# The Denatured State Ensemble Contains Significant Local and Long-Range Structure under Native Conditions: Analysis of the N-Terminal Domain of Ribosomal Protein L9

**ARTICLE** *in* BIOCHEMISTRY · MARCH 2013

Impact Factor: 3.02 · DOI: 10.1021/bi301667u · Source: PubMed

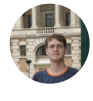


**5 AUTHORS**, INCLUDING:

Bowu Luan

**8** PUBLICATIONS **54** CITATIONS

SEE PROFILE



Nicholas James Lyle

Washington University in St. Louis

**19** PUBLICATIONS **214** CITATIONS

SEE PROFILE

Rohit V Pappu

Washington University in St. Louis

**108** PUBLICATIONS **2,847** CITATIONS

SEE PROFILE

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

# The Denatured State Ensemble Contains Significant Local and Long-Range Structure under Native Conditions: Analysis of the N-Terminal Domain of Ribosomal Protein L9

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

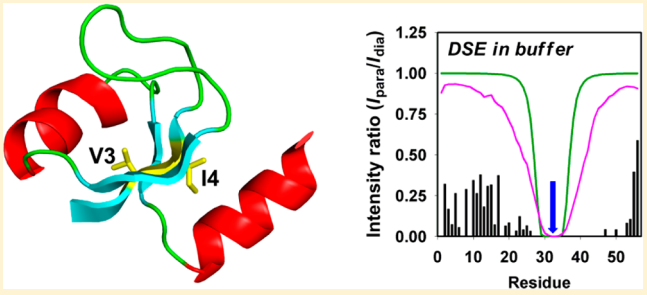
[†]Department of Chemistry, Stony Brook University, Stony Brook, New York 11794-3400, United States

[‡]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:** The denatured state ensemble (DSE) represents the starting state for protein folding and the reference state for protein stability studies. Residual structure in the DSE influences the kinetics of protein folding, the propensity to aggregate, and protein stability. The DSE that is most relevant for folding is the ensemble populated under native conditions, but the stability of proteins and the cooperativity of their folding normally prevent direct characterization of this ensemble. Indirect experiments have been used to infer residual structure in the DSE under nondenaturing conditions, but direct characterization is rare. The N-terminal domain of ribosomal protein L9 (NTL9) is a small mixed $\alpha-\beta$ domain that folds cooperatively on the millisecond time scale. A destabilized double mutant of NTL9, V3A/I4A-NTL9, populates the DSE in the absence of denaturant and is in slow exchange with the native state on the nuclear magnetic resonance time scale. The DSE populated in buffer was compared to the urea-induced DSE. Analysis of $^{1}$H and $^{13}$C chemical shifts reveals residual secondary structure in the DSE in buffer, which is stabilized by both local and long-range interactions. $^{15}$N $R_2$ relaxation rates deviate from random coil models, suggesting hydrophobic clustering in the DSE. Paramagnetic relaxation enhancement experiments show that there are transient long-range contacts in the DSE in buffer. In contrast, the urea-induced DSE has significantly less residual secondary structure and markedly fewer long-range contacts; however, the urea-induced DSE deviates from a random coil.




Quantitative characterization of the conformational propensities and residual interactions in the denatured state ensemble (DSE) of proteins is important because the DSE is the starting state for protein folding, the thermodynamic reference state for protein stability, and is important for understanding the effects of crowding on protein stability.[1−7] The DSE can be the starting state for protein aggregation, and DSE interactions can modulate aggregation propensity.[8−10] A number of proteins have been found to be intrinsically disordered in their functional state, and a detailed understanding of the conformational propensities of these intrinsically disordered proteins (IDPs) is crucial for elucidating the molecular mechanism of their biological function.[11,12]

Studies of protein DSEs are often conducted under strongly denaturing conditions, involving high concentrations of denaturant, extreme temperatures, or extreme pH values.[13−16] Expanded DSEs are usually sampled under strongly denaturing conditions, and their size, as defined by the radius of gyration ($R_g$) or the hydrodynamic radius ($R_h$), typically scales with chain length, $N$, as $N^{0.59}$. The characteristic scaling exponent arises because the DSE expands to make favorable interactions with solvent. This does not mean that such ensembles are devoid of residual structure or behave as excluded volume (EV) limit ensembles. Sequence specific preferences for local structure are compatible with the EV limit.[17−19] In some cases, transient long-range interactions have been observed in highly unfolded proteins above and beyond those expected in the EV ensemble.[20] However, the DSE that is directly relevant for protein folding is that which is populated under native conditions, and this ensemble may contain significantly more structure than the DSE populated under strongly denaturing conditions.[4,21−26] Unfortunately, the DSE under native conditions is difficult to access experimentally because the free energy of protein folding strongly favors the folded state under native conditions and folding is usually highly

**Received:** December 16, 2012
**Revised:** March 11, 2013
**Published:** March 12, 2013

   dx.doi.org/10.1021/bi301667u | *Biochemistry* 2013, 52, 2662−2671

cooperative. Therefore, direct characterization of the DSE under native conditions is still rare.

Indirect approaches have been used to characterize the DSE, including studies of peptide fragments derived from the target protein, H−D exchange measurements, and pH-dependent stability studies, but these methods are either of low resolution or incapable of providing information about long-range contacts in the DSE.[27−33] An alternative approach is to destabilize the native state so that the DSE is populated under native conditions. This can be accomplished using destabilizing mutants, reduction of disulfides, oxidation of Met residues, or truncation variants. Here we report a direct, detailed analysis of local and long-range interactions found in the DSE of the N-terminal domain of ribosomal protein L9 (NTL9) under native conditions.

NTL9 is a 56-residue globular protein with a simple $\alpha-\beta$ topology that folds cooperatively, in a reversible two-state fashion over a wide range of conditions.[27,34−37] Previous studies have demonstrated that the DSE of NTL9 populated in high concentrations of urea is expanded and largely devoid of secondary structure, but samples conformations that include long-range contacts.[20] More extensive interactions are expected in the DSE of NTL9 under native conditions, but direct characterization of this ensemble is lacking. Indirect studies show that the DSE of NTL9 is compact in the absence of denaturant and forms energetically significant non-native electrostatic interactions,[5,29,38−40] but those studies were not able to provide high-resolution secondary structure information or information about long-range contacts. The stability and rapid cooperative folding of wild-type NTL9 have prevented the direct characterization of the DSE. Here we used a destabilized double mutant of NTL9 generated by substituting the core residues V3 and I4 with Ala (V3A/I4A-NTL9) to populate the DSE in native buffer. The double mutant populates the DSE and the folded state in native buffer, and the two states are in slow exchange on the nuclear magnetic resonance (NMR) time scale. Secondary chemical shift analysis, $^{15}N$ $R_2$ relaxation rates, and paramagnetic relaxation enhancement (PRE) studies show that the DSE of V3A/I4A-NTL9 in buffer contains residual secondary structure, forms extensive hydrophobic clusters and long-range contacts, and differs significantly from the urea-induced DSE.

## EXPERIMENTAL PROCEDURES

**Protein Expression and Purification.** $^{15}N$-labeled and $^{13}C$- and $^{15}N$-labeled V3A/I4A-NTL9 and cysteine mutants were overexpressed and purified as previously described.[40] The identities of the proteins were confirmed using matrix-assisted laser desorption ionization time-of-flight mass spectroscopy, and the purity was confirmed by analytical high-performance liquid chromatography (HPLC).

**Stopped-Flow Fluorescence Measurements.** Stopped-flow fluorescence measurements were performed using an Applied Photophysics SX.18MV stopped-flow instrument. The fluorescence change of Tyr25 in V3A/I4A-NTL9 was followed using a 305 nm cutoff filter with excitation at 280 nm. Protein at pH 3.7 was mixed with 25 volumes of pH 9.0 buffer to give a final pH of 5.7, with both buffers containing 20 mM sodium acetate and 100 mM sodium chloride. The final protein concentration after mixing was around 30 $\mu$M. The measurements were repeated five times, and the five resulting fluorescence traces were averaged. The averaged curve was fit to a single-exponential function to give the observed rate constant ($k_{obs}$).

**NMR Assignments and Data Analysis.** Samples for NMR studies under native conditions were prepared in a 90% $H_2O$/10% $D_2O$ mixture with 20 mM sodium acetate and 100 mM sodium chloride (pH 5.5). For the urea-induced denatured state, urea was added to a final concentration of 8.3 M as determined by refractometry. All NMR experiments were performed at 12 °C and a protein concentration of approximately 0.5 mM. 2,2-Dimethyl-2-silapentane-5-sulfonate sodium salt (DSS) was added as an internal reference (0.00 ppm) for all samples.

NMR assignments for V3A/I4A-NTL9 under native conditions have been reported.[40] The NMR experiments used for the urea-induced denatured state assignments were performed on a 600 MHz Varian spectrometer with a cryoprobe. A set of two- and three-dimensional experiments, including $^1H-^{15}N$ HSQC, HNCO, HNCACB, and CBCA(CO)NH, were performed on $^{13}C$- and $^{15}N$-labeled samples to generate $^{13}C\alpha$, $^{13}C\beta$, $^{13}CO$, $^1HN$, and $^{15}N$ assignments. All spectra were processed with NMRPipe[41] and visualized using NMRViewJ.[42] $^1H$ chemical shifts were referenced to DSS directly, and $^{15}N$ and $^{13}C$ chemical shifts were indirectly referenced using standard methods. The assignments have been deposited in the BMRB database as entry 17675 for the assignments in buffer and entry 18918 for the assignments in 8.3 M urea. Secondary chemical shifts were calculated using the random coil values of Wishart and co-workers.[43] Sequence-dependent correction of random coil chemical shifts was performed.[44] The chemical shifts were also analyzed using the SSP method developed by Forman-Kay and co-workers.[45] A window size of three residues was used.

**Paramagnetic Relaxation Enhancement Experiment.** Spin-labeled samples were prepared by dissolving 3 mg of the V3A/I4A-NTL9 cysteine mutant in 600 $\mu$L of NMR buffer (90% $H_2O$/10% $D_2O$, 20 mM sodium acetate, and 100 mM sodium chloride); 12 $\mu$L of a 300 mM MTSL [(1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)methanesulfonate] stock solution was added for a reaction time of 10 h at room temperature, and excess MTSL was removed using a Sephadex G25 column. The extent of labeling was assessed using either HPLC or the intensity of the Cys peak in the HSQC spectrum, as this resonance is beached in the spin-labeled molecules. The percentage labeling was estimated to be ≥95%. The labeled sample was split into two equal aliquots; 300 $\mu$L of NMR buffer was added to one of the aliquots (paramagnetic form), and 270 $\mu$L of NMR buffer and 30 $\mu$L of a 100 mM TCEP [tris(2-carboxyethyl)phosphine] stock solution were added to the other aliquots to reduce the protein (diamagnetic form). The pH was then adjusted to 5.5.

For urea-containing samples, urea was added to a concentration of 10 M. The sample was split into two equal aliquots. For the paramagnetic form, only NMR buffer was added; for the diamagnetic form, 30 $\mu$L of a 100 mM TCEP stock solution and NMR buffer was added. The final concentration of urea was 8.3 M for both samples, determined from the refractive index.

$^1H-^{15}N$ HSQC spectra were collected for both the paramagnetic samples and diamagnetic samples on a Varian 600 MHz spectrometer at 12 °C. The peak intensities were measured with NMRViewJ,[42] and the intensity ratio between paramagnetic and diamagnetic forms was calculated.

The theoretical PRE intensity ratios for random coils were calculated using both a Gaussian distribution model and excluded volume (EV) simulations. For the Gaussian model, the distance between each residue and the spin-label site follows a Gaussian distribution of root-mean-square end-to-end distances:[46,47]

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

where $r$ is the end-to-end distance between a residue and the spin-label site, $n$ is the number of residues between residue $i$ and the spin-label site, $l$ (3.8 Å) is the link length of the chain, and $\alpha$ is the cosine of the bond angle supplements for the freely rotating chain model, which was set to 0.8 based on experimentally estimates of the statistical segment lengths in poly-L-alanine.[46,47] The contribution of paramagnetic relaxation enhancement to the transverse relaxation rate, $R_{2P}$, was calculated from eq 2:[46,47]

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

where $K$ is $1.23 \times 10^{-32}$ cm$^6$ s$^{-2}$, $r$ is the distance between a given residue and the spin-label site, $\tau_C$ is the effective correlation time, which is 4 ns calculated from $^{15}$N $R_1$ and $R_2$ relaxation rates, and $\omega_H$ is the Larmor frequency of a proton. The peak intensity ratios between the paramagnetic and diamagnetic forms were calculated using eq 3:

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

where $R_{2D}$ is the transverse relaxation rate of the backbone amide protons in the diamagnetic form. The average value was measured to be 14 s$^{-1}$ using a one-dimensional pulse sequence. $t$ is the total duration of the INEPT delays, which is 12 ms for the HSQC pulse sequence.

**$^{15}$N $R_2$ Relaxation Experiments.** $^{15}$N $R_2$ relaxation experiments were conducted with [$^{15}$N]V3A/I4A-NTL9 in native buffer [20 mM sodium acetate and 100 mM sodium chloride (pH 5.5)] and in 8.3 M urea on a 600 MHz spectrometer using a Carr−Purcell−Meiboom−Gill (CPMG) sequence.[48] The 180° pulse spacing in the CPMG sequence was 1 ms. The relaxation delays for [$^{15}$N]V3A/I4A-NTL9 in native buffer were set to 20 ms (twice), 40 ms, 60 ms (twice), 80 ms, and 100 ms (twice). For [$^{15}$N]V3A/I4A-NTL9 in 8.3 M urea, the relaxation delays were set to 14 ms (twice), 30 ms, 50 ms (twice), 74 ms, 86 ms (twice), 98 ms, and 110 ms. A recycle delay of 3 s was used. Data were processed and analyzed using NMRViewJ by fitting the peak intensities to a two-parameter single-exponential decay.[42] $^{15}$N $R_2$ rates were analyzed using the phenomenological model of Schwalbe[49] by fitting the experimental $R_2$ rates to eq 4.

$$R_2(i) = R_2(\text{int}) \sum_{j=1}^{N} \exp\left(-\frac{|i-j|}{\lambda}\right) \quad (4)$$

where $R_2(i)$ is the experimental $R_2$ value for residue $i$, $R_2$(int) is the intrinsic relaxation rate that depends on the temperature and the viscosity of the solution, $N$ is the total number of residues in the protein, and $\lambda$ is the apparent persistence length of the chain.

The EV model calculation for NTL9 has been reported previously and explicitly includes the spin-label.[20] The EV ensemble corresponding to the wild-type sequence was used.

**Small Angle X-ray Scattering (SAXS) Experiments.** Samples of V3A/I4A-NTL9 were prepared in 8.3 M urea at pH 5.5. Scattering experiments were performed on beamline X9 of National Synchrotron Light Source I at Brookhaven National Laboratory (Upton, NY). Protein samples were injected into a 1 mm diameter capillary continuously during the measurement at a rate of 0.67 μL/s to avoid radiation damage. The exposure time for each measurement was 30 s. Scattering data were collected for four separate protein concentrations: 7.5, 10.0, 12.5, and 15 mg/mL at 12 °C. Each sample was measured three times and then averaged before data analysis. PRIMUS[50] was used for analyzing the scattering patterns, and the radius of gyration ($R_g$) was obtained using the Guinier approximation (eq 5)

$$I(q) = I(0)\exp(-R_g{}^2q^{2/3}) \quad (5)$$

where $I(q)$ is the intensity at scattering vector $q$.[51]

## RESULTS

**The V3A/I4A Mutant of NTL9 Significantly Populates the DSE under Native Conditions.** V3 and I4 are located in the first β-strand and are buried in the hydrophobic core of wild-type NTL9 (Figure 1). Mutation of these residues to alanine destabilizes the domain significantly, and the DSE of the double mutant is ~70% populated at 12 °C and pH 5.5, with two sets of resonances observed in the HSQC spectrum.[40] pH jump fluorescence-monitored stopped-flow measurements were used to directly determine the rate of exchange between the folded state and DSE. The exchange rate ($k_{ex} = k_{fold} + k_{unfold}$) is 15.7 s$^{-1}$, confirming that the system is in slow exchange on the NMR chemical shift time scale (Figure S1 of the Supporting Information). This allows direct characterization of the DSE

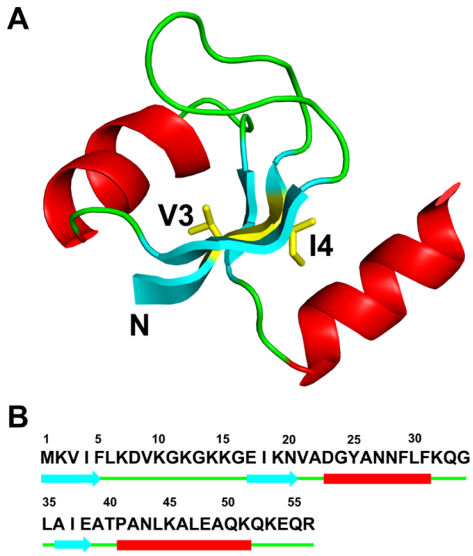



**Figure 1.** (A) Ribbon diagram of the N-terminal domain of L9 (Protein Data Bank entry 2HBB). The three β-strands are colored cyan and the two helices red. The hydrophobic residues Val3 and Ile4 as well as the N-terminus are labeled. (B) Primary sequence of NTL9 together with a schematic representation of the secondary structure elements. Arrows represent β-strands, and bars represent α-helices.


dx.doi.org/10.1021/bi301667u | Biochemistry 2013, 52, 2662−2671

under native conditions by NMR. Standard triple-resonance methods were used to obtain >95% of the $^{13}C$, $^{15}N$, and $^{1}H$ backbone assignments as well as $^{13}C_\beta$ assignments for both the folded state and the DSE in native buffer.

We also characterized V3A/I4A-NTL9 in 8.3 M urea at 12 °C and pH 5.5. The double mutant is fully unfolded under this condition as judged by CD- or NMR-monitored titrations. We also measured the $R_g$ for the urea-induced DSE using small angle X-ray scattering (SAXS). The value of 21.7 Å measured at a protein concentration of 7.5 mg/mL is within the range predicted for a highly unfolded 56-residue polypeptide chain as judged by empirical correlations between chain length and $R_g$.[13] This ensemble provides a useful benchmark for comparison to the DSE in buffer. Only one set of resonances corresponding to the DSE is visible in the $^{1}H-^{15}N$ HSQC spectrum in 8.3 M urea (Figure S2 of the Supporting Information). More than 95% of the $^{13}C$, $^{15}N$, and $^{1}H_\alpha$ backbone assignments and $^{13}C_\beta$ assignments were obtained.

**Chemical Shift Analysis Reveals Significant Secondary Structure in the DSE under Native Conditions.** It is well-known that NMR chemical shifts are sensitive probes of protein structure. Secondary shifts, defined as the deviation of observed chemical shifts from random coil values ($\Delta\delta = \delta_{observed} - \delta_{random\ coil}$), are accurate predictors of secondary structure.[52,53] For example, positive $^{13}C_\alpha$ and $^{13}CO$ secondary shifts suggest $\alpha$-helical propensity, while negative values indicate $\beta$-strand propensity. The opposite trend is observed for $^{13}C_\beta$ and $^{13}H_\alpha$ secondary shifts. The difference in $^{13}C_\alpha$ and $^{13}C_\beta$ secondary shifts, $\Delta\delta^{13}C_\alpha - \Delta\delta^{13}C_\beta$, is widely used to evaluate secondary structure propensities because it eliminates problems with chemical shift referencing. Positive values indicate $\alpha$-helical structure, and negative values indicate a propensity to populate $\beta$-strand structure.

The secondary shifts and the values of $\Delta\delta^{13}C_\alpha - \Delta\delta^{13}C_\beta$ are shown in Figure 2. The secondary shifts of the native state of V3A/I4A-NTL9 agree well with the expectations based on the structure of wild-type NTL9. Positive $^{13}C_\alpha$, $\Delta\delta^{13}C_\alpha - \Delta\delta^{13}C_\beta$, and $^{13}CO$ secondary shifts as well as negative $^{13}C_\beta$ and $^{1}H_\alpha$ secondary shifts were observed for the regions corresponding to the two native state helices. Negative $^{13}C_\alpha$, $\Delta\delta^{13}C_\alpha - \Delta\delta^{13}C_\beta$, and $^{13}CO$ secondary shifts as well as positive $^{13}C_\beta$ and $^{1}H_\alpha$ secondary shifts were observed for most of the residues in the three $\beta$-strands. This validates the use of secondary shifts to probe structure in V3A/I4A-NTL9.

Compared to those of the native state of V3A/I4A-NTL9, the secondary shifts for the DSE in native buffer are smaller in magnitude. However, the two segments corresponding to the helices in the native state exhibit continuous secondary shifts and values of $\Delta\delta^{13}C_\alpha - \Delta\delta^{13}C_\beta$ indicative of partial helical structure. The region of residues 41−51, corresponding to a part of $\alpha$-helix 2 in the native state, has positive $^{13}C_\alpha$ secondary shifts (1.13 ppm on average), $\Delta\delta^{13}C_\alpha - \Delta\delta^{13}C_\beta$ values (1.47 ppm on average), and $^{13}CO$ secondary shifts (1.19 ppm on average), as well as negative $^{13}C_\beta$ secondary shifts (−0.34 ppm on average) and $^{1}H_\alpha$ secondary shifts (−0.09 ppm on average), consistent with a strong propensity to populate $\alpha$-helix $\phi$ and $\psi$ angles. Residues 52−56 also exhibit secondary shifts consistent with $\alpha$-helical propensity, although the values gradually decline over this region as one progresses toward the C-terminus. Residues 23−31, corresponding to $\alpha$-helix 1 in the native state, exhibit smaller but non-zero secondary shifts, suggesting $\alpha$-helical propensity in this region. Positive $^{13}C_\alpha$ secondary shifts (0.45 ppm on average), positive values of $\Delta\delta^{13}C_\alpha - \Delta\delta^{13}C_\beta$ (0.61 ppm on average), and positive $^{13}CO$ secondary shifts (0.21 ppm on average) as well as negative $^{13}C_\beta$ secondary shifts (−0.13 ppm on average) and $^{1}H_\alpha$ secondary shifts (−0.03 ppm on average) are observed. The secondary shifts for the regions corresponding to the three $\beta$-strands are much closer to zero, and only $\beta$-strand 3 exhibits any systematic deviations in the DSE.

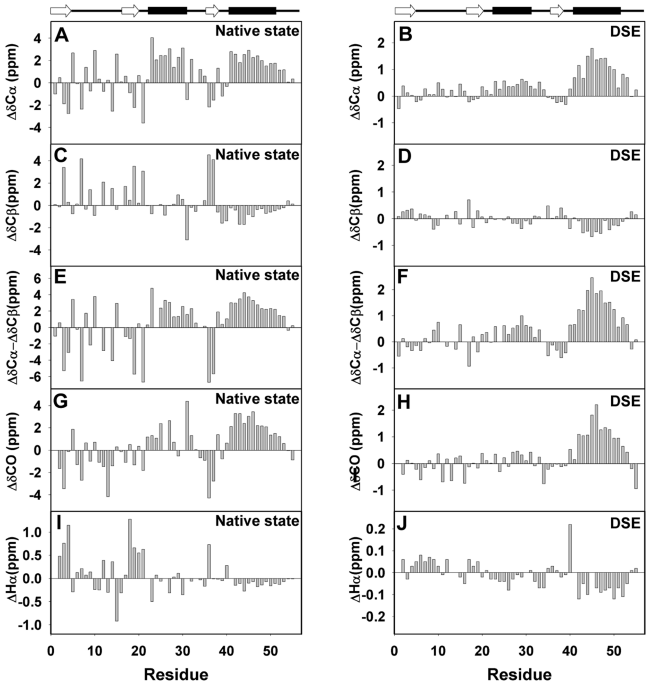



**Figure 2.** Plots of secondary shifts for V3A/I4A-NTL9 for the native state (left) and for the DSE in native buffer [20 mM sodium acetate and 100 mM sodium chloride (pH 5.5)] at 12 °C. Data are plotted as observed minus random coil: (A) native state $^{13}C_\alpha$, (B) DSE $^{13}C_\alpha$, (C) native state $^{13}C_\beta$, (D) DSE $^{13}C_\beta$, (E) native state $\Delta\delta^{13}C_\alpha - \Delta\delta^{13}C_\beta$, (F) DSE $\Delta\delta^{13}C_\alpha - \Delta\delta^{13}C_\beta$, (G) native state $^{13}CO$, (H) DSE $^{13}CO$, (I) native state $^{1}H_\alpha$ and (J) DSE $^{1}H_\alpha$. Random coil values of Wishart et al.[43] were used. Note the different scales used for the folded form and DSE. Schematic diagrams of the native state secondary structure are shown at the top of each column. $\beta$-Strands are depicted as arrows and $\alpha$-helices as bars.

The secondary shifts for the V3A/I4A-NTL9 DSE in 8.3 M urea are all very small, and no significant $\alpha$-helical or $\beta$-strand propensity is observed, indicating that the residual secondary structure in the DSE is largely abolished by the presence of high concentrations of urea (Figure S3 of the Supporting Information). This result confirms that the DSE populated under native conditions is dramatically different from the DSE populated under strongly denaturing conditions.

We calculated secondary structure propensity scores (SSP) to quantitatively analyze the secondary structure in the DSE of V3A/I4A-NTL9.[45] The analysis incorporates the chemical shifts of different nuclei to calculate a single SSP score for each residue, which is then averaged via a sliding window. An SSP score of 1 indicates fully formed $\alpha$-helical structure, and a value of −1 represents fully formed $\beta$-strand structure. The SSP scores of the V3A/I4A-NTL9 native state are congruent with the wild-type NTL9 structure (Figure 3A). For the two $\alpha$-helical regions, large positive SSP scores are observed, while residues in the three $\beta$-strand regions display significant negative SSP scores. The native state SSP scores for the $\beta$-strands are larger than −1, because of the short lengths of the


dx.doi.org/10.1021/bi301667u | Biochemistry 2013, 52, 2662−2671

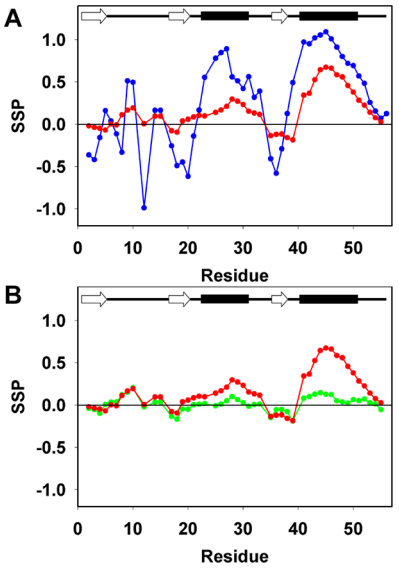



**Figure 3.** (A) SSP scores for the folded form (blue) and the DSE (red) of V3A/I4A-NTL9 in native buffer [20 mM sodium acetate and 100 mM sodium chloride (pH 5.5)] at 12 °C. (B) SSP scores for the DSE (red) of V3A/I4A-NTL9 in native buffer [20 mM sodium acetate and 100 mM sodium chloride (pH 5.5)] at 12 °C and the DSE (green) of V3A/I4A-NTL9 in 8.3 M urea at 12 °C. Positive values represent $\alpha$-helical propensity and negative values $\beta$-strand propensity. $^{13}C_\alpha$ and $^{13}C_\beta$ shifts were used in the calculations with a window size of 3. A schematic diagram of the native state secondary structure is shown at the top of the figure. $\beta$-Strands are depicted as arrows and $\alpha$-helices as bars.

three $\beta$-strands in NTL9 and the three-residue window size used in the SSP calculation. The SSP scores for the DSE in native buffer are smaller than those of the native state, but the regions corresponding to the two native state $\alpha$-helices have positive SSP scores (Figure 3). The SSP scores for the segment running from residue 23 to 31 range from 0.10 to 0.30, with an average value of 0.20, indicating a tendency to adopt $\alpha$-helical conformations in the DSE. Positive SSP scores ranging from 0.23 to 0.67, with an average value of 0.48, are observed for the segment corresponding to residues 41−51, indicating a strong $\alpha$-helix propensity in this region. Positive, albeit smaller, SSP scores are found for residues 52−56, with an average value of 0.12. The SSP scores for $\beta$-strands 1 and 2 show no systematic deviations from zero. The values for strand 3 are small but negative. There have been suggestions that $^1H_\alpha$ chemical shifts are not as reliable as $^{13}C_\alpha$ and $^{13}C_\beta$ shifts for use in the SSP analysis; therefore, we repeated the SSP analysis for the DSE in native buffer with and without $^1H_\alpha$ chemical shifts, and essentially identical results were obtained (Figure S4 of the Supporting Information). In contrast to the DSE in native buffer, the DSE in 8.3 M urea exhibits SSP scores that are very close to zero, indicating that there is no significant persistent secondary structure in this ensemble (Figure 3B).

**$^{15}$N Transverse Relaxation Studies Are Consistent with Hydrophobic Cluster Formation in the DSE in Native Buffer.** $^{15}$N transverse relaxation rates ($R_2$) are frequently used to characterize the structure and dynamics of protein denatured states because they are sensitive to backbone mobility. Completely unstructured polypeptides usually exhibit an inverted U shape profile with uniform $^{15}$N $R_2$ rates in the center of the polypeptide chain and slightly decreased rates at the termini. Positive deviations of the experimental $^{15}$N $R_2$ rates from the phenomenological random coil model are interpreted to indicate the formation of persistent local or long-range hydrophobic clusters in the denatured state.[16,49] Deviations of relaxation parameters from random coil models might also result from the formation of $\alpha$-helical structure as seen, for example, in the DSE of monomeric $\lambda$-repressor.[26] Figure 4 plots

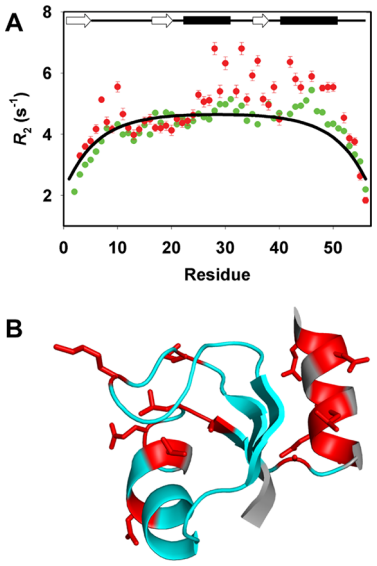



**Figure 4.** (A) $^{15}$N $R_2$ rates for the DSE of V3A/I4A-NTL9 in native buffer (red circles) at 12 °C and the DSE of V3A/I4A-NTL9 in 8.3 M urea (green circles) at 12 °C. The solid line is the best fit to the phenomenological model of Schwalbe and co-workers.[16,49] Uncertainties represent the apparent standard deviations determined using NMRViewJ fitting routines.[42] A schematic diagram of the native state secondary structure is shown at the top of the figure. $\beta$-Strands are depicted as arrows and $\alpha$-helices as bars. (B) Diagram of NTL9. Residues with $^{15}$N $R_2$ rates in the DSE in native buffer showing significant deviations (>1.0 s$^{-1}$) from the Schwalbe phenomenological model are colored red. Residues for which $^{15}$N $R_2$ values could not be determined are colored gray. Sites that are within 1.0 s$^{-1}$ of the model are colored cyan.

the measured $^{15}$N $R_2$ rates for the V3A/I4A-NTL9 DSE in native buffer and in 8.3 M urea, as well as the relaxation rates calculated on the basis of the phenomenological model developed by Schwalbe and co-workers.[49] The calculated $R_2$ rates exhibit the expected inverted U shape. For the DSE in 8.3 M urea, the $R_2$ rates for most residues show only very small deviations from the phenomenological model (<1.0 s$^{-1}$). However, the $R_2$ rates of the DSE in native buffer display significant positive deviations (>1.0 s$^{-1}$) from the random coil model for a set of residues located in the regions of $\alpha$-helix 1 and $\alpha$-helix 2, and two residues, K7 and K10, located in the loop connecting $\beta$-strand 1 and $\beta$-strand 2 in the native state. This suggests that the side chains of these residues are involved in the formation of clusters in the denatured state that restrict the mobility of the backbone.

**Paramagnetic Relaxation Enhancement Studies Indicate Extensive, Transient, Long-Range Contacts in the DSE in Native Buffer.** Paramagnetic relaxation enhancement via nitroxide spin labeling is sensitive to distances of up to 20 Å and is a valuable technique for probing long-range contacts in protein denatured states.[23,46,54−57] Nitroxide spin-labels were incorporated into V3A/I4A-NTL9 at four solvent-exposed positions (10, 32, 49, and 51). In the native state, K10 is in the


dx.doi.org/10.1021/bi301667u | Biochemistry 2013, 52, 2662−2671

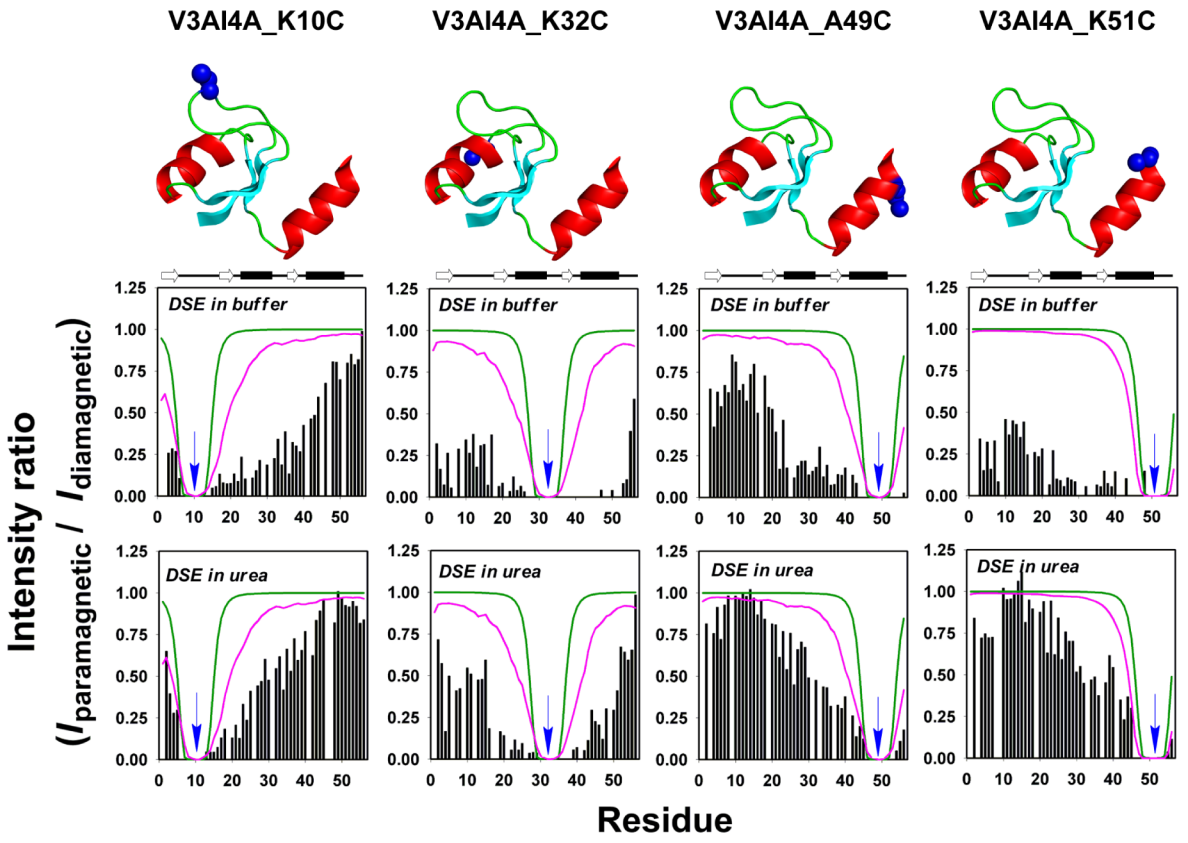



**Figure 5.** Paramagnetic relaxation enhancement data for the DSE of V3A/I4A-NTL9 in native buffer [20 mM sodium acetate and 100 mM sodium chloride (pH 5.5)] at 12 °C and the DSE of V3A/I4A-NTL9 in 8.3 M urea at 12 °C. The histograms display the intensity ratio ($I_{paramagnetic}/I_{diamagnetic}$) of the $^{1}H-^{15}N$ cross-peaks in the HSQC spectra. Values for the DSE in urea are shown in the bottom row; values for the DSE in native buffer are shown in the top row. The location of the spin-label is indicated by a blue sphere. Green lines represent the values expected from the Gaussian random coil model. The purple line represents the values calculated using the excluded volume limit model with the spin-label explicitly accounted for. An arrow indicates the location of each spin-label. Diagrams of the native state secondary structure are shown at top of each column. $\beta$-Strands are depicted as arrows and $\alpha$-helices as bars.

loop connecting $\beta$-strand 1 and $\beta$-strand 2. K32 is located right after $\alpha$-helix 1. A49 and K51 are both located in $\alpha$-helix 2, but their $C_\alpha-C_\beta$ vectors project in opposite orientations: into the solvent for residue 49 and toward the large central loop for residue 51. $^{1}H-^{15}N$ HSQC spectra were collected for both the paramagnetic form and diamagnetic form, and the peak intensity ratios ($I_{paramagnetic}/I_{diamagnetic}$) were calculated to evaluate the paramagnetic broadening. The diamagnetic form was obtained by reduction of the paramagnetic form with TCEP. The expected PRE profiles for a random coil were simulated using a Gaussian model that assumes a Gaussian distribution of distances between each residue and the spin labeling site,[46] and an excluded volume (EV) model, which is a self-avoiding random walk model generated from Monte Carlo simulations using atomistic steric interactions alone.[19,20]

The PRE profiles for the DSE of V3A/I4A-NTL9 in native buffer and in 8.3 M urea, as well as the two simulated random coil models, are displayed in Figure 5. Significant deviations from random coil models are observed for the DSE in both native buffer and urea, but the PRE effects in buffer are much more extensive than those in 8.3 M urea. For the K10C label, significant PRE effects ($I_{paramagnetic}/I_{diamagnetic} < 0.5$) are observed for sites up to 30 residues from the spin-label in buffer. The K32C labeled variant also shows significant PRE effects spread throughout the sequence except for a few residues at the C-terminus. For the A49C label, peak intensity ratios of <0.5 are observed over a range of ~30 residues from the labeling site. For the K51C site, significant PRE effects are observed throughout the sequence. It is interesting to note that there are clear differences between the A49C label and the K51C label even though the sites are only two residues apart. Such differences are not predicted by either the EV limit or the Gaussian model. The large, extensive PRE effects are not due to contributions from a small fraction of the V3A/I4A-NTL9 folded state, because it is in slow exchange with the DSE on the NMR time scale.

The DSE in 8.3 M urea exhibits significantly weaker PRE effects, but they still clearly deviate from the Gaussian chain and EV limit models. The detection of significant PRE effects in the urea-induced DSE confirms our earlier observation that long-range contacts can form in expanded DSEs even in the absence of secondary structure.[20] For K10C, A49C, and K51C, the PRE effects in the urea-induced DSE are mostly localized but are still more extensive than predicted by random coil models. For K32C, some sites up to 20 residues from the labeling site exhibit significant PRE effects, but the overall effects are still clearly weaker than those observed for this site in the DSE populated in native buffer. Notice that the PRE profiles for A49C and K51C are quite similar to each other for the DSE in urea, in contrast to the results obtained for the DSE in buffer.

In summary, the PRE profiles demonstrate that there are extensive transient long-range contacts in the DSE in native


dx.doi.org/10.1021/bi301667u | *Biochemistry* 2013, 52, 2662−2671

buffer, and these long-range contacts are weakened by adding a high concentration of urea.

## DISCUSSION

The denatured state of proteins under physiological conditions is of great interest because it is the DSE most relevant for folding, stability, and aggregation. Because of experimental difficulties, direct characterization of the DSE under native conditions is still very rare. In the study presented here, a destabilized double mutant, V3A/I4A-NTL9, provides a useful model of the protein DSE under native conditions. Secondary shift analysis, SSP scores, $^{15}N$ $R_2$ relaxation rates, and PRE studies demonstrate that the DSE under native conditions contains significant secondary structure and long-range contacts. The residual secondary structure is largely abolished by adding high concentrations of urea. The long-range contacts detected by the PRE studies are weakened in urea, but significant nonlocal effects are observed that are not predicted by simple random coil models. These observations are consistent with our previous hydrodynamic measurements of the DSE of V3A/I4A-NTL9 in native buffer and in different concentrations of urea, which showed that the DSE is compact in native buffer and becomes more extended in urea.[40]

The secondary shift analysis shows that there is substantial $\alpha$-helical propensity in the DSE of V3A/I4A-NTL9 populated in buffer. It is natural to inquire about the origin of these propensities and to ask whether the residual secondary structure is native-like or non-native-like, and if it is stabilized by local interactions. The overall SSP scores are moderately correlated for the folded state and the DSE in native buffer as shown in Figure 6 ($R^2$ = 0.60). However, this is largely due to a strong correlation for $\alpha$-helix 2. A much more significant correlation ($R^2$ = 0.81) is observed for that segment, suggesting that the $\alpha$-helical propensities of this region in the DSE are nativelike. No significant correlation is detected if $\alpha$-helix 2 is excluded from the analysis ($R^2$ = 0.34). AGADIR analysis of the V3A/I4A-NTL9 sequence,[58,59] which calculates the intrinsic helical propensity based on amino acid sequence without considering tertiary interactions, shows that the segment corresponding to $\alpha$-helix 1 has very weak helical propensity (<2%), while the segment corresponding to $\alpha$-helix 2 has moderate intrinsic helical propensity (~10−20%) (Figure S5 of the Supporting Information). Peptide fragment studies of NTL9 are consistent with the AGADIR analysis; a fragment corresponding to $\alpha$-helix 1 was shown to have no measurable helical propensity, while the fragment corresponding to $\alpha$-helix 2 did have significant helical content.[27] Therefore, the helical propensities detected in $\alpha$-helix 2 in the DSE can be rationalized by local sequence effects, while the helical propensity detected in $\alpha$-helix 1 in the DSE can be explained only by longer-range interactions. Along these lines, it is noteworthy that significant PRE effects are observed for residues in the first $\alpha$-helical region for all four spin-labeled constructs.

The paramagnetic relaxation enhancement results reveal that there are extensive long-range contacts in the DSE of V3A/I4A-NTL9 in native buffer. K32C and K51C exhibit more extensive PRE effects than K10C and A49C, although significant deviations from random coil are observed for all four spin-labeled mutants. Wright and co-workers have used a hydrophobicity analysis, the average area buried upon folding (AABUF), to correlate PRE effects to potential hydrophobic clustering.[46,55] They showed that the regions corresponding to

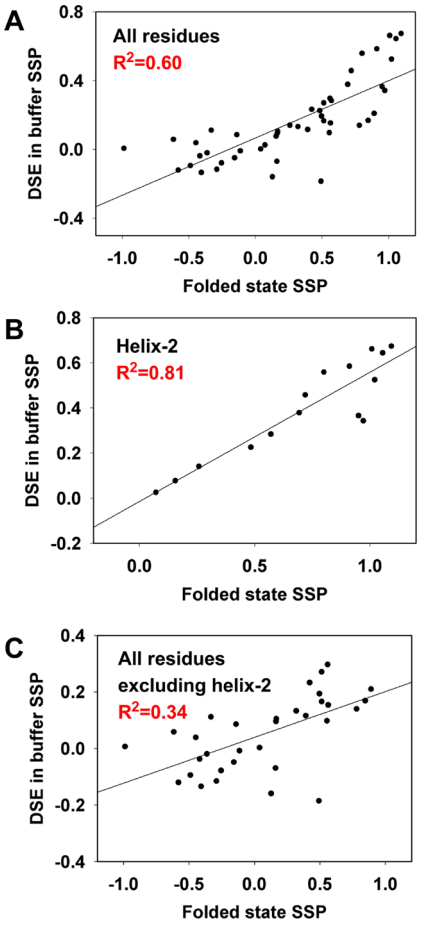



**Figure 6.** Comparison of the SSP scores of the V3A/I4A-NTL9 native state and DSE in buffer for (A) all residues, (B) helix 2, and (C) all residues excluding helix 2.

AABUF maxima in the apomyoglobin sequence exhibit more extensive PRE effects in the acid unfolded state, suggesting that they could be more buried in this state. The calculated AABUF scores for V3A/I4A-NTL9 are displayed in Figure 7. Residue 32 is located at the maximum of the AABUF plot, and extensive PRE effects are observed throughout the sequence for the K32 spin-label in the DSE in native buffer. Residue 10 is located close to the minimum of the AABUF plot, and the PRE effects of the K10C label are less extensive than those of the K32C label although still significant. Residue 49 and residue 51 are close in sequence, and their AABUF scores are both close to the average value. However, as noted, these two residues are located in $\alpha$-helix 2, which is largely formed in the DSE. Thus, the hydrophobicity of this region cannot be analyzed simply on the basis of sequence. In the native state structure, K51 is located on the same face of $\alpha$-helix 2 as two large hydrophobic residues, L44 and L47. This face is remarkably more hydrophobic (average AABUF of 148) than the other face on which A49 resides (average AABUF of 96). In the native state, the hydrophobic face of the helix, which includes K51, is oriented toward the center of the domain and packs against it. The K51C label exhibits more extensive PRE effects than the A49C label, showing that the hydrophobic face is involved in hydrophobic clustering in the DSE. We proposed that the orientation and topology of $\alpha$-helix 2 sample transient nativelike conformations in the DSE in native buffer. At high concentrations of urea, the helical propensities are abolished


dx.doi.org/10.1021/bi301667u | Biochemistry 2013, 52, 2662−2671

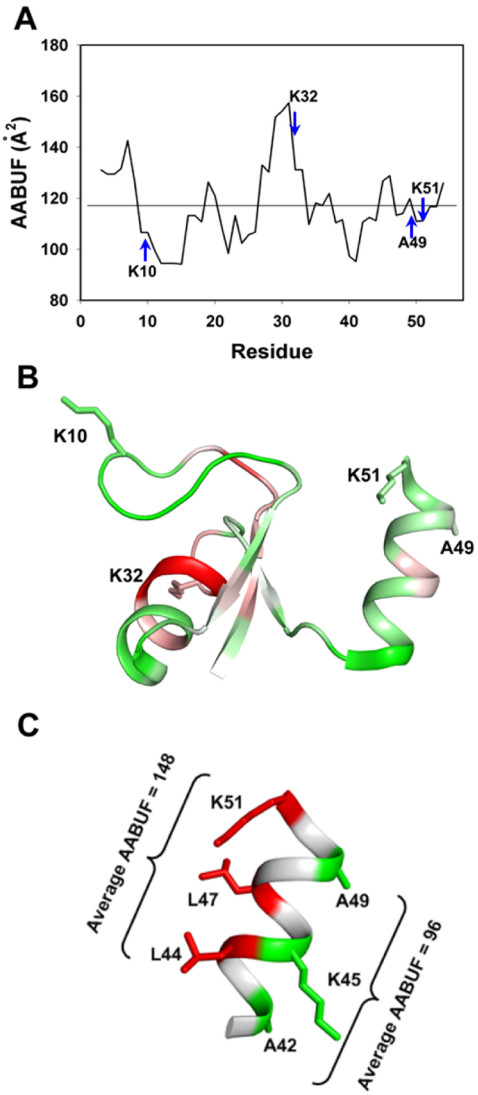



**Figure 7.** (A) Average area buried upon folding (AABUF) of V3A/I4A-NTL9 calculated with a five-residue moving average. The horizontal line is the mean value of AABUF for the entire sequence. (B) Ribbon diagram of V3A/I4A-NTL9 with residues colored according to AABUF scores from panel A; regions with high AABUF scores are colored red and regions with low AABUF scores green. (C) Ribbon diagram of $\alpha$-helix 2 in the native state of NTL9 showing that one face of the helix has high AABUF scores and the other has low AABUF scores. Note that color coding used in panel B corresponding to a five residue moving average, while it corresponds to individual residue values in panel C.

and the number of long-range contacts is reduced, and accordingly, the PRE effects from the K51C label and A49C label are similar. The $^{15}$N $R_2$ data are consistent with the PRE measurements and provide additional evidence of long-range interactions in the DSE. Enhanced relaxation rates are thought to be caused by hydrophobic clustering in the DSE, although they may also arise because of local secondary structure formation.[26] In our case, all of the residues that show enhanced $^{15}$N $R_2$ values also show significant PRE effects. Thirteen residues exhibit $R_2$ values that differ from the phenomenological model by $\geq$1 s$^{-1}$. All of these experience significant PRE effects, defined as an $I_{paramagnetic}/I_{diamagnetic}$ value of $\leq$0.5, from spin-labels at least 11 residues away (Figure 5 and Table S1 of the Supporting Information). Furthermore, the spin-labels are in different elements of secondary structure. This suggests that the $R_2$ enhancements are also reporting on hydrophobic cluster formation. Wu et al.[60] have proposed that clusters involving Val, Leu, Phe, and Ile form early in folding and are present in DSEs, and our data are consistent with this conjecture.

The DSE of the double mutant that is populated in buffer is considerably more structured than the urea-induced DSE. The analysis presented here shows that DSE in native buffer is compact, contains significant secondary structure, and makes extensive transient long-range contacts. The DSE populated by the wild-type protein is expected to contain even more residual structure because the mutation involves the truncation of two large hydrophobic residues, which should reduce the level of hydrophobic clustering in the DSE. In particular, I4 packs against residues located on one face of $\alpha$-helix 2, and we observed extensive PREs from this region in the double mutant (Figure 5).

The population of a compact DSE with significant local and nonlocal contacts may contribute to the rapid folding of wild-type NTL9. The formation of compact conformations stabilized by extensive transient nativelike intramolecular contacts should reduce the scope of the conformational search required in the early steps of folding, as should the existence of significant native secondary structure. In contrast, the formation of non-native interactions that stabilize the DSE should slow folding. NTL9 is a complicated case; this work documents the presence of nativelike DSE interactions, while earlier studies have shown that K12 is involved in energetically important non-native interactions in the DSE.[39,40,61] Furthermore, studies of wild-type NTL9 under denaturing conditions have revealed the presence of extensive, long, and midrange non-native contacts in the DSE.[20] Thus, the DSE of NTL9 contains both native and non-native interactions. The preferential formation of intramolecular contacts in the DSE may contribute to the high solubility of NTL9 by weakening the tendency to form deleterious intermolecular contacts that might promote aggregation.

It is commonly observed that the folding rates measured after dilution from denaturant, after a pH jump, or after a temperature jump are equal provided the same final conditions are reached. DSE structure clearly depends upon conditions; thus, it is natural to enquire if such studies indicate that DSE interactions are unimportant for folding. This does necessarily follow; the reequilibration of a protein chain after a pH, temperature, or denaturant jump can be very rapid relative to the folding time; thus, folding will essentially start from the same state in each case, and the observation of similar rates is not surprising. The folding of NTL9 occurs on the millisecond time scale, which is considerably faster than the time expected for a collapse transition. Modulation of DSE interactions by mutation has been shown to affect the folding rate of NTL9, indicating that DSE structure is important.[39,40,61] Perturbation of DSE energetics by the mutant can also significantly complicate $\phi$ value analysis of protein folding transition states.[62,63]

## ■ ASSOCIATED CONTENT

### Supporting Information

Table S1, Figures S1−S5. This material is available free of charge via the Internet at http://pubs.acs.org.


dx.doi.org/10.1021/bi301667u | *Biochemistry* 2013, 52, 2662−2671

## AUTHOR INFORMATION


### Corresponding Author

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



### Funding

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


### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS


We thank Dr. Jaehyun Cho for helpful discussions and Dr. Shibani Bhattacharya and Dr. Mark W. Maciejewski for help with NMR experiments. We thank Prof. Jeff Hoch for generously providing access to the NMR facility at the University of Connecticut Health Center (Farmington, CT).


## ABBREVIATIONS

CD, circular dichroism; DSE, denatured state ensemble; EV, excluded volume; HSQC, heteronuclear single-quantum coherence spectroscopy; MTSL, (1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)methanesulfonate; NTL9, the first 56 residues of ribosomal protein L9; PRE, paramagnetic relaxation enhancement; $R_2$, transverse relaxation rate; $R_h$, hydrodynamic radius; $R_g$, radius of gyration; TCEP, tris(2-carboxyethyl)-phosphine; V3A/I4A-NTL9, double mutant of NTL9 in which V3 and I4 are mutated to alanine.

## REFERENCES


(1) Baldwin, R. L. (2002) A new perspective on unfolded proteins. *Adv. Protein Chem. 62*, 361−367.

(2) Shortle, D. (1996) The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB J. 10*, 27−34.

(3) Bowler, B. E. (2007) Thermodynamics of protein denatured states. *Mol. BioSyst. 3*, 88−99.

(4) Choy, W. Y., Mulder, F. 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) 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.

(6) Anil, B., Song, B., Tang, Y., 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.

(7) Zhou, H. X., Rivas, G., and Minton, A. P. (2008) Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys. 37*, 375−397.

(8) Jahn, T. R., and Radford, S. E. (2008) Folding versus aggregation: Polypeptide conformations on competing pathways. *Arch. Biochem. Biophys. 469*, 100−117.

(9) 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.

(10) Polverino de Laureto, 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.

(11) Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol. 6*, 197−208.

(12) Uversky, V. N. (2002) Natively unfolded proteins: A point where biology waits for physics. *Protein Sci. 11*, 739−756.

(13) 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.

(14) Yao, J., Chung, J., Eliezer, D., Wright, P. E., and Dyson, H. J. (2001) NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. *Biochemistry 40*, 3561−3571.

(15) 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.

(16) 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.

(17) 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.

(18) Tran, H. T., Wang, X., and Pappu, R. V. (2005) Reconciling observations of sequence-specific conformational propensities with the generic polymeric behavior of denatured proteins. *Biochemistry 44*, 11369−11380.

(19) 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.

(20) 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.

(21) Tollinger, M., Skrynnikov, N. R., Mulder, F. A., Forman-Kay, J. D., and Kay, L. E. (2001) Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc. 123*, 11341−11352.

(22) 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.

(23) 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.

(24) Zhang, O., and Forman-Kay, J. D. (1995) Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer. *Biochemistry 34*, 6784−6794.

(25) Fleming, P. J., and Rose, G. D. (2008) Conformational Properties of Unfolded Proteins. *Protein Science Encyclopedia*, 710−736.

(26) Chugha, P., and Oas, T. G. (2007) Backbone dynamics of the monomeric $\lambda$ repressor denatured state ensemble under non-denaturing conditions. *Biochemistry 46*, 1141−1151.

(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) 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.

(29) Cho, J. H., and Raleigh, D. P. (2005) Mutational analysis demonstrates that specific electrostatic interactions can play a key role in the denatured state ensemble of proteins. *J. Mol. Biol. 353*, 174−185.

(30) Wright, P. E., Dyson, H. J., and Lerner, R. A. (1988) Conformation of peptide fragments of proteins in aqueous solution:

Implications for initiation of protein folding. *Biochemistry 27*, 7167−7175.
(31) Viguera, A. R., Jimenez, M. A., Rico, M., and Serrano, L. (1996) Conformational analysis of peptides corresponding to $\beta$-hairpins and a $\beta$-sheet that represent the entire sequence of the $\alpha$-spectrin SH3 domain. *J. Mol. Biol. 255*, 507−521.
(32) Ladurner, A. G., Itzhaki, L. S., de Prat Gay, G., and Fersht, A. R. (1997) Complementation of peptide fragments of the single domain protein chymotrypsin inhibitor 2. *J. Mol. Biol. 273*, 317−329.
(33) Mayor, U., Guydosh, N. R., Johnson, C. M., Grossmann, J. G., Sato, S., Jas, G. S., Freund, S. M., Alonso, D. O., Daggett, V., and Fersht, A. R. (2003) The complete folding pathway of a protein from nanoseconds to microseconds. *Nature 421*, 863−867.
(34) Kuhlman, B., Luisi, D. L., Evans, P. A., and Raleigh, D. P. (1998) Global analysis of the effects of temperature and denaturant on the folding and unfolding kinetics of the N-terminal domain of the protein L9. *J. Mol. Biol. 284*, 1661−1670.
(35) Luisi, D. L., and Raleigh, D. P. (2000) pH-dependent interactions and the stability and folding kinetics of the N-terminal domain of L9. Electrostatic interactions are only weakly formed in the transition state for folding. *J. Mol. Biol. 299*, 1091−1100.
(36) Anil, B., Sato, S., Cho, J. H., and Raleigh, D. P. (2005) Fine structure analysis of a protein folding transition state; distinguishing between hydrophobic stabilization and specific packing. *J. Mol. Biol. 354*, 693−705.
(37) Taskent-Sezgin, H., Marek, P., Thomas, R., Goldberg, D., Chung, J., Carrico, I., and Raleigh, D. P. (2010) Modulation of p-cyanophenylalanine fluorescence by amino acid side chains and rational design of fluorescence probes of $\alpha$-helix formation. *Biochemistry 49*, 6290−6295.
(38) Anil, B., Li, Y., Cho, J. H., and Raleigh, D. P. (2006) The unfolded state of NTL9 is compact in the absence of denaturant. *Biochemistry 45*, 10110−10116.
(39) 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.
(40) Meng, W., and Raleigh, D. P. (2011) Analysis of electrostatic interactions in the denatured state ensemble of the N-terminal domain of L9 under native conditions. *Proteins 79*, 3500−3510.
(41) 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.
(42) Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol. 278*, 313−352.
(43) Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) $^{1}$H, $^{13}$C and $^{15}$N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR 5*, 67−81.
(44) Schwarzinger, S., Kroon, G. J., Foss, T. R., Chung, J., Wright, P. E., and Dyson, H. J. (2001) Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc. 123*, 2970−2978.
(45) Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006) Sensitivity of secondary structure propensities to sequence differences between $\alpha$- and $\gamma$-synuclein: Implications for fibrillation. *Protein Sci. 15*, 2795−2804.
(46) 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.
(47) Sung, Y. H., and Eliezer, D. (2007) Residual structure, backbone dynamics, and interactions within the synuclein family. *J. Mol. Biol. 372*, 689−707.
(48) Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by $^{15}$N NMR relaxation. *Biochemistry 33*, 5984−6003.
(49) Schwalbe, H., Fiebig, K. M., Buck, M., Jones, J. A., Grimshaw, S. B., Spencer, A., Glaser, S. J., Smith, L. J., and Dobson, C. M. (1997) Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry 36*, 8977−8991.
(50) 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.
(51) Guinier, A., and Fournet, G. (1955) *Small Angle Scattering of X-rays*, Wiley, New York.
(52) Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol. 222*, 311−333.
(53) Spera, S., and Bax, A. (1991) Empirical correlation between protein backbone conformation and C$\alpha$ and C$\beta$ $^{13}$C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc. 113*, 5490−5492.
(54) 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.
(55) 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.
(56) Dedmon, M. M., Lindorff-Larsen, K., Christodoulou, J., Vendruscolo, M., and Dobson, C. M. (2005) Mapping long-range interactions in $\alpha$-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *J. Am. Chem. Soc. 127*, 476−477.
(57) 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.
(58) Munoz, V., and Serrano, L. (1995) Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides. *J. Mol. Biol. 245*, 275−296.
(59) Munoz, V., and Serrano, L. (1994) Elucidating the folding problem of helical peptides using empirical parameters. *Nat. Struct. Biol. 1*, 399−409.
(60) 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.
(61) Cho, J. H., and Raleigh, D. P. (2006) Electrostatic interactions in the denatured state and in the transition state for protein folding: Effects of denatured state interactions on the analysis of transition state structure. *J. Mol. Biol. 359*, 1437−1446.
(62) Cho, J. H., and Raleigh, D. P. (2006) Denatured state effects and the origin of nonclassical $\phi$ values in protein folding. *J. Am. Chem. Soc. 128*, 16492−16493.
(63) Raleigh, D. P., and Plaxco, K. W. (2005) The protein folding transition state: What are $\phi$-values really telling us? *Protein Pept. Lett. 12*, 117−122.