

# **The Unfolded State of the C-Terminal Domain of L9 Expands at Low Temperature but not at Elevated Temperatures**

N.E. Stenzoski<sup>1</sup>, B. Luan<sup>2</sup>, A.S. Holehouse<sup>3</sup>, D.P. Raleigh<sup>1,2,4\*</sup>

<sup>1</sup>Graduate Program in Biochemistry & Structural Biology, <sup>2</sup>Department of Chemistry Stony Brook University, Stony Brook, New York 11794-3400, United States, <sup>3</sup>Center for Biological Systems Engineering, Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA <sup>4</sup>Institute of Structural and Molecular Biology, University of College London. Gower St, London WC1E.6BT, United Kingdom

\*Author to whom correspondence should be address

Email: Daniel.raleigh@stonybrook.edu

Phone +1 631 632 4547

**Key words:** Unfolded State, Denatured State Ensemble, SAXS, Protein Folding

**Running Title:** Compaction of Protein Unfolded States

## **Abstract:**

The temperature dependence of the overall dimensions of the denatured state ensemble (DSE) of proteins remains unclear. Some studies indicate compaction of the DSE at high temperatures, while others argue that dimensions do not decrease. The degree of compaction or expansion in the cold denatured state has been less studied. To investigate the temperature dependence of unfolded state dimensions, SAXS measurements were performed in native buffer in the absence of denaturant for a designed point mutant of the C-terminal domain of L9 (CTL9), a small cooperatively folded  $\alpha$ - $\beta$  protein, at 14 different temperatures over the range of 5-60°C. The I98A mutation destabilizes the domain such that both the DSE and the folded state are populated at 25°C in the absence of denaturant or extreme pH. Thermal unfolding as well as cold unfolding can thus be observed in the absence of denaturant, allowing a direct comparison of these regimes for the same protein using the same technique. The temperature of maximum stability,  $T_s$ , is 30°C. There is no detectable change in  $R_g$  of the unfolded state as the temperature is increased above  $T_s$ , but a clear expansion is detected as the temperature is decreased below  $T_s$ . The  $R_g$  of the DSE populated in buffer was found to be  $27.8 \pm 1.7 \text{ \AA}$  at 5°C,  $21.8 \pm 1.9 \text{ \AA}$  at 30°C, and  $21.7 \pm 2.0 \text{ \AA}$  at 60°C. In contrast, no significant temperature dependence was observed for the value of  $R_g$  measured in 6M GdnHCl. The SAXS data reported here indicate clear differences between the cold and thermal unfolded states and shows that there is no significant compaction at elevated temperatures.

## Introduction:

The denatured state ensemble (DSE) of a protein is the starting state for folding and the reference state for protein stability. The DSE is an important factor in protein design and can be the starting point of aggregation, both *in vitro* and *in vivo* (1-6). As well as providing information relevant for protein folding, studies of the DSE of globular proteins can also provide insight into the properties of intrinsically disordered proteins (IDPs). A plethora of studies have identified IDPs as key players in many biological processes, with the relationship between sequence and conformational behaviour being of particular interest (7-9).

The properties of the DSE are sensitive to the local environment and can vary considerably depending upon solvent properties, pH and temperature. The sensitivity to perturbations can affect folding kinetics and the pathways that lead the protein to its native fold. A protein can be driven to unfold and primarily populate its DSE through a variety of means including the addition of chemical denaturant, extreme pH, pressure, or heat (10). Temperature denaturation occurs at both high and low temperatures, as predicted by the Gibbs-Helmholtz equation:

$$\Delta G_u^\circ(T) = \Delta H^\circ(T_m) - \frac{T\Delta H^\circ(T_m)}{T_m} + \Delta C_p^\circ [T - T_m - T \ln(\frac{T}{T_m})] \quad [1]$$

Where  $\Delta H^\circ$  is the change in enthalpy and  $\Delta C_p^\circ$  is the heat capacity. This equation identifies two distinct transition temperatures: the midpoint of cold denaturation,  $T_c$ , and the midpoint of thermal denaturation,  $T_m$ . Both are defined as the temperature at which the free energy ( $\Delta G^\circ$ ) of unfolding is equal to zero. Cold denaturation, which is less commonly studied experimentally, can be rationalized by a decrease in the strength of hydrophobic effect as the temperature is decreased below the temperature of maximum stability,  $T_s$  (11-19).

The effects of temperature on the dimensions of the DSE remain unclear. Recent studies indicate the unfolded state expands as temperature is decreased below  $T_s$ , but only a limited number of studies have been performed, and the behavior at temperatures above  $T_s$  is still uncertain. Single molecule Förster Resonance Energy Transfer (smFRET) studies of the small cold shock protein from *Thermotoga maritima* (CspTm) (20), the IDP prothymosin alpha (ProT $\alpha$ ) (20), and yeast frataxin (Yfh1) suggest continuous compaction as the temperature increases above  $T_s$  (21). However, small angle x-ray scattering (SAXS) studies on Yfh1 suggest the thermally unfolded state does not undergo a compaction as the temperature is increased above  $T_s$  (22). Early SAXS

temperature-dependent studies on reduced ribonuclease r-RNase A at pH 2.5 showed that the radius of gyration ( $R_g$ ) of the highly unfolded chain did not significantly change with increasing temperature from 20°-90°C (23). Reduced RNase A does not refold to its native state, thus this study is best viewed as an example of the properties of an IDP with significant net-charge. Possible differences in the tendency to form secondary structure along with differences in protein hydration could account for the different behavior at the high and low temperature regimes of a protein, but any relationship between compaction and the level of secondary structure is not completely understood (22).

Some of the differences in the findings of the various experimental studies may be due to the different methods used. Although smFRET is a highly sensitive technique, it is not without disadvantages. Proteins must be labeled with dyes that may or may not perturb the underlying conformational behavior via dye-dye or protein-dye interactions, and some form of theoretical models—typically taken from homopolymer physics—is required both to extract the end-to-end distance from transfer efficiencies and to convert that end-to-end distance data into values of  $R_g$ . Although these models often appear to work reasonably well, they are not necessarily applicable to heteropolymers and the relationship between a FRET derived end-to-end distance and  $R_g$  is not always clear (24-26). On the other hand, smFRET does allow the simultaneous observation of the folded and DSE states under a range of conditions where it would not be possible to do so using ensemble-based methods. SAXS measurements, although requiring considerably higher protein concentrations and sometimes being limited in signal-to-noise with a level of uncertainty in the analysis, can provide  $R_g$  directly from the scattering data in absence of models of the DSE. If the folded and unfolded populations are known, a direct measurement of the  $R_{g,DSE}$  can be taken after subtraction of the known folded population signal.

The role sequence plays in modulating unfolded state dimensions continues to be investigated (27). Applying simple polymer models can provide a useful first approximation and allow for some predictions regarding the role that sequence can play with respect to parameters such as hydrophathy, net charge per residue (NCPR), and the patterning of oppositely charged residues (28, 29).

Ideally, the heat and cold denatured state of a protein would be studied under the same conditions (pH, ionic strength, pressure, etc.) with minimal perturbation to the system, as close to

physiologically relevant conditions as possible, using the same experimental technique. Oftentimes, the  $T_c$  of proteins is below the freezing point of water, presenting challenges in accessing the cold denatured state without the addition of perturbations (30-33). There are a handful of proteins that have  $T_c$  above  $0^\circ\text{C}$ , allowing such studies to be achieved without the use of additional perturbations, such as added urea (22, 34-36). Direct observation of the DSE above the temperature of maximum stability,  $T_s$ , but below  $T_m$  can also be difficult, especially near  $T_s$  where the population of the DSE will be low.

To avoid the need for added perturbations, point mutations can be introduced to destabilize a protein. In our case we use a mutant of the protein CTL9 (I98A-CTL9). CTL9 is a 92-residue single domain protein, that folds in a two-state fashion, containing a rare mixed  $\alpha$ - $\beta$  topology, whose denaturant- and pH-induced unfolding has been previously described (Figure 1) (37). Previous characterization of I98A-CTL9 show that  $T_c$  is  $12^\circ\text{C}$  in 20mM NaOAc and 100 mM NaCl buffer at pH 5.7 and its temperature of maximum stability is  $30^\circ\text{C}$ , and  $T_m$  is  $53.3^\circ\text{C}$  (38). The cold unfolding of I98A-CTL9 appears to be cooperative and initial SAXS experiments at a limited number of temperatures suggest that the DSE expands as the temperature is lowered (39). Diffusion measurements have also shown the cold unfolded state at  $12^\circ\text{C}$  in 50 mM sodium phosphate and 100 mM NaCl, pD 6.6 is more compact than the highly charged DSE of CTL9 WT populated at  $12^\circ\text{C}$ , pD 3.8. The cold unfolded state is also more compact than the urea unfolded state (40).

Here we report a detailed investigation of the dimensions of the DSE traversing the cold denaturation transition all the way through thermal denaturation. SAXS scattering profiles were collected for I98A-CTL9 in the absence of denaturant at pD 5.6 at fourteen temperature points from  $5$ - $60^\circ\text{C}$ , a range which includes the protein's  $T_c$ ,  $T_s$ , and  $T_m$ . Complimentary temperature-dependent CD studies were also performed to investigate the residual secondary structure present at both high and low temperatures.

## **Methods:**

### *Protein Expression and Purification*

I98A-CTL9 was expressed in *Escherichia coli* BL21(DE3) cells in LB media using standard methods (38-41). The cells were grown at  $37^\circ\text{C}$  until the optical density at 600 nm reached 0.8,

then induced with 1mM IPTG for 4 hours at 37°C. Cells were harvested and then lysed by sonication or by cell disrupter. The lysate was centrifuged for 1hr at 10,000 g to remove the cell debris. The supernatant was then loaded onto a SP-Sepharose fast flow ion-exchange column (GE Healthcare) equilibrated in 20mM Tris-HCl, pH 7.5 and the protein eluted with a 0-2 M NaCl gradient. Protein was further purified using reverse phase HPLC with a Higgins Analytical, Inc. C18 column (250x20mm). An A-B gradient was used where buffer A was 99.9% distilled de-ionized H<sub>2</sub>O (DDI) and 0.1% trifluoroacetic acid (TFA), and buffer B was 90% acetonitrile, 9.9% DDI, and 0.1% TFA. The gradient used was 25-55% buffer B in 60 minutes. The proteins eluted at ~42% buffer B.

The identity of the protein was confirmed by DNA sequencing and MALDI-TOF mass spectrometry. The observed molecular weight (M+H) for I98A CTL9 was 9940.9 Da and the expected molecular weight was 9940.5 Da. The yield of the protein was ~70 mg/L in LB media. The observed molecular weight for CTL9 WT was 9982.0 Da and the expected molecular weight was 9982.0 Da. A similar yield was observed for CTL9 WT in LB. The purity was confirmed by analytical HPLC.

*Circular Dichroism:*

CD experiments were performed on a Chirascan CD spectrometer (Applied Photophysics). The protein was dissolved in 10 mM DMG (3,3-dimethylglutaric acid), 120 mM NaCl buffer in 100% D<sub>2</sub>O, at a protein concentration of ~25uM, and at a pD of 5.6 (uncorrected pH reading). D<sub>2</sub>O was used to allow comparison with <sup>1</sup>H NMR and FTIR studies (39) . DMG was chosen because it has a small heat of ionization and this aids in maintaining constant pD (pH) value over wide temperature ranges.

Far-UV wavelength scans were recorded in a 1mm cuvette from 196-260 nm with a 0.2 nm increment, repeated in triplicate and averaged. Wild type CTL9 is folded over the temperatures studied, providing a baseline for the folded CD signal. A sample of I98A-CTL9 at pD 8.0 was used to provide a folded baseline. The population-weighted CD spectrum of the folded state was subtracted from the experimental I98A-CTL9 CD spectrum at pD 5.6 in order to obtain the unfolded I98A-CTL9 CD spectrum, using the following equation:

$$\theta_{I98A,unfolded} = \frac{(\theta_{I98A-pfolded} - \theta_{CTL9WT})}{(1-pfolded)} \quad [2]$$

Where  $\theta_{I98A}$  is the CD signal collected for I98A-CTL9 at pD 5.6, containing both the folded and unfolded ensembles;  $\theta_{CTL9WT}$  is the CD signal collected for CTL9 WT at pD 5.6, only the folded state is present in this sample, providing a spectrum of the folded state for subtraction;  $p_{\text{folded}}$  is the fractional population of folded I98A-CTL9 present under the conditions of the CD measurements.  $p_{\text{folded}}$  was determined by thermal denaturation monitored by CD;  $\theta_{I98A,un\text{folded}}$  is the calculated I98A-CTL9 DSE CD signal.

The mean residue ellipticity at 222 nm, calculated for the I98A-CTL9 DSE ( $[\theta]_{\text{DSE}}$ ), was used to estimate the fraction helical structure:

$$f_h = \frac{([\theta]_{\text{DSE}} - [\theta]_C)}{([\theta]_H - [\theta]_C)} \quad [3]$$

Where  $[\theta]_C$  corresponds to the signal for a random coil and  $[\theta]_H$  corresponds to the signal expected for a 100% helix, calculated from:

$$[\theta]_H = -40,000 \times \left(1 - \frac{2.5}{n}\right) + 100 \times T \quad [4]$$

$$[\theta]_C = 640 - 45 \times T \quad [5]$$

Where  $n$  is the number of residues in the peptide, and  $T$  is the temperature ( $^{\circ}\text{C}$ ) (42).

Thermal denaturation measurements were conducted using a 1cm cuvette, by monitoring ellipticity at 222 nm, from 4-98 $^{\circ}$  C in 2 $^{\circ}$  C steps, at a heating rate of 1 $^{\circ}$  C/min, as a function of pD between 4.0 and 8.0. The protein is unfolded at pD 4.0 and the thermal denaturation data was fit to a quadratic equation to obtain the unfolded signal as a function of temperature, I98A-CTL9 has been shown to populate the DSE for all temperatures at this pD [28, 34]. The completely folded baseline was estimated from fitting thermal denaturation data for I98A-CTL9 at pD 8.0, and extrapolating the folded state signal. Data was fit to the following equation to obtain thermodynamic parameters for unfolding. The observed CD signal,  $\theta(T)$ , was fit to the equation:

$$\theta(T) = \frac{(a_n + b_n T) + (a_d + b_d T) \exp\left(-\frac{\Delta G_u^{\circ}(T)}{RT}\right)}{1 + \exp\left(-\frac{\Delta G_u^{\circ}(T)}{RT}\right)} \quad [6]$$

Where  $\Delta G_u^{\circ}(T)$  is the free energy change upon thermal unfolding described by the Gibbs-Helmholtz Equation (Equation 1), where  $a_n$ ,  $b_n$ ,  $a_d$ , and  $b_d$  are parameters which define the

signals of the native state  $n$  and the denatured state ensembles  $d$  as a function of temperature.  $T_m$  is the thermally induced unfolding midpoint temperature, and  $\Delta C_p^\circ$  is the change in heat capacity upon unfolding. The signal expected for a fraction folded of 0.5 was estimated by taking the average of the native state baseline and the unfolded state baseline as a function of temperature.

The population of the folded state and the DSE were estimated from the temperature dependent CD curves using native and DSE temperature-dependent baselines described in Figure 2, fit to:

$$\theta_{obs} = (Fraction\ Folded) * \theta(T)_{folded} + (1 - Fraction\ Folded) * \theta(T)_{DSE} \quad [7]$$

Where  $\Theta(T)_{obs}$  is the measured CD signal,  $\Theta(T)_{folded}$  is the temperature-dependent CD signal of the folded state and  $\Theta(T)_{DSE}$  is the temperature-dependent DSE CD signal.

SAXS:

Samples of CTL9 WT and I98A-CTL9 were prepared in buffer containing 10 mM DMG and 120 mM NaCl in 100% D<sub>2</sub>O, pD 5.6. Small angle x-ray scattering was performed at X9 Beamline at Brookhaven National Lab in their National Synchrotron Light Source I (NSLS I) in Upton, NY. D<sub>2</sub>O was used to allow comparison to previously collected CD, FTIR, and NMR data.

Scattering data was collected for I98A-CTL9 and CTL9 WT at a concentration of 3.75 mg/mL with an exposure time of 30 seconds. Samples were injected into a 1mm capillary at a rate of 0.67  $\mu$ L/s to avoid radiation damage. Each SAXS profile was collected at 14 different increments over 5-60°C, each being measured three times and averaged prior to data analysis. Additional data was collected at 7°C at 2.5, 3.75, and 5.0 mg/mL protein (MOPS buffer). The values of  $R_g$  of the DSE were identical to within experimental uncertainty.

CTL9 WT provided a folded scattering profile at each temperature point, as only the folded state is populated over the temperature range of interest, used for subtraction of the folded state signal. The I98A-CTL9 DSE scattering profiles were obtained after subtracting the population-weighted folded signal according to:

$$I_{I98A,unfolded} = (I_{I98A} - p_{folded} \cdot I_{CTL9,WT}) / (1 - p_{folded}) \quad [7]$$



Where  $I_{I98A}$  is the SAXS data collected for the sample of I98A-CTL9 at pD 5.6, containing both the folded and unfolded ensembles.  $I_{CTL9,WT}$  is the scattering profile collected for CTL9 WT at pD 5.6, where only the folded state is populated, providing a folded scattering profile for subtraction.  $p_{folded}$  is the population of folded I98A-CTL9 present under these conditions.  $I_{I98A,unfolded}$  is the calculated I98A-CTL9 DSE scattering profile. The  $R_g$  of the unfolded state was determined from this data using the Guinier approximation.

The program pyXS was used for buffer subtraction. The program PRIMUS was used to calculate the radius of gyration ( $R_g$ ) based on the scattering profiles, using the Guinier approximation:

$$I(q) = I(0) \cdot \exp(-R_g^2 q^2) \quad [8]$$

Where  $I(q)$  is the intensity at scattering vector  $q$  (43).

The range of  $q^2$  for the Guinier approximation used was 0.002-0.1.

To obtain the apparent  $R_g$  values,  $R_{g,app}$ , at each temperature point, Guinier analysis was performed on I98A-CTL9. The  $R_g$  of the I98A-CTL9 DSE was also estimated using a two-state approximation:

$$R_{g,app}^2 = p_{folded} \cdot R_{g,folded}^2 + (1 - p_{folded}) \cdot R_{g,DSE}^2 \quad [9]$$

Where  $p_{folded}$  and  $(1-p_{folded})$  are the fraction populations of folded ensemble and the DSE, respectively.

### *Analysis of Primary Structures*

Using the publicly available software, CIDER (44), the sequences of 8 IDPs and r-RNaseA at pH 2.5 were compared to I98A-CTL9. The analysis is based on attributes including hydrophathy, net charge per residue (NCPR), and patterning of polar and nonpolar residues (captured by  $\kappa$ ). Each protein is plotted on a phase diagram predicting its subclass (e.g. polyampholyte, polyelectrolyte, globule, and swollen coil).

## **Results**

**The I98A mutation allows the direct observation of the DSE at  $T_s$  as well as at low and elevated temperatures.**

The point mutant, I98A-CTL9 is an excellent system for directly studying the temperature dependence of the dimensions of the DSE in the absence of denaturant. CTL9 stability is strongly pH-dependent, due to three histidines that can readily undergo protonation at low pH. To avoid potential effects from any temperature-dependence behavior of the buffer, (i.e. shifts in pH as a function of temperature), 3,3-dimethylglutaric acid (DMG) was used.

The temperature dependence of the structure of the I98A-CTL9 construct was monitored by CD at 222 nm (Figure 2), as a function of pD from 5-60°C in DMG buffer. It is important to conduct the experiments using the same buffer composition as used for the SAXS studies since this increases the precision of the calculated values of the fraction folded as a function of temperature.

The calculated folded and unfolded baselines, shown in Figure 2, allow the fraction folded to be estimated at any relevant temperature and pD since the observed CD signal is the sum of the signal from the DSE and the folded state. Assuming that equilibrium unfolding is two-state the relative populations are easily obtained (See Methods). At pD 5.6 (uncorrected pH-meter reading), an equal amount of the folded and unfolded ensembles is populated at 30°C, the temperature of maximum stability ( $T_s$ ). At 5°C, the folded state population is decreased to only 13%, while at the highest temperature experimentally accessible, 60°C, it is only ~12%. The estimated fraction folded at each temperature where SAXS scattering profiles were collected are listed in Table 1.

Temperature (°C)	$R_{g,app}$ (Å) <sup>a</sup>	$R_{g,folded}$ (Å) <sup>b</sup>	$P_{folded}$ (%) <sup>c</sup>	$R_{g,DSE}$ (Å) <sup>d</sup>	$R_{g,DSE}$ (Å) <sup>e</sup>
5	26.0±1.7	14.9±0.3	13.2	27.3±1.7	27.8±1.7
7	24.4±1.4	15.1±0.4	14.8	25.7±1.5	27.0±1.6
10	23.7±1.4	15.1±0.4	16.3	25.0±1.5	26.5±1.6
12	24.0±1.9	15.1±0.4	21.2	25.9±1.6	26.6±1.3
15	22.3±1.2	15.2±0.4	26.2	24.3±1.4	24.1±1.5
20	20.5±1.4	15.1±0.4	36.4	23.0±0.9	22.7±1.4
25	19.7±1.6	15.1±0.3	44.9	22.8±1.0	22.8±1.8
30	18.7±1.4	14.9±0.3	50.7	21.9±0.9	21.8±1.9
35	18.6±1.2	14.8±0.4	50.9	21.9±0.8	20.0±2.1
40	19.3±1.6	15.0±0.3	46.8	22.4±1.0	22.6±1.9
45	20.0±1.5	15.0±0.3	38.0	22.5±0.9	20.7±1.8
50	19.6±1.8	14.9±0.4	28.1	21.2±1.1	21.4±1.7
55	21.4±1.8	14.8±0.3	18.0	22.6±1.2	21.9±1.7
60	20.8±1.6	15.2±0.2	12.1	21.5±0.9	21.7±2.0

**Table 1.  $R_g$  values and fraction folded of I98A-CTL9, in 10 mM DMG and 120 mM NaCl, 100% D<sub>2</sub>O, pD 5.6**

a) Apparent  $R_g$  values obtained directly by the Guinier analysis of the observed I98A-CTL9 scattering profile. b)  $R_g$  values of wild type CTL9, used as a control for the folded state over the temperature range used in the study. c) the population of the folded ensemble for I98A-CTL9 as a function of temperature, estimated from the CD monitored thermal denaturation data. d)  $R_g$  values of the I98A-CTL9 DSE calculated by the two-state approximation. e)  $R_g$  values of the I98A-CTL9 DSE obtained by the Guinier analysis on the DSE scattering profile after subtracting the population weighted signal from the folded state contribution. The uncertainty represents the linear fitting in the Guinier approximation.

In these conditions, in the absence of denaturant or extreme pH, the low temperature and high temperature DSE populated where  $\Delta G^{\circ}_{\text{unfolding}}$  is equal, can be directly observed and compared.

**I98A-CTL9 in denaturant has an  $R_g$  value consistent with a random coil model.**

To establish a basis of comparison for the I98A-CTL9 temperature induced DSE in the absence of denaturant, SAXS studies in high concentrations of chemical denaturant were performed. The  $R_g$  value of I98A-CTL9 populated in 6M GuHCl (Figure S1) from the Guinier analysis is estimated to be between 29-30 Å and is in good agreement with the  $R_g$  value predicted for a random coil in good solvent, ~29 Å (45). Also worth noting is the lack of any detectable temperature dependence of the  $R_g$  value measured in the presence of denaturant (Figure S1).

**The  $R_g$  of the I98A-CTL9 DSE shows different trends during heating vs during cooling.**

Having established I98A-CTL9 is highly unfolded in the presence of chemical denaturants and that the  $R_g$  does not vary with temperature under these conditions, we next examined the properties of the DSE under more native-like conditions as a function of temperature. To monitor the overall dimensions of the DSE, we collected SAXS data as a function of temperature and extracted  $R_g$  values.

Because the samples contain both folded and unfolded molecules, the observed scattering profiles include contributions from both states. Two procedures were adopted to deconvolve the DSE signal as described below, however a clear trend in the apparent  $R_g$  obtained from directly fitting the temperature-dependent experimental data was seen even without deconvolving folded state contributions. Despite the folded population being very similar at 5° and 60°C, the observed  $R_{g,\text{app}}$  is significantly smaller at the highest recorded temperature as compared to the lowest temperature. Furthermore, no significant change in  $R_{g,\text{app}}$  is seen as temperature is increased above 20°C (Figure 3).

These observations indicate that the dimensions of the cold and thermally unfolded states must be different. Control experiments were performed to analyze the dependence of the derived  $R_g$  values to variations in pD, by conducting additional experiments 0.6 pD units above and below the pD of the sample used for the experiments (Table S1, Figure S2). Any uncertainty in the pD of the protein samples is much less than this range. Lowering the pD by 0.6 units had no effect upon the derived values of  $R_g$ , and increasing the pD by 0.6 units leads to a variation in the

calculated  $R_g$  of the DSE of between 7 and 13%. The analysis indicates that the temperature dependence of the measured  $R_g$  values are not due to variations in pD.

Estimates of the  $R_g$  values of I98A-CTL9 DSE,  $R_{g,DSE}$ , were first obtained by subtracting the population weighted folded state signal of CTL9 WT over the range of 5-60°C, from the observed scattering profile of I98A-CTL9 (Equation 7). Wild type CTL9 was used to directly measure  $R_g$  of the folded state at each temperature, and it remains fully folded over this temperature range, with its  $R_g$  values varying by less than 3%. The difference scattering profiles were analyzed using the Guinier analysis. For the second approach, Guinier analysis was performed directly on I98A-CTL9 scattering data to obtain apparent  $R_g$  values,  $R_{g,app}$ , at each temperature point. The value of  $R_{g,DSE}$  was then estimated using a two-state approximation by including the weight of the known folded population (Equation 9). The fraction folded of I98A-CTL9 at each temperature point is listed in Table 1 and plotted as a function of temperature in Figure 4. Guinier plots of the CTL9 WT control, and of the observed scattering profile before subtracting the folded contribution, as well as after subtraction are shown in Figure S3.

The  $R_{g,DSE}$  of I98A-CTL9 at the  $T_s$  (30°C) is 22Å, a 45% increase compared to the CTL9 WT folded state control of ~15Å. A continuous increase in  $R_{g,DSE}$  is observed as the temperature decreases from 30° to 5°C. The  $R_g$  of the cold unfolded state is ~28 Å at 5°C, representing an 87% increase over the native state value. The values of  $R_g$  from 30°C to 60°C are scattered between  $20.2 \pm 2.1$  Å and  $21.9 \pm 1.7$  Å, and in contrast to the cold unfolded data, exhibit no clear trend. The absence of any clear expansion or compaction of the  $R_{g,DSE}$  above  $T_s$  is in agreement with the trend noticed by simply plotting  $R_{g,app}$  as a function of temperature. It is interesting to note that the simple two-state system approximation applied to the I98A-CTL9 scattering profiles gives calculated  $R_{g,DSE}$  values that are in strong agreement with the analysis using the subtraction of the population-weighted signal, providing additional confidence in the analysis. The temperature dependence of the apparent  $R_g$  and calculated  $R_g$  of the DSE indicates that the cold and high temperature unfolded states of I98A-CTL9 behave markedly differently.

**There are only modest differences in the overall  $\alpha$ -helical content in the DSE of I98A-CTL9 at low and high temperature.**

Far-UV CD spectra were recorded as a function of temperature for I98A-CTL9 to compare the cold and heat-induced DSE (Figure S2). CD data for wild type CTL9 was recorded to provide a

folded baseline spectrum. The folded spectrum of CTL9 WT can be combined with the estimate population of the folded state of I98A-CTL9 to calculate the expected CD signal of I98A-CTL9 DSE using equation 3. The fractional folded population is calculated to be ~13% and ~12% at 5°C and 60°C, respectively. These temperatures allow for comparison of cold and heat unfolded states where the fraction folded is similar. The  $\alpha$ -helical content of the cold denatured state is estimated to be low, 11% at 5°C, and the helical content, of the heat-denatured DSE is calculated to be 3%, at 60°C (Table S2, Figure S4). Given the uncertainty inherent in deconvolving CD spectra or in the analysis of  $\Theta_{222}$  in terms of helical content particularly when difference spectra are required, we believe it is not possible to interpret any apparent differences. However, it is clear that the helical content is modest at best and any small differences are very unlikely to account for the differences in  $R_g$  values.

### **Discussion:**

The temperature dependent studies reported here confirm the expansion of the I98A-CTL9 DSE at low temperature but provide considerably more data and also monitor the thermally unfolded state. The clear expansion of the cold DSE is in contrast with the lack of significant variation of the thermally unfolded state dimensions. This lack of significant variation in the value of  $R_g$  of the DSE of I98A-CTL9 at higher temperatures presumably reflects competing contributions from conformational entropy ( $-T\Delta S_{\text{confor}}$ ), the temperature dependence of the hydrophobic effect, the increased thermal energy at higher temperature, and the temperature dependence of the solvation free energy (46). Some unfolded protein chains have been reported to be more compact at higher temperatures relative to their behavior near room temperature, a result which is different from that expected for a polymer chain with temperature-independent monomer-monomer interactions (20, 46, 47). Wuttke reported a smFRET study in which temperature-induced collapse was seen for five IDPs, despite the proteins having varying amounts of hydrophilicity, hydrophobicity, and net charge (46). Wuttke et al. also examined the unfolded state  $R_g$  of a non-IDP, monomeric  $\lambda$ -repressor. The sequence properties of their  $\lambda$ -repressor construct are closest to those of I98A-CTL9 in terms of the fraction of positively charged and negatively charged residues (Figure 1). Comparison of the temperature-dependent  $R_g$  values derived from the FRET studies of the  $\lambda$ -repressor construct and these directly measured for I98A-CTL9 via SAXS shows remarkably similar behavior (Figure 5). Moreover, similar trends were observed for Yhf1 and CspTm (as

measured by smFRET), both of which exist in a folded-unfolded equilibrium and have broadly similar sequence properties to I98A-CTL9. All four proteins exhibit a clear expansion as the temperature is lowered below  $T_s$ , and show no significant change in  $R_g$  at higher temperatures. The low temperature behavior of these four proteins suggests that expansion at temperatures below  $T_s$  could be a general phenomenon and further highlights that cold and heat denatured states are not equivalent. In contrast, the *bona fide* IDPs studied by Wuttke *et al.* show a continuous contraction at high temperatures, and in general have a higher fraction of charged residues. This is consistent with the proposal that different trends in the temperature-dependent solvation free energies of hydrophobic vs polar groups plays an important role in the temperature-dependence of the unfolded state dimensions (46). However, it is worth noting that the N-terminal domain of HIV Integrase (IN) has general sequence features more similar to I98A-CTL9 than Yhf1, but in contrast to I98A-CTL9 undergoes continuous contraction at higher temperatures (46). Unlike Yhf1 and I98A-CTL9, IN does not undergo appreciable folding across the temperature range studied in the absence of zinc. With this in mind, the determinants of the temperature dependence may also depend on the formation of transient structure in the unfolded state, although additional work will be required to understand this relationship.

Using the approach of Das and Pappu, the sequence analysis of I98A-CTL9, 8 IDPs and r-RNaseA (proteins for which temperature dependent studies of the unfolded state dimensions have been reported) reveal that I98A-CTL9 is not an outlier in terms of sequence properties (Figure 6, Table S3), and falls within the same region of the plot as several of the IDPs. This indicates that these properties alone do not predict the temperature dependence of  $R_g$  and additional information is required. It is important to note that the analysis was developed based on sequence properties at 25°C, and was not explicitly designed to predict the temperature-dependent behavior. The data presented here should provide a useful test case for the development of methods to analyze temperature-dependent changes in the DSE and in IDPs and is useful as a benchmark for molecular dynamics (MD) simulations of unfolded states and IDPs (29, 48).

Recent efforts have shown that existing MD protocols predict unfolded states that are more compact than is experimentally observed. One contribution to this issue is believed to result from the strength of water-water versus water-protein interactions (49). The temperature dependence

of  $R_g$  of the unfolded state provides a more rigorous test of refinements in these methods than does the ability to reproduce  $R_g$  at just a single temperature.



### **Author Contributions**

NES, BL designed and performed experiments and analyzed data. ASH analyzed data.

DPR designed and directed research. NES, ASH and DPR wrote manuscript.

**Acknowledgements:**

We thank Marc Allaire and Lin Yang at Brookhaven National Lab for assistance in collection of the SAXS profiles. We thank Professor Rohit Pappu for helpful discussions. This work was supported through NSF Grant MCB-1330259.

## References

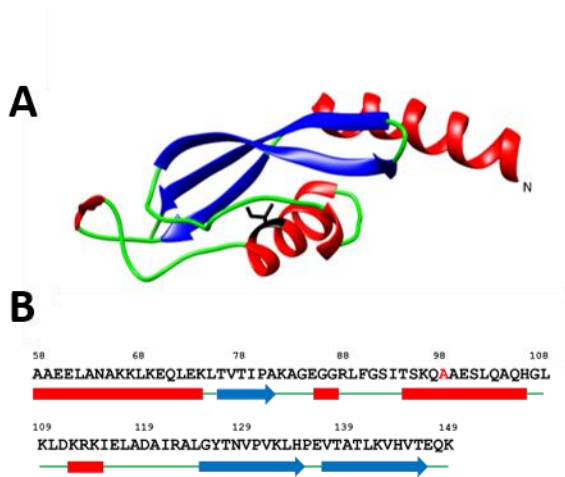
1. Zou, J., B. Song, C. Simmerling, and D. Raleigh. 2016. Experimental and Computational Analysis of Protein Stabilization by Gly-to-d-Ala Substitution: A Convolution of Native State and Unfolded State Effects. *Journal of the American Chemical Society* 138:15682-15689.
2. Akter, R., P. Cao, H. Noor, Z. Ridgway, L.-H. Tu, H. Wang, A. G. Wong, X. Zhang, A. Abedini, A. M. Schmidt, and D. P. Raleigh. 2016. Islet Amyloid Polypeptide: Structure, Function, and Pathophysiology. *Journal of Diabetes Research* 2016:18.
3. Chiti, F., and C. M. Dobson. 2017. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annual Review of Biochemistry* 86:27-68.
4. Eisenberg, D. S., and M. R. Sawaya. 2017. Structural Studies of Amyloid Proteins at the Molecular Level. *Annual Review of Biochemistry* 86:69-95.
5. Sathasivam, K., A. Neueder, T. A. Gipson, C. Landles, A. C. Benjamin, M. K. Bondulich, D. L. Smith, R. L. M. Faull, R. A. C. Roos, D. Howland, P. J. Detloff, D. E. Housman, and G. P. Bates. 2013. Aberrant splicing of *HTT* generates the pathogenic exon 1 protein in Huntington disease. *Proceedings of the National Academy of Sciences* 110:2366-2370.
6. Spillantini, M. G., M. L. Schmidt, V. M. Y. Lee, J. Q. Trojanowski, R. Jakes, and M. Goedert. 1997.  $\alpha$ -Synuclein in Lewy bodies. *Nature* 388:839.
7. Wright, P. E., and H. J. Dyson. 2014. Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology* 16:18.
8. Dunker, A. K., J. D. Lawson, C. J. Brown, R. M. Williams, P. Romero, J. S. Oh, C. J. Oldfield, A. M. Campen, C. M. Ratliff, K. W. Higgs, J. Ausio, M. S. Nissen, R. Reeves, C. Kang, C. R. Kissinger, R. W. Bailey, M. D. Griswold, W. Chiu, E. C. Garner, and Z. Obradovic. 2001. Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling* 19:26-59.
9. Das, R. K., K. M. Ruff, and R. V. Pappu. 2015. Relating sequence encoded information to form and function of intrinsically disordered proteins. *Current Opinion in Structural Biology* 32:102-112.
10. Dill, K. A., and D. Shortle. 1991. Denatured states of proteins. *Annu Rev Biochem* 60:795-825.
11. Davidovic, M., C. Mattea, J. Qvist, and B. Halle. 2009. Protein cold denaturation as seen from the solvent. *J Am Chem Soc* 131:1025-1036.
12. Dias, C. L. 2012. Unifying microscopic mechanism for pressure and cold denaturations of proteins. *Phys Rev Lett* 109:048104.
13. Dias, C. L., T. Ala-Nissila, M. Karttunen, I. Vattulainen, and M. Grant. 2008. Microscopic mechanism for cold denaturation. *Phys Rev Lett* 100:118101.
14. Dias, C. L., T. Ala-Nissila, J. Wong-ekkabut, I. Vattulainen, M. Grant, and M. Karttunen. 2010. The hydrophobic effect and its role in cold denaturation. *Cryobiology* 60:91-99.
15. Franks, F. 1995. Protein destabilization at low temperatures. *Adv Protein Chem* 46:105-139.
16. Graziano, G. 2010. On the molecular origin of cold denaturation of globular proteins. *Phys Chem Chem Phys* 12:14245-14252.

17. Lopez, C. F., R. K. Darst, and P. J. Rossky. 2008. Mechanistic elements of protein cold denaturation. *J Phys Chem B* 112:5961-5967.
18. Privalov, P. L. 1990. Cold denaturation of proteins. *Crit Rev Biochem Mol Biol* 25:281-305.
19. Tsai, C. J., J. V. Maizel, Jr., and R. Nussinov. 2002. The hydrophobic effect: a new insight from cold denaturation and a two-state water structure. *Crit Rev Biochem Mol Biol* 37:55-69.
20. Nettels, D., S. Müller-Späh, F. Küster, H. Hofmann, D. Haenni, S. Rügger, L. Reymond, A. Hoffmann, J. Kubelka, B. Heinz, K. Gast, R. B. Best, and B. Schuler. 2009. Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins. *Proceedings of the National Academy of Sciences* 106:20740-20745.
21. Aznauryan, M., D. Nettels, A. Holla, H. Hofmann, and B. Schuler. 2013. Single-molecule spectroscopy of cold denaturation and the temperature-induced collapse of unfolded proteins. *J Am Chem Soc* 135:14040-14043.
22. Adrover, M., G. Martorell, S. R. Martin, D. Urosev, P. V. Konarev, D. I. Svergun, X. Daura, P. Temussi, and A. Pastore. 2012. The role of hydration in protein stability: comparison of the cold and heat unfolded states of Yfh1. *J Mol Biol* 417:413-424.
23. Jacob, J., R. S. Dothager, P. Thiyagarajan, and T. R. Sosnick. 2007. Fully Reduced Ribonuclease A Does not Expand at High Denaturant Concentration or Temperature. *Journal of Molecular Biology* 367:609-615.
24. Song, J., G.-N. Gomes, T. Shi, C. C. Gradinaru, and H. S. Chan. 2017. Conformational Heterogeneity and FRET Data Interpretation for Dimensions of Unfolded Proteins. *Biophysical Journal* 113:1012-1024.
25. Ruff, K. M., and A. S. Holehouse. SAXS versus FRET: A Matter of Heterogeneity? *Biophysical Journal* 113:971-973.
26. Fuertes, G., N. Banterle, K. M. Ruff, A. Chowdhury, D. Mercadante, C. Koehler, M. Kachala, G. Estrada Girona, S. Milles, A. Mishra, P. R. Onck, F. Gräter, S. Esteban-Martín, R. V. Pappu, D. I. Svergun, and E. A. Lemke. 2017. Decoupling of size and shape fluctuations in heteropolymeric sequences reconciles discrepancies in SAXS vs. FRET measurements. *Proceedings of the National Academy of Sciences* 114:E6342-E6351.
27. Holehouse, A. S., and R. V. Pappu. 2018. Collapse Transitions of Proteins and the Interplay Among Backbone, Sidechain, and Solvent Interactions. *Annual Review of Biophysics* 47:null.
28. Holehouse, A. S., K. Garai, N. Lyle, A. Vitalis, and R. V. Pappu. 2015. Quantitative Assessments of the Distinct Contributions of Polypeptide Backbone Amides versus Side Chain Groups to Chain Expansion via Chemical Denaturation. *Journal of the American Chemical Society* 137:2984-2995.
29. Das, R. K., and R. V. Pappu. 2013. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proceedings of the National Academy of Sciences* 110:13392-13397.
30. Kitahara, R., A. Okuno, M. Kato, Y. Taniguchi, S. Yokoyama, and K. Akasaka. 2006. Cold denaturation of ubiquitin at high pressure. *Magn Reson Chem* 44 Spec No:S108-113.
31. Pometun, M. S., R. W. Peterson, C. R. Babu, and A. J. Wand. 2006. Cold denaturation of encapsulated ubiquitin. *J Am Chem Soc* 128:10652-10653.

32. Van Horn, W. D., A. K. Simorellis, and P. F. Flynn. 2005. Low-temperature studies of encapsulated proteins. *J Am Chem Soc* 127:13553-13560.
33. Whitten, S. T., A. J. Kurtz, M. S. Pometun, A. J. Wand, and V. J. Hilser. 2006. Revealing the nature of the native state ensemble through cold denaturation. *Biochemistry* 45:10163-10174.
34. Adrover, M., V. Esposito, G. Martorell, A. Pastore, and P. A. Temussi. 2010. Understanding cold denaturation: the case study of Yfh1. *J Am Chem Soc* 132:16240-16246.
35. Martin, S. R., V. Esposito, P. De Los Rios, A. Pastore, and P. A. Temussi. 2008. Cold denaturation of yeast frataxin offers the clue to understand the effect of alcohols on protein stability. *J Am Chem Soc* 130:9963-9970.
36. Pastore, A., S. R. Martin, A. Politou, K. C. Kondapalli, T. Stemmler, and P. A. Temussi. 2007. Unbiased cold denaturation: low- and high-temperature unfolding of yeast frataxin under physiological conditions. *J Am Chem Soc* 129:5374-5375.
37. Li, Y., F. Picart, and D. P. Raleigh. 2005. Direct Characterization of the Folded, Unfolded and Urea-denatured States of the C-terminal Domain of the Ribosomal Protein L9. *Journal of Molecular Biology* 349:839-846.
38. Shan, B., S. McClendon, C. Rospigliosi, D. Eliezer, and D. P. Raleigh. 2010. The Cold Denatured State of the C-terminal Domain of Protein L9 Is Compact and Contains Both Native and Non-native Structure. *Journal of the American Chemical Society* 132:4669-4677.
39. Luan, B., B. Shan, C. Baiz, A. Tokmakoff, and D. P. Raleigh. 2013. Cooperative cold denaturation: the case of the C-terminal domain of ribosomal protein L9. *Biochemistry* 52:2402-2409.
40. Li, Y., B. Shan, and D. P. Raleigh. 2007. The cold denatured state is compact but expands at low temperatures: hydrodynamic properties of the cold denatured state of the C-terminal domain of L9. *J Mol Biol* 368:256-262.
41. Sato, S., and D. P. Raleigh. 2002. pH-dependent stability and folding kinetics of a protein with an unusual alpha-beta topology: the C-terminal domain of the ribosomal protein L9. *J Mol Biol* 318:571-582.
42. Rohl, C. A., and R. L. Baldwin. 1997. Comparison of NH Exchange and Circular Dichroism as Techniques for Measuring the Parameters of the Helix-Coil Transition in Peptides. *Biochemistry* 36:8435-8442.
43. Rice, S. A. 1956. Small angle scattering of X-rays. A. Guinier and G. Fournet. Translated by C. B. Wilson and with a bibliographical appendix by K. L. Yudowitch. Wiley, New York, 1955. *Journal of Polymer Science* 19:594-594.
44. Holehouse, A. S., R. K. Das, J. N. Ahad, M. O. G. Richardson, and R. V. Pappu. 2017. CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophysical Journal* 112:16-21.
45. Kohn, J. E., I. S. Millett, J. Jacob, B. Zagrovic, T. M. Dillon, N. Cingel, R. S. Dothager, S. Seifert, P. Thiyagarajan, T. R. Sosnick, M. Z. Hasan, V. S. Pande, I. Ruczinski, S. Doniach, and K. W. Plaxco. 2004. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proceedings of the National Academy of Sciences of the United States of America* 101:12491-12496.

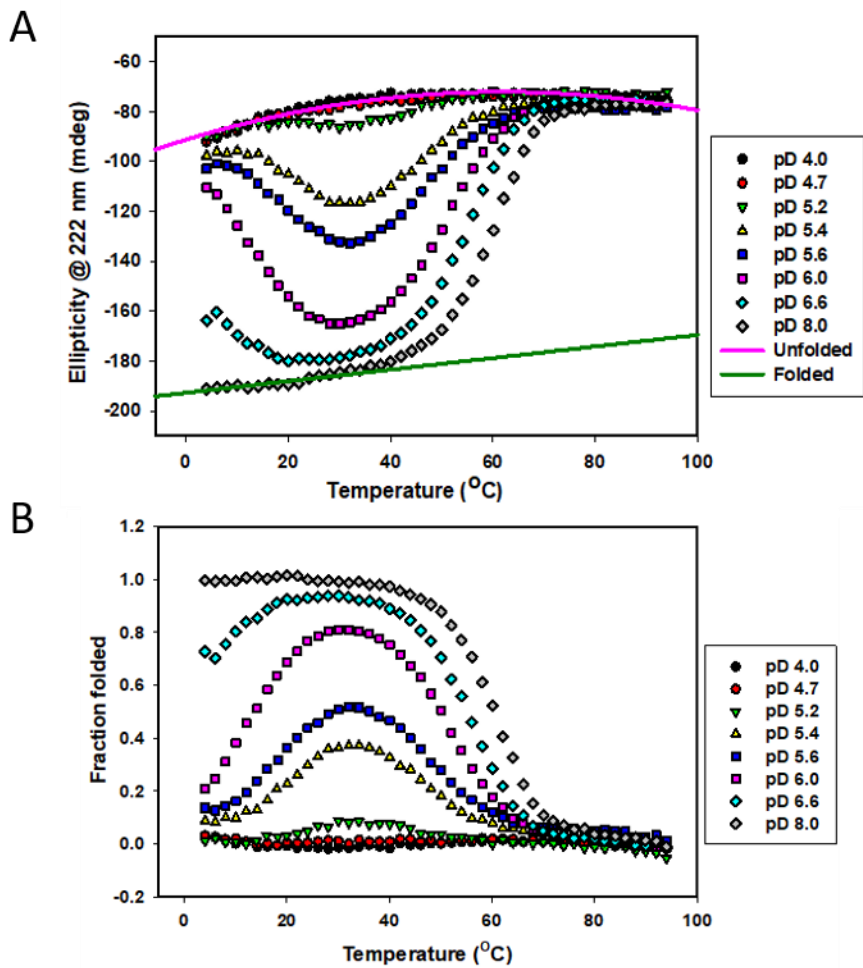
46. Wuttke, R., H. Hofmann, D. Nettels, M. B. Borgia, J. Mittal, R. B. Best, and B. Schuler. 2014. Temperature-dependent solvation modulates the dimensions of disordered proteins. *Proceedings of the National Academy of Sciences* 111:5213-5218.
47. Dunker, A. K., M. M. Babu, E. Barbar, M. Blackledge, S. E. Bondos, Z. Dosztányi, H. J. Dyson, J. Forman-Kay, M. Fuxreiter, J. Gsponer, K.-H. Han, D. T. Jones, S. Longhi, S. J. Metallo, K. Nishikawa, R. Nussinov, Z. Obradovic, R. V. Pappu, B. Rost, P. Selenko, V. Subramaniam, J. L. Sussman, P. Tompa, and V. N. Uversky. 2013. What's in a name? Why these proteins are intrinsically disordered. *Intrinsically Disordered Proteins* 1:e24157.
48. Firman, T., and K. Ghosh. 2018. Sequence charge decoration dictates coil-globule transition in intrinsically disordered proteins. *The Journal of Chemical Physics* 148:123305.
49. Piana, S., J. L. Klepeis, and D. E. Shaw. 2014. Assessing the accuracy of physical models used in protein-folding simulations: quantitative evidence from long molecular dynamics simulations. *Current Opinion in Structural Biology* 24:98-105.

## Figure Captions



### Figure 1. Schematic of I98A-CTL9

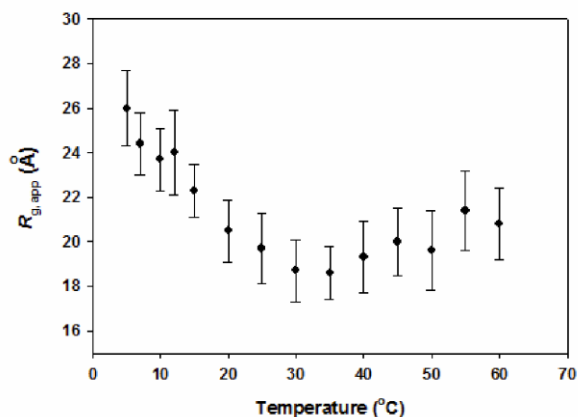
A) Ribbon diagram of I98A-CTL9 with I98 shown in black in stick format. Figure is constructed from PDB ID: 1DIV. B) Primary sequence of I98A-CTL9 (residues 58-149 of L9) with schematic diagram of secondary structural elements where arrows represent  $\beta$ -strands and bars represent  $\alpha$ -helices. The location of the I98A mutation is shown in red in the primary sequence.



**Figure 2. pH and temperature dependent unfolding studies reveal cold and thermally induced unfolding transitions**

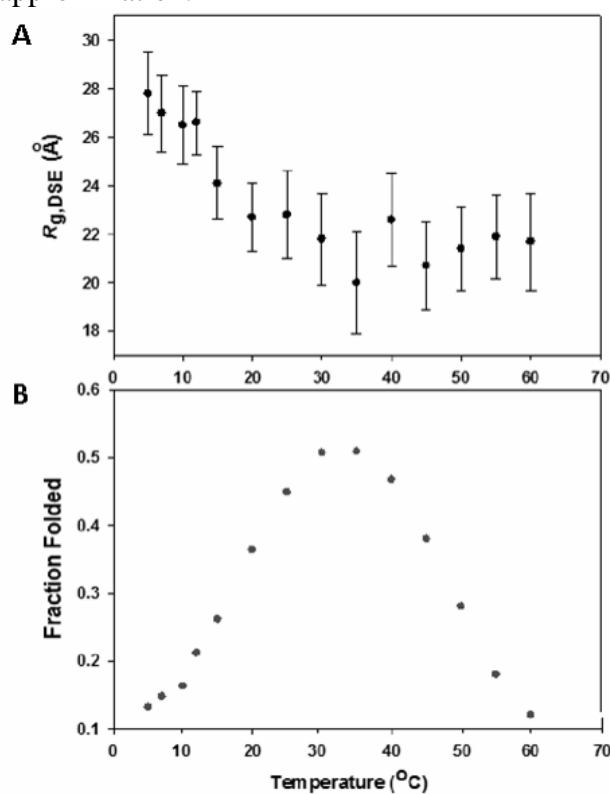
(A) CD monitored unfolding curves of I98A-CTL9 as a function of temperature at different pD values (uncorrected pH reading), in 10 mM DMG and 120 mM NaCl, 100% D<sub>2</sub>O. The estimated folded (green) and unfolded (pink) baselines are shown. The folded baseline was determined by an extrapolation of the pre-unfolding transition of the thermal denaturation curve at pD 8.0. The unfolded baseline was determined by a quadratic fit to the pD 4.0 data. The fraction folded at any given temperature and pD value can be estimated from these plots. (B) Plots of fraction folded as a function of temperature at different pD values.





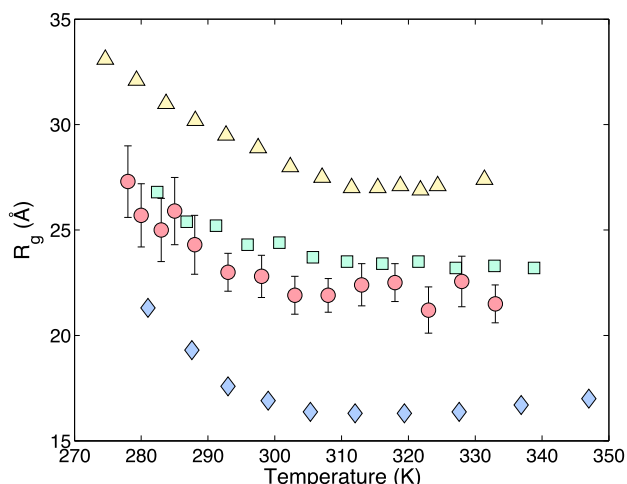
**Figure 3. Apparent  $R_g$  of I98A-CTL9 as a function of temperature**

Apparent  $R_g$  values at different temperatures from SAXS studies of I98A-CTL9 in 10 mM DMG and 120 mM NaCl, 100%  $D_2O$ .  $R_{g,app}$  values were calculated from a Guinier analysis of the observed I98A-CTL9 SAXS data. The error bars represent the linear fitting in the Guinier approximation.

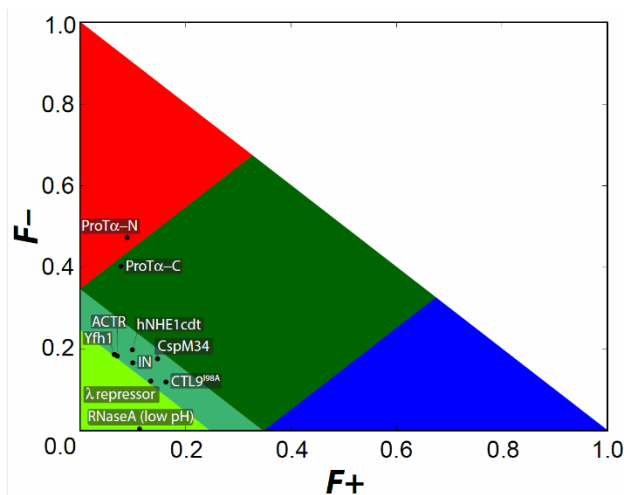


**Figure 4. The  $R_g$  of I98A-CTL9 DSE plotted as a function of temperature**

(A) The  $R_{g,DSE}$  of I98A-CTL9 as a function of temperature, in 10 mM DMG and 120 mM NaCl, 100%  $D_2O$  calculated by equation 7. Error bars were determined from the linear fitting in the Guinier analysis. (B) A plot of the fraction of folded I98A-CTL9 over the same range of temperature used for the SAXS experiments.



**Figure 5. Comparison of the temperature-dependent  $R_g$  values of the DSE of I98A-CTL9 (red circles) monomeric  $\lambda$ -repressor (blue diamonds), CspTm (green squares) and Yhf1 (yellow triangles).** The vertical axis is the radius of gyration of the unfolded state, calculated from SAXS data (I98A-CTL9 in red circles), or extrapolated from smFRET data ( $\lambda$ -repressor in blue diamonds, CspTm in green squares, and Yhf1 in yellow triangles) I98A-CTL9 data were taken from column  $d$  in table 1. The smFRET data were extracted from references 19,20, and 44 in the main text. Although the magnitude of the radii of gyration vary from protein to protein, the trends observed are highly similar. In all four cases, these proteins coexist in a folded-unfolded equilibrium over at least part of the temperature range examined. The reduced extent of expansion at low temperature for CsptTm may reflect the fact that these experiments were done under 0.49 M GdmCl, while all others were done under native buffer conditions.



**Figure 6. Diagram of states generated by CIDER.** A comparison of 10 proteins are plotted with respect to various sequence attributes using the analysis developed by Das and Pappu (44). The x-axis represents fraction of positively charged residues ( $F_+$ ) and the y-axis is the fraction of negatively charged residues ( $F_-$ ). Light green refers to weak polyampholytes and polyelectrolytes (sequences with a low fraction of charged residues). Medium green refers to the sequences with an intermediate fraction of charged residues. Dark green refers to strong polyampholytes. Red

and blue refer to negatively and positively charged strong polyelectrolytes, both of which are expected to behave as swollen coils.

## Supporting Material

### **The Unfolded State of the C-Terminal Domain of L9 Expands at Low Temperature but not at Elevated Temperatures**

N. Stenzoski<sup>1</sup>, B. Luan<sup>2</sup>, A.S. Holehouse<sup>3</sup>, D. Raleigh<sup>1,2,4\*</sup>

<sup>1</sup>Graduate Program in Biochemistry & Structural Biology, <sup>2</sup>Department of Chemistry Stony Brook University, Stony Brook, New York 11794-3400, United States, <sup>3</sup>Center for Biological Systems Engineering, Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA <sup>4</sup>Institute of Structural and Molecular Biology, University of College London. Gower St, London WC1E.6BT, United Kingdom

<b>pD</b>	<b>Folded %</b>	<b><math>R_{g,app}(\text{\AA})^a</math></b>	<b><math>R_{g,DSE}(\text{\AA})^b</math></b>	<b><math>R_{g,DSE}(\text{\AA})^c</math></b>
5.0	3.0	25.5±1.9	24.0±1.6	25.8±1.2
5.6	12.1	25.0±2.1	23.4±1.9	26.1±1.3
6.2	34.2	21.7±1.8	21.7±2.1	22.8±1.4

**Table S1. The pD-dependence of the  $R_g$  values of I98A-CTL9 DSE at 60°C, in 10 mM DMG and 120 mM NaCl, 100% D<sub>2</sub>O**

a) Apparent  $R_g$  values obtained directly by the Guinier analysis from the observed I98A-CTL9 scattering profile. b)  $R_g$  values of the I98A-CTL9 DSE obtained by Guinier analysis of the DSE scattering profile after subtracting the signal from the weighted folded state contribution. c)  $R_g$  values of the I98A-CTL9 DSE calculated by the equation describing a two-component system.

Temperature (°C)	$P_{\text{folded}}^{\text{a}}$	Helix (%) <sup>b</sup>
5	13.2	11
15	26.2	9
50	28.1	7
60	12.1	3

**Table S2. Calculated helical fraction of I98A-CTL9**

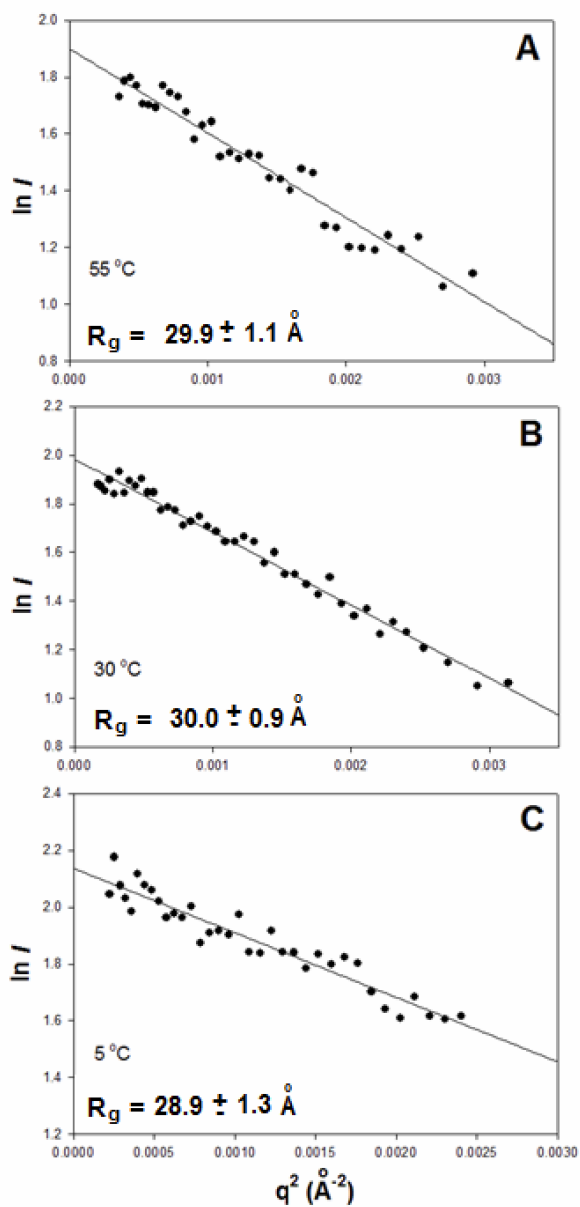
Helical fraction estimated for the I98A-CTL9 DSE using the mean residue ellipticity monitored at 222 nm by CD, in 10 mM DMG and 120 mM NaCl, 100% D<sub>2</sub>O, pD 5.6

a) Fraction folded ensemble at given temperature. b) Estimated by equation 3, using  $[\theta]_c$  calculated by equation 5.

<b>Protein</b>	<b>Kappa</b>	<b>Hydropathy</b>	<b>NCPR</b>
rRNaseA	0.140	3.84	0.032
Yfh1	0.224	4.07	-0.122
$\lambda$ -Repressor	0.178	4.25	0.012
CspM34	0.193	4.11	-0.029
IN	0.107	3.97	-0.067
ACTR	0.149	3.83	-0.113
hNHE1cdt	0.266	3.68	-0.099
ProT $\alpha$ N	0.399	2.60	-0.384
ProT $\alpha$ C	0.416	2.87	-0.326
I98A CTL9	0.117	4.13	0.043

**Table S3. Sequence analysis of unfolded proteins and IDPs**

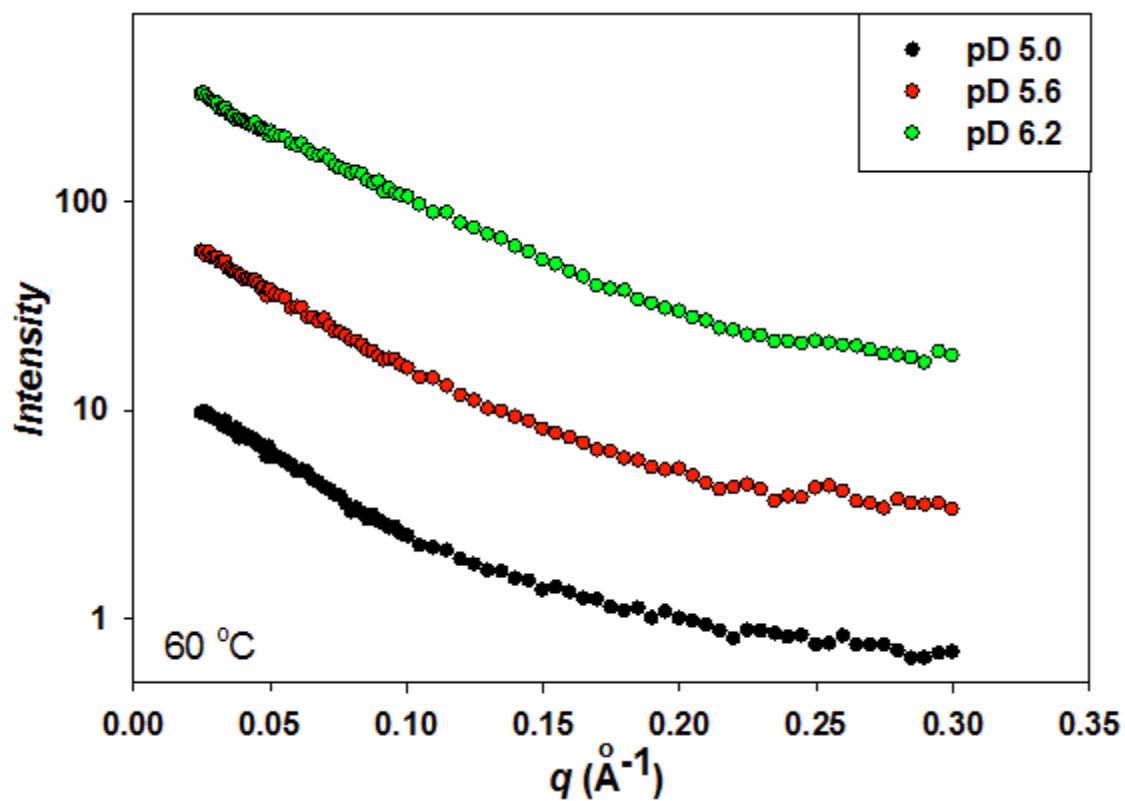
Comparison of I98A-CTL9 to seven other unfolded proteins or IDPs. The sequence of each protein was analyzed to determine various sequence attributes including hydropathy, net charge per residue (NCPR), and patterning of oppositely charged residues (kappa) using the publicly available CIDER software (44).



**Figure S1. Guinier analysis of I98A-CTL9 in the presence of denaturant**

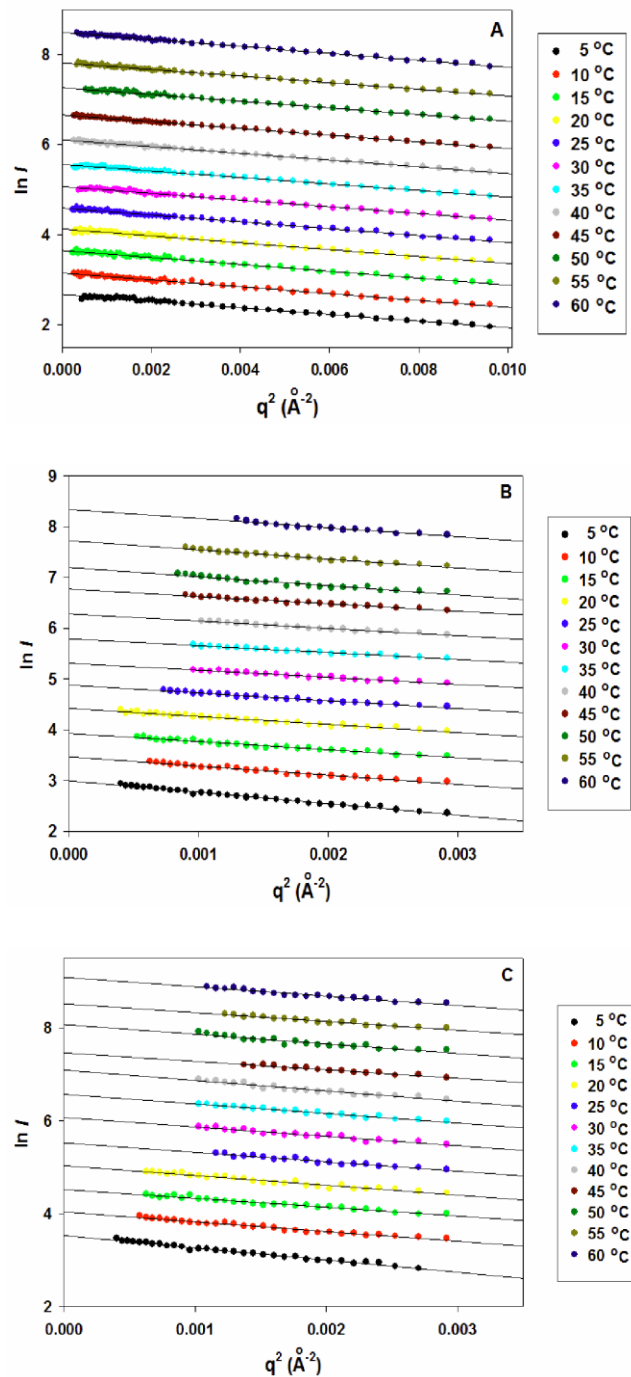
Guinier analysis of I98A-CTL9 SAXS data collected in 6M GdnHCl, pD 5.6, at (A)  $55^\circ\text{C}$ , (B)  $30^\circ\text{C}$ , (C)  $5^\circ\text{C}$ . The  $R_g$  values obtained from the analysis agree with the  $R_g$  value expected for a random coil model. The buffer consists of 10 mM DMG and 120 mM NaCl.





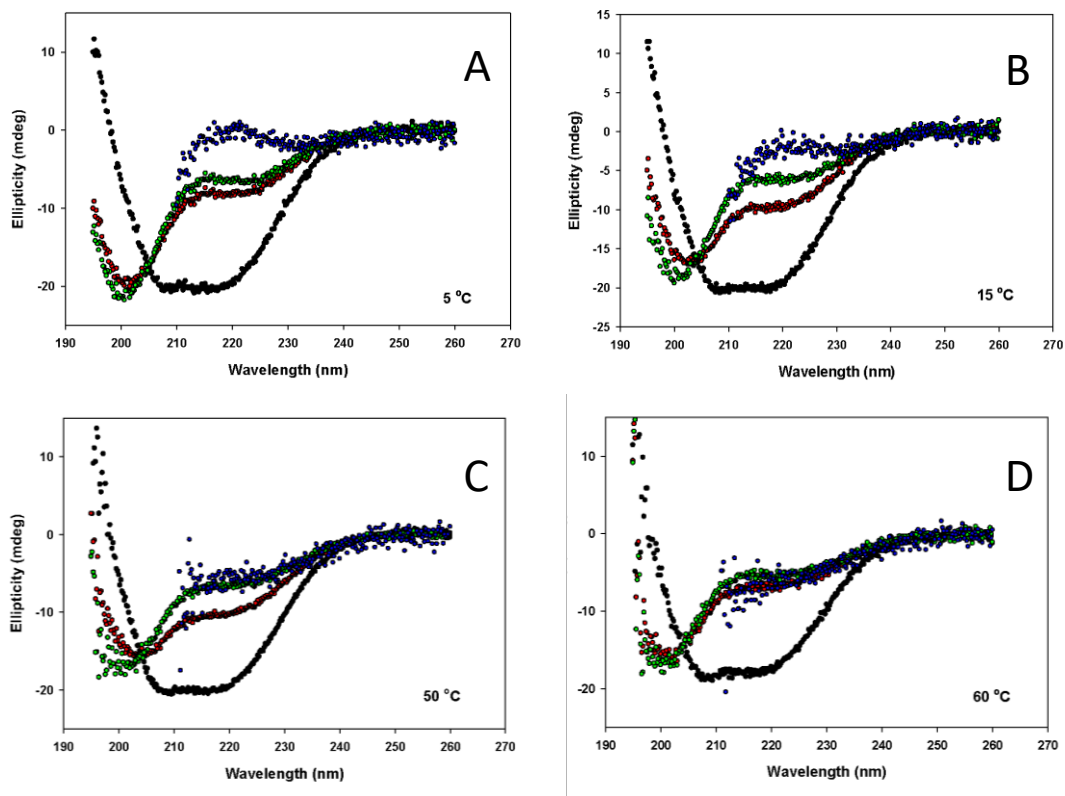
**Figure S2. Scattering profiles of I98A-CTL9 at 60°C**

Data were collected at three pD values: 5.0, 5.6, and 6.2. The buffer consists of 10 mM DMG and 120 mM NaCl, in 100% D<sub>2</sub>O.



**Figure S3. Temperature dependent analysis of the DSE of I98A-CTL9 by SAXS.**

Guinier plots of (A) wild type CTL9 control scattering data, (B) the I98A-CTL9 observed scattering profile before subtracting the folded contribution, and (C) the I98A-CTL9 DSE SAXS scattering profile from 5 to 60°C, in 10 mM DMG and 120 mM NaCl, 100% D<sub>2</sub>O at pD 5.6 obtained after subtracting the contribution of the folded state. The curves are offset for clarity. Note the different x-axis for the wild type protein.



**Figure S4. CD spectra at different temperatures**

The CD spectra of I98A-CTL9 (red), and wild type CTL9 (black) in 10 mM DMG and 120 mM NaCl, 100% D<sub>2</sub>O at pD 5.6 (uncorrected pH-meter reading). The CD spectrum of the DSE of I98A-CTL9 in buffer (green) was calculated by subtracting the population-weighted CD signal of the folded state (using the CD spectrum of CTL9 WT as the folded spectrum). The CD spectrum of I98A-CTL9 in 8M urea (blue) is shown for comparison. A) 5 °C; B) 15 °C; C) 25 °C, D) 60 °C.