

Evolutionarily Conserved Pattern of Interactions in a Protein Revealed by Local Thermal Expansion Properties

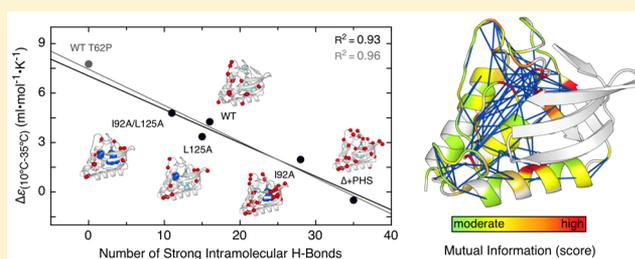
Mariano Dellarole,^{†,§} Jose A. Caro,^{‡,#} Julien Roche,^{†,||} Martin Fossat,^{†,⊥} Philippe Barthe,[†] Bertrand García-Moreno E.,[‡] Catherine A. Royer,^{*,†,⊥} and Christian Roumestand^{*,†}

[†]Centre de Biochimie Structurale, CNRS UMR5048, INSERM U554, Université Montpellier 1, 29 rue de Navacelles, Montpellier, France 34090

[‡]T. C. Jenkins Department of Biophysics, Johns Hopkins University, 3400 N. Charles St., Baltimore, Maryland 21218, United States

S Supporting Information

ABSTRACT: The way in which the network of intramolecular interactions determines the cooperative folding and conformational dynamics of a protein remains poorly understood. High-pressure NMR spectroscopy is uniquely suited to examine this problem because it combines the site-specific resolution of the NMR experiments with the local character of pressure perturbations. Here we report on the temperature dependence of the site-specific volumetric properties of various forms of staphylococcal nuclease (SNase), including three variants with engineered internal cavities, as measured with high-pressure NMR spectroscopy. The strong temperature dependence of pressure-induced unfolding arises from poorly understood differences in thermal expansion between the folded and unfolded states. A significant inverse correlation was observed between the global thermal expansion of the folded proteins and the number of strong intramolecular hydrogen bonds, as determined by the temperature coefficient of the backbone amide chemical shifts. Comparison of the identity of these strong H-bonds with the co-evolution of pairs of residues in the SNase protein family suggests that the architecture of the interactions detected in the NMR experiments could be linked to a functional aspect of the protein. Moreover, the temperature dependence of the residue-specific volume changes of unfolding yielded residue-specific differences in expansivity and revealed how mutations impact intramolecular interaction patterns. These results show that intramolecular interactions in the folded states of proteins impose constraints against thermal expansion and that, hence, knowledge of site-specific thermal expansivity offers insight into the patterns of strong intramolecular interactions and other local determinants of protein stability, cooperativity, and potentially also of function.



INTRODUCTION

The folded states of proteins are marginally stable relative to their unfolded states. The stability and conformational landscape of each protein has evolved to enable function appropriate to the organism and environment in which the protein functions. A fundamental understanding of how the sequence of amino acids determines the folding and functional properties of a protein requires detailed characterization of conformational landscapes and the effects of mutations thereon. This usually involves perturbation of the protein in a controlled fashion to attempt to identify the conformational states that constitute the ensemble under a variety of conditions.

Folding equilibria can be perturbed by changing the fundamental thermodynamic variables, temperature and pressure, by changing a fundamental physiological variable, pH, or by the addition of chemical compounds known to either destabilize or stabilize the folded state. Temperature, pH, and chemical denaturants have been widely used to probe the physical mechanism and sequence determinants of protein folding and stability. The mechanism of action of temperature and of chemical denaturants is governed by the amount of

surface area that is exposed to either solvent or denaturant upon unfolding.¹ pH effects are governed by differences in pK_a values of ionizable groups in the different conformational states in the ensemble.²

Pressure has been used also to unfold proteins, although to a much lesser extent, and until recently its mechanism of action was not well-understood. Differences in solvent density related to the hydration of surface area exposed by unfolding were thought to contribute significantly to the volume change upon unfolding. Such effects, like those of denaturants, should scale with the size of the protein. In contrast to this expectation, the deletion of 3 out of 7 repeats of the ankyrin repeat domain of the Notch receptor did not decrease the magnitude of the volume change for folding, ΔV_f . Indeed, deletion of the first two repeats actually increased ΔV_f .³ These studies demonstrate the lack of significant contribution of hydration effects to ΔV_f . This may arise from compensation of effects of different sign from the polar backbone and hydrophobic moieties. In contrast

Received: April 26, 2015

Published: July 2, 2015

to the lack of effect of the size of the protein, single amino acid substitutions that created cavities in the interior of a globular protein, staphylococcal nuclease (SNase), could double ΔV_f .⁴ Together our results led to the conclusion that the internal solvent-excluded void volume or packing defects in the folded state is the major contributing factor to the magnitude of pressure effects on thermodynamic stability.^{3,4}

Packing defects are local features, specific to the structure of each protein. Thus, the effects of pressure on stability are exerted locally, and hence, the extent to which pressure disrupts the structure globally depends on the internal network of interactions unique to the fold of the individual protein. Characterization of the volumetric properties of proteins with the site-specific resolution afforded by NMR spectroscopy should provide insight into the structural basis and sequence determinants of stability and folding cooperativity.

Since Bridgman first reported on the effects of pressure on a protein,⁵ it has been known that volume changes for unfolding are strongly dependent on temperature. Protein stability diagrams in the pressure–temperature plane, assuming cooperative two-state equilibrium between folded and unfolded states and energetically equivalent unfolded states, can be described by an ellipse,^{6–9} where the first- and second-order parameters, ΔH , ΔS , ΔC_p , ΔV , $\Delta\kappa'$, and $\Delta\varepsilon$ correspond to the differences in molar enthalpy, entropy, heat capacity, volume, compressibility ($d\Delta V/dp$) and thermal expansivity ($d\Delta V/dT$) between folded and unfolded states, and where T_o and p_o are the chosen reference points in temperature and pressure (298 K and 1 bar).

$$\begin{aligned} \Delta G(p, T) = & \Delta G^\circ - \Delta C_p \left[T \left(\ln \frac{T}{T_o} - 1 \right) + T \right] - \Delta S(T - T_o) \\ & + \Delta V(p - p_o) - \frac{1}{2} \Delta\kappa' (p - p_o)^2 \\ & - \Delta\varepsilon(T - T_o)(p - p_o) \end{aligned} \quad (1)$$

In contrast to the rather small differences in compressibility, $\Delta\kappa'$ (more than 10-fold smaller than the ΔV and within the uncertainty of our measurement),⁸ the differences in thermal expansivity, $\Delta\varepsilon$, between folded and unfolded states of proteins are significant relative to ΔV and are the cause of the strong temperature dependence of pressure effects. Indeed, the magnitude of ΔV for unfolding of proteins decreases significantly with increasing temperature and can even change sign. The molar thermal expansivity of the unfolded states of proteins can be modeled empirically based on sequence alone, assuming additivity of the expansivity values of the individual amino acid residues.^{10,11} On the other hand, the molecular contributions to the thermal expansion of folded states are not known. Substitution of a single amino acid in the core of a protein can change the molar expansivity of their folded state significantly (up to 3-fold), whereas significant changes to the polarity at the protein–water interface make very little difference.^{12,13} Earlier studies using the temperature dependence of the intrinsic tryptophan fluorescence anisotropy^{14–16} suggest that the intramolecular interactions in the folded state may act as constraints against thermal expansion. Hence, characterization of thermal expansivity with site-specific resolution could yield significant insight into the organization of intramolecular interactions of the folded protein and how they depend on sequence.

To probe systematically the molecular determinants of the thermal expansion of a protein, we examined the temperature

dependence of the volumetric properties SNase using the highly stable form of this protein known as Δ +PHS, and variants of Δ +PHS in which cavities were engineered by substitution of core residues by alanine. Some of these variants were used previously to demonstrate the importance of packing defects in the magnitude of the pressure effect.⁴ In these prior studies, the three-dimensional (3-D) structures of the single substitution cavity variants were determined by X-ray crystallography, and the existence of the expected cavities was confirmed. The structure of a double variant (I92A/L125A) can be found in Supporting Information. Negligible rearrangements in the structures of the cavity containing variants were observed relative to the reference protein. Moreover, no penetration of water molecules into these cavities could be inferred from the electron densities, although this does not rule out penetration and the presence of transient or disordered water molecules in the cavities. Likewise, NMR chemical shift perturbations by the mutations were minimal,¹⁷ particularly for the variants bearing cavities in the OB-fold region of the protein.

In the present work, the temperature response of the folded states of several of these cavity containing variants and the pressure-induced unfolding at different temperatures was monitored using 2-D NMR spectroscopy. Results from these experiments were compared with volume changes at the folding transition temperature, T_m , and with the expansivity values of their folded states at low temperature, measured previously by pressure perturbation calorimetry (PPC).¹² The temperature dependence of the amide proton chemical shifts revealed a strong inverse correlation between the number of amide resonances exhibiting particularly small temperature coefficients, which are equated with strong intramolecular hydrogen bonds (H-bonds),¹⁸ and the thermal expansion of the folded states determined by PPC. The specific patterns of H-bonds and the perturbations affected by mutations compared with the co-evolution of pairs of residues in the SNase family suggest how the protein's sequence evolved with respect to function.

MATERIALS AND METHODS

NMR. All proteins were produced as described earlier.⁴ Uniform ¹⁵N labeling was obtained from overexpression of recombinant protein in *Escherichia coli* grown in M9 medium containing ¹⁵NH₄Cl as the sole nitrogen source, as described for SNase previously.¹⁹ Uniformly ¹⁵N-labeled Δ +PHS SNase and its variants with either I92A, L125A, or I92A/L125A substitutions were dissolved at approximately 1 mM concentration in 300 μ L 50 mM Tris buffer at pH7. True wild-type SNase was dissolved at a similar concentration but in 50 mM Bis-Tris buffer at pH5.5. 10% of D₂O was added for the lock procedure. In all experiments the ¹H carrier was centered on the water resonance, and a WATERGATE sequence^{20,21} was incorporated to suppress solvent resonances. All NMR spectra were processed and analyzed with GIFA.²² ¹H and ¹⁵N resonance assignments were available for the wild-type SNase and for the Δ +PHS protein and its variants with either I92A or L125A.⁴ Amide resonances of the I92A/L125A variant were assigned at atmospheric pressure from 3D [¹H,¹⁵N] NOESY-HSQC (mixing time 150 ms) and 3D [¹H,¹⁵N] TOCSY-HSQC (isotropic mixing 60 ms) double-resonance experiments^{23,24} recorded on a Bruker Avance III 700 MHz spectrometer equipped with a 5 mm Z-gradient ¹H–¹³C–¹⁵N cryogenic probe, using the standard sequential procedure. ¹H chemical shifts were directly referenced to the methyl resonance of DSS, and ¹⁵N chemical shifts were referenced indirectly to the absolute frequency ratios ¹⁵N/¹H = 0.101329118.

Variable pressure experiments were recorded at four different temperatures (288, 293, 298, and 303 K) on a 600 MHz Bruker Avance III spectrometer equipped with a 5 mm Z-gradient ¹H-X

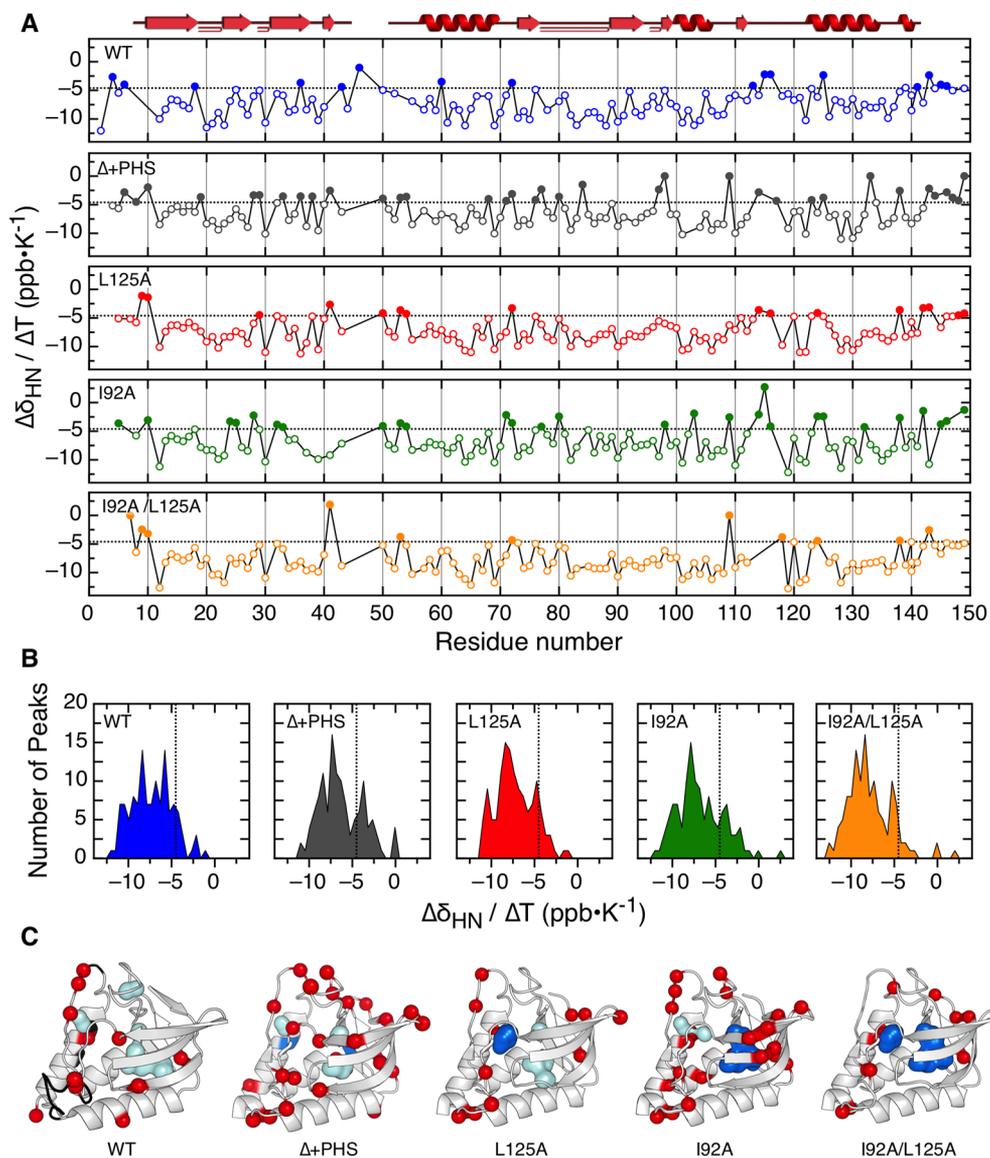


Figure 1. Intramolecular H-bonds variations among SNase proteins. (A) Amide proton temperature coefficients of WT, Δ +PHS SNase, and of cavity enlarging variants with L125A, I92A, and I92A/L125A substitutions. Values above the solid line ($-4.5 \text{ ppb}\cdot\text{K}^{-1}$) are represented with filled symbols. (B) Histogram of $\Delta\delta_{\text{NH}}/\Delta T$ for each protein. The solid line corresponds to $-4.5 \text{ ppb}\cdot\text{K}^{-1}$; see Materials and Methods. (C) Structures of WT, Δ +PHS, and variants with substitutions L125A, I92A or I92A/L125A. Locations of amides with $\Delta\delta_{\text{NH}}/\Delta T$ values above $-4.5 \text{ ppb}\cdot\text{K}^{-1}$ are red spheres. Residues that were deleted or mutated from WT to engineer Δ +PHS SNase are colored black. Residues L125 and I92 in Δ +PHS are colored blue. The molecular surface of the internal void volumes is shown in (light blue) and highlighted in dark blue when additional void is present due to mutation.

double-resonance broadband inverse (BBI) probe. Commercial zirconia ceramic high-pressure NMR tubes connected to an Xtreme 60 syringe automatic pump (Daedalus Innovations, Philadelphia, PA) were used to vary the pressure in the 1–2500 bar range. Fourteen 2D [^1H , ^{15}N] HSQC spectra were recorded at variable pressure to monitor protein unfolding. Subdenaturing 1.8, 0.5, and 0.75 M guanidinium chloride concentrations were added to the NMR sample to achieve the complete denaturation of Δ +PHS and variants with L125A and I92A, respectively, in the pressure range accessible by the instrumentation. The $^{15}\text{N}/^1\text{H}$ cross-peak maximal intensities were used to construct the fractional intensity vs pressure plots for each residue. Cross peaks were picked using a box size of 0.3/0.03 ppm ($^{15}\text{N}/^1\text{H}$) with the PARIS algorithm²⁵ included in the GIFA software. Cross peak line width changes were minimal and uniform. Thermal coefficients were calculated as the slope of linear regression fits of ^1H chemical shifts recorded as a function of temperature at 1 bar.¹⁸ Only residues with thermal coefficients calculated from linear regressions with $R^2 > 0.95$

were considered (a total of 622 amides). Apparent residue-specific ΔV were obtained as described previously.^{4,26} Briefly, the HSQC peak intensity pressure profiles were fit to a two-state model, for the ΔG° and ΔV° values for unfolding (and the asymptotic intensity values) at each temperature, assuming that the difference in compressibility between folded and unfolded states was negligible. The residue-specific apparent folded-state expansion corresponds to the slope of the thermal dependence of residue-specific apparent ΔV calculated from linear regression fits. For more detail on our previously published data analysis procedures see Supporting Information. Only residues with ΔV for more than three temperatures and with linear regression coefficients, $R^2 > 0.7$ were considered (a total of 477 amides with an R^2 median of 0.94 ± 0.13). The magnitude distribution of the residue-specific apparent folded-state expansion was independently normalized among variants and reported on the corresponding crystal structures via an in-house Python tool. Experimental data were analyzed using

ProFit (QuantumSoft) and Prism (Graphpad) software packages and plotted using ProFit (QuantumSoft).

Equilibrium Thermodynamics. Differential scanning calorimetry (DSC), PPC, and high-pressure fluorescence experiments were performed as described earlier.^{12,27} Refer to Figures S4 and S5 for details.

Sequence Analysis. Sequences of 3872 staphylococcal nuclease homologues obtained from the NCBI PFAM database (PF00565)²⁸ were reduced to 3196 nonredundant sequences using Duplicate Finder Java standalone application. A subgroup of 1976 homologues sequences of amino acid length equal or higher than 100 were selected using Jalview v 15.0 to discard partial and or incomplete sequences. The selected subgroup was aligned using the online multiple sequence alignment (MSA) tool kalign.²⁹ Residual co-evolution among the MSA was estimated using mutual information (MI) using MISTIC online source.³⁰ The network of co-evolved pairs having MI score >9 were represented on 1SNC pdb structure using MISTIC interactive network view tool.³⁰ Refer to Figure S7 for MI statistical analysis details.

RESULTS

Temperature Dependence of Amide Proton Chemical Shifts.

HSQC spectra were collected over a range of temperatures for wild-type (WT) and Δ +PHS SNase and for the Δ +PHS L125A, Δ +PHS I92A, and Δ +PHS L125A/I92A variants. The crystal structures of these proteins are nearly identical except for size and location of the internal cavities⁴ and Table S1. The differences between the structures of WT and the Δ +PHS protein are limited to the substitutions (G50F, F51N, P117G, H124L, S128A) and the deletion of the Ω -loop (44–49) used to engineer the Δ +PHS variants. The structure of SNase consists of an OB-fold domain (5 stranded β -barrel and α -helix 1), a C-terminal α -helical domain (helix 3), and an interfacial domain with several loops and α -helix 2.

HSQC peak assignments of the backbone amides for all the variants have been previously reported,⁴ except for the double variant (Figure S1).

As noted by Baxter and Williamson,¹⁸ amide proton chemical shifts depend linearly upon temperature below the temperature of unfolding, and different amide protons in protein structures shift to different extents upon heating. This is the result of increased thermal motions. Small temperature coefficients ($\Delta\delta_{\text{NH}}/\Delta T$) for the chemical shifts of particular amide protons (more positive than -4.5 ppb/K) were interpreted by these authors as arising from constraints against expansion due to strong intramolecular hydrogen bonding for those residues. We used the histogram of $\Delta\delta_{\text{NH}}/\Delta T$ for all the residues of the SNase variants to establish the strong H-bond cutoff (Figure S2) at -4.5 ppb/K, as was done by Baxter and Williamson.¹⁸ The temperature coefficients of the amide proton chemical shifts for the SNase variants (Figure 1A,B) revealed an increase in the number of the presumably strong intramolecular H-bonds in the Δ +PHS variant, compared to WT SNase. Interactions were reinforced particularly in the region linking the C-terminal helix to the core of the protein, where the Ω -loop was deleted in the Δ +PHS variant (Figure 1C). The destabilizing cavity creating substitutions I92A and L125A in the Δ +PHS protein lead to a nearly identical loss in global stability relative to the Δ +PHS reference protein,⁴ yet they exhibited very different perturbations to their H-bond pattern. Interestingly, the I92A substitution, deep in the core of the protein, had only a moderate effect on the number and distribution of strong H-bonds. In fact, the H-bonds in the core β -barrel region appear to have been reinforced by the enlargement of the cavity (Figure 1C). In contrast, the

L125A substitution leads to a much larger decrease in the number of strong intramolecular H-bonds compared to the Δ +PHS reference protein, with a pattern resembling that of WT SNase. These observations are consistent with the previously reported chemical shift perturbations between the reference protein and its I92A and L125A variants.¹⁷ In those studies, very small chemical shift perturbations were observed for the I92A mutation, whereas considerable and long-range perturbations were apparent in the HSQC spectra of the L125A variant, relative to the reference protein. The I92A/L125A variant exhibits only a few residual amide protons with small temperature coefficients, indicating only a few residual strong H-bonds remain in the structure of this highly destabilized variant.

Temperature Dependence of Pressure-Induced Unfolding. The pressure-induced unfolding of WT SNase was characterized at four temperatures by observing the pressure dependent decrease in the intensity of the folded-state HSQC resonances (Figure 2). Each of the over 100 residue-specific pressure unfolding profiles at each temperature was fit to a two-state unfolding model to recover residue-specific apparent values for $\Delta G_f^\circ(T)$ and $\Delta V_f^\circ(T)$, the free energy and volume changes for folding, respectively. The distributions for the

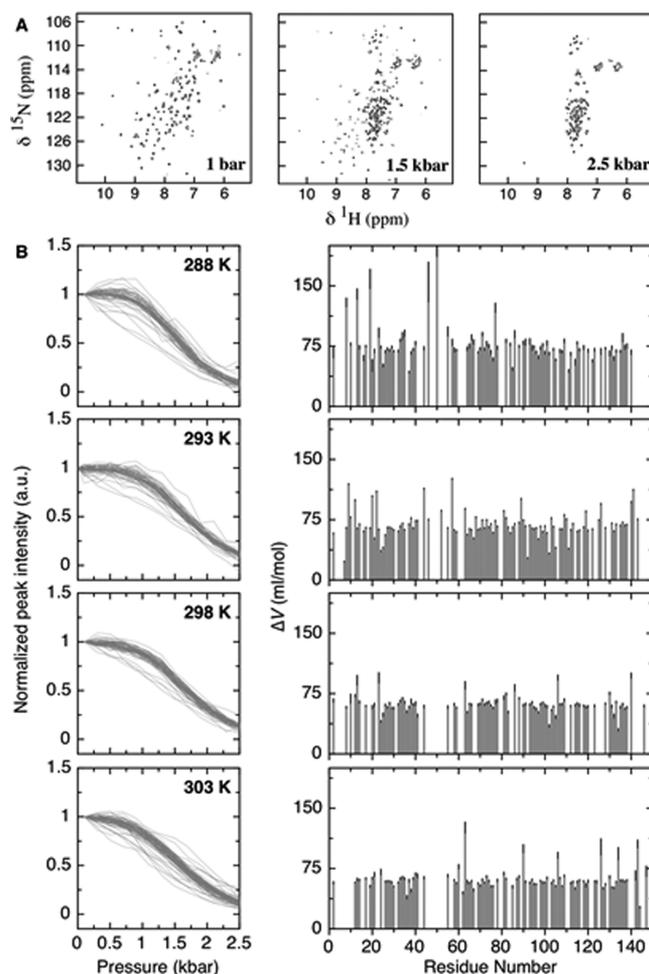


Figure 2. Temperature dependence of pressure induced unfolding of WT SNase monitored by HSQC peak intensity. (A) Representative HSQC spectra recorded at 1 bar, 1.5 kbar, and 2.5 kbar and 293 K. (B) Normalized intensity profiles and fitted ΔV_f values of individual amides as a function of pressure recorded at 288, 293, 298, and 303 K.

apparent $\Delta V_f^\circ(T)$ values (Figure 3) were fairly narrow for the WT protein, close to the uncertainty of the measurement,

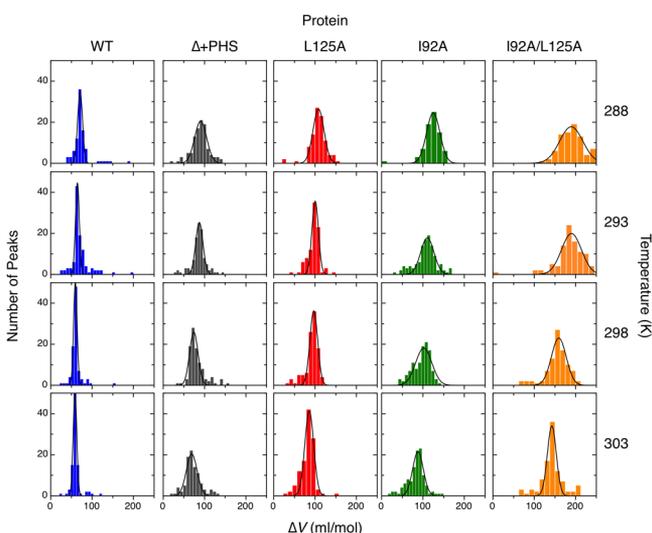


Figure 3. Temperature dependence of the distribution of ΔV_f values among SNase protein and cavity enlarging mutants. Bins of ΔV_f values of SNase WT, the highly stable Δ +PHS, and the corresponding cavity enlarging variants L125A, I92A, and I92A/L125A from all 2D ^{15}N - ^1H HSQC fitted peaks recorded at 288, 293, 298, and 300 K. The line is a fit to a Gaussian function.

although a few outliers are apparent in the distributions, and some broadening occurred at low temperature. This indicates that equilibrium pressure unfolding of WT SNase can be modeled accurately as a two-state transition. It is known that the SNase WT folding mechanism follows a “foldon” type scheme, with the β -strand 5 and α -helix 2 of the OB-fold representing the most stable and first folding unit.³¹ A kinetic intermediate after the major folding barrier involves an ordered OB-fold region and disorder in the interface and much of the C-terminal helix.³² We have observed both the foldon behavior and the population of this intermediate in p-jump NMR³³ and in pressure-dependent hydrogen-exchange experiments.²⁶ However, the equilibrium population of these intermediates under pressure is negligible for WT SNase, leading to the narrow distribution of the over 100 apparent volume changes obtained from the high-pressure NMR curves and allowing analysis according to a two-state model.

NMR-detected pressure-induced unfolding of the Δ +PHS variant (Figure S3A) yielded site-specific unfolding profiles and site specific apparent values of $\Delta G_f^\circ(T)$ and $\Delta V_f^\circ(T)$ at each temperature (Figure 3). The distributions of the residue-specific apparent $\Delta V_f^\circ(T)$ for the Δ +PHS variant (Figure 3) were somewhat broader than for the WT SNase. At the higher temperatures they were rather asymmetric. This heterogeneity in the recovered parameters is consistent with some departure from a two-state transition for the highly stable Δ +PHS variant, as previously reported.⁴ Small apparent values for ΔV_f° obtained from the fits of the pressure dependence of the HSQC peak intensity of a given residue are indicative of partial unfolding involving that residue. Nonetheless, the distributions of apparent volume changes remain reasonably narrow, allowing us to consider the peak of the distribution to be a reasonable estimation of the thermodynamic $\Delta V_f^\circ(T)$. For both proteins the average value of $\Delta V_f^\circ(T)$ decreased as a function of increasing temperature (Table 1). This is a general phenomenon observed for all proteins and is due to the smaller thermal expansivity of the folded state relative to that of the unfolded state, although deviations from two-state behavior can contribute as well to the temperature-dependent decrease in $\Delta V_f^\circ(T)$; thermal expansion of the folded state of the Δ +PHS variant is smaller than that of the WT SNase, as previously observed directly by PPC.¹²

The difference in apparent expansivity between the folded and unfolded states, $\Delta\epsilon$, calculated from the pressure dependence of the average volume change of unfolding, was twice as large for the Δ +PHS variant than for WT SNase. Because the differences in sequence are relatively modest, the unfolded-state expansivities are expected to be very similar for these two proteins. Hence the differences in $\Delta\epsilon$ derived from the HSQC experiments must arise from a much smaller average folded-state expansivity for the Δ +PHS variant.

The pressure-induced unfolding of the three cavity-containing variants of the Δ +PHS reference protein was also monitored by 2D-NMR HSQC (Figures S3B–D) and Trp fluorescence (Figure S4) at four temperatures. The profiles were fit for apparent values of $\Delta G_f^\circ(T)$ and $\Delta V_f^\circ(T)$ at each temperature. The distributions of the NMR-detected site-specific ΔV_f° values at each temperature (Figure 3) all shifted to lower values as a function of increasing temperature. The ΔV_f° distributions for the L125A variant were relatively narrow and symmetric, although some broadening indicates minor deviation from two-state behavior. Nonetheless it was not unreasonable in this case to assume that the peak of the

Table 1. Volume Changes upon Folding ΔV_f at Different Temperatures and Their Temperature Dependence, $\Delta\epsilon^a$

protein	parameter	temperature (K)			
		288	293	298	303
SNase WT	ΔV_f (ml/mol)	70 (12)	65 (11)	60 (7)	58 (7)
	$\Delta\epsilon$ (ml/mol/K)			−0.86 (0.67)	
Δ PHS	ΔV_f (ml/mol)	93 (18)	87 (13)	76 (14)	68 (16)
	$\Delta\epsilon$ (ml/mol/K)			−1.98 (0.95)	
I92A	ΔV_f (ml/mol)	126 (16)	109 (20)	100 (17)	86 (18)
	$\Delta\epsilon$ (ml/mol/K)			−2.37 (1.76)	
L125A	ΔV_f (ml/mol)	109 (15)	99 (11)	95 (12)	84 (11)
	$\Delta\epsilon$ (ml/mol/K)			−1.52 (0.93)	
I92A/L125A	ΔV_f (ml/mol)	195 (27)	186 (22)	158 (20)	142 (19)
	$\Delta\epsilon$ (ml/mol/K)			−4.47 (2.11)	

^aA linear temperature dependence of the ΔV_f over the range probed is assumed. Uncertainties are given in parentheses for each parameter value.

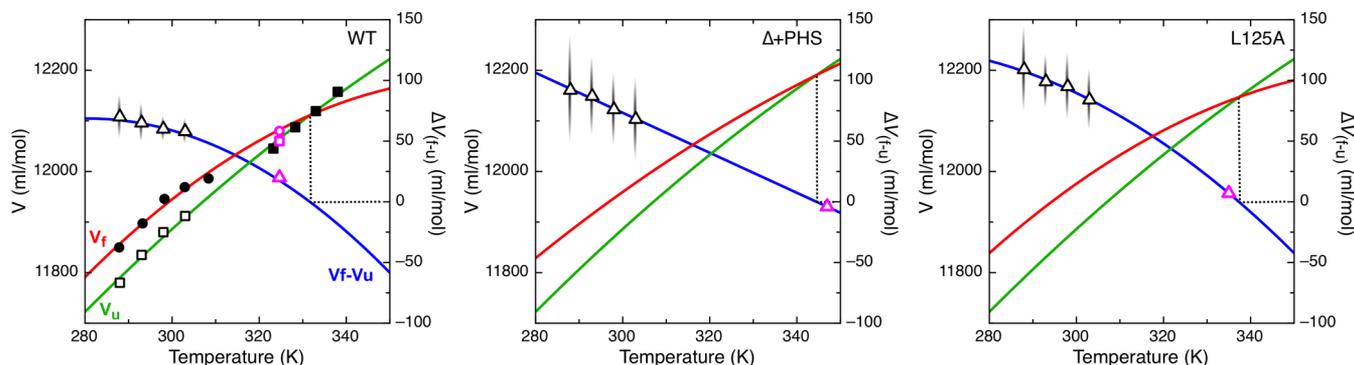


Figure 4. Temperature dependence of volumetric properties of SNase variants. Left axis, partial molar volumes of folded (red line) and unfolded (green line) states of SNase variants folded (black circles). For WT SNase, V_u (black closed squares) is taken from densitometry measurements of WT SNase³⁷ at high temperature and extrapolates to the V_u values (black open squares) obtained from subtracting ΔV_f (NMR) from V_f (filled black circles), measured also previously by densitometry at low temperatures. The V_f value for WT SNase at T_m (pink open circle) was obtained by adding ΔV_f (pink open triangle) obtained at T_m by PPC to the V_u value (pink open square) obtained at T_m by interpolation of the V_u values from densitometry at high temperature. For Δ +PHS and L125A, V_u is assumed to be identical to WT and V_f is calculated by adding to that the present ΔV_{NMR} values. Right axis: ΔV_{NMR} (black open triangles) and ΔV_{PPC} (pink open triangles) of WT, Δ +PHS, and L125A. Blue line is a quadratic fit to $V_f - V_u$ for WT or L125A, and a linear fit of ΔV values for Δ +PHS. Vertical shadows correspond to the distributions of site specific ΔV_{NMR} values.

distribution represents a reasonable approximation of the true thermodynamic volume change. The average apparent difference in expansivity, $\Delta\epsilon$, calculated from the temperature dependence of the average ΔV_f° (Table 1) was intermediate between that observed for the WT and the Δ +PHS reference protein, indicating that the folded-state expansivity of L125A is larger than that of Δ +PHS, yet smaller than that of WT SNase.

In contrast to the near two-state behavior of WT SNase, the L125A variant, and Δ +PHS, the variants with I92A and I92A/L125A substitutions both exhibited very broad ΔV_f° distributions, with those for the I92A variant at intermediate temperatures exhibiting bimodal character. Small apparent values for ΔV_f° obtained from the fits of the pressure dependence of the HSQC peak intensity of a given residue are indicative of partial unfolding involving that residue. The positions of these residues in the I92A structure (Figure S5) show that partial unfolding occurs in the C-terminal helix and its interface to the OB-fold domain, in addition to some disruption in the vicinity of the I92A substitution. This folding intermediate of SNase and several of its variants, disrupted in the C-terminal helix, has been reported previously.^{26,34–36}

Thermal Expansion and Volumetric Profiles. Complete volumetric profiles of the folded and unfolded states of the SNase variants can be derived from the temperature dependence of the averages of the NMR based ΔV_f° values (Figure 4), as done previously for WT SNase based on $\Delta V_f^\circ(T)$ values obtained from fluorescence detected unfolding.¹³ This was possible only for those variants with quasi-two-state behavior (WT, Δ +PHS, and L125A), for which the peaks of the $\Delta V_f^\circ(T)$ distributions are assumed to approximate the thermodynamic values reasonably well. The black triangles in Figure 4 with dispersed points represent the distribution of values of $\Delta V_f^\circ(T)$ obtained from the present NMR experiments. The values of the volume of the folded WT SNase, V_f , were measured directly by densitometry as a function of temperature and reported previously.³⁷ The values for the volume of the unfolded WT SNase, V_u , below 320 K were calculated by subtracting the average value of $\Delta V_f^\circ(T)$ from V_f . The values of the molar volume of unfolded WT SNase above 320 K were measured previously directly by densitometry,³⁷ since WT SNase unfolds at ~ 325 K. The pink triangle represents the value of $\Delta V_f(T_m)$,

the volume change for folding at the transition temperature, obtained directly from PPC measurements^{12,38} and Figure S6. Similar plots were constructed for Δ +PHS and its L125A variant, assuming by convention that the unfolded-state volume is the same for all proteins. Different unfolded-state volumes would simply shift both curves up or down on the y-axis relative to WT SNase. We exclude the possibility of residual volume in the unfolded states of the variants obtained by pressure denaturation because pressure, by nature, favors the state of least volume.

Over the limited temperature range for which pressure-induced unfolding profiles were measured by NMR, the temperature dependence of $\Delta V_f^\circ(T)$ ($=\Delta\epsilon$) is approximately linear for all three variants. However, taking into account the volume change obtained from PPC, the temperature dependence of ΔV_f° deviates from linearity over a broader temperature range for WT and the L125A variant, although not for the Δ +PHS protein. Deviation from linearity for $\Delta V_f^\circ(T)$ is not surprising since the expansivities for the folded states of WT and L125A decrease significantly with increasing temperature. In contrast, the expansivity of the folded state of Δ +PHS is rather low and constant over a broad temperature range,¹³ and given that the expansivities of unfolded states are not strongly temperature dependent,³⁹ the linearity observed for $\Delta V_f^\circ(T)$ of the hyper-stable variant is expected. The apparent expansivity of the unfolded states of the Δ +PHS protein and of its L125A variant was considered to be the same as for the WT protein. This assumption is supported by several lines of evidence. First, identical expansivities of the unfolded states of these variants were measured directly by pressure perturbation calorimetry at temperatures above the unfolding transition.¹² Second, the zero point temperature for the $\Delta V_f^\circ(T)$ plots for the three variants corresponds well to the crossover temperature of the V_f and V_u curves. That is the point above which pressure would favor folding rather than unfolding (dotted lines in Figure 4). Finally, the expansivity values for the folded state (slope of the red lines in Figure 4) deduced based on the assumption of equivalent unfolded-state expansivities and the $\Delta V_f^\circ(T)$ values from the high-pressure NMR experiments (~ 6 mL/(mol·K) for Δ +PHS and between ~ 8 mL/(mol·K) at low temperature and 4 mL/(mol·K) at high temperature for the L125A variant) are in

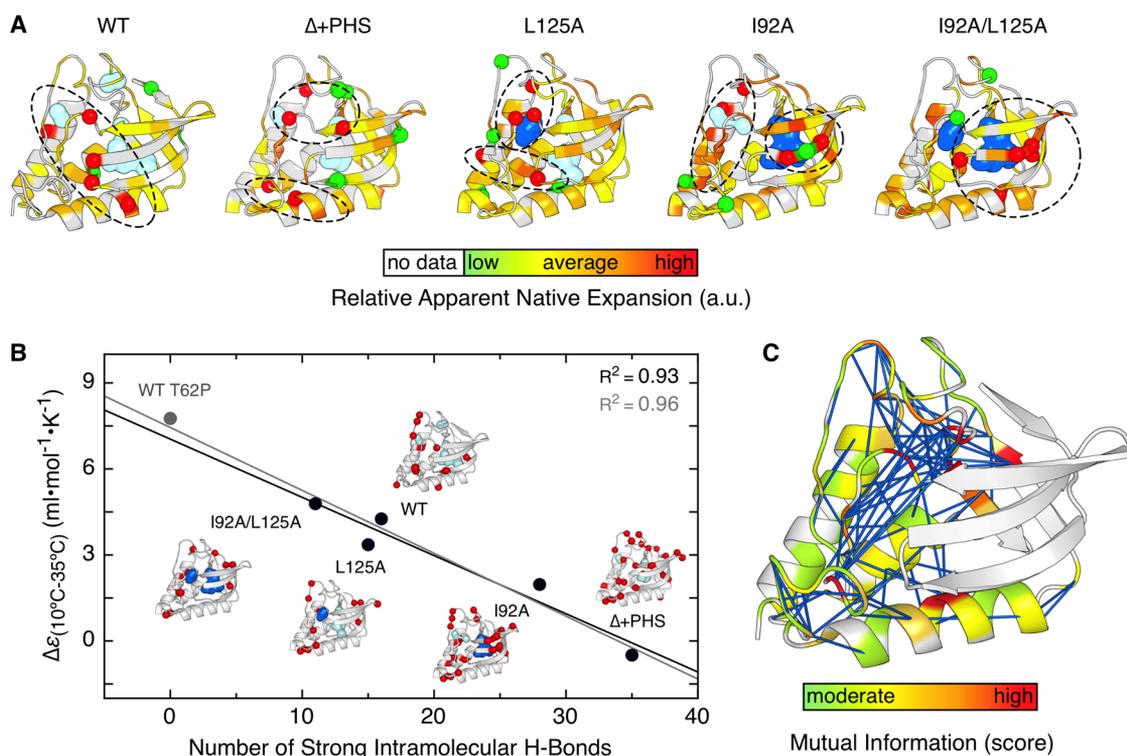


Figure 5. Correlation of native expansion with intramolecular interactions. (A) Cartoon representation of the thermal dependence of site specific folded-state expansivity values, see Materials and Methods, of WT and Δ PHS SNase and of cavity enlarging variants with I92A, L125A and I92A/L125A substitutions. (B) Negative correlation between folded-state expansivity¹² obtained from PPC¹² and the number of strong intramolecular H-bonds derived from Figure 1. Gray and black lines are linear regression fits including or excluding SNase WT T62P data, respectively. Cartoons highlight in red spheres the backbone amide positions of residues implicated in strong intramolecular H-bonds. The molecular surface of the internal void volumes is shown in light blue and highlighted in dark blue when the size is altered by the mutation. (C) Mutual information interaction network of SNase family with MI values >6.5 (significant) (low = 9; high = 26.5) represented as straight yellow sticks and colored according to their score. PDB codes for SNase, Δ PHS, I92A, L125A, and I92A/L125A proteins correspond to 1SNC, 3BDC, 3MEH, 3NXW, and 4DGZ, respectively.

reasonable agreement with the expansivities of the folded state measured for the Δ +PHS and L125A variants directly by pressure perturbation calorimetry (5 mL/(mol·K) for Δ +PHS and between 11 and 7 mL/(mol·K) for L125A.¹²

DISCUSSION

Pressure-dependent NMR spectroscopy has the potential to contribute unprecedented site-specific structural insight concerning folding mechanisms and structural origins of stability and cooperativity under a variety of solution conditions (temperature, denaturant, pH). Assuming quasi-two-state behavior and that the expansion of the unfolded form is the same across all variants, one can interpret the temperature dependence of the site specific ΔV_f° (which provide residue specific values for $\Delta\epsilon$) as a reflection of the local apparent thermal expansivity ($\Delta\epsilon$) in the folded state. High values for differences in expansion between the folded and unfolded states correspond to low values of expansion of the folded state.

We showed previously by PPC that the global thermal expansivity of WT SNase in the folded state at 283 K is rather high, 16 mL/mol·K, approximately 3-fold higher than that of the Δ +PHS variant, 5.8 mL/mol·K. The L125A and I92A variants have intermediate expansivity at 283, 12.2, and 11.5 mL/mol·K, respectively,¹² while the double mutant increases to 13.8 mL/mol·K (Figure S6), nearly that of WT SNase.

From the present high-pressure NMR data (Figure 5A) we observe that locally, WT SNase exhibits the highest thermal expansion values for residues at the interface between the C-

terminal helix 3 and the core of the protein and the nearby helix 1, all in the vicinity of the flexible Ω -loop. The deletions and substitutions used to engineer the highly stable Δ +PHS variant not only lower the overall thermal expansion 3-fold but they also disrupt the interfacial pattern of residues with the highest expansivity. For the L125A variant, the pattern of residues with the highest expansivity resembles that of WT SNase, with the exception of helix-1, indicating that creation of a cavity at the interface between helix-3 and the core by the L-to-A substitution disrupts interactions that constrain the Δ +PHS protein against expansion. The strong deviation from two-state behavior of the I92A and I92A/L125A variants precluded establishing this type of correlation between structure and thermodynamics. However, the pattern of “apparent expansivity” for these variants, which is distributed throughout the structure, is consistent with temperature-dependent population of multiple intermediate states and a complex folding landscape, as previously shown for the I92A variant.⁴

The difference in thermal expansion between 285 and 308 K (5–35 °C), $\Delta\epsilon_{10-35}$, provides another good global measure of folded-state expansivity for all the SNase variants.^{12,13} (Figure S6). We find a strong inverse correlation between the thermodynamic expansivity of the folded state of the variants and the number of strong intramolecular H-bonds in their structure, as deduced from the temperature dependence of the amide proton chemical shifts (Figure 5B). Including results for a T62P variant of WT SNase, which is constitutively unfolded in water, assuming it contains no strong H-bonds, reinforces

this negative correlation. These observations support the notion that these strong intramolecular H-bonds impose local constraints against thermal volumetric expansion; the stronger the interactions, the lower the expansivity.

Multiple sequence alignment of the SNase family of proteins and extraction of the mutual information (MI) network (Figures S5 and S7) revealed a large number of co-evolving residues at the interface between the core OB-fold domain and the C-terminal helix, whereas co-evolution within the OB fold domain was more limited. The strongest MI scores linked residues in the three interfacial loops to each other and to the N-terminus of helix-1. Interestingly, these are the same regions for which the stabilizing deletions and substitutions used to engineer the Δ +PHS protein reinforce H-bonding, compared to WT SNase. These inverse correlations between H-bonding and co-evolution suggest that SNase has evolved to retain considerable flexibility in and around its active site (near the Ω -loop). This flexibility is evident as significant local thermal expansion in these regions, measured by high-pressure NMR spectroscopy.

Data about co-evolving residues are often used to infer physical contact between two residues. In this case, the mutual information informs on a more complex requirement for the appropriate balance of interaction and dynamics. Indeed, in addition to flexibility, certain key interactions such as the active site clamp between D77-T120 have been conserved,⁴⁰ although here D77 only exhibits low native-state expansivity in the context of the Δ +PHS variant. The notion that these patterns of conserved flexibility levels are key to tuning the functional properties of the protein is supported by the fact that the mutual information is strongest in and around the substrate binding site at the interface between the two subdomains (Figure S6).

CONCLUSIONS

Two-dimensional NMR spectroscopy was used to examine structural origins of the temperature dependence of the volumetric properties of some proteins. The thermal expansivity of the protein appears to be anticorrelated with the presence of strong intramolecular H-bonds. Stabilizing substitutions that locally reinforce H-bonding impose constraints against thermal expansion, whereas local disruption of interactions via cavity creating mutations increases the thermal expansion of the native state in a site-specific manner. The effects of these perturbations to the hydrogen-bonding patterns of protein correlate well with the network of mutual information obtained from examination of the co-evolution of sequence within this protein family. This analysis revealed a link between the high thermal expansivity near the active site of this enzyme, with the known functional requirements for conformational flexibility in this region. This study shows how the characterization of volume and volumetric expansion with the site-specific resolution afforded by NMR spectroscopy can yield detailed information about intramolecular interactions essential for folding and function.

ASSOCIATED CONTENT

Supporting Information

Figure S1, Residue Specific assignments for backbone amide group of I92A/L125A double mutant. Figure S2, Histogram of $\Delta\delta\text{NH}/\Delta T$ for all measured protein variants. Figure S3, Thermal dependences of pressure induced unfolding of Δ +PHS protein and the corresponding variants L125A, I92A,

I92A/L125A as followed by 2D ^{15}N - ^1H HSQC. Figure S4, Pressure unfolding curves followed by the center of spectral mass of Trp fluorescence of I92A/L125A protein recorded at 288, 293, 298, 303, and 308 K. Figure S5, Representation of the residues of I92A which exhibit low volume changes. Figure S6, DSC and PPC molar expansivity vs temperature calorimetric profiles of I92A, L125A, and I92A/L125A proteins. Figure S7, Co-evolving network of the SNase protein family. Figure S8, Temperature dependence of SNase WT protein ΔV as calculated by NMR, fluorescence, and PPC. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b04320.

AUTHOR INFORMATION

Corresponding Authors

*royerc@rpi.edu

*christian.roumestand@cbs.cnrs.fr

Present Addresses

[§]Structural Virology Department, Pasteur Institute, 28 rue du Dr. Roux, Paris, France 75015

^{||}Laboratory of Physical Chemistry, National Institutes of Health, Rockville Pike, Bethesda, Maryland 20892, United States

[†]Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, New York 12218, United States

[#]904 Stellar-Chance Labs Department of Biochemistry & Biophysics University of Pennsylvania Perelman School of Medicine 422 Curie Blvd. Philadelphia, PA 19104-6059, United States

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by grants from the Agence National pour la Recherche PiriBio no. 09-455024 to CAR, from the National Science Foundation (MCB-0743422) to BGME and from the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01 to the Center for Structural Biochemistry in Montpellier, France.

REFERENCES

- (1) Myers, J. K.; Pace, C. N.; Scholtz, J. M. *Protein Sci.* **1995**, *4*, 2138.
- (2) Bell-Upp, P.; Robinson, A. C.; Whitten, S. T.; Wheeler, E. L.; Lin, J.; Stites, W. E.; Garcia-Moreno E., B. *Biophys. Chem.* **2011**, *159*, 217.
- (3) 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.
- (4) Roche, J.; Caro, J. A.; Norberto, D. R.; Barthe, P.; Roumestand, C.; Schlessman, J. L.; Garcia, A. E.; Garcia-Moreno E., B.; Royer, C. A. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6945.
- (5) Bridgman, P. W. *J. Biol. Chem.* **1914**, *19*, 511.
- (6) Brandts, J. F.; Oliveira, R. J.; Westort, C. *Biochemistry* **1970**, *9*, 1038.
- (7) Meersman, F.; Smeller, L.; Heremans, K. *Biochim. Biophys. Acta, Proteins Proteomics* **2006**, *1764*, 346.
- (8) Winter, R.; Dzwolak, W. *Philos. Trans. R. Soc., A* **2005**, *363*, 537.
- (9) Zipp, A.; Kauzmann, W. *Biochemistry* **1973**, *12*, 4217.
- (10) Lin, L. N.; Brandts, J. F.; Brandts, J. M.; Plotnikov, V. *Anal. Biochem.* **2002**, *302*, 144.
- (11) Tsamaloukas, A. D.; Pyzocha, N. K.; Makhatazde, G. I. *J. Phys. Chem. B* **2010**, *114*, 16166.
- (12) Dellarole, M.; Kobayashi, K.; Rouget, J. B.; Caro, J. A.; Roche, J.; Islam, M. M.; Garcia-Moreno E., B.; Kuroda, Y.; Royer, C. A. *J. Phys. Chem. B* **2013**, *117*, 12742.

- (13) Mitra, L.; Rouget, J. B.; Garcia-Moreno E., B.; Royer, C. A.; Winter, R. *ChemPhysChem* **2008**, *9*, 2715.
- (14) Rholam, M.; Scarlata, S.; Weber, G. *Biochemistry* **1984**, *23*, 6793.
- (15) Scarlata, S.; Rholam, M.; Weber, G. *Biochemistry* **1984**, *23*, 6789.
- (16) Weber, G.; Scarlata, S.; Rholam, M. *Biochemistry* **1984**, *23*, 6785.
- (17) Roche, J.; Caro, J. A.; Dellarole, M.; Guca, E.; Royer, C. A.; Garcia-Moreno E., B.; Garcia, A. E.; Roumestand, C. *Proteins: Struct., Funct., Genet.* **2013**, *81*, 1069.
- (18) Baxter, N. J.; Williamson, M. P. *J. Biomol. NMR* **1997**, *9*, 359.
- (19) Castaneda, C. A.; Fitch, C. A.; Majumdar, A.; Khangulov, V.; Schlessman, J. L.; Garcia-Moreno E., B. *Proteins: Struct., Funct., Genet.* **2009**, *77*, 570.
- (20) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661.
- (21) Sklenar, V. *Basic Life Sci.* **1990**, *56*, 63.
- (22) Pons, J. L.; Malliavin, T. E.; Delsuc, M. A. *J. Biomol. NMR* **1996**, *8*, 445.
- (23) Bax, A.; Pochapsky, S. S. *J. Magn. Reson.* **1992**, *99*, 638.
- (24) Marion, D.; Driscoll, P. C.; Kay, L. E.; Wingfield, P. T.; Bax, A.; Gronenborn, A. M.; Clore, G. M. *Biochemistry* **1989**, *28*, 6150.
- (25) Stoven, V.; Mikou, A.; Lallemand, J.-Y.; Piveteau, D.; Guittet, E. *J. Magn. Reson.* **1989**, *82*, 163.
- (26) Roche, J.; Dellarole, M.; Caro, J. A.; Guca, E.; Norberto, D. R.; Yang, Y.; Garcia, A. E.; Roumestand, C.; Garcia-Moreno E., B.; Royer, C. A. *Biochemistry* **2012**, *51*, 9535.
- (27) Dellarole, M.; Royer, C. A. *Methods Mol. Biol.* **2014**, *1076*, 53.
- (28) Finn, R. D.; Bateman, A.; Clements, J.; Coghill, P.; Eberhardt, R. Y.; Eddy, S. R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; Sonnhammer, E. L.; Tate, J.; Punta, M. *Nucleic Acids Res.* **2014**, *42*, D222–D230.
- (29) Lassmann, T.; Sonnhammer, E. L. *BMC Bioinf.* **2005**, *6*, 298.
- (30) Simonetti, F. L.; Teppa, E.; Chernomoretz, A.; Nielsen, M.; Marino, B. C. *Nucleic Acids Res.* **2013**, *41*, W8.
- (31) Bedard, S.; Mayne, L. C.; Peterson, R. W.; Wand, A. J.; Englander, S. W. *J. Mol. Biol.* **2008**, *376*, 1142.
- (32) Maki, K.; Cheng, H.; Dolgikh, D. A.; Roder, H. *J. Mol. Biol.* **2007**, *368*, 244.
- (33) Roche, J.; Dellarole, M.; Caro, J. A.; Norberto, D. R.; Garcia, A. E.; Garcia-Moreno E., B.; Roumestand, C.; Royer, C. A. *J. Am. Chem. Soc.* **2013**, *135*, 14610.
- (34) Jacobs, M. D.; Fox, R. O. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 449.
- (35) Kalnin, N. N.; Kuwajima, K. *Proteins: Struct., Funct., Genet.* **1995**, *23*, 163.
- (36) Walkenhorst, W. F.; Green, S. M.; Roder, H. *Biochemistry* **1997**, *36*, 5795.
- (37) Seemann, H.; Winter, R.; Royer, C. A. *J. Mol. Biol.* **2001**, *307*, 1091.
- (38) Mitra, L.; Smolin, N.; Ravindra, R.; Royer, C.; Winter, R. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1249.
- (39) Schweiker, K. L.; Makhataдзе, G. I. *Methods Enzymol.* **2009**, *466*, 527.
- (40) Carra, J. H.; Anderson, E. A.; Privalov, P. L. *Biochemistry* **1994**, *33*, 10842.