

High Pressure ZZ-Exchange NMR Reveals Key Features of Protein Folding Transition States.

Y. Zhang^{1,‡}, S. Kitazawa^{2,†,‡}, I. Peran³, N. Stenzoski⁴, S. A. McCallum⁵, D. P. Raleigh^{3,&} and C. A. Royer^{1,2,&}

¹Department of Chemistry & Chemical Biology, Rensselaer Polytechnic Institute, Troy NY 12180; ²Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy NY 12180; ³Department of Chemistry, Stony Brook University, Stony Brook, NY 11794; ⁴Graduate Program in Biochemistry and Structural Biology, Stony Brook University, Stony Brook, NY 11794; ⁵NMR Core Facility, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy NY 12180

KEYWORDS Protein folding, Ribosomal protein L9, NMR, Pressure, Activation volume.

ABSTRACT: Understanding protein folding mechanisms and their sequence dependence requires the determination of residue-specific apparent kinetic rate constants for the folding and unfolding reactions. Conventional two dimensional NMR, such as HSQC experiments, can provide residue-specific information for proteins. However, folding is generally too fast for such experiments. ZZ-exchange NMR spectroscopy allows determination of folding and unfolding rates on much faster timescales, yet even this regime is not fast enough for many protein folding reactions. The application of high hydrostatic pressure slows folding by orders of magnitude due to positive activation volumes for the folding reaction. We combined high pressure perturbation with ZZ-exchange spectroscopy on two autonomously folding protein domains derived from the ribosomal protein, L9. We obtained residue-specific apparent rates at 2500 bar for the N-terminal domain of L9 (NTL9), and rates at atmospheric pressure for a mutant of the C-terminal domain (CTL9) from pressure dependent ZZ-exchange measurements. Our results revealed that NTL9 folding is almost perfectly two-state, while small deviations from two-state behavior were observed for CTL9. Both domains exhibited large positive activation volumes for folding. The volumetric properties of these domains reveal that their transition states contain most of the internal solvent excluded voids that are found in the hydrophobic cores of the respective native states. These results demonstrate that by coupling it with high pressure, ZZ-exchange can be extended to investigate a large number of protein conformational transitions.

Introduction

The observation of residue specific folding kinetics is required for detailed descriptions of protein folding mechanisms and their sequence dependence. Knowledge of the sequence dependence of folding mechanisms illuminates how they may have evolved, and moreover, should provide key insights for protein design. Unfortunately high resolution methods such as NMR have insufficient time resolution to follow protein folding in real time expect for very slow folding proteins. Methods for rapid kinetic measurements, stopped-flow and continuous-flow, rely largely on fluorescence detection and do not provide residue specific information. Rapid temperature jump methods combined with IR detection and specific labeling provide residue specific information, but are limited by the requirement for site specific ¹³C/¹⁸O labeling. Over the past few years, pressure has garnered increasing interest as a thermodynamic variable, since it was demonstrated to perturb protein folding equilibria due to a loss of internal solvent excluded void volume upon unfolding ¹. Combining high pressure (HP)

with 2D NMR allows for observation of protein folding on a residue specific basis, and has been used at equilibrium to explore deviations from two-state behavior and to map protein folding landscapes ¹⁻³.

The major destabilizing effect of pressure on protein folding equilibria derives from a large positive activation volume for folding. This slows the folding reaction considerably ⁴. In cases for which folding is intrinsically sufficiently slow and folding activation volumes sufficiently large ^{4,5}, pressure-jump can be combined with classical 2D NMR experiments, such as ¹H-¹⁵N HSQC, to study kinetic mechanisms of folding, yielding folding and unfolding rates for nearly every residue in a protein. Pressure-jump 2D HSQC experiments on staphylococcal nuclease and a series of variants was possible because folding relaxation times at high pressure are extremely slow, ranging between 20 minutes and 24 hours ⁶. Results from these studies revealed significant changes in the structural properties of the SNase folding transition state and the complexity of its folding pathway brought about by single site mutations.

In contrast to these slow folding nuclease variants, most small globular proteins fold rather quickly, many on time-scales of milliseconds or even less. For a protein with a folding time of 1 ms ($k_f = 1000 \text{ s}^{-1}$) at atmospheric pressure, and a positive activation volume of 100 ml/mol, folding will be slowed to about 1 min at a pressure of 2500 bar. Unfortunately, this timescale still remains too fast for acquisition of classical 2D NMR spectra as a function of time. Conformational exchange phenomena that are slow on the NMR time-scale (100 ms– ~1 second) can be accurately quantified by 2D exchange NMR⁷. One such approach to chemical exchange is ZZ-exchange spectroscopy, which was first used to determine tyrosine ring flipping⁸ and was later extended to measure the folding and unfolding rates of the N-terminal SH3 domain of the protein *Drk*⁹. However, most small proteins fold on faster timescales than this SH3 domain, and hence ZZ exchange spectroscopy has not been widely applied to characterizing folding reactions.



Figure 1. Ribbon diagram of L9 from *Bacillus stearothermophilus* (Protein Data Bank entry 1DIV). The N-terminal domain (NTL9), consisting of residues 1–56, is colored blue. The C-terminal domain (CTL9), corresponding to residues 58–149, is colored red. The hydrophobic core residue 198, colored green, and the two termini are labeled. The diagram was created with PyMOL.¹⁰

Here we sought to rigorously test the utility of combining high pressure with ZZ-exchange NMR (HPZZEx) for measuring protein folding kinetics on a residue specific basis. We tested HPZZEx using two small globular proteins as model systems, the N and C-terminal domains of the ribosomal protein L9, from *Bacillus stearothermophilus*, (Figure 1) whose folding properties at atmospheric pressure have been extensively characterized^{11–21}. The N-terminal domain of L9, NTL9, is made up of residues 1–56 of L9, and folds into a three-stranded anti-parallel beta sheet sandwiched between two α -helices, one of which is part of the central helix connecting the two domains in the full-length protein. The C-terminal domain, CTL9, is made up of residues 58–149 and adopts an interesting 3-stranded mixed parallel and anti-parallel beta-sheet buttressed by two α -helices, one of which includes the remaining residues from the central helix of the full-length protein. Both NTL9 and CTL9 fold cooperatively in experiments conducted at atmospheric pressure, in an apparent two-state manner^{11,21}. At room temperature NTL9 folds in approximately 1 ms²¹. We used a destabilized variant of CTL9, CTL9 I98A, for the studies reported here. This variant has been shown to fold to the same structure as WT CTL9 and the folding time of CTL9 I98A is ~70 ms at 25 °C and pH 8¹¹. We predicted

that the activation volumes for the folding of these domains should be sufficiently large such that their folding/unfolding relaxation times at high pressure would be in a time regime appropriate for ZZ-exchange spectroscopy. The results demonstrate the utility of the approach and importantly provide unique insight into the packing of sidechains in the transition state for folding.

Materials and Methods

Protein Expression and Purification

¹⁵N-labeled NTL9 and the ¹⁵N-labeled CTL9 I98A variant were overexpressed and purified as previously described¹⁷. The identity of the proteins was confirmed with MALDI-TOF mass spectrometry and the purity was checked using analytical HPLC.

Nuclear Magnetic Resonance (NMR) Experiments

¹⁵N-labeled NTL9 protein samples for NMR experiments were prepared in 90% H₂O and 10% D₂O at pH 5.1, with 20 mM sodium acetate, 100 mM NaCl, and 1 M GuHCl. High pressure ZZ-exchange spectra on NTL9 were recorded at 285 K. ¹⁵N-labeled CTL9 I98A protein samples were dissolved in 90% H₂O and 10% D₂O at pH 6.6, with 20 mM Bis-tris and 100 mM NaCl. High pressure ZZ-exchange experiments on CTL9 I98A were recorded at 308 K. The protein concentration was ~1 mM for both samples. DSS was used as a reference for the chemical shift. All NMR experiments were recorded on Bruker Avance III 600MHz/14T Wide-bore NMR spectrometer equipped with a 5mm TBI probe optimized for ¹H detection with ¹H, ²H, ¹³C fixed frequency RF channels and a broadband tunable channel that was configured to ¹⁵N for these studies. A commercial ceramic zirconia high-pressure NMR cell^{22,23} and an automatic pump system (Daedalus Innovations, Philadelphia, PA) were used to vary the pressure in the 1 to 2500 bar range. The ¹H–¹⁵N ZZ-exchange spectra were collected using a variation of the ¹⁵N magnetization exchange pulse sequence reported by the Kay group⁹ and provided by Bruker (hsqctexf3gp). For NTL9, the spectra were recorded at 2500 bar, with 8 scans, 2048 × 256 complex points with a 9615 and 2190 Hz spectral width for the ¹⁵N and ¹H dimensions, respectively. In order to quantify the exchange rates, the following mixing times were acquired in random order: 10, 20, 40, 60, 80, 100, 150, 200, 250, 300, 350, 400, 500 ms. For CTL9 I98A, the spectra were recorded at 300, 450, and 600 bar. At each pressure, mixing times of 10, 40, 60, 80, 100, 120, 150, 200, 250, 300, 400, 500 ms were collected in random order. Every spectrum was collected using 32 scans, 2048 × 256 complex points and a 9615 and 1886 Hz spectral width for the ¹⁵N and ¹H dimensions, respectively. A series of ¹H–¹⁵N HSQC spectra for the CTL9 I98A mutant were collected 15 mins after each positive 150 bar pressure jump from 1 bar to 2100 bar. The spectra were recorded with 2048 × 256 complex points using 4 scans and spectral widths of 9615 and 1886 Hz for the ¹⁵N and ¹H dimensions, respectively. All the spectra were processed with Topspin and analyzed with NMRViewJ or CcpNmr²⁴.

To calculate the folding and unfolding rate constants, peak heights were used to characterize the evolution of the magnetization of peaks as a function of mixing time. The equations, for a two-state model, describing the intensity decay due to longitudinal relaxation for the folded (F, Equation 1) and unfolded (U, Equation 2) peaks, and the emergence due to chemical exchange and longitudinal relaxation of the cross peaks from F to U (Equation 3) and from U to F (Equation 4), have been described before⁹ and are given below.

$$I_{FF}(\tau) = I_F(0)(-\lambda_2 - a_{11})e^{-\lambda_1\tau} + (\lambda_1 - a_{11})e^{-\lambda_2\tau}/(\lambda_1 - \lambda_2) \quad (1)$$

$$I_{UU}(\tau) = I_U(0)(-\lambda_2 - a_{22})e^{-\lambda_1\tau} + (\lambda_1 - a_{22})e^{-\lambda_2\tau}/(\lambda_1 - \lambda_2) \quad (2)$$

$$I_{FU}(\tau) = I_F(0)(a_{21}e^{-\lambda_1\tau} - a_{21}e^{-\lambda_2\tau})/(\lambda_1 - \lambda_2) \quad (3)$$

$$I_{UF}(\tau) = I_U(0)(a_{12}e^{-\lambda_1\tau} - a_{12}e^{-\lambda_2\tau})/(\lambda_1 - \lambda_2) \quad (4)$$

$I_F(0)$ and $I_U(0)$ denote the equilibrium magnetization of the folded and unfolded peaks, τ the mixing time, and $\lambda_{1,2}$ the eigenvalues of the 2x2 dynamics matrix ($a_{1,2}$) defined by equation 5,

$$\lambda_{1,2} = \frac{1}{2}\sqrt{(a_{11} + a_{22}) \pm [(a_{11} - a_{22})^2 + 4k_f k_u]} \quad (5)$$

where

$$a_{11} = R_F + k_u; a_{22} = R_U + k_f; a_{21} = -k_u; a_{12} = -k_f.$$

R_F and R_U are the longitudinal relaxation rates for the folded and unfolded signal. k_f and k_u are the folding and unfolding rate constants, respectively.

The pressure-dependent ZZ-exchange experiments were performed on CTL9 I98A at 300, 450, and 600 bar. To extrapolate the rate constants to atmospheric pressure ($k_{f,u}^0$) and obtain the activation volumes ($\Delta V_{f,u}^*$), the natural logarithm of the rate constants ($\ln k_{f,u}(p)$) were fit to the following linear equations with the pressure (p).

$$\ln k_f(p) = \ln k_f^0 - (p - p_0)\Delta V_f^*/RT \quad (6)$$

$$\ln k_u(p) = \ln k_u^0 - (p - p_0)\Delta V_u^*/RT \quad (7)$$

where p_0 is atmospheric pressure (1 bar). The folding free energy (ΔG_f^0) and volume change for unfolding (ΔV_u^0) were calculated using the following relations:

$$\Delta G_f^0 = -RT \ln \frac{k_f^0}{k_u^0} \quad (8)$$

$$\Delta V_u^0 = \Delta V_u^* - \Delta V_f^* \quad (9).$$

Results

Residue-specific apparent folding and unfolding rates were measured for NTL9 using ZZ-exchange experiments conducted in 20 mM sodium acetate, 100 mM NaCl buffer containing 1 M GuHCl at pH 5.1, 285 K and 2500 bar. The buffer has a low pressure coefficient²⁵ and under these conditions NTL9 is near its unfolding midpoint. Figure 2 shows the ZZ-exchange spectrum of NTL9 at 350 ms mixing time. The figure features peaks corresponding to the amide groups of several residues in the folded and unfolded states, as well as their exchange cross peaks. ZZ-exchange cross-peaks could be quantified as a function of mixing time for 36 of the 56 residues of NTL9 at 13 mixing times between 10 and 500 ms. The ZZ-exchange profiles for each of these residues (Figure 3) were fit to determine the residue-specific apparent folding and unfolding rate constants and the longitudinal relaxation rates as described in the Material and methods section.

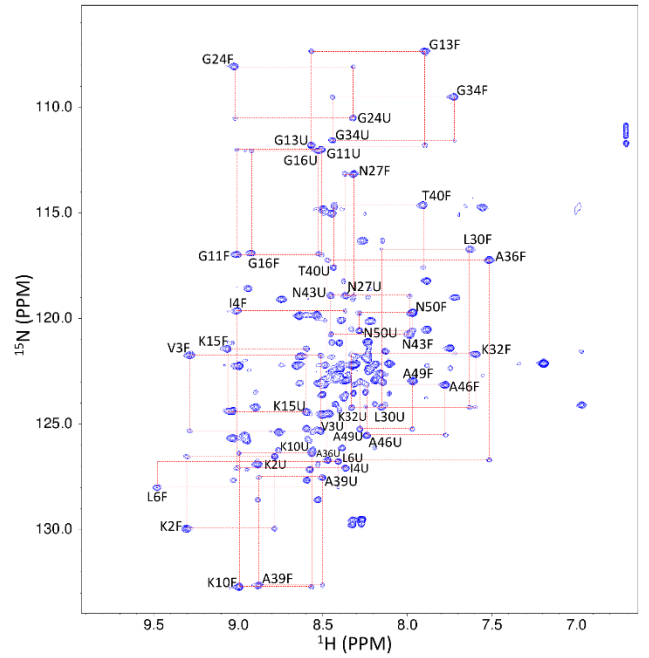


Figure 2. ZZ-exchange spectrum of NTL9 taken at 600 MHz and 350 ms mixing time for 1 mM ¹⁵N labeled NTL9 in 20 mM sodium acetate, 100 mM NaCl and 1 M GuHCl, at pH 5.1, 285 K and 2500 bar. Peaks for the folded and unfolded forms of the residue are marked F and U, respectively and dotted lines indicate the exchange cross-peaks.

Figure 4 shows the residue-specific apparent folding and unfolding rates obtained from the ZZ-exchange data at 2500 bar. This pressure is near the folding midpoint, hence the average values for the folding and unfolding rate constants are similar (1.1 and 1.3 s⁻¹, respectively). Most residues monitored (32 out of the 36) exhibited apparent rates that deviated from the average k_u or k_f by less than 2-3 fold the average standard deviation. Only residues 17 and 38 exhibited apparent unfolding rates that were significantly higher than the average, and residues 27, 38 and 54 exhibited

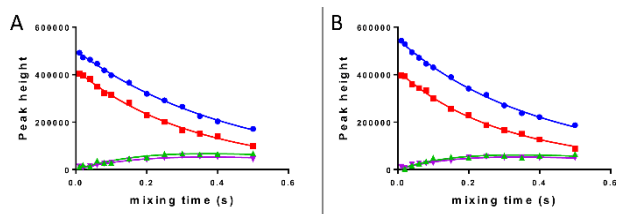


Figure 3. Plots of the peak height and fitted intensity curves for residues G13 (A) and A36 (B) of NTL9. The top two sets of data indicate the peak intensity decay of the folded state (F, blue) and the unfolded state (U, red), due to longitudinal relaxation. The green line and points represent the emergence of the cross peaks as a result of F to U exchange, and the purple represents that of U to F exchange. Data were obtained under conditions as noted in Figure 2.

larger than average apparent folding rates. Residue 54 is in the C-terminal helix which is partially formed in the unfolded state ^{26,27}. The role of E38 in folding has been previously probed by mutation and no significant effect upon the folding or the unfolding rate was observed for the E38A mutant, arguing that interactions made by E38 are not critical for folding or unfolding ²⁸. The reason for these deviations is unclear, however, and only one set of rate constants, at a single pressure, were obtained in these experiments.

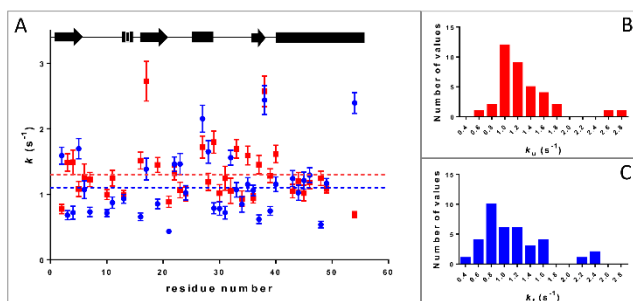


Figure 4. Residue specific folding and unfolding rate constants for NTL9 at 285 K and 2500 bar. The blue dashed line corresponds to the average value of k_f at 1.1 s^{-1} , and the red dashed line represents that of k_u at 1.3 s^{-1} . A schematic representation of the elements of secondary structure of the folded state of NTL9 is shown at the top of the plot A (arrows represent β -strands and filled cylinders α -helices, dashed cylinders 3_{10} -helices, and single lines loop regions).

Pressure-dependent ZZ-exchange on the CTL9 I98A variant

What is of most interest in protein folding studies are the rate constants for folding and unfolding under native conditions, rather than values of rate constants under perturbing conditions. Residue-specific values for apparent folding and unfolding rate constants at 1 bar can be obtained by measuring the pressure dependence of ZZ-exchange and extrapolating the results to atmospheric pressure as described in the materials and methods section. A better estimation of the uncertainty is obtained for these extrapolated rates compared to the single pressure values for NTL9 because they are deduced from more than one measurement. The stability of the NTL9 domain and the WT CTL9 domain (data not shown) at high pressure did not allow us to establish the pressure-dependence of the kinetics, however a cavity containing variant of the C-terminal

domain of L9, CTL9 I98A, ²⁹ exhibited the appropriate stability and kinetic properties for ZZ-exchange measurements at multiple pressures.

The pressure dependence of the ^1H - ^{15}N HSQC spectra of CTL9 I98A in 20 mM Bis-tris, 100 mM NaCl buffer, at 308 K and pH 6.6 (Figure S1) revealed that, in addition to pressure dependent shifts in the amide resonance frequencies due to linear compression, pressure leads to unfolding,

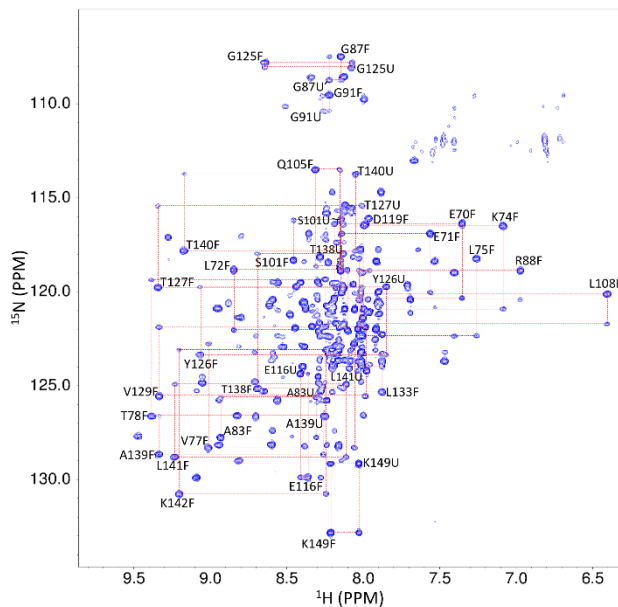


Figure 5. ZZ-exchange spectrum taken at 600 MHz and 200 ms mixing time for 1 mM ^{15}N labeled CTL9 I98A in 100 mM NaCl, 20 mM Tris-bis at pH 6.6, 308K and 450 bar.

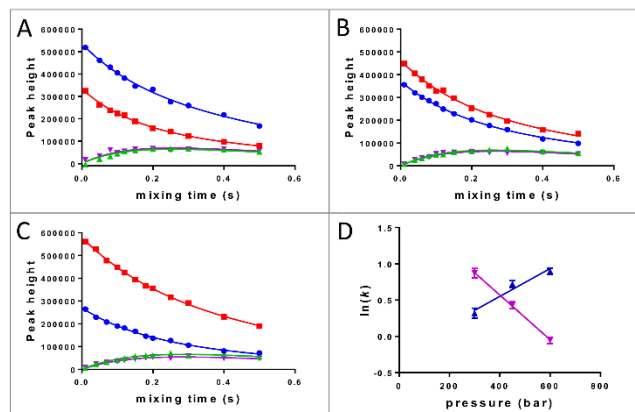


Figure 6. ZZ-exchange data for residue L72 of the CTL9 I98A variant at different pressures. Peak height and fitted curves are represented in the same fashion as in Figure 3. (A) 300 bar, (B) 450 bar, and (C) 600 bar. (D) The natural logarithm of folding rate constants (magenta) and unfolding rate constants (navy blue) are plotted as a linear function of pressure.

with a loss of native state peak intensities, concomitant with an increase in the unfolded state peak intensities. This demonstrates that the folded and unfolded states of CTL9 I98A are in slow exchange on the NMR timescale under pressure. At 308 K and pressures of 300, 450 and 600 bar, the folding reaction is close to the midpoint, allowing reli-

able quantification of the ZZ-exchange signal for many amide resonances of CTL9 I98A (Figure 5, Figure S2) at all three pressures. Plots of peak intensities vs mixing time at each of the three pressures (Figure 6a-c) were fit for the folding and unfolding rate constants as described in the materials and methods section. The pressure dependence of the natural logarithm of these rate constants was linear in all cases (Figure 6d, and Figure S3). The intercepts yielded the residue-specific apparent folding and unfolding rate constants of CTL9 I98A at atmospheric pressure. Residue-specific apparent activation volumes for folding and unfolding were calculated from the slopes.

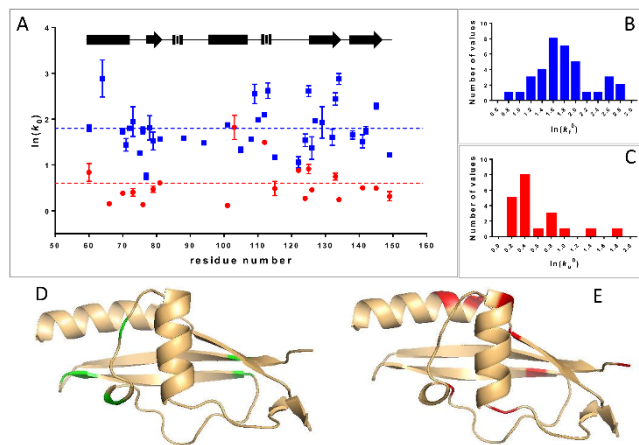


Figure 7. Exchange rates at ambient pressure of CTL9 I98A. A: The natural logarithm of the apparent folding rates ($\ln(k_f^0)$) and unfolding rates ($\ln(k_u^0)$) for each analyzed residue are illustrated with blue and red data points, respectively. Only those values for which the experimental uncertainty was less than 30% of the value were retained. The blue dashed line corresponds to the average value of $\ln(k_f^0)$ at 1.8, and the red dashed line represents that of $\ln(k_u^0)$ at 0.6. A schematic representation of the elements of secondary structure of the folded state of CTL9 is shown at the top of the plot in A (arrows represent β -strands and filled cylinders α -helices, dashed cylinders 3_{10} -helices, and single lines loop regions). B: the histogram of residue specific values for $\ln(k_f^0)$. C: the histogram of residue specific values for $\ln(k_u^0)$. D and E: Schematic representation of the largest folding rates (green) (D) and smallest folding rates (red) (E) on the CTL9 structure. Unfolding rates were less broadly distributed and not presented on the structure.

The average folding rate at atmospheric pressure, 303 K and pH 6.6 calculated for CTL9-I98A from the extrapolated values of the HPZZEx for all residues was found to be $6.0 (+3.9/-2.4) \text{ s}^{-1}$, while the average unfolding rate constant at 1 bar was $1.8 (+1.0/-0.6) \text{ s}^{-1}$. Here we report only those rates for which the uncertainty was less than 30% of the extrapolated value (Tables S1 and S2). Denaturant based kinetics experiments¹¹ at pH 8 and 293 K, conditions under which the protein is more stable, observing tryptophan fluorescence yielded average values of 14 and 0.6 s^{-1} for the folding and unfolding rate constants, respectively. Thus, the values obtained by HPZZEx are consistent with the known kinetics for folding/unfolding of CTL9-I98A. The average folding free energy calculated from the average pressure-derived rate constants, 0.75 kcal/mol , differs by about 1.0 kcal/mol from that obtained from the rate constants measured by stopped-flow at pH 8 and 293 K. This difference is expected and is consistent with the decrease in stability of the protein at the higher temperature and lower pH used in the present experiments.

In the case of CTL9, different residues exhibited different values for k_f^0 and k_u^0 , beyond experimental uncertainty (Figure 7). The apparent folding rates appeared to be more heterogeneous in the C-terminal half of CTL9, and the fastest rates were found in this C-terminal region. Interestingly four of the six residues exhibiting apparent folding rates that were larger than the average ($> 2\sigma$ of their individual standard deviations) are positioned at the ends of the two β -strands that are known to be ordered in the transition state for folding¹¹. Most of the residues exhibiting significantly slower apparent folding rates were positioned at the ends of helical segments of CTL9. The analysis also yielded residue specific values for the apparent activation volumes of folding and unfolding, ΔV_f^* and ΔV_u^* (Figure 8). The average value for ΔV_f^* was $65 \pm 6.4 \text{ ml/mol}$, (\pm the average of the standard deviations of each measurement) whereas the average ΔV_u^* was $-24 \pm 7.2 \text{ ml/mol}$. The distribution of activation volumes for folding was bimodal, and the standard deviation of the distribution ($\pm 20 \text{ ml/mol}$) was much larger than the average uncertainty of the measurements. The largest values of ΔV_f^* were found in the core region of the protein near the position of the cavity formed by substitution of alanine for isoleucine at position 98 (Figure S4). Interestingly, the residues exhibiting the smallest values of ΔV_f^* are located on the periphery of the core region and correspond, in many cases, to those residues with the slowest apparent folding rates.

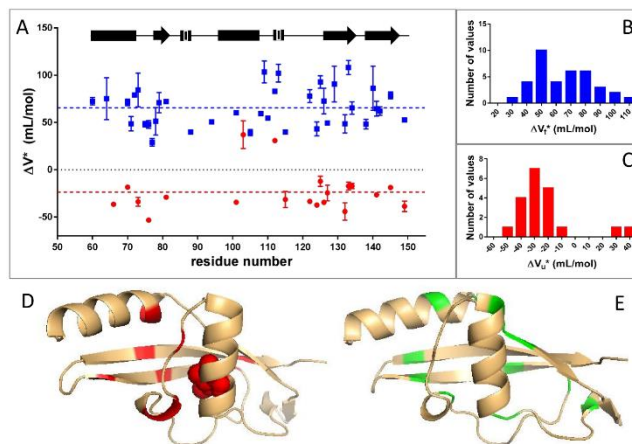


Figure 8. Residue specific apparent activation volumes of CTL9 I98A. A: The activation volume upon folding (ΔV_f^*) and unfolding (ΔV_u^*) for each analyzed residue are illustrated with blue and red data points, correspondingly. Only those values for which the experimental uncertainty was less than 30% of the value were retained. The blue dashed line corresponds to the average value of ΔV_f^* at 65 mL/mol , and the red dashed line represents that of ΔV_u^* at -24 mL/mol . A schematic representation of the secondary structural elements of the folded state of CTL9 is shown at the top of the plot. B: histogram of residue-specific values of ΔV_f^* . C: histogram of residue-specific values of ΔV_u^* . D and E: Schematic representation of largest activation volumes (red) (D) and lowest activation volumes (green) (E) on the CTL9 structure. Residue I98 is shown in red spheres in D).

Discussion

NTL9

The folding and unfolding rate constants for NTL9 have been previously determined under a variety of conditions at atmospheric pressure²¹. Here the residue specific values of the apparent folding and unfolding rates reveals a nearly perfect two-state transition for NTL9. Interpolating the temperature dependence of the rate constants under similar conditions of pH and denaturant concentration as those used in the present study, a folding rate at atmospheric pressure of around 55 s^{-1} and an unfolding rate of near 0.5 s^{-1} can be estimated. Thus increasing pressure decreases the folding rate significantly and modestly increases the unfolding rate. The sign and magnitude of the pressure perturbation of a folding or unfolding reaction rate depends upon the sign and magnitude of the associated volume of activation⁵. Protein folding generally proceeds with a rather large and positive change in volume, whereas the activation volumes for unfolding have been found to be positive, negative or null, depending upon the protein and the conditions^{4,6,30,31}. Pressure slows folding of NTL9 at 2500 bar by a factor of ~ 50 and speeds up unfolding by a factor of ~ 2.6 . This would require an activation volume for folding (ΔV_f^*) of $+36 \text{ ml/mol}$ and a value of -8.6 ml/mol for unfolding (ΔV_u^*). These values, in turn, correspond to a volume change for unfolding at equilibrium of -45 ml/mol . We have measured an equilibrium volume change of $-39 \pm 6 \text{ ml/mole}$ for NTL9 unfolding in sodium acetate at pH 5.1 with 3.5 M urea at 298 K, conditions under which we could obtain complete unfolding profiles for multiple resonances (data not shown). Although denaturants do not generally affect the equilibrium volume change of protein unfolding^{1,32}, temperature effects can be significant³³, depending on the difference in the thermal expansivity of the folded and unfolded states^{34,35}. In any case, the total volume change of unfolding estimated from the present pressure perturbed ZZ-exchange kinetics is in the range of the value obtained from equilibrium high pressure unfolding measurements.

The increase in volume upon protein folding has been demonstrated to arise due to the creation of solvent excluded void volumes in the protein interior¹. We have defined a parameter called the V_i value as the ratio ($\Delta V_f^*/\Delta V_f^0$)³¹, which defines the fraction of the total increase in volume that occurs in the transition state for folding. This parameter is the volumetric equivalent of the β_T value, which reports on the fraction of surface area buried in the transition state relative to the total amount that is buried upon folding. Our estimates of a large positive activation volume for folding and a rather small negative activation volume for unfolding for NTL9 are consistent with a folding transition state in which most of the internal solvent excluded voids have been created upon folding of the hydrophobic core of the protein. From denaturant based kinetics experiments, the position of the transition state of NTL9 in terms of burial of surface area, β_T , is about 65% folded, while temperature dependent studies indicate that a significant amount of hydrophobic surface area is buried in the transition state for folding²¹. In those studies, a clear Hammond effect was observed, in which the transition

state becomes more native-like at higher denaturant concentrations. The equivalent volumetric parameter, V_i , from the present results is 80%. Hence the transition state of NTL9 appears to be somewhat more native-like volumetrically than in regards to surface area. This is not surprising, since the volumetric changes upon folding arise from the formation of internal void volumes, which are distributed non-uniformly throughout the folded structure. Moreover, pressure may be less disruptive than denaturant to transition state structure. Our results are consistent with a transition state structure for NTL9 in which the main hydrophobic core is formed, creating most of the internal void volume, but in which the peripheral regions of the protein, which account for 35% of the surface area, remain disordered. Consistent with this hypothesis, pH dependent studies have shown that electrostatic interactions involving surface residues are not well developed in the transition state for folding, while amide isotope effect studies have shown that a significant fraction of hydrogen bonded secondary structure is^{16,36}. Interestingly, a multi-mutational ϕ -value analysis at atmospheric pressure of 6 major hydrophobic core residues shows that all of these sites exhibit low ϕ -values, arguing that the final specific sidechain interactions are formed on the downhill side of the folding barrier and not in the transition state for folding, even though a significant fraction of the hydrophobic surface area has been sequestered from solvent. ϕ -value analysis can lead to underestimation of transition state structure in some instances^{37,38}, and even be misleading in cases where the mutation changes the transition state structure⁶. Nonetheless, the present pressure dependent studies support the previous ϕ -value analyses and indicate that internal void volumes are formed even though the final sidechain interactions are not wholly developed³⁹. The picture of the transition state for NTL9 which emerges from the β_T and V_i values and sidechain ϕ -value analysis is reminiscent of the dry molten globule transition state ensemble hypothesized by Shakhnovich and Finkelstein⁴⁰ and later by Baldwin and Rose⁴¹.

CTL9 I98A

In the case of CTL9, HPZZEx measurements at multiple pressures yielded residue specific apparent rate constants at atmospheric pressure, as well as residue specific values for the apparent activation volumes of folding and unfolding. Hence ZZ-exchange in general, presents an advantage over relaxation kinetics, in which the folding and unfolding rate constants, along with their pressure (or denaturant) dependence, must be extracted from the same chevron plot. Comparison of the individual residue specific apparent rate constants at atmospheric pressure and apparent activation volumes for several residues of CTL9 I98A revealed a few residues, notably at the ends of the two C-terminal β -strands, that exhibited faster apparent folding rates, suggesting that these might be initiation sites for folding.

Differences in the residue-specific apparent rate constants observed using HPZZEx arise from deviations from 2-state

behavior. Each residue exists in slow exchange between an ensemble of conformations in which it exhibits its “unfolded state” resonance frequency and an ensemble of conformations in which it exhibits the “folded state” frequency. If partially folded (or partially unfolded) intermediates are populated to any significant extent, then these ensembles will not be identical for all residues. This can lead to different “apparent” rate constants for folding or unfolding obtained from different residues. How different the apparent rate is for a given residue compared to the average rate will depend upon the relative flux through the intermediate(s). The observed “apparent” residue-specific rates correspond to averages for each residue of the rates from all of the molecules in the ensemble.

The apparent faster folding region in CTL9 I98A overlaps with that in which residual structure has been observed in the unfolded state^{18–20}, and which has been proposed, based on ϕ -value analysis, to be more ordered in the transition state of CTL9¹¹. Residue-specific values of apparent activation volumes for folding revealed larger values for the region surrounding the original CTL9 cavity (Figure S4) which is enlarged by the I98A mutation. The average activation volumes for folding and unfolding obtained from the pressure-dependent ZZ-exchange data on CTL9-I98A yielded a V_i value of 0.73. This indicates that volumetrically, the transition state of CTL9 I98A lies much closer to the folded state than the unfolded state. Burial of 75% of the surface area in the transition state was observed in denaturant dependent kinetics experiments on this variant¹¹. Thus the volumetric and surface area indications of transition state structure are similar for CTL9, and describe a largely native-like structure at the barrier, as was the case for the NTL9 domain.

Conclusions

We have established the use of HPZZEx NMR spectroscopy for the determination of residue-specific apparent folding and unfolding rate constants for proteins which fold too rapidly for conventional real time NMR approaches. Pressure perturbation, which slows protein folding due to significant activation volumes, drove the folding rates of these proteins into the time regime appropriate for exchange spectroscopy. ZZ-exchange provides independent measurements of the apparent folding and unfolding rate constants at nearly every residue in the chain, an advantage over relaxation kinetics. The volumetric information gained from these experiments provides a structural parameter for the transition state which is complementary to that obtained from denaturant and thermal based kinetics experiments. For both NTL9 and CTL9, the volumetric properties of the transition state are consistent with the notion of a dry molten globule, in which the core of the protein is collapsed, but lacks the fixed tertiary interactions that stabilize the folded state. Importantly, the residue-specific nature of the kinetic parameters obtained by HPZZEx is shown to yield crucial information concerning

protein structure in the transition state, thus providing insight into mechanistic details of the folding reaction. This information is critical for benchmarking atomistic simulations of protein folding and provides significant constraints to test theory and simulation. While many small proteins fold too quickly for ZZ-exchange at atmospheric pressure, the combination of this approach with high pressure dramatically extends the applicability of exchange spectroscopy in protein folding, and as such should be very useful for determining the folding mechanisms of a large number of proteins.

ASSOCIATED CONTENT

Supporting Information.

Figure S1. Pressure dependent ¹H-¹⁵N HSQC spectra of I98A CTL9.

Figure S2. Expanded region of the ZZ-exchange spectrum of CTL9 I98A.

Figure S3. Linear regression of lnk as a function of pressure for several residues of CTL9 I98A.

Figure S4. Schematic Ribbon diagram of the structure of CTL9 WT with the central cavity.

Table S1. Residue specific folding rates at atmospheric pressure and activation volumes for CTL9 I98A.

Table S2. Residue specific unfolding rates at atmospheric pressure and activation volumes for CTL9 I98A

This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

C.A.R.; email, royerc@rpi.edu; Phone, (518) 276-3796. D.P.R.; email, Daniel.Raleigh@Stonybrook.edu; Phone, (631) 632-9547.

Present Addresses

† Department of Pharmaceutical Sciences, College of Pharmaceutical Sciences, Ritsumeikan University, 1-1-1 Noji-Higashi, Kusatsu, Shiga, 525-8577, Japan

Author Contributions

The manuscript was written through contributions of all authors. / All authors have given approval to the final version of the manuscript. / ‡ These authors contributed equally.

Funding Sources

This work was supported by grants NSF MCB-1514575 to CAR and NSF MCB-1330259 to DPR.

ACKNOWLEDGMENT

The authors would like to acknowledge Gaetano Montelione for suggesting we try ZZ-exchange under pressure.

ABBREVIATIONS

NTL9 – N-terminal domain of the ribosomal protein L9.
CTL9 – C-terminal domain of the ribosomal protein L9.

REFERENCES

- (1) Roche, J.; Caro, J. A.; Norberto, D. R.; Barthe, P.; Roumestand, C.; Schlessman, J. L.; Garcia, A. E.; Garcia-Moreno, B. E.;

- Royer, C. A. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6945–6950.
- (2) Nucci, N. V.; Fuglestad, B.; Athanasoula, E. A.; Wand, A. J. *Proc. Natl. Acad. Sci.* **2014**, *111*, 13846–13851.
- (3) Fossat, M.; Dao, T. P.; Jenkins, K.; Dellarole, M.; Yang, Y.; McCallum, S. A.; Garci, A. E.; Barrick, D.; Roumestand, C.; Royer, C. A. *Biophys. J.* **2016**, in press.
- (4) Vidugiris, G. J.; Markley, J. L.; Royer, C. A. *Biochemistry* **1995**, *34*, 4909–4912.
- (5) Royer, C. A. *Arch. Biochem. Biophys.* **2008**, *469*, 34–45.
- (6) Roche, J.; Dellarole, M.; Caro, J. a; Norberto, D. R.; Garcia, A. E.; Garcia-Moreno, B.; Roumestand, C.; Royer, C. A. *J. Am. Chem. Soc.* **2013**, *135*, 14610–14618.
- (7) Palmer, A. G. *Chem. Rev.* **2004**, *104*, 3623–3640.
- (8) Montelione, G. T.; Wagner, G. J. *J. Am. Chem. Soc.* **1989**, *111*, 3096–3098.
- (9) Farrow, N. A.; Zhang, O. W.; Forman-Kay, J. D.; Kay, L. E. *J. Biomol. Nmr* **1994**, *4*, 727–734.
- (10) Schrodinger LLC. 2015.
- (11) Li, Y.; Gupta, R.; Cho, J. H.; Raleigh, D. P. *Biochemistry* **2007**, *46*, 1013–1021.
- (12) Li, Y.; Shan, B.; Raleigh, D. P. *J. Mol. Biol.* **2007**, *368*, 256–262.
- (13) Luan, B.; Shan, B.; Baiz, C.; Tokmakoff, A.; Raleigh, D. P. *Biochemistry* **2013**, *52*, 2402–2409.
- (14) Meng, W.; Lyle, N.; Luan, B.; Raleigh, D. P.; Pappu, R. V. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 2123–2128.
- (15) Kuhlman, B.; Luisi, D. L.; Evans, P. A.; Raleigh, D. P. *J. Mol. Biol.* **1998**, *284*, 1661–1670.
- (16) Sato, S.; Raleigh, D. P. *J. Mol. Biol.* **2007**, *370*, 349–355.
- (17) Sato, S.; Raleigh, D. P. *J. Mol. Biol.* **2002**, *318*, 571–582.
- (18) Shan, B.; Bhattacharya, S.; Eliezer, D.; Raleigh, D. P. *Biochemistry* **2008**, *47*, 9565–9573.
- (19) Shan, B.; Eliezer, D.; Raleigh, D. P. *Biochemistry* **2009**, *48*, 4707–4719.
- (20) Shan, B.; McClendon, S.; Rospigliosi, C.; Eliezer, D.; Raleigh, D. P. *J. Am. Chem. Soc.* **2010**, *132*, 4669–4677.
- (21) Taskent, H.; Cho, J. H.; Raleigh, D. P. *J. Mol. Biol.* **2008**, *378*, 699–706.
- (22) Peterson, R. W.; Wand, A. J. *Rev. Sci. Instrum.* **2005**, *76*, 94101.
- (23) Erlach, M. B.; Munte, C. E.; Kremer, W.; Hartl, R.; Rochelt, D.; Niesner, D.; Kalbitzer, H. R. *J. Magn. Reson.* **2010**, *204*, 196–199.
- (24) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. *Proteins Struct. Funct. Bioinforma.* **2005**, *59*, 687–696.
- (25) Dellarole, M.; Royer, C. A. *High-Pressure Fluorescence Applications, Chapter 4; Fluorescence Spectroscopy and Microscopy: Methods and Protocols, Methods in Molecular Biology, 1076, Y Engelborghs and Antonie J.W.G. Visser, eds., Humana Press, 2014, 53-74.*
- (26) Meng, W.; Luan, B.; Lyle, N.; Pappu, R. V.; Raleigh, D. P. *Biochemistry* **2013**, *52*, 2662–2671.
- (27) Meng, W.; Lyle, N.; Luan, B.; Raleigh, D. P.; Pappu, R. V. *Proc. Natl. Acad. Sci.* **2013**, *110*, 2123–2128.
- (28) Cho, J.-H.; Raleigh, D. P. *J. Mol. Biol.* **2006**, *359*, 1437–1446.
- (29) Li, Y.; Gupta, R.; Cho, J.-H.; Raleigh, D. P. *Biochemistry* **2007**, *46*, 1013–1021.
- (30) Rouget, J.-B.; Aksel, T.; Roche, J.; Saldana, J.-L.; Garcia, A. E.; Barrick, D.; Royer, C. A. *J. Am. Chem. Soc.* **2011**, *133*, 6020–6027.
- (31) Mitra, L.; Hata, K.; Kono, R.; Maeno, A.; Isom, D.; Rouget, J.; Winter, R.; Akasaka, K.; Garcia-Moreno, B.; Royer, C. A. *J. Am. Chem. Soc.* **2007**, *129*, 14108–14109.
- (32) Rouget, J.-B.; Schroer, M. a; Jeworrek, C.; Pühse, M.; Saldana, J.-L.; Bessin, Y.; Tolan, M.; Barrick, D.; Winter, R.; Royer, C. A. *Biophys. J.* **2010**, *98*, 2712–2721.
- (33) Heremans, K.; Smeller, L. *Biochim. Biophys. Acta* **1998**, *1386*, 353–370.
- (34) Mitra, L.; Smolin, N.; Ravindra, R.; Royer, C.; Winter, R. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1249–1265.
- (35) Dellarole, M.; Caro, J. A.; Roche, J.; Fossat, M.; Barthe, P.; Royer, C. A.; Roumestand, C. *J. Am. Chem. Soc.* **2015**, *137*, 9354–9362.
- (36) Luisi, D. L.; Raleigh, D. P. *J. Mol. Biol.* **2000**, *299*, 1091–1100.
- (37) Bulaj, G.; Goldenberg, D. P. *Nat Struct Mol Biol* **2001**, *8*, 326–330.
- (38) Pandit, A. D.; Jha, A.; Freed, K. F.; Sosnick, T. R. *J. Mol. Biol.* **2006**, *361* (4), 755–770.
- (39) Anil, B.; Sato, S.; Cho, J.-H.; Raleigh, D. P. *J. Mol. Biol.* **2005**, *354* (3), 693–705.
- (40) Shakhnovich, E. I.; Finkelstein, A. V. *Biopolymers* **1989**, *28*, 1667–1680.
- (41) Baldwin, R. L.; Rose, G. D. *Curr. Opin. Struct. Biol.* **2013**, *23*, 4–10.

Insert Table of Contents artwork here

