there is a modest, but non-zero, propensity to preferably populate the helical region of the Ramachandran plot for those residues that are helical in the native state. In contrast, this propensity to form secondary structure was significantly reduced in the low-pH urea unfolded state.[40]

SAXS Experiments Show That the Urea-Induced and Acid-Induced DSEs Are Expanded. The R_g measured for the native state of CTL9 is 14.5 ± 0.3 Å.[51] The values of R_g for the acid and urea DSEs, determined from the Guinier plot, are 30.8 ± 1.6 and 32.9 ± 1.5 Å, respectively. The difference between the two R_g values is 6% and is statistically insignificant given the intrinsic uncertainty associated with each value. The value of R_g predicted for a fully unfolded 92-residue polypeptide based on empirical scaling relationships is 28.9 ± 4.6 Å.[16] By this criterion, both ensembles are classified as expanded unfolded states. We also fit the scattering patterns using the ensemble optimization method (EOM).[43] The EOM algorithm builds a pool of structures based on the primary sequence of the target protein, and a series of theoretical scattering curves are generated. A combination of the generated structures is used to generate an ensemble of representative structures that reproduce the experimental data. The average R_g of the DSE can be deduced from the EOM fitting of the experimental data. The average R_g values estimated using this method are within 8% of each other, 30.8 and 33.5 Å for the acid- and urea-induced DSE, respectively (Figure 2), and the widths of the distribution are 8.0 and 9.0 Å, respectively.

PRE Studies Reveal Long-Range Contacts in the Acid-Induced DSE but Not in the Urea-Induced DSE. We used paramagnetic relaxation enhancement (PRE) experiments to

dx.doi.org/10.1021/bi4008337 | Biochemistry 2014, 53, 39−47

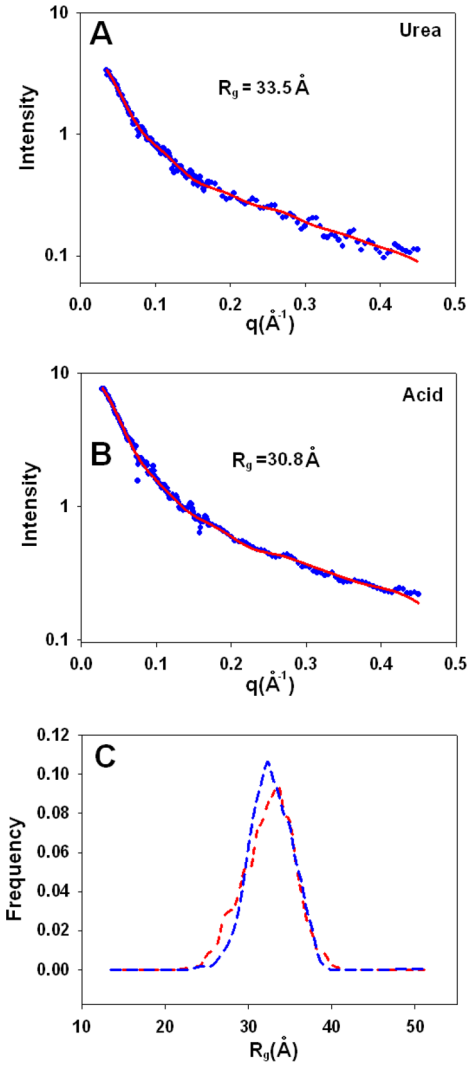


Figure 2. Acid and urea unfolded states of CTL9 are expanded. (A) SAXS scattering curve from a sample of CTL9 in 8 M urea at pH 2.5. (B) SAXS scattering curve from a sample of CTL9 at pH 2.0. The experimental data are shown as blue dots, and the fit using the Ensemble Opitimiztion Method (EOM) is shown as a red line. (C) Distribution of R_g for the DSE in urea (red) and acid (blue), calculated using the EOM. The average R_g values are 33.5 and 30.8 Å, respectively.

obtain information about long-range contacts. PRE_{NMR} is frequently applied in probing long-range interactions in the DSE and is sensitive to distances of up to 20 Å.[33,35,48,54−69] Spin-labels were attached to six sites using single-cysteine mutants. The sites are all surface-exposed in the native state (Figure 1): E61 near the N-terminus of the first α-helix, K74 at the C-terminus of the same α-helix, K96 near the N-terminus of the second α-helix, K109 and D119 in the loop region between the second α-helix and the second β-strand, and K149, which lies at C-terminus. The mutants are denoted E61C-CTL9, K74C-CTL9, K96C-CTL9, K109C-CTL9, D119C-CTL9, and K149C-CTL9, respectively. The CD spectra of the mutants are similar to that of wild-type CTL9 (Figure S2 of the Supporting Information). Comparison of the thermodynamic properties determined from the thermal unfolding shows that the cysteine mutants have similar values of T_m and $\Delta H°$ with respect to those of wild-type CTL9 (Table S1 and Figure S3 of the Supporting Information), although the T_m of D119C-CTL9 is somewhat depressed compared to that of the wild type. The peaks in the $^1H-^{15}N$ HSQC spectra of the acid- and urea-induced DSE of the mutants do not shift their positions, relative to those of the wild type, aside from the mutated residue (Figures S4−S10 of the Supporting Information). Thus, all of the available data suggest that the mutations do not alter the properties of the native state or the DSE.

$^1H-^{15}N$ HSQC spectra were collected for the six mutants with spin-labels (paramagnetic state) and without spin-labels (diamagnetic state), in 8 M urea. The cross-peak intensity ratios (I_{para}/I_{dia}) were calculated and plotted against the corresponding residue number (Figure 3). The expected PRE profile for a random coil model was generated using two different models. We used a Gaussian model, in which there is a Gaussian distribution of distances between the spin-label site and each residue of the protein,[33,35] and an excluded volume (EV) model. The Gaussian chain model has been widely used to benchmark PRE studies of unfolded states because of its simplicity. However, the model lacks any detail. Quantitative descriptions of chain statistics for polymers in good solvents rely on the so-called EV limit as an important reference state, and this is true for denatured state ensembles, as well.[70−76] The EV model is generated using an all atom representation of the chain with only excluded volume interactions, which also takes into account the size and flexibility of the side chain-linked spin-label. The EV ensemble corresponding to the wild-type sequence was used.[72] EV ensembles were generated for CTL9 using atomistic descriptions of proteins and all nonbonded interactions except steric repulsions were ignored.

The observed PRE effects for all the mutants in 8 M urea correlate very well with the random coil models, especially the EV model, indicating the urea-induced DSE of CTL9 in 8 M urea at pH 2.5 behaves like a highly unfolded chain.

The same strategy was used to study the acid-induced DSE. However, in this case, clear deviations from both models were observed for four of the six spin-labeled mutants (Figure 3). The two exceptions are the labels near the N- and C-termini. The other four PRE profiles display significant differences between the experimental and calculated profiles (Figure 3). Significant PRE effects, defined here as a value of $\leq$0.5 for the I_{para}/I_{dia} ratio, are detected for sites as many as 30 residues from the spin-label. Even longer-range effects are observed for some of the spin-labels, including residues 74 and 109.

DISCUSSION

The acid- and urea-induced DSEs of CTL9 differ significantly in the pattern of long-range contacts. The differences in the intensity ratios between the PRE profiles for the two DSEs are compared in Figure 4 as difference plots, where positive values indicate stronger PRE effects in the acid-induced DSE of CTL9. For the N- and C-termini, residues 61 and 149, respectively, the difference in the intensity ratios is close to zero, indicating the there are no obvious changes in the acid- and urea-induced PRE profiles. Positions 74, 96, 109, and 119 display positive values and clearly illustrate the deviations between the two DSEs. The analysis of the acid-induced DSE of CTL9 clearly shows that long-range contacts can form in a highly expanded DSE.

There are also differences in the local secondary structure propensities of the urea and acid unfolded DSEs of CTL9 despite their similar R_g values. NMR studies have shown that the urea-induced DSE of CTL9 contains very little residual secondary structure, while in the acid unfolded DSE, there are

dx.doi.org/10.1021/bi4008337 | Biochemistry 2014, 53, 39−47

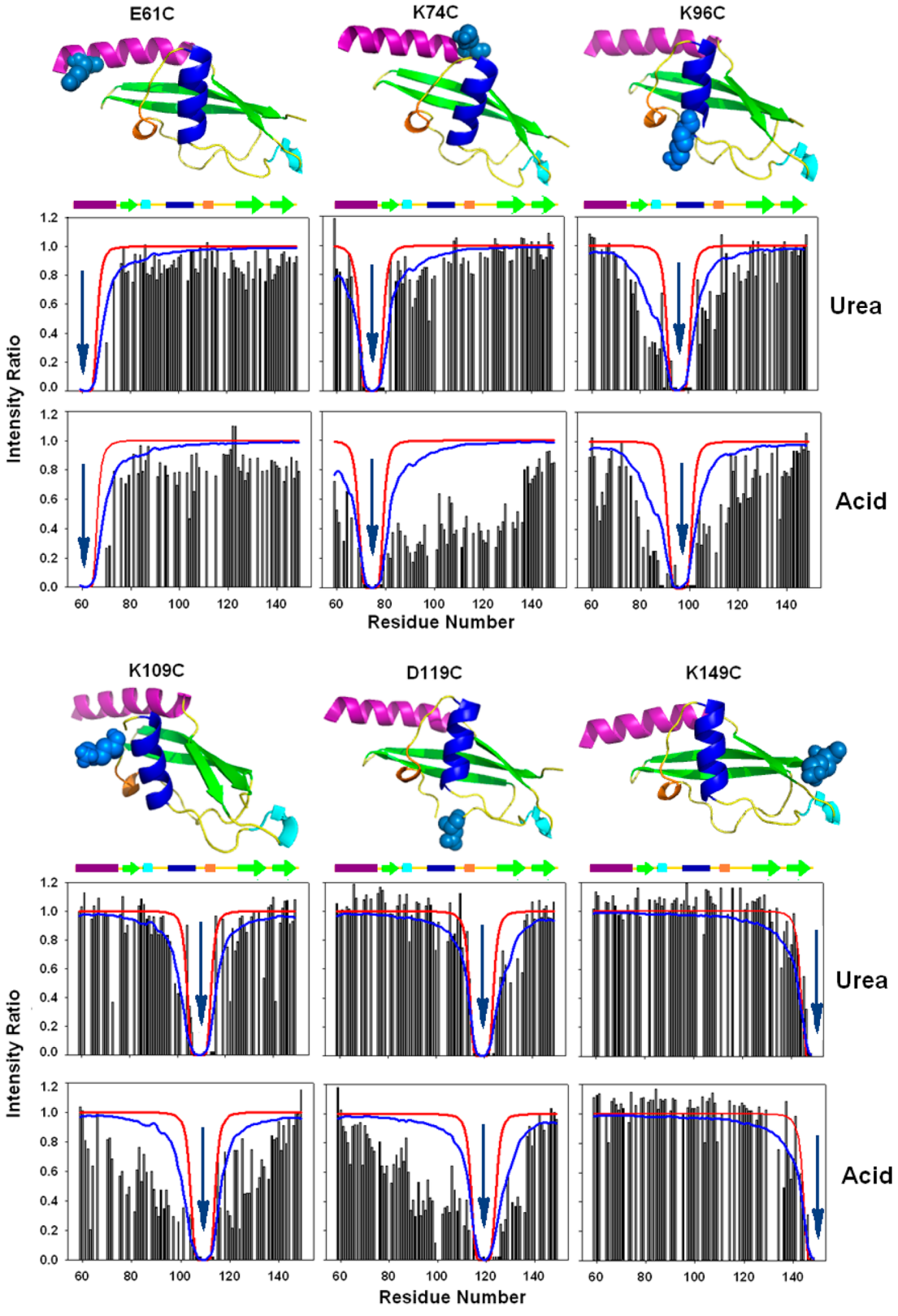


Figure 3. Paramagnetic relaxation enhancement experiments show that there are significant deviations from random coil behavior for the acid unfolded state, but not for the urea unfolded state. The histograms display the intensity ratio of the $^{1}H-^{15}N$ cross-peaks in the HSQC spectra. The dark blue arrow (↓) indicates the site of attachment of the spin-label. The solid red curve represents the values predicted by the Gaussian model and the solid blue curve the values predicted by the excluded volume model. Ribbon diagrams of CTL9 are shown for each mutant, indicating the site of attachment of the spin-label. Experiments were conducted in 8 M urea at pH 2.5 and at pH 2.0 in the absence of urea.

two regions that have a modest propensity to preferentially populate helical ϕ and ψ angles that are helical in the native state.[40] However, there is no obvious significant correlation between the regions of the polypeptide chain that exhibit an

dx.doi.org/10.1021/bi4008337 | *Biochemistry* 2014, 53, 39−47

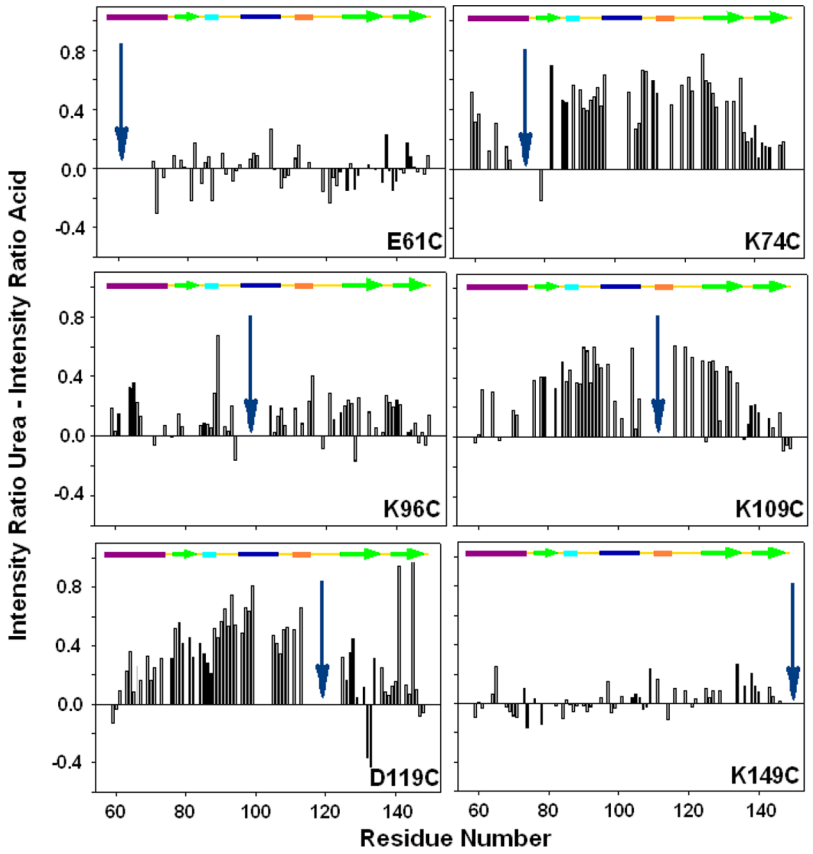


Figure 4. Plots of the differences between the PRE effects in the urea-induced DSE and the acid-induced DSE. Data are plotted as the PRE ratio in the urea-induced DSE minus the PRE ratio in the acid-induced DSE. The dark blue arrow (↓) indicates the site of attachment of the spin-label. Positive values indicate larger PRE effects in the acid unfolded state. Data were collected at 25 °C, with the pH adjusted to 2.0, by adding HCl for the DSE in acid and in 8 M urea at pH 2.5 for the DSE in urea.

increased propensity for helical ϕ and ψ angles and strong PRE effects. This study, together with the previous work, shows that DSEs with similar R_g values can differ significantly in their patterns of long-range contacts as well as their propensities to form local structure.

The data obtained for CTL9 further demonstrate that R_g, while a very useful measure of unfolded state dimensions, should not be used as the sole criterion to judge if a protein is "fully unfolded", or if longer-range contacts are absent. This may be relevant for SAXS and FRET studies of DSEs. There are examples in which SAXS indicates a highly unfolded state but FRET reveals apparent long-range contacts.[77] Our analysis together with other studies suggests that these observations can be comparable.[48,73,74,78]

The data collected on the urea-induced DSE of CTL9 show that it lacks detectable long-range contacts, but this should not be interpreted to mean that all expanded states that lack secondary structure are devoid of long-range contacts. The N-terminal domain of L9 provides a counter example.[48] The urea-induced DSE of that domain transiently populates long- and medium-range contacts; however, the distribution of internal distances is still consistent with $N^{0.59}$ scaling, and the value of R_g is consistent with an expanded DSE. Reduced hen egg white lysozyme offers another example. The urea-induced DSE appears to contain transient clusters of hydrophobic residues, as judged by ^{15}N NMR relaxation measurements, and they can be modulated by mutation.[34] In contrast, ^{15}N NMR relaxation measurements of the urea-induced DSE of other proteins suggest that these sorts of contacts can be less populated in other proteins.[79] Collectively, the available data in the literature argue that the formation of long-range contacts in expanded unfolded states depends on the protein primary sequence and is not a generic property of all polypeptides. There is a connection with emerging studies of IDPs. Recent work has revealed that the patterning of residues, i.e., the specific distribution of polar and hydrophobic residues, significantly influences the properties of IDPs and provides a more precise description of their behavior than simple correlations based on mean hydrophobicity and average net charge.[80]

■ ASSOCIATED CONTENT

Supporting Information

A table summarizing the $\Delta H°(T_m)$ and T_m for wild-type CTL9 and its single-cysteine mutants (Table S1) and figures showing the comparison of wild-type CLT9 and the six cysteine mutants by CD and thermal denaturation curves monitored by CD, the acid- and urea-induced unfolding of CTL9 monitored by CD at 222 nm, and ^{1}H−^{15}N HSQC spectra of wild-type CTL9 and six cysteine mutants recorded at pH 2.0 in the absence of urea and in 8 M urea at pH 2.5 (Figures S1−S10). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: daniel.raleigh@stonybrook.edu. Phone: (631) 632-9547. Fax: (631) 632-7960.

Funding
Supported by National Science Foundation Grants MCB-0919860 and MCB-1330259 to D.P.R. and MCB-1121867 to R.V.P.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Bing Shan and Dr. Wenli Meng for helpful discussions about NMR data collection and analysis and Dr. Marc Allaire for help with the SAXS experiments. Use of the National Synchrotron Light Source, Brookhaven National Laboratory, was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract DE-AC02-98CH10886.

■ ABBREVIATIONS

CD, circular dichroism; CTL9, C-terminal domain of ribosomal protein L9 from *Geobacillus stearothermophilus*; DSE, denatured state ensemble; EOM, ensemble optimization method; EV, excluded volume; HSQC, heteronuclear single-quantum coherence spectroscopy; MOPS, 3-(*N*-morpholino)-propanesulfonic acid; MTSL, (1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-yl)methyl methanesulfonate (thio-reactive methanethiosulfonate spin-label); R_h, radius of hydration; R_g, radius of gyration; SAXS, small angle X-ray scattering; TCEP, tris(2-carboxyethyl)phosphine.

■ REFERENCES

(1) Wu, Y., Kondrashkina, E., Kayatekin, C., Matthews, C. R., and Bilsel, O. (2008) Microsecond acquisition of heterogeneous structure in the folding of a TIM barrel protein. *Proc. Natl. Acad. Sci. U.S.A. 105*, 13367−13372.
(2) Baldwin, R. L. (2002) A new perspective on unfolded proteins. *Adv. Protein Chem. 62*, 361−367.
(3) Shortle, D. (1996) The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB J. 10*, 27−34.
(4) Choy, W. Y., Mulder, F. A. A., Crowhurst, K. A., Muhandiram, D. R., Millett, I. S., Doniach, S., Forman-Kay, J. D., and Kay, L. E. (2002) Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. *J. Mol. Biol. 316*, 101−112.
(5) Bowler, B. E. (2007) Thermodynamics of protein denatured states. *Mol. BioSyst. 3*, 88−99.
(6) Cho, J. H., Sato, S., Horng, J. C., Anil, B., and Raleigh, D. P. (2008) Electrostatic interactions in the denatured state ensemble: Their effect upon protein folding and protein stability. *Arch. Biochem. Biophys. 469*, 20−28.
(7) Anil, B., Song, B. B., Tang, Y. F., and Raleigh, D. P. (2004) Exploiting the right side of the Ramachandran plot: Substitution of glycines by D-alanine can significantly increase protein stability. *J. Am. Chem. Soc. 126*, 13194−13195.
(8) Jahn, T. R., and Radford, S. E. (2008) Folding versus aggregation: Polypeptide conformations on competing pathways. *Arch. Biochem. Biophys. 469*, 100−117.
(9) de Laureto, P. P., Taddei, N., Frare, E., Capanni, C., Costantini, S., Zurdo, J., Chiti, F., Dobson, C. M., and Fontana, A. (2003) Protein aggregation and amyloid fibril formation by an SH3 domain probed by limited proteolysis. *J. Mol. Biol. 334*, 129−141.
(10) Mishima, T., Ohkuri, T., Monji, A., Imoto, T., and Ueda, T. (2006) Amyloid formation in denatured single-mutant lysozymes where residual structures are modulated. *Protein Sci. 15*, 2448−2452.
(11) Pavitt, G. D., and Ron, D. (2012) New insights into translational regulation in the endoplasmic reticulum unfolded protein response. *Cold Spring Harbor Perspect. Biol. 4*, DOI: 10.1101/cshperspect.a012278.
(12) Ma, Y. J., and Hendershot, L. M. (2004) The role of the unfolded protein response in tumour development: Friend or foe? *Nat. Rev. Cancer 4*, 966−977.
(13) Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol. 6*, 197−208.
(14) Uversky, V. N. (2002) Natively unfolded proteins: A point where biology waits for physics. *Protein Sci. 11*, 739−756.
(15) Guinier, A., and Fournet, G. (1955) *Small Angle Scattering of X-Rays*, Wiley, New York.
(16) Kohn, J. E., Millett, I. S., Jacob, J., Zagrovic, B., Dillon, T. M., Cingel, N., Dothager, R. S., Seifert, S., Thiyagarajan, P., Sosnick, T. R., Hasan, M. Z., Pande, V. S., Ruczinski, I., Doniach, S., and Plaxco, K. W. (2004) Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. U.S.A. 101*, 12491−12496.
(17) Millett, I. S., Doniach, S., and Plaxco, K. W. (2002) Toward a taxonomy of the denatured state: Small angle scattering studies of unfolded proteins. *Adv. Protein Chem. 62*, 241−262.
(18) Candotti, M., Esteban-Martin, S., Salvatella, X., and Orozco, M. (2013) Toward an atomistic description of the urea-denatured state of proteins. *Proc. Natl. Acad. Sci. U.S.A. 110*, 5933−5938.
(19) Huang, J. R., Gabel, F., Jensen, M. R., Grzesiek, S., and Blackledge, M. (2012) Sequence-specific mapping of the interaction between urea and unfolded ubiquitin from ensemble analysis of NMR and small angle scattering data. *J. Am. Chem. Soc. 134*, 4429−4436.
(20) Voets, I. K., Cruz, W. A., Moitzi, C., Lindner, P., Areas, E. P. G., and Schurtenberger, P. (2010) DMSO-induced denaturation of hen egg white lysozyme. *J. Phys. Chem. B 114*, 11875−11883.
(21) Tanford, C., Kawahara, K., and Lapanje, S. (1966) Proteins in 6 M guanidine hydrochloride. Demonstration of random coil behavior. *J. Biol. Chem. 241*, 1921−1923.
(22) Chen, L. L., Wildegger, G., Kiefhaber, T., Hodgson, K. O., and Doniach, S. (1998) Kinetics of lysozyme refolding: Structural characterization of a non-specifically collapsed state using time-resolved X-ray scattering. *J. Mol. Biol. 276*, 225−237.
(23) Flory, P. J. (1953) *Principles of Polymer Chemistry*, Cornell Universtiy Press, Ithaca, NY.
(24) Fersht, A. R., Matouschek, A., and Serrano, L. (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J. Mol. Biol. 224*, 771−782.
(25) Pace, C. N., Alston, R. W., and Shaw, K. L. (2000) Charge-charge interactions influence the denatured state ensemble and contribute to protein stability. *Protein Sci. 9*, 1395−1398.
(26) Tollinger, M., Crowhurst, K. A., Kay, L. E., and Forman-Kay, J. D. (2003) Site-specific contributions to the pH dependence of protein stability. *Proc. Natl. Acad. Sci. U.S.A. 100*, 4545−4550.
(27) Kuhlman, B., Luisi, D. L., Young, P., and Raleigh, D. P. (1999) pKa values and the pH dependent stability of the N-terminal domain of L9 as probes of electrostatic interactions in the denatured state. Differentiation between local and nonlocal interactions. *Biochemistry 38*, 4896−4903.
(28) Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C., and Obradovic, Z. (2001) Intrinsically disordered protein. *J. Mol. Graphics Modell. 19*, 26−59.
(29) Mok, Y. K., Kay, C. M., Kay, L. E., and Forman-Kay, J. (1999) NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions. *J. Mol. Biol. 289*, 619−638.
(30) Cho, J. H., Sato, S., and Raleigh, D. P. (2004) Thermodynamics and kinetics of non-native interactions in protein folding: A single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state. *J. Mol. Biol. 338*, 827−837.
(31) Swint-Kruse, L., and Robertson, A. D. (1995) Hydrogen bonds and the pH dependence of ovomucoid third domain stability. *Biochemistry 34*, 4724−4732.
(32) Mok, Y. K., Elisseeva, E. L., Davidson, A. R., and Forman-Kay, J. D. (2001) Dramatic stabilization of an SH3 domain by a single

substitution: Roles of the folded and unfolded states. *J. Mol. Biol. 307*, 913−928.
(33) Lietzow, M. A., Jamin, M., Dyson, H. J., and Wright, P. E. (2002) Mapping long-range contacts in a highly unfolded protein. *J. Mol. Biol. 322*, 655−662.
(34) Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L. J., Dobson, C. M., and Schwalbe, H. (2002) Long-range interactions within a nonnative protein. *Science 295*, 1719−1722.
(35) Sung, Y., and Eliezer, D. (2007) Residual structure, backbone dynamics, and interactions within the synuclein family. *J. Biol. Chem. 372*, 689−707.
(36) Neri, D., Billeter, M., Wider, G., and Wuthrich, K. (1992) NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science 257*, 1559−1563.
(37) Sato, S., and Raleigh, D. P. (2002) pH-dependent stability and folding kinetics of a protein with an unusual α-β topology: The C-terminal domain of the ribosomal protein L9. *J. Mol. Biol. 318*, 571−582.
(38) Horng, J. C., Cho, J. H., and Raleigh, D. P. (2005) Analysis of the pH-dependent folding and stability of histidine point mutants allows characterization of the denatured state and transition state for protein folding. *J. Mol. Biol. 345*, 163−173.
(39) Li, Y., Gupta, R., Cho, J. H., and Raleigh, D. P. (2007) Mutational analysis of the folding transition state of the C-terminal domain of ribosomal protein L9: A protein with an unusual β-sheet topology. *Biochemistry 46*, 1013−1021.
(40) Shan, B., Bhattacharya, S., Eliezer, D., and Raleigh, D. P. (2008) The low-pH unfolded state of the C-terminal domain of the ribosomal protein L9 contains significant secondary structure in the absence of denaturant but is no more compact than the low-pH urea unfolded state. *Biochemistry 47*, 9565−9573.
(41) Li, Y., Picart, F., and Raleigh, D. P. (2005) Direct characterization of the folded, unfolded and urea-denatured states of the C-terminal domain of the ribosomal protein L9. *J. Mol. Biol. 349*, 839−846.
(42) Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr. 36*, 1277−1282.
(43) Bernado, P., Mylonas, E., Petoukhov, M. V., Blackledge, M., and Svergun, D. I. (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc. 129*, 5656−5664.
(44) 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.
(45) Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol. 278*, 313−352.
(46) Vitalis, A., and Pappu, R. V. (2009) ABSINTH: A new continuum solvation model for simulations of polypeptides in aqueous solutions. *J. Comput. Chem. 30*, 673−699.
(47) Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L. (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B 105*, 6474−6487.
(48) Meng, W., Lyle, N., Luan, B., 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.
(49) Li, Y., Horng, J. C., and Raleigh, D. P. (2006) pH dependent thermodynamic and amide exchange studies of the C-terminal domain of the ribosomal protein L9: Implications for unfolded state structure. *Biochemistry 45*, 8499−8506.
(50) Sato, S., Kuhlman, B., Wu, W. J., and Raleigh, D. P. (1999) Folding of the multidomain ribosomal protein L9: The two domains fold independently with remarkably different rates. *Biochemistry 38*, 5643−5650.
(51) Luan, B., Shan, B., Baiz, C., Tokmakoff, A., and Raleigh, D. P. (2013) Cooperative cold denaturation: The case of the C-terminal domain of ribosomal protein L9. *Biochemistry 52*, 2402−2409.
(52) Shan, B., McClendon, S., Rospigliosi, C., Eliezer, D., and Raleigh, D. P. (2010) The cold denatured state of the C-terminal domain of protein L9 is compact and contains both native and non-native structure. *J. Am. Chem. Soc. 132*, 4669−4677.
(53) Manning, M. C., and Woody, R. W. (1991) Theoretical CD studies of polypeptide helices: Examination of important electronic and geometric factors. *Biopolymers 31*, 569−586.
(54) Salmon, L., Nodet, G., Ozenne, V., Yin, G., Jensen, M. R., Zweckstetter, M., and Blackledge, M. (2010) NMR characterization of long-range order in intrinsically disordered proteins. *J. Am. Chem. Soc. 132*, 8407−8418.
(55) Xue, Y., Podkorytov, I. S., Rao, D. K., Benjamin, N., Sun, H. L., and Skrynnikov, N. R. (2009) Paramagnetic relaxation enhancements in unfolded proteins: Theory and application to drkN SH3 domain. *Protein Sci. 18*, 1401−1424.
(56) Perez, Y., Gairi, M., Pons, M., and Bernado, P. (2009) Structural characterization of the natively unfolded N-terminal domain of human c-Src kinase: Insights into the role of phosphorylation of the unique domain. *J. Mol. Biol. 391*, 136−148.
(57) Lee, C. W., Arai, M., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2009) Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP. *Biochemistry 48*, 2115−2124.
(58) Franzmann, M., Otzen, D., and Wimmer, R. (2009) Quantitative use of parramagnetic relaxation enhancements for determining orientations and insertion depths of peptides in micelles. *ChemBioChem 10*, 2339−2347.
(59) Allison, J. R., Varnai, P., Dobson, C. M., and Vendruscolo, M. (2009) Determination of the free energy landscape of α-synuclein using spin label nuclear magnetic resonance measurements. *J. Am. Chem. Soc. 131*, 18314−18326.
(60) Felitsky, D. J., Lietzow, M. A., Dyson, H. J., and Wright, P. E. (2008) Modeling transient collapsed states of an unfolded protein to provide insights into early folding events. *Proc. Natl. Acad. Sci. U.S.A. 105*, 6278−6283.
(61) Iwahara, J., and Clore, G. M. (2006) Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature 440*, 1227−1230.
(62) Kristjansdottir, S., Lindorff-Larsen, K., Fieber, W., Dobson, C. M., Vendruscolo, M., and Poulsen, F. M. (2005) Formation of native and non-native interactions in ensembles of denatured ACBP molecules from paramagnetic relaxation enhancement studies. *J. Mol. Biol. 347*, 1053−1062.
(63) Dedmon, M. M., Lindorff-Larsen, K., Christodoulou, J., Vendruscolo, M., and Dobson, C. M. (2005) Mapping long-range interactions in α-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *J. Am. Chem. Soc. 127*, 476−477.
(64) Lindorff-Larsen, K., Kristjansdottir, S., Teilum, K., Fieber, W., Dobson, C. M., Poulsen, F. M., and Vendruscolo, M. (2004) Determination of an ensemble of structures representing the denatured state of the bovine acyl-coenzyme A binding protein. *J. Am. Chem. Soc. 126*, 3291−3299.
(65) Teilum, K., Kragelund, B. B., and Poulsen, F. M. (2002) Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling. *J. Mol. Biol. 324*, 349−357.
(66) Battiste, J. L., and Wagner, G. (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear Overhauser effect data. *Biochemistry 39*, 5355−5365.
(67) Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J. Mol. Biol. 268*, 170−184.
(68) Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease I.

Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol. 268*, 158−169.
(69) Kosen, P. A., Scheek, R. M., Naderi, H., Basus, V. J., Manogaran, S., Schmidt, P. G., Oppenheimer, N. J., and Kuntz, I. D. (1986) Two-dimensional ^{1}H NMR of three spin-labeled derivatives of bovine pancreatic trypsin inhibitor. *Biochemistry 25*, 2356−2364.
(70) Zhou, H. X. (2004) Polymer models of protein stability, folding, and interactions. *Biochemistry 43*, 2141−2154.
(71) Bernado, P., and Blackledge, M. (2009) A self-consistent description of the conformational behavior of chemically denatured proteins from NMR and small angle scattering. *Biophys. J. 97*, 2839−2845.
(72) Tran, H. T., and Pappu, R. V. (2006) Toward an accurate theoretical framework for describing ensembles for proteins under strongly denaturing conditions. *Biophys. J. 91*, 1868−1886.
(73) Tran, H. T., Wang, X. L., and Pappu, R. V. (2005) Reconciling observations of sequence-specific conformational propensities with the generic polymeric behavior of denatured proteins. *Biochemistry 44*, 11369−11380.
(74) Jha, A. K., Colubri, A., Freed, K. F., and Sosnick, T. R. (2005) Statistical coil model of the unfolded state: Resolving the reconciliation problem. *Proc. Natl. Acad. Sci. U.S.A. 102*, 13099−13104.
(75) Ding, F., Jha, R. K., and Dokholyan, N. V. (2005) Scaling behavior and structure of denatured proteins. *Structure 13*, 1047−1054.
(76) Fitzkee, N. C., and Rose, G. D. (2004) Reassessing random-coil statistics in unfolded proteins. *Proc. Natl. Acad. Sci. U.S.A. 101*, 12497−12502.
(77) Yoo, T. Y., Meisburger, S. P., Hinshaw, J., Pollack, L., Haran, G., Sosnick, T. R., and Plaxco, K. (2012) Small-angle X-ray scattering and single-molecule FRET spectroscopy produce highly divergent views of the low-denaturant unfolded state. *J. Mol. Biol. 418*, 226−236.
(78) Bernado, P., Blanchard, L., Timmins, P., Marion, D., Ruigrok, R. W. H., and Blackledge, M. (2005) A structural model for unfolded proteins from residual dipolar couplings and small-angle X-ray scattering. *Proc. Natl. Acad. Sci. U.S.A. 102*, 17002−17007.
(79) Wirmer, J., Peti, W., and Schwalbe, H. (2006) Motional properties of unfolded ubiquitin: A model for a random coil protein. *J. Biomol. NMR 35*, 175−186.
(80) 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.