PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Structural, energetic, and dynamic responses of the native state ensemble of staphylococcal nuclease to cavity-creating mutations

Julien Roche,[1] Jose A. Caro,[2] Mariano Dellarole,[1] Ewelina Guca,[1] Catherine A. Royer,[1] Bertrand García-Moreno E.,[2] Angel E. Garcia,[3] and Christian Roumestand[1]*

[1] Centre de Biochimie Structurale, INSERM U1054, CNRS UMR 5048, Universités de Montpellier, France

[2] Department of Biophysics, Johns Hopkins University, Baltimore, Maryland

[3] Department of Physics and Applied Physics and Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, New York

ABSTRACT

The effects of cavity-creating mutations on the structural flexibility, local and global stability, and dynamics of the folded state of staphylococcal nuclease (SNase) were examined with NMR spectroscopy, MD simulations, H/D exchange, and pressure perturbation. Effects on global thermodynamic stability correlated well with the number of heavy atoms in the vicinity of the mutated residue. Variants with substitutions in the C-terminal domain and the interface between α and β subdomains showed large amide chemical shift variations relative to the parent protein, moderate, widespread, and compensatory perturbations of the H/D protection factors and increased local dynamics on a nanosecond time scale. The pressure sensitivity of the folded states of these variants was similar to that of the parent protein. Such observations point to the capacity of the folded proteins to adjust to packing defects in these regions. In contrast, cavity creation in the β-barrel subdomain led to minimal perturbation of the structure of the folded state, However, significant pressure dependence of the native state amide resonances, along with strong effects on native state H/D exchange are consistent with increased probability of population of excited state(s) for these variants. Such contrasted responses to the creation of cavities could not be anticipated from global thermodynamic stability or crystal structures; they depend on the local structural and energetic context of the substitutions.

Proteins 2013;81:1069–1080.

Key words: protein folding and stability; high-pressure; packing defects; conformational fluctuations; cavities.

INTRODUCTION

The ability of proteins to fluctuate and to sample wide swaths of conformational space is central to their biological and functional roles.[1–6] Volume calculations[7,8] and compressibility measurements[9–13] have shown that globular proteins are not uniformly packed. As demonstrated recently for lysozyme[14] and myoglobin,[15] the presence of evolutionary conserved cavities may play a crucial role in function by modulating structural flexibility, promoting conformational fluctuations, and enabling intramolecular communication. Packing defects can be introduced artificially in proteins by the substitution of internal hydrophobic side chains with alanine. Systematic alanine scanning studies have been performed over the past 20 years to examine the role of packing on thermodynamic properties of proteins.[16–18] Detailed X-ray crystallographic analyses have shown that, depending on the location of the mutated side chain, the structural consequences of alanine substitution range from minor modifications to significant side chain rearrangements.[19] The L99A substitution in T4 lysozyme is a good example of a substitution in a very rigid environment that has no

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Agence National de la Recherche grant PiriBio; Grant number: 09-455024; Grant sponsor: National Science Foundation; Grant numbers: MCB-0743422, MCB-0543769 and MCB-1050966; Grant sponsor: French Ministry of Research and Higher Education; Grant sponsor: Fulbright International Graduate Fellowship.

*Correspondence to: Christian Roumestand, Centre de Biochimie Structurale, 29 rue de Navacelles, 34090 Montpellier Cedex 5, France. E-mail: christian.roumestand@cbs.cnrs.fr

Received 9 August 2012; Revised 19 October 2012; Accepted 21 November 2012
Published online 12 December 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24231

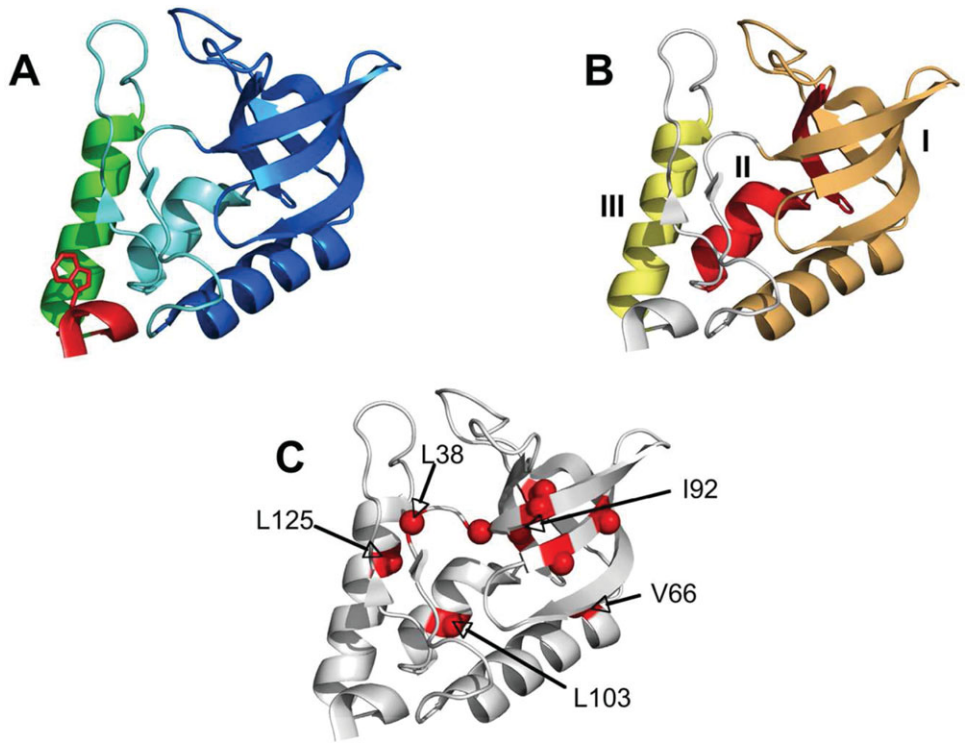

Figure 1

SNase architecture and location of the cavity-creating mutations. **A**: Putative subdomain organization of SNase: the subD1 is colored in blue, IntD in cyan, SubD2 in green, and the C-term helical turn in red. Trp-140 is depicted in stick representation in red. **B**: Foldon organization in SNase according to:[23] foldon I (orange, β1–β4 + α1), foldon II (red, β5 + α2), and foldon III (yellow, α3). **C**: Positions that were substituted with Ala to create internal cavities (red spheres). The location of L38 and the sites of the four Ala substitutions studied with NMR spectroscopy and MD simulations are labeled.

observable structural consequence. High-pressure crystallography and MD simulations have shown that under high pressure the cavity introduced in the L99A variant fills with water molecules instead of collapsing.[20,21] Little is known about the consequences of cavity creating substitutions on conformational fluctuations. The goal of this study was to examine in detail how packing defects in different regions of a model protein, Staphylococcal Nuclease (SNase) affect its flexibility, local stability, and dynamics.

The subdomain organization of SNase [Fig. 1(A)] has been amply demonstrated by mutagenesis,[22] H/D exchange[23,24] and mechanical unfolding experiments.[25] Subdomain 1 (SubD1), formed by the β-barrel and helix 1, corresponds to a canonical OB-fold domain.[26] The main hydrophobic core of the protein is defined primarily by this subdomain. Helix-3 constitutes subdomain 2 (SubD2). Helix-2, as well as the mini β sheet composed of residues 39–40 and 110–111, have been described as an interface (IntD) between SubD1 and SubD2. The C-terminal α-helical turn (L137 through S141) where Trp-140, the single Trp in SNase is found, is known to stabilize the protein via multiple long-range contacts. An alternative, more energetically based classification of SNase architecture resulting from native state H/D exchange experiments[23] describes SNase in terms of three main foldons [Fig. 1(B)]. Foldon 2, the most stable, encompasses elements from both SubD1 and the IntD.

Ten cavity-containing variants of the highly stable form of SNase known as Δ + PHS SNase were described previously [Fig. 1(C)].[27] Compared with WT SNase, the Δ + PHS reference protein bears stabilizing substitutions in the C-terminal helix (G50F, V51N, P117G, H124L, and S128A), and a deletion of the mobile Ω loop (residues 44–49), which is part of the active site. Five of these variants have Leu-to-Ala substitutions, three have Val-to-Ala substitutions, one has an Ile-to-Ala substitution and one has a Phe-to-Ala substitution. High-resolution crystal structures of these variants are already available.[27] The backbone atoms of the variants and of the parent proteins are almost superimposable and the cavities in the structures are found at the expected locations. No electron density that could be ascribed to internal water molecules was found in any of the cavities.

In this work, we sought to characterize in detail the consequences of cavity creation. We measured the global thermodynamic stability of the 10 cavity-containing

variants. Four of them (V66A, I92A, L103A, and L125A) with similar global stability (within 1.2 kcal/mol) were selected for detailed study. In addition to comparison of their X-ray structures and global stability measurements, chemical shift data measured with NMR spectroscopy under atmospheric and high pressure, H/D exchange and MD simulations were used to explore in detail how the introduction of internal cavities into different subdomains affected local stability, flexibility, and dynamics. Despite similar global thermodynamic effects, the introduction of cavities resulted in two distinct responses. In the least stable region of the protein and in the interface between this region and the OB-fold core, the structure of the folded protein in solution exhibited significant adjustments to the cavity, and increased dynamics. In contrast, in the most stable β-barrel subdomain, almost no structural adjustment was observed. Instead, the creation of the cavity led to a significant increase in the probability of populating more open, partially unfolded states involving residues remote from the site of the mutation. The observation or not of such long-range consequences points to the importance of intrinsic cooperativity of protein structures and suggests how core packing could modulate, specifically and locally, protein function, turnover or aggregation propensity.

MATERIAL AND METHODS

Protein preparation

The highly stable Δ + PHS form of SNase and the cavity-containing variants were expressed and purified as described previously by Shortle and Meeker.[28] Uniform ^{15}N labeling was obtained from over-expression of recombinant protein in *E. coli* grown in M9 medium containing $^{15}NH_4Cl$ as the sole nitrogen source, as described for SNase previously.[29]

Thermodynamic stability

Stability measurements were performed with guanidine hydrochloride (GuHCl) titrations using an Aviv Automated Titration Fluorimeter 105 as described previously.[30] All stability measurements were performed at 20°C and pH 7.

NMR experiments

Uniformly, ^{15}N-labeled protein samples were dissolved at 1 m*M* concentration in 10 m*M* Tris buffer at pH7. 10% of D_2O was added for the lock procedure. In all experiments, the ^{1}H carrier was centered on the water resonance and a WATERGATE sequence[31,32] was incorporated to suppress solvent resonances. All NMR spectra were processed and analyzed with GIFA.[33] Amide resonances were assigned at atmospheric pressure from 3D [^{1}H,^{15}N] NOESY-HSQC (mixing time 100 ms) and 3D [^{1}H,^{15}N] TOCSY-HSQC (isotropic mixing 60 ms) double-resonance experiments[34,35] recorded on a Bruker AVANCE 700 MHz spectrometer equipped with a 5 mm Z-gradient ^{1}H-^{13}C-^{15}N cryogenic probe, using the standard sequential procedure. ^{1}H chemical shifts were directly referenced to the methyl resonance of DSS, while ^{15}N chemical shifts were referenced indirectly to the absolute frequency ratios $^{15}N/^{1}H = 0.101329118$. The weighted $^{15}N/^{1}H$ averaged amide chemical shift differences between the Δ + PHS reference protein and Ala variants were calculated as:

$$\Delta\delta = \sqrt{\Delta\delta(^{1}H)^{2} + \left(\Delta\delta(^{15}N)\frac{\gamma_N}{\gamma_H}\right)^{2}}$$

where γ_N and γ_H represent the gyromagnetic ratio of ^{15}N and ^{1}H nuclei, respectively.

High-pressure heteronuclear 2D ^{15}N—^{1}H HQSC spectra[36] were recorded at 293 K on a 600 MHz Bruker Avance III spectrometer equipped with a 5 mm Z-gradient ^{1}H-X double-resonance broadband inverse probe. Commercial ceramic high-pressure NMR cell and an automatic pump system (Daedalus Innovations, Philadelphia, PA) were used to vary the pressure in the 0–2.5 kbar range. For each protein, the pressure dependence of the weighted $^{15}N/^{1}H$ amide chemical shift differences from the reference 1 bar HSQC spectra were fitted to a second order polynomial function: $\Delta\delta(p) = a + bp + cp^2$, where P represents the pressure in bar.

H/D exchange experiments were performed as previously described[37] with freshly lyophilized samples dissolved in D_2O at a concentration of 1 m*M*. Time series of ^{15}N—^{1}H HQSC spectra were recorded at 600 MHz (700 MHz when testing the exchange regime on I92A) with a common dead time of 15 min and a time limit of 4–6 days, depending on the variant. Protection factors (PF)[38] were calculated from the exchange rate constants deduced from the time dependence of peak intensities fitted to a monoexponential decay model.

Molecular dynamics simulations

MD simulations were performed as previously described,[39] using the GROMACS 4.4 package[40] and Amber99sb[41] protein force field and starting from the crystal structures: 3BDC for Δ + PHS SNase, 3NXW for the L125A variant, 3MZ5 for the L103A variant, 3MEH for the I92A variant, and 3NQT for the V66A variant. Proteins were inserted in a cubic box with a minimal distance of 1 nm from any protein atom to the box boundary. The box was hydrated with TIP3P[42] water molecules: 12547 for Δ + PHS, 12514 for L125A, 11938 for L103A, 12227 for I92A and 11979 for V66A. His residues were assigned to neutral state and all Glu, Asp, Lys, and Arg were modeled as charged. 6 Cl^- ions were added in all five-protein boxes to neutralize the system. After steepest descent energy minimization and 10 ns equilibration, the resulting conformation was used as

Table I
Thermodynamic Consequences of Cavity-Creating Substitutions in SNase

Variant	ΔG_f	$\Delta\Delta G_f$	$\Delta\Delta G_f$[a]	m_{GuHCl}	m_{GuHCl}[a]
Δ+ PHS	11.9 (0.1)	–	–	4.9	6.9
V23A	8.7 (0.1)	3.2 (0.2)	2.9 (0.1)	5.2	8.1
L25A	8.9 (0.1)	3.0 (0.2)	2.7 (0.1)	5.0	7.7
F34A	8.3 (0.1)	3.6 (0.2)	3.7 (0.1)	5.2	7.7
L36A	8.6 (0.1)	3.3 (0.2)	3.5 (0.1)	5.4	7.7
L38A	10.5 (0.1)	1.4 (0.2)	1.7 (0.1)	4.8	6.4
V66A	9.0 (0.2)	2.9 (0.3)	2.2 (0.1)	4.8	7.5
V74A	8.5 (0.2)	3.4 (0.3)	3.1 (0.1)	5.0	7.7
I92A	7.9 (0.3)	4.0 (0.4)	4.0 (0.1)	5.4	8.1
L103A	7.8 (0.1)	4.1 (0.2)	4.6 (0.1)	5.2	6.6
L125A	8.1 (0.1)	3.8 (0.2)	4.9 (0.1)	5.5	4.5

Compilation of global thermodynamic stability and m values for 10 variants. $\Delta\Delta G_f$ = ΔG_f(variant) − ΔG_f(Δ + PHS) Global stability was measured with GuHCl titrations at 20°C, pH = 7.[a]$\Delta\Delta G_f$ values from Shortle *et al.*[17] These experiments were performed in similar conditions, for the same set of variants but using the true wt SNase background instead of the Δ + PHS for of SNase. Errors in m-values are ≤0.2. ΔG_f and $\Delta\Delta G_f$ values are in kcal/mol while m_{GuHCl} are expressed in kcal/mol/M.

starting point for 100 ns long MD simulation with 2 fs time step. Simulations were performed at 293 K and 1 bar with system compressibility set to 4.6×10^{-5} bar^{-1}.

N—H bond order parameters were calculated as previously described.[43] All sampled conformations were aligned and superimposed on the reference structure using backbone heavy atoms. The N—H vector autocorrelation function was calculated using:

$$C_{int}(t) = \langle P_2(a(\tau - t) \times a(t)) \rangle$$

where $a(\tau)$ is the orientation of the N—H vector and $P_2(x)$ is the second order Legendre polynomial. The internal correlation function converges if $|C_{int}(\infty) - C_{tail}| \leq 0.005$, where C_{tail} represents the tail of the correlation function averaged over the last 1 ns and $C_{int}(\infty)$ represents the infinite limit of a monoexponential decay fit of the correlation function.

RESULTS

Consequences of cavity-creating substitutions on global thermodynamic stability

Thermodynamic stability (ΔG_f) was measured for the 10 variants [Fig. 1(C)] by guanidine hydrochloride (GuHCl) denaturation experiments monitored with Trp fluorescence at 20°C and pH 7 (Table I). Of the 10 $\Delta\Delta G_f$ values measured in this study, nine are in the 3–4 kcal/mol stability difference range with a mean value of 3.3 ± 0.2 kcal/mol, in good agreement with the average value (3.15 kcal/mol) obtained from a compilation of previous studies of 14 different model proteins.[44] With one exception, the $\Delta\Delta G_f$ values showed significant correlations with both the volume of the engineered cavities and the number of heavy atoms within 10 Å radius of the mutated Cβ (Supporting Information Fig. S1). In addition, for the majority of these variants, the $\Delta\Delta G_f$ values are in good agreement with the original data from Shortle *et al.*[17] obtained for the same substitutions in the context of the true wild-type SNase (Table I). We noted, however that the m_{GuHCl} values reflecting the slope of the GuHCl dependence of ΔG_f were significantly smaller for variants of the Δ + PHS proteins than for those in the wt SNase background (Table I).

Structural consequences of V66A, I92A, L103A, and L125A substitutions

Structural consequences of substitutions on the cavity-containing variants were examined by comparing $^{15}N—^{1}H$ chemical shifts under native conditions for a subset of four Ala variants (L103 and L125A located in the IntD and SubD2, and V66A and I92A in SubD1) and for the Δ + PHS reference protein (see Supporting Information Fig. S2 for HSQC assignment of the five proteins). The weighted $^{15}N—^{1}H$ amide group chemical shift differences between Δ + PHS and the four Ala variants (Fig. 2) revealed significant long-range perturbations only for the L125A and L103A variants. For L125A, local chemical shift perturbations were observed for residues in helix-3 and its close environment (residues 111–129). These local perturbations were propagated to IntD (residues 102–105) and to SubD1 (residues 59 and 92). For L103A, located within the IntD, in addition to the local perturbations (residues 102–104 and 110), the main consequences on chemical shifts were observed in SubD1, mostly in helix-1 (residues 58–66), including a large perturbation for residue 62 whose side chain faces the L103A mutation site. In contrast, the two Ala variants located in the SubD1, V66A, and I92A, show only a small number of chemical shift perturbations, which are almost exclusively restricted to the vicinity of the mutated sites (with the notable exception of residue L103 in the case of I92A).

Effect of subdenaturing pressures

The response of the cavity containing variants folded structure to hydrostatic pressure reveals information concerning conformational plasticity. The pressure dependence of the native-state amide chemical shifts was measured for the four Ala variants and the Δ + PHS reference protein at 20°C and pH 7 over a pressure range (1–2500 bar). In this range of pressures, the five proteins remained fully folded. Large variations in both the magnitude and the linearity of the native state chemical shift changes were observed over this pressure range, as illustrated for Δ + PHS in Figure 3. For each protein and at each pressure, the weighted $^{15}N—^{1}H$ averaged amide chemical shift difference Δδ was calculated relative to the reference 1 bar HSQC spectra. Following a large body of work by Akasaka and coworkers,[11–13] the pressure dependence of Δδ for each

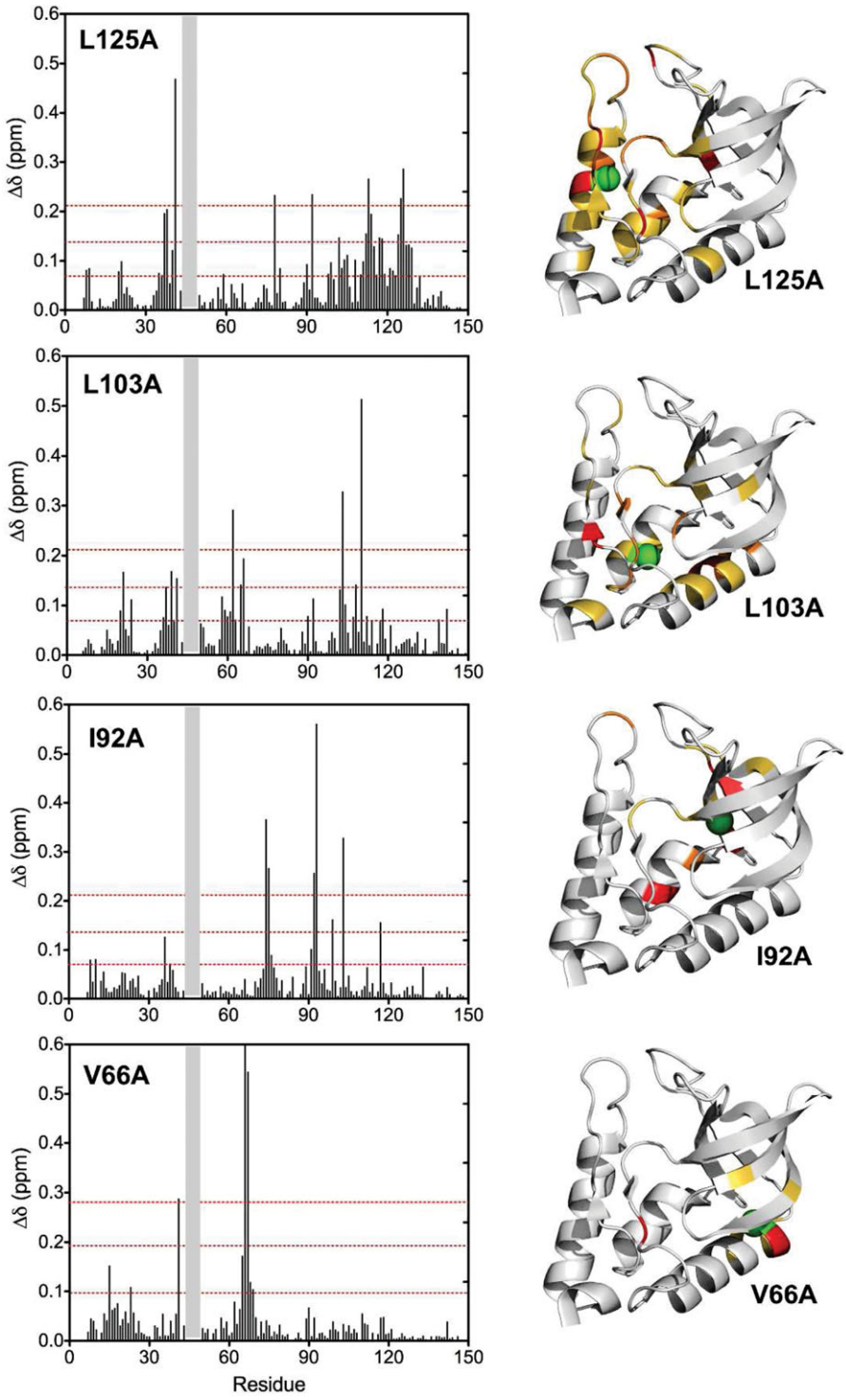

Figure 2

Structural consequences of cavity creation assessed by NMR spectroscopy. (Left) Weighted $^{15}N-^{1}H$ averaged chemical shift differences between the Δ + PHS background protein and the four Ala-containing variants. For each variant, the values of 1, 2, and 3 standard deviations are indicated with red dashed lines. For the sake of comparison, the sequence numbering of WT SNase (149 residues) was used on the abscissa axis; the gray bar indicates the location of the deletion in the Δ + PHS protein (residues 44–49). (Right) Structural mapping of chemical shift differences. Residues with $\sigma < \Delta\delta < 2\sigma$, $2\sigma < \Delta\delta < 3\sigma$, and $\Delta\delta > 3\sigma$ are colored in yellow, orange, and red, respectively. The location of the mutated side chain is indicated with a green sphere.

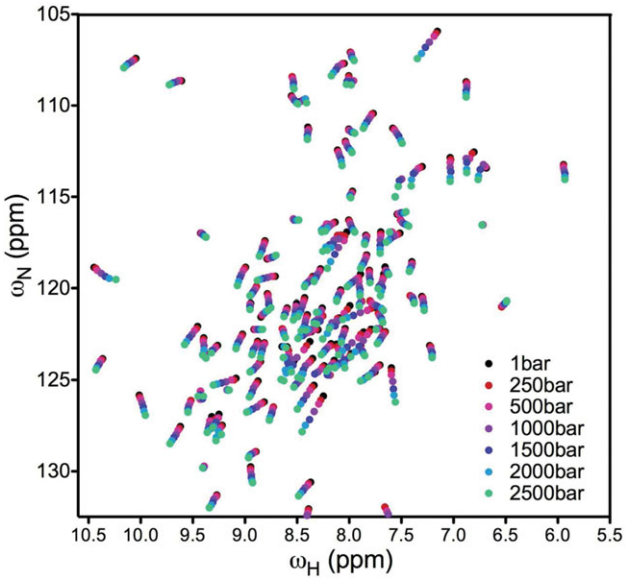

Figure 3

Pressure effects on the reconstituted $^{15}N—^{1}H$ HSQC spectra of Δ + PHS SNase. Reconstituted $^{15}N—^{1}H$ HSQC spectra of Δ + PHS recorded at 20°C, pH 7 at hydrostatic pressures ranging from 1 bar (black) to 2500 bar (green).

cross peak was fitted empirically with a second order polynomial function (see Supporting Information Table S3, and Supporting Information Figs. S4 and S5). These authors have interpreted in a number of protein systems the linear pressure response of the chemical shift as a manifestation of the local compressibility, and the nonlinear responses as an indication of pressure induced population of low lying excited states. Figure 4 highlights the position of residues with a significantly higher second order factor (more than 1 standard deviation) compared with the Δ + PHS background protein. In the L125A and L103A variants only a few of the residues exhibited increases of the second order factor, most of which residues were located in the immediate vicinity of the mutated sites. These include the residues directly facing the mutated sites (residues 38 and 62 for L125A and L103A, respectively), with very few remote positions affected by the substitution (Fig. 4). In contrast to the L125A and L103A mutations, cavities created by the I92A and V66A substitutions resulted in a broad effect of pressure on the native state conformational plasticity, with increased second order factors, indicating pressure-induced population of lower volume excited states for a large number of residues, spanning SubD1 in almost its entirety. The situation is the same if one examines the linear coefficients (Supporting Information Fig. S5); significant pressure effects with respect to the reference protein are measured throughout the protein's structure for the I92A variant, whereas the effects are more subdued and more local to the mutation for the L125A variant.

Local stability measurements

H/D exchange experiments were performed at pH 7, under native conditions to explore the consequences of the cavity-creating mutations on the local stability of Δ + PHS. We have shown in a previous study for Δ + PHS and a variant (V66K) under similar conditions, that the hydrogen exchange rates were strongly pH-dependant, suggesting an EX2 exchange regime.[37] Here, we carefully checked the exchange conditions on the I92A variant. We chose this variant because it appeared from the pressure susceptibility of the chemical shifts to be the most perturbed in terms of the population of local excited stated states. Moreover it is among the least stable of the variants studied here (ΔG= 7.9 ± 0.3 kcal/mol). We reasoned therefore that this variant, if any, would be the most likely to exhibit a change in exchange regime. H/D exchange was measured at 20°C and at three different pH values 6.8, 7.15, and 7.4), chosen around the nominal pH value used for the rest of the NMR study (pH 7). At 20°C, 80 amide cross-peaks were still observable on the first HSQC spectrum recorded at pH 6.8 after the experimental dead time (15 mn), 77 at pH 7.1, and 73 at pH 7.4. Due to extremely slow H/D exchange at 20°C, peak intensity decreases can be accurately fitted to an exponential decay only for 35 residues at pH 6.8, 42 residues at pH 7.1, and 51 residues at pH 7.4 (Supporting Information Fig. S6): most of the others residues do not exhibit any significant change in their corresponding amide cross peak intensity, even after 6 days of experiment. Exchange rate constant ratios were then measured between two pH: an average ratio of 3.5 ± 1.1 was measured between pH 6.8 and 7.15 (30 residues), of 3.2 ± 1.2 between 7.15 and 7.4 (37 residues), in good agreement with the theory predicting a 10 fold increase for the exchange rate constant per pH unit in case of EX2 regime. This is also in good agreement with results previously reported by Bedard *et al.*[23] where native-state hydrogen exchange were measured for 34 residues of a stabilized double mutant (P117G/H124L) of SNase: EX2-to-EX1 transitions were observed at 20°C for some residues but only at pH > 8. The stability of this double mutant (ΔG = 7.7 kcal/mol) is lower than that measured for Δ + PHS, and comparable to that of I92A. From this along and the present pH dependence studies it is reasonable to assume that EX2 is the prevalent regime for amide protons in Δ + PHS and the variants reported in this study at neutral pH.

In the EX2 regime, the hydrogen exchange reaction occurs in a much slower time scale than the refolding reaction, such that the measured PF report on the local stability of the protein.[38] For the variants examined here, cavity creation led to both increases and decreases in PF factors with respect to the Δ+PHS reference protein (Fig. 5). In the case of the L125A and L103A variants, local destabilization by the cavity was generally moderate,

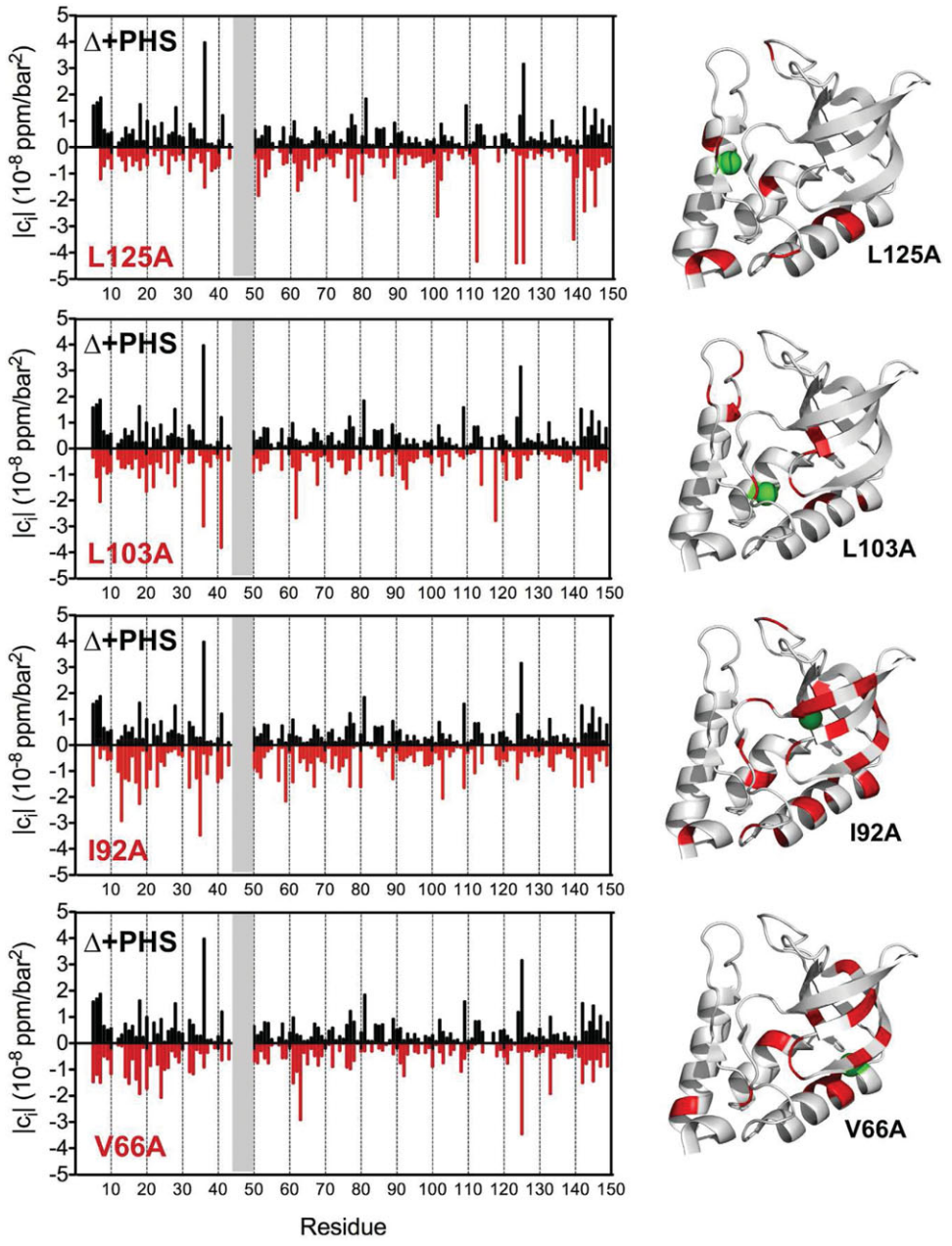

Figure 4

Pressure effects on the $^{15}N-^{1}H$ chemical shifts for Δ + PHS and four Ala variants. (Left) The absolute value of the second order factors of the chemical shift pressure dependence (c_i) as a function of protein sequence, for Δ + PHS SNase (black) and the four Ala variants: L125A, L103A, I92A, and V66A (red). For the sake of clarity, the absolute values of the second order factors are reported for the variants on an inverse axis. The sequence numbering of WT SNase (149 residues) was used on the abscissa axis; the gray bar indicates the location of the deletion in the Δ + PHS protein (residues 44–49). (Right) Residues with significantly larger second order factor (more than 1 standard deviation), in comparison to Δ + PHS, are indicated in red on the protein structure. The location of the side chains substitutes with Ala is indicated with a green sphere.

with strong affects observed for only a few residues in the vicinity of the mutation. Interestingly, for both of these variants, larger PFs were measured for residues proximal to those that exhibited faster exchange (for example in helix 3 and the IntD for L125A and helices 1 and 3 for L103A). Such behavior is indicative of compensating local energetic adjustments. The V66A variant exhibited very different behavior, with significant local destabilization of many residues throughout the protein; in SubD1 (residues 10–40 and 64–66), in the IntD (residues 97–113) and in the C-terminal loop (130–141), and very limited increases in PF values. Finally, alanine

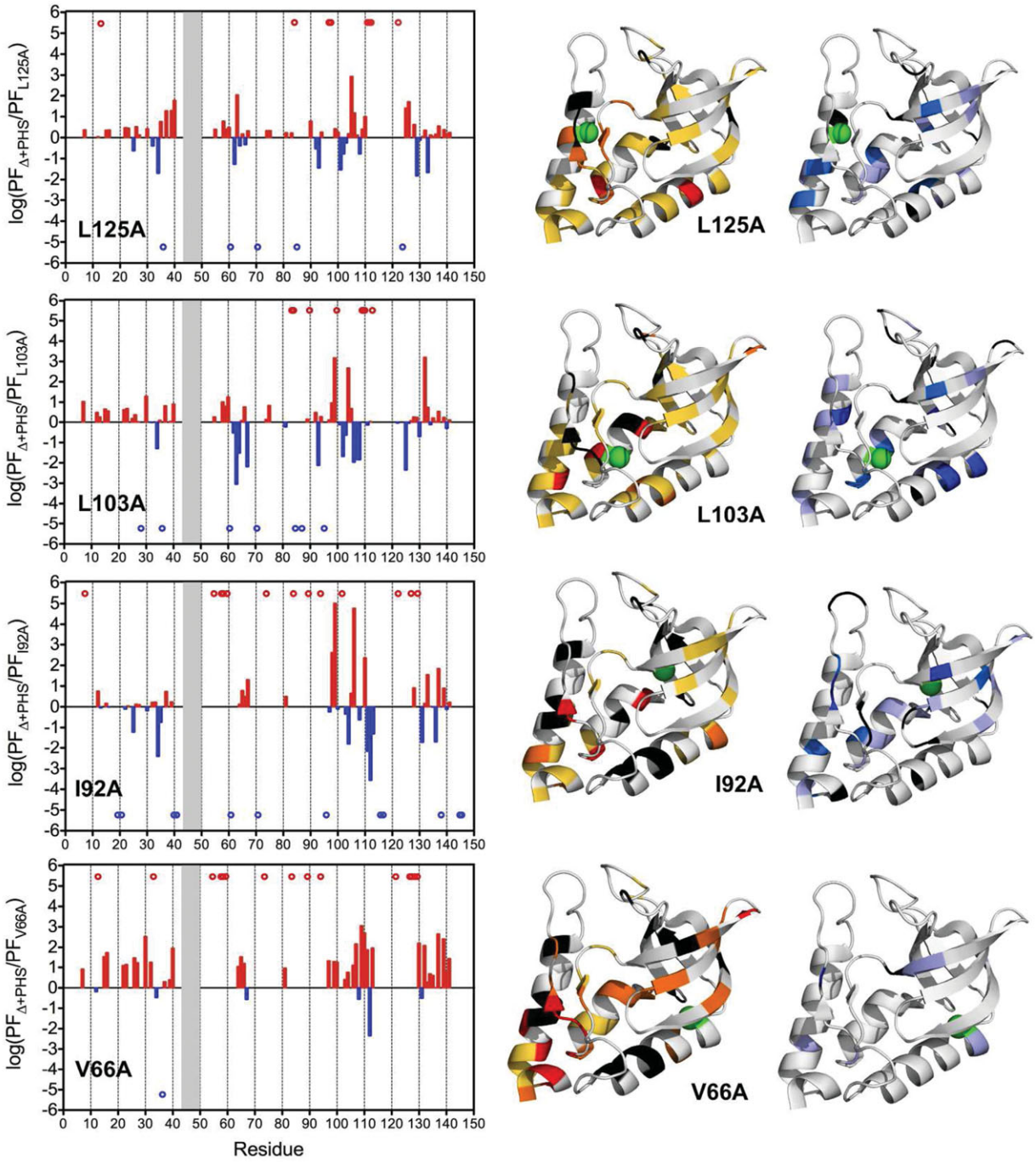

Figure 5

Effects of Ala substitutions on H/D exchange kinetics. (Left) Logarithms of the PF ratios (Δ + PHS/variant) are plotted as a function of the protein sequence. Positive ratios (red) indicate that the NH bond is more protected in Δ + PHS compared with the variants, while negative ratios (blue) indicate that the NH bond is less protected in Δ + PHS compared with the variants. Amide protons that are exchanging within the dead time of the experiments (15 mn) in the variants but that have measurable exchange rates in Δ + PHS are indicated with red dots. Amide protons with the opposite behavior—i.e., exchanging within the experimental dead time in Δ + PHS but not in the variant—are indicated with blue dots. The sequence numbering of WT SNase (149 residues) was used on the abscissa axis; the gray bar indicates the location of the deletion in the Δ + PHS protein (residues 44–49). (Right) Residues are colored as a function of the PF ratios, from red ($R > 2$), to orange ($1 < R < 2$), yellow ($0 < R < 1$), pale blue ($-1 < R < 0$), blue ($-2 < R < -1$), and dark blue ($R < -2$). For the sake of clarity, residues with positive ratios are reported on the left structures, those with negative ratios on the right structures. Residues in black correspond to amide protons where exchange rates can be measured in the variants but not in Δ + PHS (right) or to amide protons where exchange rates can be measured in the Δ + PHS but not in the variants (left). The location of the mutated side chain is indicated with a green sphere.

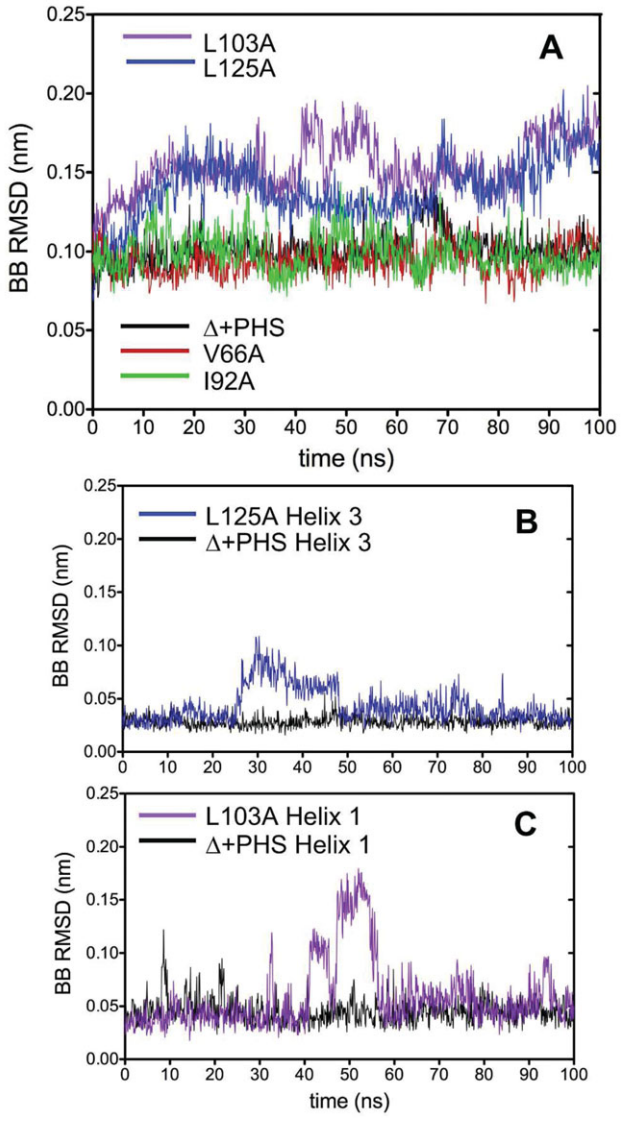

Figure 6

Dynamic consequences of cavity-creating substitutions. **A**: Full protein backbone RMSD calculated over the total simulation length for the four cavity-containing variants and for the Δ + PHS reference protein. **B**: Backbone RMSD of helix-3 calculated for the L125A variant (blue) and Δ + PHS SNase (black). **C**: Backbone RMSD of helix-1 calculated for the L103A variant (purple) and the Δ + PHS protein (black).

substitution at position 92 led to drastic local destabilization in the immediate vicinity of the mutation. Amide protons at and adjacent to position 92 exchanged too rapidly to be measured. In addition, very strong perturbations to the PF, both stabilizing and destabilizing, were observed for this variant in the IntD and, surprisingly, in the C-terminal region, which packs against the IntD. Moderate perturbations were observed as well for this variant for a few residues in the β-sheet opposite position 92 in SubD1.

Molecular dynamics simulations

All-atom MD simulations with explicit solvent were performed on Δ + PHS SNase and on the four Ala variants to probe the dynamic consequences of the cavities on the nanosecond time scale. Although H/D exchange experiments probe residue dynamics on time scales that are related to local stability, nanosecond dynamics offer complementary information about local structural flexibility. Backbone root mean square deviations (RMSD) relative to the reference crystal structures were calculated as a function of simulation time [Fig. 6(A)]. A clear difference was observed between the reference structure and the L125A and L103A variants. In contrast, the V66A and I92A variants and the Δ + PHS protein remain remarkably stable. Independent simulations starting from the same reference structures were performed twice to confirm these results (Supporting Information Fig. S7). A reversible perturbation of helix-3 was noticed for the L125A variant (between t = 25 ns and t = 48 ns) [Fig. 6(B)]. Helix-1 exhibited a similar, large reversible perturbation (between t = 40 ns and t = 56 ns) in the L103A variant [Fig. 6(C)]. This perturbation is characterized by a reversible kink of helix-1, illustrated in Figure 7 with snapshots at t = 53 ns showing the penetration of Thr-62 into the cavity created by the L103A substitution. Comparison of simulated B-factors for the five proteins also show enhanced structural fluctuations in the vicinity of the site of substitution L125A and L103A (Supporting Information Fig. S8).

To obtain a more complete picture of nanosecond dynamics, the N—H bond order parameter S^2 were calculated for the reference Δ + PHS protein using the internal correlation method (see Materials and Methods), and compared with the experimental values obtained from Kitahara *et al.*[37] for the reference Δ + PHS protein under similar conditions (Supporting Information Fig. S9). The simulated order parameters are in good agreement with the experimental ones, with an absolute difference mean value of 0.033 and a standard deviation of 0.027, indicating that the parameters used in the simulation reproduce reasonably well the intrinsic dynamics of the protein in solution. As described recently,[43,45,46] residues involved in slower motions will show a continuous decrease of the N—H bond internal correlation function because simulations weight equally the motions that are both faster and slower than the overall protein tumbling. Such residues are defined as "nonconverging" in this calculation (see Material and Methods). Interestingly, for L103A a significantly larger number (63) of nonconverging residues was found than for Δ + PHS[46] or L125A.[45] In particular, the N—H vector internal correlation function of residues 10, 59, 64, and 68 converge within a few ns in the simulations with Δ + PHS SNase, but exhibit a continuous decrease in the case of L103A (Supporting Information Fig. S10A). The residues

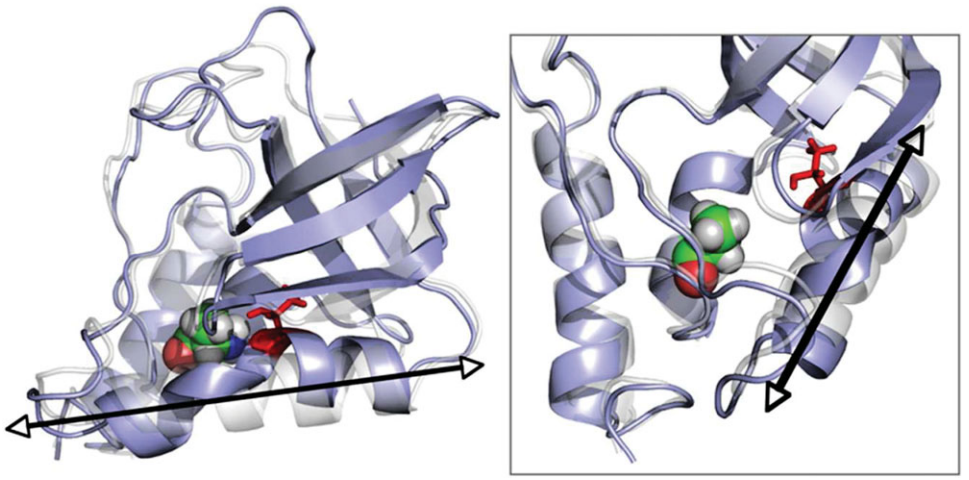

Figure 7

Snapshots of L103A simulation at $t = 53$ ns illustrating the kink of helix 1. The structure of L103A at $t = 53$ ns (blue) is superimposed to the starting structure (gray). Position of Ala 103A is indicated with vdW spheres and Thr 62 is shown in red stick. The native axis of helix 1 is shown (black arrow).

possibly involved in slower motions are located primarily in SubD1, with residues 14–30 (β-barrel) and 59–66 (helix 1), notably affected in the case of L103A (Supporting Information Fig. S10B). This observation suggests that in addition to the local destabilization of helix-1, apparent in the reversible kink and in the motion of Thr-62, the L103A substitution also affects the dynamics of several residues in the β-barrel.

DISCUSSION

For mutations that introduce cavities the correlation between $\Delta\Delta G$ values and the size of the cavity or the number of heavy atoms in the vicinity of the mutation, as observed with the SNase variants, are interpreted classically as reflecting the additive contribution of a hydrophobic solvation term (∼ 2 kcal/mol for a Leu to Ala substitution) and the loss of van der Waals interactions.[16–18,44] Only in the case of the L38A variant was the $\Delta\Delta G_f$ value [1.4 kcal/mol in Δ + PHS, (1.7 kcal/mol in WT SNase[17])] significantly lower than the typical range of stability differences measured for Ala substitutions[44] (3–4 kcal/mol). Considering that L38A is the only variant located in a loop, it is possible to invoke additional entropic compensation effects to explain this discrepancy. A similar conclusion was reached by Bothelo *et al.*[47] concerning the effects of two substitutions on the global stability of β-lactoglobulin isoforms A and B. For 7 of the 10 variants, we observed a remarkable agreement between the present $\Delta\Delta G_f$ values and those obtained for the same variants in the true wild type background.[17] The largest discrepancy was observed for the L125A mutation, which showed a significantly smaller $\Delta\Delta G_f$ in the Δ + PHS background than in the true wild type background (3.8 kcal/mol and 4.9 kcal/mol, respectively). Presumably, this reflects stabilization of helix-3 via three mutations (P117G, H124L, and S128A) in the Δ + PHS variant. The lower m_{GuHCl} values observed for the Δ + PHS variants compared to those in WT SNase[17] may be interpreted in terms of a smaller amount of surface area exposed upon unfolding, or alternatively, in terms of a lower unfolding cooperativity owing to the population of intermediate species for variants of the more stable Δ + PHS background.

Various consequences of cavity creating mutations in the crystal structures of proteins have been reported, from major rearrangement of the hydrophobic core in the Sso7d F31A variant,[48] limited local rearrangement of side chains in T4 lysozyme L121A[16] or even the absence of any structural adjustment in the case of T4 lysozyme L99A.[20] The I92A and V66A variants in the β-barrel core of SNase clearly fall into the last category since no structural rearrangement was observed for these two variants, even in solution. It is likely that, as for the L99A variant of T4 lysozyme, topological and steric constraints render the core in SubD1 of SNase structurally rigid and unable to adjust to the presence of the engineered cavity. Native state high-pressure NMR experiments suggest that the nonoptimal accommodation of the engineered cavities in the I92A and V66A variants leads to destabilization of the native states with respect to lower volume, high-energy conformers that can be populated by application of pressure. Along these lines, a recent relaxation dispersion NMR study of the L99A variant of T4 lysozyme[49] has revealed the presence of a high-energy conformer that minimizes the volume of the engineered cavity. A coherent picture emerges for cavity-creating mutations in a

rigid environment. Such perturbations do not affect the native structure, nor do they affect the global stability any differently than the introduction of cavities in other regions of the protein. Rather, they alter the local stability and hence the probability of populating higher energy states in the folded state ensemble.

The H/D exchange measurements reported here, as well as the pressure dependence of the amide chemical shifts reveal that substitutions in the rigid central core of SNase (I92A and V66A) lead to significant increases in the probability of populating locally open (partially unfolded) species at positions far removed from the site of the mutation. We suggest that these long-range consequences arise from the cooperative disruption of multiple and widespread interactions by the mutation. At the same time, these cavity creating mutations in the rigid central core of SNase result in only very limited changes in the native state HSQC chemical shifts. In contrast, substitution of large hydrophobic residues by alanine in more flexible regions of the protein (i.e., L125A) has few long-range consequences for partial unfolding. This is because the folded structure can adjust locally to accommodate the mutation, and these structural adjustments of the folded state are apparent in the significant changes in the amide chemical shifts of the folded form.

Consistent with the notion of adjustments in their folded states, all-atom, MD simulations suggested increased dynamics on a nanosecond time scale for L125A and L103A variants. In particular, a large concerted motion involving a kink of helix-1 and the penetration of Thr-62 into the protein core was observed with the L103A variant. Perturbation of Thr-62 because of the L103A substitution was also noticed in our atmospheric and high-pressure NMR experiments and likely reflects a local destabilization of helix-1. The OH group of Thr62 forms an H-bond to a buried backbone amide in the turn between β1 and β2, Perturbation of this residue is likely responsible for the increased dynamics in SubD1 for this variant. We hypothesize that, in comparison to the rigid core of SubD1, the more flexible environments of the IntD and SubD2 allow subtle structural rearrangements of the native structure and a better accommodation of the engineered cavity.

CONCLUSION

This study illustrates how, depending on their location, the introduction of cavities differentially perturb local stability, flexibility, and dynamics. These complex consequences are not predicted by simple $\Delta\Delta G_f$ measurements. Indeed, two variants, L125A and I92A, exhibited highly distinct behavior in terms of flexibility, local stability and dynamics, whereas their effects on global thermodynamic stability were equivalent within experimental uncertainty. Such distinct effects cannot be predicted by examination of crystal structures, as no significant rearrangements were apparent in any of the variants.[27] We find two limiting responses to cavity creating mutations. Cavities engineered in inherently flexible environments can be accommodated structurally in solution. This structural accommodation is reflected by significant changes in the native state NMR spectra and in the rapid folded state dynamics probed by MD simulations, but only very subtle changes in slower higher energy fluctuations as revealed by native state H/D exchange, and pressure sensitivity of the NMR spectra. In contrast, in a rigid environment, structural adjustments to cavities are not possible. Instead, the cavities destabilize the native state relative to higher energy conformers that are observed in H/D exchange and pressure perturbation experiments. Further, we have identified a perturbation propagation pathway between subdomains of SNase, which suggests a role for SubD2 as conformational sensor, perhaps involved in regulating the enzymatic activity of the protein or its interactions with substrate.[50,51] More generally, this study illustrates how natural mutations to protein cores could modulate the accessibility of functionally important or detrimental excited states. Our approach also constitutes a novel strategy for probing structural plasticity and free energy landscapes of proteins, which should be useful to gain new perspectives into subtle aspects of the structural basis of protein function.

REFERENCES

1. Goodey NM, Benkovic SJ. Allosteric regulation and catalysis emerge via a common route. Nat Chem Biol 2008;4:474–482.
2. Taylor SS, Kim C, Vigil D, Haste NM, Yang J, Wu J, Anand GS. Dynamics of signaling by PKA. Biochim Biophys Acta 2005; 1754:25–37.
3. Nashine VC, Hammes-Schiffer S, Benkovic SJ. Coupled motions in enzyme catalysis. Curr Opin Chem Biol 2010;14:644–651.
4. Volkman BF, Lipson D, Wemmer DE, Kern D. Two-state allosteric behavior in a single-domain signaling protein. Science 2001;291: 2429–2433.
5. Fraser JS, Clarkson MW, Degnan SC, Erion R, Kern D, Alber T. Hidden alternative structures of proline isomerase essential for catalysis. Nature 2009;462:669–673.
6. Pan H, Lee C, Hilser VJ. Binding sites in *Escherichia coli* dihydrofolate reductase communicate by modulating the conformational ensemble. Proc Natl Acad Sci USA 2000;97:12020–12025.
7. Richards FM. Packing defects, cavities, volume fluctuations and access to the interior of proteins. Including some general comments on surface area and protein structure. Carlsberg Res Commun 1979;44:47–63.
8. Liang J, Edelsbrunner H, Woodwards C. Anatomy of protein pockets and cavities: measurements of binding site geometry and implication for ligand design. Protein Sci 1998;7:1884–1897.
9. Kundrot CE, Richards CM. Crystal structure of hen egg-white lysozyme at a hydrostatic pressure of 1000 atmospheres. J Mol Biol 1987;193:157–170.
10. Fu Y, Kasinath V, Moorman VR, Nucci NV, Hilser VJ, Wand AJ. Coupled motion in proteins revealed by pressure perturbation. J Am Chem Soc 2012;134:8543–8550.
11. Lassalle MW, Akasaka K. The use of high-pressure nuclear magnetic resonance to study protein folding. Methods Mol Biol 2007;350:21–38.

12. Li H, Akasaka K. Conformational fluctuations of proteins revealed by variable pressure NMR. Biochim Biophys Acta 2006;1764:331–345.
13. Akasaka K, Li H. Low-lying excited states of proteins revealed from nonlinear pressure shifts in ^{1}H and ^{15}N NMR. Biochemistry 2001;40:8665–8671.
14. Kamatari YO, Smith LJ, Dobson CM, Akasaka K. Cavity hydration as a gateway to unfolding: an NMR study of hen lysozyme at high pressure and low temperature. Biophys Chem 2011;156:24–30.
15. Tomita A, Sato T, Ichiyanagi K, Nozawa S, Ichikawa H, Chollet M, Kawai F, Park S-Y, Tsuduki T, Yamoto T, Koshihara S, Adachi S. Visualizing breathing motion of internal cavities in concert with ligand migration in myoglobin. Proc Natl Acad Sci USA 2009;106:2612–2616.
16. Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, Matthews BW. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. Science 1992;255:178–183.
17. Shortle D, Stites WE, Meeker AK. Contributions of the large hydrophobic amino acids to the stability of Staphylococcal Nuclease. Biochemistry 1990;29:8033–8041.
18. Serrano L, Kellis JT, Jr., Cann P, Matouschek A, Fersht AR. The folding of an enzyme II. Substructure of barnase and the contribution of different interactions to protein stability. J Mol Biol 1992;224:783–904.
19. Collins MD, Kim CU, Gruner SM. High-pressure protein crystallography and NMR to explore protein conformations. Annu Rev Biophys 2011;40:81–98.
20. Collins MD, Quillin ML, Hummer G, Matthews BW, Gruner SM. Structural rigidity of a large cavity-containing protein revealed by high-pressure crystallography. J Mol Biol 2007;367:752–763.
21. Collins MD, Hummer G, Quillin ML, Matthews BW, Gruner SM. Cooperative water filling of a nonpolar protein cavity observed by high-pressure crystallography and simulation. Proc Natl Acad Sci USA 2005;102:16668–16671.
22. Ye K, Jing G, Wang J. Interactions between subdomains in the partially folded state of staphylococcal nuclease. Biochim Biophys Acta 2000;1479:123–134.
23. Bedard S, Mayne LC, Petersen RW, Wand AJ, Englander SW. The foldon substructure of staphylococcal nuclease. J Mol Biol 2008;376:1142–1154.
24. Watson E. Matousek WM, Irimies EL, Alexandrescu AT. Partially folded states of staphylococcal nuclease highlight the conserved structural hierarchy of OB-fold proteins. Biochemistry 2007;46: 9484–9494.
25. Ishii T, Murayama Y, Katano A, Maki K, Kuwajima K, Sano M. Probing force-induced unfolding intermediates of a single staphylococcal nuclease molecule and the effect of ligand binding. Biochem Biophys Res Commun 2008;375:586–591.
26. Murzin AG. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. EMBO J 1993;12:861–867.
27. Roche J, Caro JA, Norberto DR, Barthe P, Roumestand C, Schlessman JL, Garcia AE, Garcia-Moreno BE, Royer CA. Cavities determine the pressure unfolding of proteins. Proc Natl Acad Sci USA 2012;109:6945–6950.
28. Shortle D, Meeker AK. Residual structure in large fragments of staphylococcal nuclease: effects of amino acids substitutions. Biochemistry 1998;28:936–944.
29. Castaneda CA, Fitch CA, Majumbar A, Khangulov V, Schlessman JL, Garcia-Moreno BE. Molecular determinants of the pKa values of Asp and Glu residue in staphylococcal nuclease. Proteins 2009;77:570–588.
30. Whitten ST, Garcia-Moreno BE. pH dependence of stability of staphylococcal nuclease: evidence of substantial electrostatic interactions in the denaturated state. Biochemistry 2009;39:14292–14304.
31. Piotto M, Saudek V, Sklenar V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J Biomol NMR 1992;2:661–665.
32. Sklenar V. Selective excitation techniques for water suppression in one- and two-dimensional NMR spectroscopy. Basic Life Sci 1990;56:63–84.
33. Pons JL, Malliavin TE, Delsuc MA. Gifa V. 4: a complete package for NMR data set processing. J Biomol NMR 1996;8:445–452.
34. Bax A, Pochapsky SS. Optimized recording of heteronuclear multidimensional NMR spectra using pulsed field gradients. J Magn Reson 1992;99:638–643.
35. Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, Clore GM. Overcoming the overlap problem in the assignment of 1H NMR spectra of larger proteins by use of three-dimensional heteronuclear 1H-15N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta. Biochemistry 1989;28:6150–6156.
36. Bodenhausen G, Ruben DJ. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. Chem Phys Lett 1980;69:185–189.
37. Kitahara R, Hata K, Maeno A, Akazaka K, Chimenti MS, Garcia-Moreno BE, Schroer MA, Jeworrek C, Tolan M, Winter R, Roche J, Roumestand C, Montet de Guillen K, Royer CA. Structural plasticity of staphylococcal nuclease probed by perturbation with pressure and pH. Proteins 2011;79:1293–1305.
38. Bai Y, Milne JS, Mayne L, Englander SW. Primary structure effects on peptide group hydrogen exchange. Proteins 1993;17:75–86.
39. Rouget JB, Aksel T, Roche J, Saldana JL, Garcia AE, Barrick D, Royer CA. Size and sequence and the volume change of protein folding. J Am Chem Soc 2011;133:6020–6027.
40. van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible and free. J Comput Chem 2005;26:1701–1718.
41. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 2006;65:712–725.
42. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
43. Sgourakis NG, Day R, McCallum SA, Garcia AE. Pressure effects on the ensemble dynamics of ubiquitin inspected with molecular dynamics simulations and isotropic reorientational eigenmode dynamics. Biophys J 2008;95:3943–3955.
44. Loladze VV, Ermolenko DN, Makhatadze GI. Thermodynamic consequences of burial of polar and non-polar amino acid residues in the protein interior. J Mol Biol 2002;320:343–357.
45. Maragakis P, Lindorff-Larsen K, Eastwood MP, Dror RO, Klepeis JL, Arkin IT, Jensen MO, Xu H, Trbovic N, Friesner RA, Palmer AG, III, Shaw DE. Microsecond molecular dynamics simulation shows effect of slow loop dynamics on backbone amide order parameters of proteins. J Phys Chem B 2008;112:6155–6158.
46. Tian J, Garcia AE. Simulations of the confinement of ubiquitin in self-assembled reverse micelles. J Chem Phys 2011;134:225101–225110.
47. Botelho MM, Valente-Mesquita VL, Oliveira KM, Polikarpov I, Ferreira ST. Pressure denaturation of β-lactoglobulin. Eur J Biochem 2000;267:2235–2241.
48. Consonmi R, Santomo L, Fusi P, Tortora P, Zetta L. A single-point mutation in the extreme heat and pressure-resistant Sso7d protein from *Sulfolobus solfataricus* leads to a major rearrangement of the hydrophobic core. Biochemistry 1999;38:12709–12717.
49. Bouvignies G, Vallurupalli P, Hansen DF, Correia BE, Lange O, Bah A, Vernon RM, Dahlquist FW, Baker D, Kay LE. Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. Nature 2011;477:111–116.
50. Arcus V. OB-fold domains: a snapshot of the evolution of sequence, structure and function. Curr Opin Struct Biol 2002;12:794–801.
51. Alexandrescu AT, Gittis AG, Abeygunawardana C, Shortle D. NMR structure of a stable "OB-fold" sub-domain isolated from Staphylococcal nuclease. J Mol Biol 1995;250:134–143.