Local Structural Preferences and Dynamics Restrictions in the Urea-Denatured State of SUMO-1: NMR Characterization

Ashutosh Kumar,* Sudha Srivastava,* Ram Kumar Mishra,† Rohit Mittal,† and Ramakrishna V. Hosur*
*Department of Chemical Sciences and †Department of Biological Sciences, Tata Institute of Fundamental Research, Mumbai, India

ABSTRACT We have investigated by multidimensional NMR the structural and dynamic characteristics of the urea-denatured state of activated SUMO-1, a 97-residue protein belonging to the growing family of ubiquitin-like proteins involved in post-translational modifications. Complete backbone amide and ^{15}N resonance assignments were obtained in the denatured state by using HNN and HN(C)N experiments. These enabled other proton assignments from TOCSY-HSQC spectra. Secondary H^{α} chemical shifts and ^{1}H-^{1}H NOE indicate that the protein chain in the denatured state has structural preferences in the broad β-domain for many residues. Several of these are seen to populate the (ϕ,ψ) space belonging to polyproline II structure. Although there is no evidence for any persistent structures, many contiguous stretches of three or more residues exhibit structural propensities suggesting possibilities of short-range transient structure formation. The hetero-nuclear ^{1}H-^{15}N NOEs are extremely weak for most residues, except for a few at the C-terminal, and the ^{15}N relaxation rates show sequence-wise variation. Some of the regions of slow motions coincide with those of structural preferences and these are interspersed by highly flexible residues. The implications of these observations for the early folding events starting from the urea-denatured state of activated SUMO-1 have been discussed.

INTRODUCTION

Characterization of the unfolded/denatured states of proteins is important for many reasons; here, the term ''unfolded'' refers to a state completely devoid of any structural preferences and the term ''denatured'' refers to the lowest energy nonnative state under a given set of conditions (1). The unfolded and the denatured states may coincide under extreme denaturing conditions. Firstly, the denatured state represents the reference state for estimating the stability of a protein. Next, the denatured state is the starting point for the folding process in vivo. In this context, it has been a major area of research to understand how the unfolded chain finds its final folded state destination (2–4). There is now ample evidence to believe that the different members of the unfolded ensemble fold along different pathways on a so-called folding funnel, whose broad end represents the unfolded state, and the narrow end, the folded state (5–7). Although it is almost impossible to know at this point of time, what the exact nature of the denatured state in vivo is, it is believed that the chain is not entirely random but has at least some local preferences on the average that restrict the conformational search by the folding chain. The local preferences could vary depending upon the environmental conditions. In view of this, there is vigorous effort in the literature to characterize the denatured states created by a variety of means, which include use of chemical denaturants (2,8–14), change of pH conditions (15–17), or thermal denaturation (18,19). It has been observed that different conditions populate different conformational states in the ensemble (20,21). Among the various means used for creating denatured states, guanidine denaturation is known to cause the maximum degree of unfolding. This is followed by urea and other denaturing agents. pH denaturation appears to be more tolerant and so is temperature. Sodium dodecyl sulphate micelles cause mostly removal of tertiary interactions (22). Consequently, as must be expected, the equilibrium unfolding pathways in many of these cases have also been found to be different (23–26); some follow a two-state process, whereas others follow multistate processes which have implications for protein stability studies. Thus, study of all such denatured states is expected to provide insights into the different pathways of folding, starting from the different denatured states in the ensemble.

The unfolded state and the partially unfolded states have also been thought to be playing important roles in protein function. Recently, the structure-function paradigm has been revisited and a model called the ''protein-quartet model'' has been put forward, according to which a protein function arises as a consequence of an interplay between folded states, unfolded states, molten globules, and premolten globules (27,28). A direct role of protein unfolding has been seen in transport of proteins across the membranes (29). Lastly, the denatured states of proteins have often been seen to cause protein aggregation (30), and such aggregation is known to be one of the reasons for many diseases in vivo (31,32).

One type of structure reported by NMR in denatured proteins is a hydrophobic cluster, usually formed by local side-chain interactions (33–36). Another type of structure reported is a fluctuating secondary structure. In the case of

Submitted July 30, 2005, and accepted for publication December 29, 2005.

Address reprint requests to Prof. Ramakrishna V. Hosur, Dept. of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India. Tel.: 91-22-2280-4545; Fax: 91-22-2280-4610; E-mail: hosur@tifr.res.in.

Abbreviations used: HSQC, hetero-nuclear single quantum coherence; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; SUMO, small ubiquitin related modifier.

0006-3495/06/04/2498/12 $2.00

doi: 10.1529/biophysj.105.071746

barnase (37,38), the urea-denatured protein shows regions of fluctuating secondary structure, which correspond fairly well to the moderately and highly structured regions of the intermediates and the transition states of folding (39). Native- and nonnative-type secondary structural preferences were seen in the urea-denatured state of barstar (9), guanidine-denatured state of HIV-1 protease (2,8,10), and very recently in the acid-denatured state of hUBF HMG Box 1 (15). A nonnative α-helical structure was observed in the guanidine denatured state of β-lactoglobulin, an all-β-protein in the native state (40). Similarly, β-type preferences were seen in denatured apomyoglobin, a largely helical protein (41). In some cases, multiple types of structures have been reported in the denatured state. For example, for the drkN SH3 domain, detailed NMR studies (42–44) indicate the presence of multiple structures ranging from conformers with non-native structure possessing long-range contacts to those with more compact structures maintaining native-like secondary structure.

There is growing evidence in the literature, mostly based on circular dichroism (CD) spectroscopy, that chemically denatured proteins populate polyproline (PP_{II}) structure which lies in the β-domain of (ϕ,ψ) space of the Ramachandran map. This is a left-handed helical structure, generally observed in proline-rich peptides (45–48) and has a characteristic positive CD band at 225 nm. Further, where residual structures in the β-domain have been detected by NMR in denatured proteins, it is suggested that PP_{II} structure may be populated to a significant extent (49,50).

In view of the above, we have initiated investigation of the structural and dynamic characteristics of the variety of denatured states of the protein, SUMO-1, a 101-residue protein belonging to the growing family of ubiquitin-like proteins involved in post-translational modifications (51). It attaches itself to many target proteins, by a process called ''sumoylation''. Before sumoylation the protein gets activated by cleaving off four residues from the C-terminal end. There have been two reports in the literature on the structure of the 101-residue SUMO-1 (52,53). We reported the NMR structure of the activated SUMO-1 at pH 7.4 and 27°C (54). All three structures were essentially similar and the average structure of the molecule contains the following secondary structural elements: β1: 22–28; β2: 33–39; β3: 62–66; β4: 69–72; β5: 87–93; α1: 45–55, and α2: 77–82. The β-strands form a bent sheet. The regions intervening between secondary structures form flexible loops. The amino-terminal 20 residues are highly flexible. In this article, we report NMR investigations on the characteristics of the denatured state of SUMO-1 (1–97) created by 8 M urea at pH 5.6 and 27°C. Hereafter, we refer to this protein as SUMO-1 only for sake of brevity. Our results indicate that in the urea-denatured state, a large number of residues have slight (ϕ,ψ) preferences in the broad β-domain, which encompasses the extended structures as well. The dynamic motions at the high frequency scale appear to be fairly restricted. There are sequence-wise small variations in the milli- to microsecond timescale motions. These results provide a description of the denatured state on the one hand, and also provide clues to the possible nucleation sites for the folding reaction of the protein, on the other.

MATERIALS AND METHODS

Protein preparation

SUMO-1 was prepared and purified as described elsewhere (54).

NMR spectroscopy

Isotopically labeled protein samples for NMR experiments were prepared as described in Mishra et al. (54). The proteins were concentrated to ~1 mM and exchanged with 0.1 M phosphate buffer (pH 5.6) containing 150 mM NaCl, 5 mM EDTA, 1 mM DTT, and 8 M urea. The NMR experiments were started after keeping the solutions for ~3 h so as to reach equilibrium.

All the NMR experiments were performed on a Varian Inova 600 MHz spectrometer at 27°C (Varian, Cary, NC). A series of two-dimensional and three-dimensional experiments (see below) were carried out which lasted for six days. At the end of these the HSQC spectrum was again recorded to check for the stability of the protein. We observed no change in the HSQC spectra indicating that the protein was very stable and also had reached equilibrium state at the beginning of the experiments. H^N and ^{15}N resonance assignments were obtained using HNN and HN(C)N triple resonance experiments (55–57). HNCA, HN(CO)CA, TOCSY-HSQC, and NOESY-HSQC experiments (reviewed in (58,59)) facilitated the assignments by providing additional checks. Standard experimental parameters were used: 24–30 ms for N-C^{α} and N-C′ transfers in HNN, HN(C)N, HNCA, and HN(CO)CA; and 60–80 ms for TOCSY transfer and 150–200 ms for NOESY transfers. The 4–16 scans were used for each *fid* in the different experiments. The 40–50 and 80 complex increments were used along the ^{15}N and C′ dimensions respectively; 60–80 complex increments were used along the aliphatic carbon dimensions; 80–96 complex increments were used along the indirect ^1H dimension. Temperature coefficients of amide protons were measured by recording HSQC spectra at 3°C temperature increments in the range 15–36°C. ^{15}N transverse relaxation rates (R_2) were measured using CPMG delays, 10, 50, 70*, 90, 130, 170*, 210, and 250 ms, where asterisks indicate duplicate measurement. ^{15}N longitudinal relaxation rates (R_1) were measured using the inversion recovery delays 10, 50, 120*, 220, 350, 700*, and 950 ms. Steady-state ^1H-^{15}N heteronuclear NOE measurements were carried out with a proton saturation time of 3 s and a relaxation delay of 2 s. For the experiment without proton saturation the relaxation delay was 5 s. All the relaxation experiments were carried out using the pulse sequences described by Farrow et al. (60). H^N-H^{α} coupling constants were measured from a high resolution HSQC spectrum recorded with 8192 complex t_2 points and 512 complex t_1 points.

Circular dichroism

Far-UV circular dichroism (CD) spectra of the protein at 27°C, were recorded on a JASCO–J810 spectropolarimeter (Jasco, Hachioji, Japan), at 0 and 8 M urea concentrations using 0.2 cm cell. The protein concentration was 20.7 μM. The samples at the appropriate conditions were equilibrated for at least 10–12 h before CD measurements. Each spectrum was an average of eight scans (slit width of 2 nm).

RESULTS AND DISCUSSION

Resonance assignments

Resonance assignment in denatured proteins by the standard triple resonance experiments such as HNCA, HN(CO)CA,

HNCACB, CBCANH, CBCA(CO)NH, etc., which have proved extremely successful for folded proteins, is severely hampered because of the high degeneracy of the amide and carbon chemical shifts (61). Consequently there have been relatively few proteins on which detailed studies at residue level have been carried out (61–65). In such situations, recent experiments that exploit the ^{15}N or CO chemical shifts, which have good dispersions in both folded and unfolded proteins, would be extremely valuable (65,66). Presently, we were able to obtain complete amide proton and backbone ^{15}N assignments for SUMO-1, at pH 5.6 and 27°C, in 8 M urea using the HNN and HN(C)N experiments described by us a few years ago (55–57). These exploit the good dispersion of the ^{15}N chemical shifts along two of the three dimensions, and also provide many checkpoints along the sequential walk by way of special patterns of peaks around Gly and Pro residues. Thus side-chain identification does not become very essential during the sequential walk, and this enhances the speed of assignment. However, under rare occasions of overlaps, TOCSY-HSQC spectrum, which provides the spin system information, helps resolve the ambiguities. An illustrative sequential walk through the HNN spectrum of SUMO-1 at 27°C, pH 5.6 is shown in Fig. 1 *A* and the summary of all the sequential connectivities is shown in Fig. 1 *B*. Fig. 2 shows the ^{1}H-^{15}N HSQC spectrum with all the peak assignments. During the course of these sequential assignments, many of the side-chain proton assignments were obtained from the TOCSY-HSQC spectrum.

Structural propensities in the denatured state

Deviations of observed chemical shifts from the random coil values (secondary shifts) are very good indicators of secondary structures in proteins. The random coil values are generally measured using short 5–6 residue peptides, as they do not possess any structural preferences in solution. These are particularly useful in denatured proteins where NOEs are rather scarce. However, there have been more than one set of random coil values reported in the literature and these differ due to slightly different experimental conditions. We have calculated these shifts using two sets, one by Wishart and Sykes (67,68) and the other by Schwarzinger et al. (69). The former uses pH 5 and 1 M urea, while the latter uses 8 M urea and pH 2.3 for the measurements on peptides, GGXGG, where *X* is the residue of interest for random coil chemical shift and *G* is glycine. In our case the two calculations made only a marginal difference and we decided to use the data given by Schwarzinger et al. (70), since these authors have also described a set of sequence corrections to the random coil values for all the nuclei.

Among the various secondary chemical shifts ($\Delta\delta$), those of H^{α}, $^{13}C^{\alpha}$, and ^{13}CO are the most diagnostic of structural information in the polypeptide chain (63,64). Thus if some residues show positive (downfield) H^{α} and negative (upfield) $^{13}C^{\alpha}$ and ^{13}CO structure-dependent chemical shift deviations (secondary shifts, $\Delta\delta$), then those residues are taken to have (ϕ,ψ) preferences in the β-domain of the Ramachandran map

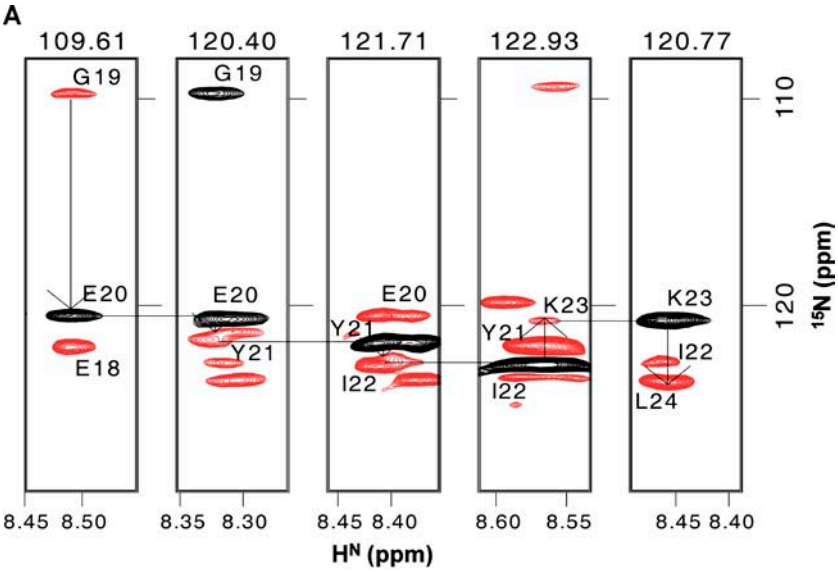

FIGURE 1 (*A*) Sequential walk through the F_1–F_3 planes of HNN spectrum of SUMO-1 in 8 M urea at pH 5.6 and 27°C. Sequential connectivities are shown for Gly-19 to Leu-24 stretch. The F_2 (^{15}N) values are shown at the top for each strip. The black and red contours indicate positive and negative peaks. The distinct Gly plane serves as the start point in the sequential assignment. (*B*) Summary of sequential assignment obtained from HNN walk along the primary sequence of SUMO-1. Gly shown in green worked as start/check point; Pro shown in red worked as break point.

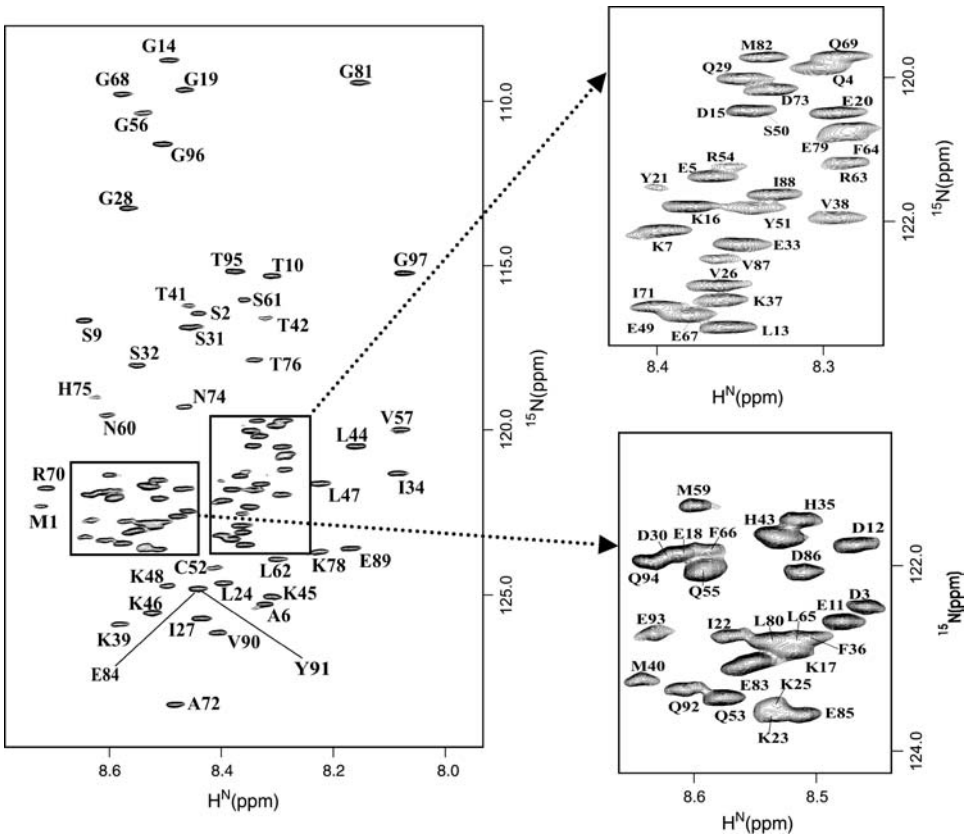

FIGURE 2 (^{1}H-^{15}N) HSQC spectrum of SUMO-1 in 8 M urea at pH 5.6 and 27°C. Residue-specific assignment for each peak is marked on the spectrum.

((ϕ,ψ) = (−70 to −150°, 90–180°)). Residues with α-helical propensities ((ϕ,ψ) = (−40 to −80°, −30 to −60°)) show the opposite pattern.

We have used H^{α} secondary shifts to characterize the residual structure in the protein. Fig. 3 *A* shows a plot of the secondary shifts against the sequence for SUMO-1 at pH 5.6 and 27°C. Clearly, from the pattern of the deviations which show a positive bias, the chain does not appear to be a random coil. Though the shifts are small, overall there seems to be some tendency for a large number of residues across the chain to populate the broad β-region of the (ϕ,ψ) space; this includes the extended conformations and PP_{II} structures ((ϕ,ψ) = (−70 to −80°, 140 to 150°)) as well. There are a few contiguous stretches of three or more residues showing positive deviations of >0.08 ppm and these may be taken to indicate formation of β-structures, perhaps transiently, along the chain. They have been indicated by solid horizontal bars inside the figure.

Next, we measured ^{3}J (H^N- H^{α}) coupling constants for all the residues. Three bond H^N-H^{α} coupling constants are useful indicators of ϕ-torsion angles. This coupling constant is of the order of 3–5 Hz for a helical structure (including PP_{II}) and 8–11 Hz for a β-structure (64). For a random coil, a weighted average of these values would be observed, which is typically between 6.0 and 8.0 Hz for most residues. It is also observed that the random coil value for any residue is influenced by its N-terminal neighbor and thus two sets of values have been reported for each residue, depending upon whether the N-terminal neighbor belongs to one of the two groups of residues (group I: F, W, H, Y, I, T, V, and group II: remaining residues except glycine) (71). Thus, under any given experimental conditions, one can calculate deviation of the observed coupling constant from the sequence-dependent random coil value, (J_{obs}–J_{rc}), for every residue. These ''secondary coupling constants,'' as we call them, throw valuable light on the secondary structural propensities along the polypeptide chain. Negative secondary coupling constants would indicate helical propensities (including PP_{II}) and positive secondary coupling constants would indicate β-propensities (71). This led us to the idea that a combined use of secondary chemical shifts and secondary coupling constants would enable distinction between β and PP_{II} structures. For a truly β-structure, both the secondary shifts and the secondary coupling constants would be positive, whereas for the PP_{II} helix the secondary coupling constant would be negative, as in the common right-handed α-helices, and the secondary shifts would be positive, like in the β-structures.

The measured values of the secondary coupling constants from a high resolution HSQC spectrum of the protein (an illustrative region to show the quality of the spectral resolution is shown in Fig. 3 *B*), are shown in Fig. 3 *C*. The deviations are mostly <1.0 Hz for most residues. We may mention here that the accuracy of secondary coupling constant estimation is ~0.5 Hz for the positive values, and

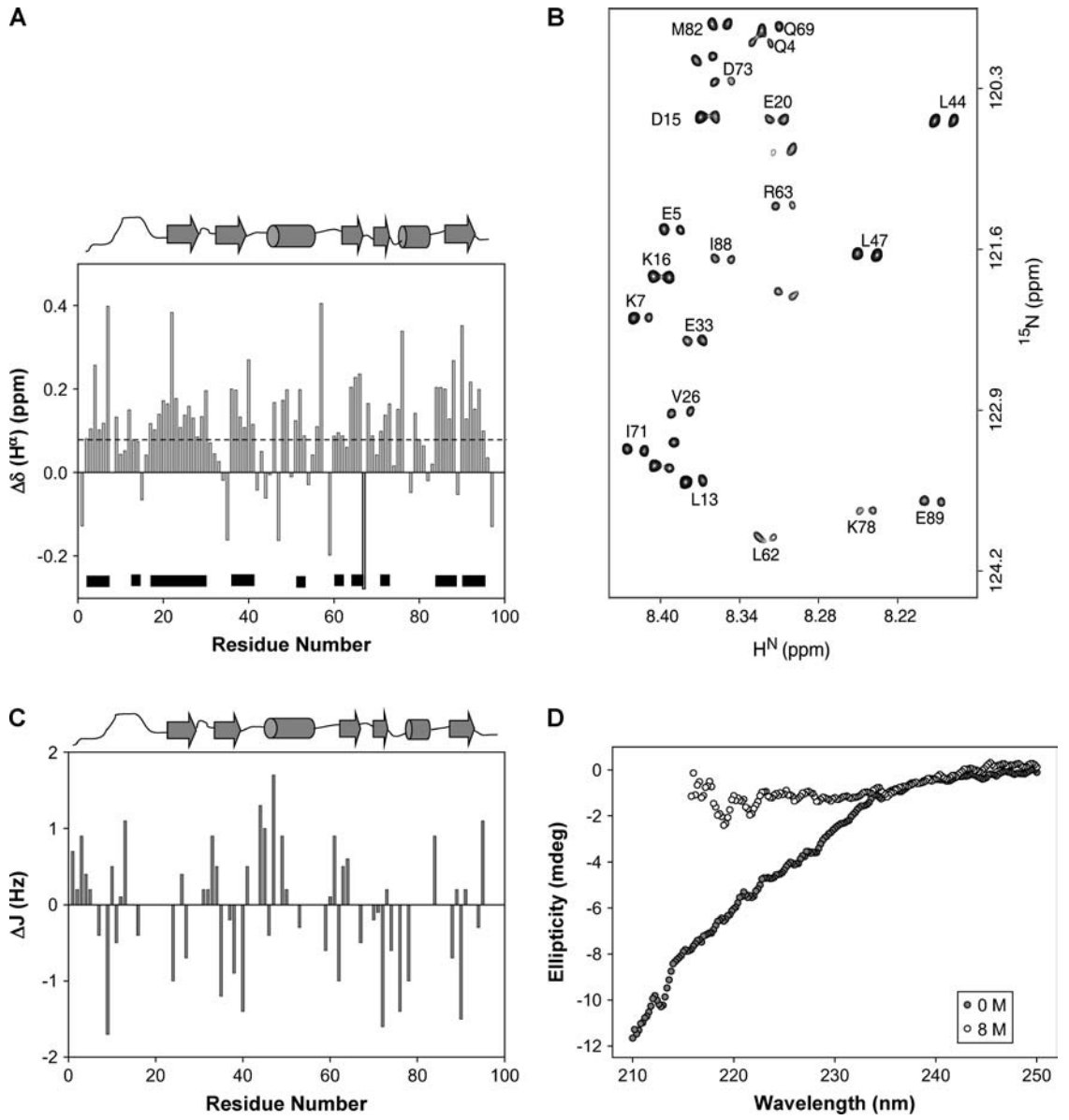

FIGURE 3 (*A*) Sequence-corrected secondary chemical shift for H^{α} ($\Delta\delta(H^{\alpha})$) in SUMO-1 at 8 M urea at pH 5.6 and 27°C. The sequence-corrected random coil values used are those determined by Schwarzinger et al. (70). The native secondary structural elements are shown on the top of the figure and the segments showing structural propensities are indicated by solid bars. (*B*) Selected region from a high resolution (^{1}H-^{15}N) HSQC spectrum of SUMO-1 in 8 M urea at pH 5.6 and 27°C. The splitting in the peaks was used to measure ^{3}J (H^{N}-H^{α}) coupling constants. (*C*) Secondary coupling constants (see text) are plotted against sequence for SUMO-1. (*D*) Far-UV CD spectra of SUMO-1 at pH 5.6 and 27°C, in the absence of urea (*shaded circles*), and in 8 M urea (*open circles*).

slightly worse but $<\sim 1$ Hz for negative deviations; this is because the negative deviations occur when the coupling constants themselves are small. Thus the deviations beyond these values may indicate slight (ϕ,ψ) preferences. Interestingly, most of the residues exhibiting negative secondary coupling constants have positive secondary shifts, suggesting that their (ϕ,ψ) values may populate the PP_{II} structure to a reasonable extent. A residue-wise assessment of the β or extended, and PP_{II} preferences, from a combined use of secondary shifts and secondary coupling constants, is given in Table 1 of the Supplementary Material. The fact that the residues exhibiting PP_{II} type propensities are not contiguous along the chain indicates that there is no stable PP_{II} helix formed in the ensemble. To check on this further, we recorded the CD spectrum at 27°C in 8 M urea and this is shown along with a similar spectrum in the absence of urea to serve as a reference, in Fig. 3 *D*. Clearly, there is no characteristic positive band at 225 nm confirming the absence of a stable PP_{II} helix in the denatured state ensemble.

The amide proton temperature coefficients (negative values with magnitudes $<\sim 4.5$ ppb/K) indicate hydrogen bonding and thus report on the presence of persistent structures

(72). For random coils the temperature coefficients are roughly in the range −6 to −9.5 ppb/K for the different residues (73). We measured the residue-wise amide proton temperature coefficients of the protein from HSQC spectra recorded in the temperature range 15–36°C. Since all the measured values were negative we calculated their deviations from the random coil values by taking magnitudes only, and these results are shown in Fig. 4. We see that the observed deviations $\Delta(\Delta\delta_{NH}/\Delta T)$ are largely positive. From the absence of significant negative deviations, it is clear that there are no stable intramolecular H-bonds formed in most regions of the polypeptide chain. This rules out the possibility of persistent well-defined structural elements like α-helix, β-sheet, β-turn, etc., in the denatured state.

^{1}H-^{1}H NOEs in folded proteins exhibit secondary structure specific patterns. β-structures including PP_{II} and type II turns are characterized by strong $H^{\alpha}(i) - H^{N}(i+1)$ NOEs, whereas α-helices are characterized by strong $H^{N}(i) - H^{N}(i+1)$ NOEs (74). The α-helices also produce medium range NOEs, from H^{N} (i) to H^{α} $(i-3)$ and $H^{\alpha}(i-4)$. In addition, several other NOEs will also be seen from the amide protons to other intraresidue protons. In unfolded proteins, however, the NOEs are generally weak in the first place due to greater flexibility of the chain, and the above selectivity with regard to secondary structure may be lost due to measurably good population of the different secondary structure types in the ensemble. However, if one does find any kind of preference in the NOE patterns, then that can be taken to indicate a higher population of those specific structures. In SUMO-1 denatured by 8 M urea, the ^{1}H-^{1}H NOEs in the NOESY-HSQC spectra recorded with a mixing time of 150 ms, are rather sparse, but, interestingly, do show some pattern. These are shown in Fig. 5, which seem to support the conclusions derived above from the secondary shifts. A major portion of the chain shows fairly intense sequential $H^{\alpha}(i) - H^{N}(i+1)$ NOEs, and only a few of them also show self-NOEs. In a dipeptide, the sequential distance, $d_{\alpha N}(i, i+1)$ is sensitive to torsion angle ψ_i and varies from 2.2 to 3.5 A°, the shortest distance occurring for ψ_i in the β-domain. Likewise, the distance between H^{α} and H^{N} within the same residue, $d_{\alpha N}(i, i)$, which is sensitive to the torsion angle ϕ_i, varies within a narrow range from 2.2 to 2.9 A° (74). Thus, for those residues where only the sequential peaks are seen, the (ϕ,ψ) propensity has a preponderance of the β-structure, which has the shortest sequential distance and the longest intraresidue distance. For those residues where both the peaks are observed, the ϕ_i may show slight preference toward a shorter intraresidue distance, still within the β-domain. From the NOE data, we conclude that the regions, 9–15, 17–33, 43–48, 61–68, 71–74, 82–87, and 90–95 have a higher population of β-structures in the ensemble.

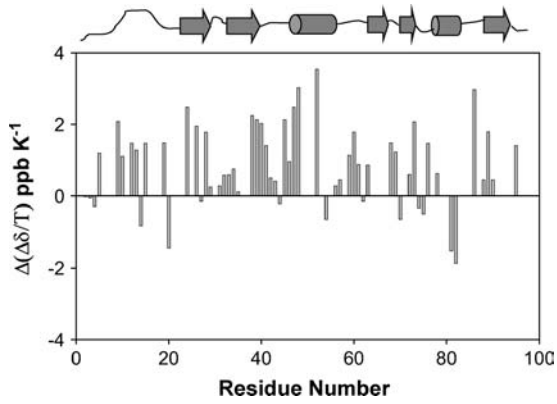

FIGURE 4 Secondary temperature coefficients. Deviation of amide proton temperature coefficients from the random coil temperature coefficient (magnitude mode) at pH 5.6 plotted against the sequence. The native secondary structure elements are indicated at the top.

All the above observations taken together indicate that the SUMO-1 polypeptide chain in 8 M urea does not possess any stable secondary structures, but has some propensity for formation of transient structures with (ϕ,ψ) preferences in the broad β-domain, in 4–5 regions. Several residues in these regions seem to populate the (ϕ,ψ) space belonging to the PP_{II} structure. The structural propensities presently observed in the 8 M urea-denatured state encompass both native and nonnative type preferences for the individual residues, and this would have implications for the folding mechanisms of the protein.

Dynamics restrictions

Since the protein chain exhibited some local structural preferences in the denatured ensemble, it is intuitively natural to expect some variations in the motional characteristics along the length of the chain as well. To probe this explicitly, we carried out ^{15}N R_1, R_2 and heteronuclear ^{1}H-^{15}N NOE measurements, which throw light on the motions over a wide range of timescales (75–77). Among these, the NOEs are very sensitive to picosecond timescale motions, R_2 values are sensitive to milli- to microsecond timescale motions (conformational transitions) and the R_1 values are sensitive to both low and high frequency motions (nanosecond-to-picosecond timescale motions). The heteronuclear NOE intensity varies from ~−4 to 0.9 depending upon the motional correlation times (75,78). Negative values indicate high frequency (picosecond timescale) large amplitude motions, whereas positive values indicate nanosecond timescale motions. In the denatured states, where there is heterogeneity of the conformational states and of the correlation times, these relaxation and NOE data can only be interpreted qualitatively to derive relative motional trends along the chain. Reduced spectral density analysis based on ^{15}N R_1, R_2 and heteronuclear ^{1}H-^{15}N NOE is the best suited method for studying motions at residue level in denatured proteins (15,64). However, in the present case this was hampered by the absence of measurable ^{1}H-^{15}N NOEs due to their low intensities (see below); this of course must have a bearing on the motional characteristics themselves. Nevertheless, we have been able

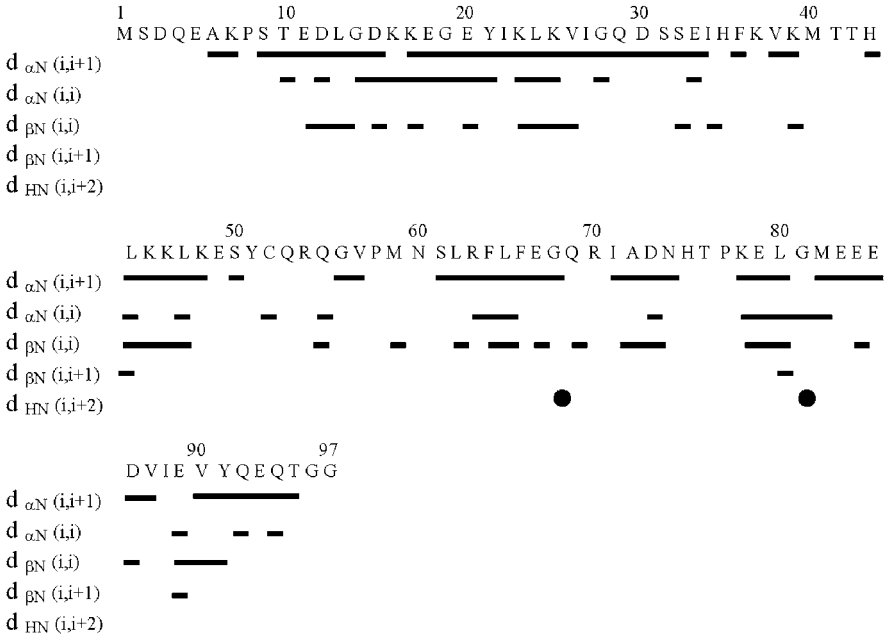

FIGURE 5 Summary of sequence-specific ^{1}H-^{1}H connectivities obtained from the NOESY-^{15}N-^{1}H HSQC spectrum.

to obtain some useful insights into these motions, especially for the nanosecond timescale components and the conformational transitions, from a qualitative analysis of the R_1, R_2 relaxation rates.

Fig. 6 shows ^{1}H-^{15}N correlation spectra recorded during the ^{1}H-^{15}N steady-state NOE measurements. The spectrum in Fig. 6 *A* is obtained without irradiation of the protons and the spectrum in Fig. 6 *B* is the one recorded with proton irradiation of 3 s to reach the steady state. Although the normal spectrum is similar to the typical HSQC spectrum, the spectrum in Fig. 6 *B* shows only four peaks; the peaks have also good intensities, as can be seen from the cross section for the weakest of the peaks. This is somewhat unusual compared to similar studies on other denatured proteins where the HSQC spectra recorded in the presence of proton saturation have good peak intensities and a distribution of positive and negative NOEs has been seen. As a reference, we show in Fig. 6 *C* the spectrum on folded SUMO-1 (pH 7.4, 27°C), which was recorded with proton irradiation and the same protein concentration as in *A* and *B*. This spectrum has clearly good intensities. Thus the present observation of only four discernible peaks in spectrum *B* in Fig. 6 must have dynamic implications. Firstly, two of the four peaks have greater intensity than the corresponding ones in spectrum *A*, and this can happen only if the ^{1}H-^{15}N NOEs are negative. All the peaks in spectrum *B* have the same phase and therefore the other two peaks also must be negative in sign. As marked, these peaks belong to the C-terminal of the protein, and the negative sign indicates that this segment is highly mobile, exhibiting high amplitude picosecond timescale motions. This, however, is not surprising, since the terminal residues generally have been seen to exhibit a greater flexibility than the other residues. More importantly, the fact that no other peak is seen in spectrum *B*, suggests that the NOEs are very small, perhaps below the noise level and a simple explanation for this could be that the rest of the entire

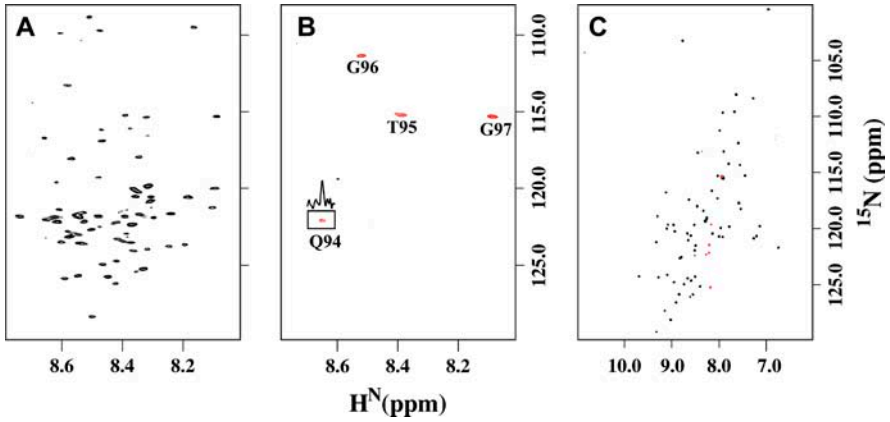

FIGURE 6 Spectra showing ^{15}N-^{1}H heteronuclear NOE of SUMO-1 in 8 M urea at pH 5.6 and 27°C. (*A*) Reference spectrum without irradiation of protons; (*B*) spectrum with irradiation of protons for 3 s; and (*C*) spectrum of the folded protein (pH 7.4, 27°C) with proton irradiation as in *B*. In *B*, a cross section through the peak for Q94 (Q94 is enclosed in a box) is shown to demonstrate the good sensitivity in the spectrum. The red and black contours indicate negative and positive peaks.

chain is relatively more restricted, i.e., the amplitudes of the high frequency motions are small. And, the correlation times for the individual N-H vectors (in the nanosecond regime) may also be such that the steady-state NOEs are nearly zero.

Fig. 7 *A* shows residue-wise R_1 values in the SUMO-1 protein at 27°C. The average relaxation rate is $\sim$1.69 s^{-1}, which is typically the value one sees for this size protein. Interestingly, however, we see some sequence-dependent variations along the chain, some of which seem to coincide with the expected motional behaviors in the native state. For example, the segments/residues (residues 29–31, 40, 60) belonging to the loops between the native secondary structures (β1: 22–28; β2: 33–39; β3: 62–66; β4: 69–72; β5: 87–93; α1: 45–55; and α2: 77–82), which should be more flexible than other regions, have indeed significantly larger than average R_1 values, indicating larger nanosecond-timescale motions; these residues are identified by solid circles in the figure. However, this is not true for other flexible regions of the native protein, especially at the long N-terminal segment. These seem to suggest that just as one can have native- and nonnative-type structural preferences, there can also be native and nonnative type motional characteristics in the denatured state.

The R_2 data shown in Fig. 7 *B* also displays interesting sequence-wise variations. In the denatured states, almost every residue would have some contribution from conformational exchange (R_{ex}) to its transverse relaxation rate,

$$R_2 = R_{2,int} + R_{ex},$$

where $R_{2,int}$ is the intrinsic transverse relaxation rate. However, there can be subtle variations in these due to possibilities of differences in short-range interactions because of sequence effects and consequent different local structural preferences. Thus deviations from the average values, though small as they would be, can be interpreted to suggest different degrees of conformational exchange contributions to the transverse relaxation rates. In Fig. 7 *B*, among the many residues/short stretches that have higher than average R_2 (3.66 s^{-1}) values, we see that the stretches, 19–26 (average R_2 is 4.17 s^{-1}), and 89–92 (average R_2 is 4.12 s^{-1}), are at least four residues long and coincide with the regions of structural propensity derived from the secondary chemical shifts (this data from Fig. 3 *A* is included in the figure here for ready comparison). Hence, these regions may be considered to execute slow segmental motions. The segment 19–26 overlaps with the native β1 strand and the segment 89–92 overlaps with the native β5 strand. Furthermore, the loop 29–31 between the β1 and β2 strands that exhibited high R_1 value also shows high R_2 values. This indicates, firstly, that the slow motions representing conformational exchange and the nanosecond-scale motions of the N-H vectors are independent. Secondly, since the β1 and β2 strands interact in the native structure to form the antiparallel β-sheet, we are tempted to think that the slow motions in this region help to facilitate such an interaction in the denatured state as well. However, as both the strands have some residues with PP_{II} propensities, the transient structure formation may not be entirely native-like.

Among the residues that have lower than average R_2 values signifying greater flexibilities, special mention must be made of K37 and S61, which are in the interior of the chain and have conspicuously low R_2 ($\sim$2 s^{-1}) values. Such observations have been made in the past in at least two cases, namely, apomyoglobin (36) and HIV-1 protease (2) and they have been taken to indicate molecular hinges along the polypeptide chain. Interestingly, in SUMO-1 these residues are seen to be near the edges of those areas that have exhibited higher β-propensities along the chain, consolidating their role as molecular hinges. Another interesting feature noticeable in the R_2 data in Fig. 7 *B* is that the N-terminal and the C-terminal residues have distinctly different motional behaviors, the C-terminal residues having relatively greater flexibility.

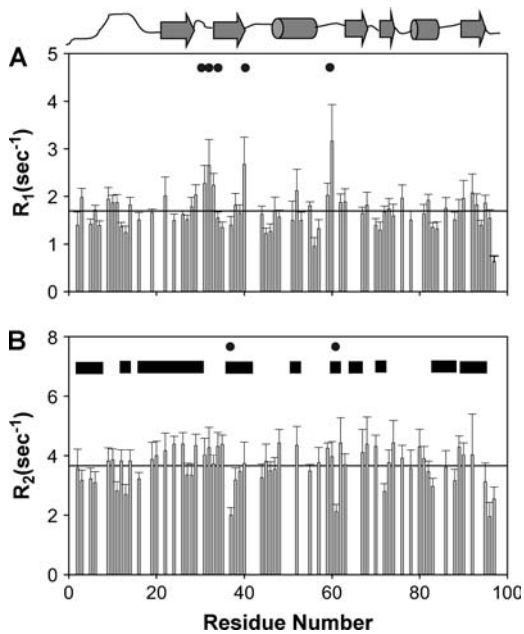

FIGURE 7 ^{15}N relaxation rates plotted against sequence for SUMO-1 in 8 M urea at pH 5.6 and 27°C. (*A*) Longitudinal (R_1) and (*B*) transverse (R_2) relaxation rates. The structural propensities seen in Fig. 3 *A* are also indicated in *B* with solid bars for ready reference. Solid circles in *A* and *B* identify residues with high flexibility.

Implications to protein folding

As pointed out by several authors, the characteristics of the denatured state have a significant influence in directing the folding process of a protein (79–83). A denatured state under some conditions may correspond to an intermediate along the refolding pathway of a protein (1). The residual structural propensities in the denatured state are suggestive of initial

folding events along the chain and its dynamic characteristics throw further light on these events. There may be local and nonlocal interactions and therefore, native and nonnative propensities as well. Although the local interactions essentially define intrinsic preferences, the nonlocal interactions would define the topologies. The existence and significance of nonlocal preferences under strong denaturing conditions, however, remains a controversy (41,50). For those proteins, where there is an initial collapse of the polypeptide chain when refolding is initiated by dilution of denaturant, the collapse is likely to be a consequence of transient interactions between hydrophobic clusters that persist in the denatured state (33,34,36). In the case of barnase, which exhibits a fluctuating secondary structure in the denatured state, it is suggested that folding is initiated around the native-like local structures (38). In β-lactoglobulin, it was observed that the early folding intermediate contained a nonnative helix and this was attributed to the intrinsic high helical preference of the amino acid residues involved in the denatured state (40,84).

In SUMO-1 we observe that the protein chain in the urea-denatured state has (ϕ,ψ) preferences in the β-domain of the Ramachandran map in many segments. Several of these are native-type preferences. Many residues exhibit (ϕ,ψ) preferences corresponding to the PP_{II} helix. However, there is no indication of stable secondary structure in the denatured ensemble. From the dynamics data it appears that in the urea-denatured state the protein chain is not randomly fluctuating, but exhibits sequence dependent motional variations. Interestingly, the segment (residues 29–31) exhibiting larger nanosecond timescale motions coincides with the flexible loop between β1 and β2 strands in the native structure. Although this is not true for all the flexible regions of the native structure, it seems to suggest the existence of some native-type motional signature in the denatured state. The structural and dynamics results provide extremely valuable clues to the initial folding events in the protein. Although the structural preferences indicate possible nucleation sites for the folding process, the relaxation data indicate how motions can facilitate interactions between segments in directing the folding process. Specifically, the slow conformational exchange exhibited by the loop between β1 and β2 strands may promote an interaction between the two strands; these are involved in the formation of an antiparallel β-sheet in the native state. However, few residues in each of these strands have PP_{II} propensities and thus the resultant interaction may not be entirely native-like in the denatured state. Further, residues K37 and S61, which exhibit greater flexibilities, are seen to flank regions of higher structural propensities. Thus, they may serve as molecular hinges for motions of the neighboring segments that may be transiently ordered in the ensemble. All these lead to an implication that, possibly, the conformational transitions and the hinged motions facilitate short range structure formations, though transiently, in the denatured state, and thus may serve as early folding events. Parallel folding events from other regions such as β3, β5, which also exhibit structural propensities and slow conformational transitions, are also conceivable.

SUMO-1 belongs to the ubiquitin family of proteins, and shares the same fold topology and the secondary structural architecture, $\beta\beta\alpha\beta\beta\alpha\beta$. Thus, it would be interesting to compare the above results with the folding studies on fragments of ubiquitin reported in the recent past (85–87). It was observed that the peptide, 1–17, representing the β1-loop-β2 hairpin in ubiquitin, formed a stable nativelike β-hairpin, which suggested that the peptide segment has an intrinsic tendency to form such a structure (86). On this basis, the authors proposed that β1-loop-β2 hairpin formation may be an initial folding event in the folding mechanism of ubiquitin. Interestingly, the structural and dynamic preferences in the denatured state of full-length SUMO-1 protein described in this work also indicate the above segment as the folding initiation site. On the other hand, the experiments with the C-terminal peptide (residues 36–76) of ubiquitin exhibited nonnative α-helical preferences (87). In SUMO-1 we found some nonnative structural preferences in the broad β-domain in the N-terminal (residues 2–6, 17–21), and in the center (residues 48–53) along the protein chain. Thus it appears that the initial folding events in both the proteins have a combination of native- and nonnative-like events, but the finer details have some events in common as well as some differences. This may not be too surprising considering that the two proteins have only 18% sequence homology and consequently the interactions directing the folding process could have differences.

CONCLUSIONS

In conclusion, we have successfully investigated here, using a variety of NMR probes, the nature of the urea-denatured state of SUMO-1, an important protein involved in post-translational modifications. This became possible because of the availability of new experimental protocols recently developed by us for obtaining resonance assignments in unfolded proteins where the chemical shift dispersion of amide protons and carbons is poor. Our data indicates that the protein in the urea-denatured state exhibits structural preferences in the broad β-domain of the Ramachandran map, over the major length of the chain. Several residues populate the (ϕ,ψ) space corresponding to the PP_{II} helix. The polypeptide chain has significant motional restrictions and these show sequence-dependent variations. As the structural propensities and the motional characteristics in the denatured state have a major influence on the folding pathways of proteins, the present observations suggest that, in SUMO-1, folding from the urea-denatured state may get initiated around the β1-loop–β2-region along the polypeptide chain. This region has not only a contiguous stretch of β-propensities, but also contains many residues exhibiting significant conformational transitions. A few residues at the edges of these segments

with structural propensities exhibit high flexibility and thus may contribute to what one may call hinge motions facilitating segment movements and transient structure formations. Early folding events in a few other regions such as $\beta 3$, $\beta 5$ strands are also conceivable, which would represent parallel folding processes.

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

We thank the Government of India for providing financial support to the National Facility for High Field NMR at the Tata Institute of Fundamental Research.

REFERENCES

1. Religa, T. L., J. S. Markson, U. Mayor, S. M. Freund, and A. R. Fersht. 2005. Solution structure of a protein denatured state and folding intermediate. *Nature*. 437:1053–1056.
2. Bhavesh, N. S., R. Sinha, P. M. Mohan, and R. V. Hosur. 2003. NMR elucidation of early folding hierarchy in HIV-1 protease. *J. Biol. Chem.* 278:19980–19985.
3. Plotkin, S. S., and J. N. Onuchic. 2002. Understanding protein folding with energy landscape theory. Part I: Basic concepts. *Q. Rev. Biophys.* 35:111–167.
4. Plotkin, S. S., and J. N. Onuchic. 2002. Understanding protein folding with energy landscape theory. Part II: Quantitative aspects. *Q. Rev. Biophys.* 35:205–286.
5. Chan, H. S., and K. A. Dill. 1998. Protein folding in the landscape perspective: chevron plots and non-Arrhenius kinetics. *Proteins.* 30: 2–33.
6. Dill, K. A., and H. S. Chan. 1997. From Levinthal to pathways to funnels. *Nat. Struct. Biol.* 4:10–19.
7. Onuchic, J. N., H. Nymeyer, A. E. Garcia, J. Chahine, and N. D. Socci. 2000. The energy landscape theory of protein folding: insights into folding mechanisms and scenarios. *Adv. Protein Chem.* 53:87–152.
8. Bhavesh, N. S., S. C. Panchal, R. Mittal, and R. V. Hosur. 2001. NMR identification of local structural preferences in HIV-1 protease tethered heterodimer in 6 M guanidine hydrochloride. *FEBS Lett.* 509:218–224.
9. Bhavesh, N. S., J. Juneja, J. B. Udgaonkar, and R. V. Hosur. 2004. Native and nonnative conformational preferences in the urea-unfolded state of barstar. *Protein Sci.* 13:3085–3091.
10. Chatterjee, A., P. Mridula, R. K. Mishra, R. Mittal, and R. V. Hosur. 2005. Folding regulates autoprocessing of HIV-1 protease precursor. *J. Biol. Chem.* 280:11369–11378.
11. Schwalbe, H., K. M. Fiebig, M. Buck, J. A. Jones, S. B. Grimshaw, A. Spencer, S. J. Glaser, L. J. Smith, and C. M. Dobson. 1997. Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry.* 36:8977–8991.
12. Shortle, D., and M. S. Ackerman. 2001. Persistence of native-like topology in a denatured protein in 8 M urea. *Science.* 293:487–489.
13. Griko, Y., N. Sreerama, P. Osumi-Davis, R. W. Woody, and A. Y. Woody. 2001. Thermal and urea-induced unfolding in T7 RNA polymerase: calorimetry, circular dichroism and fluorescence study. *Protein Sci.* 10:845–853.
14. Li, Y., F. Picart, and D. P. Raleigh. 2005. Direct characterization of the folded, unfolded and urea-denatured states of the C-terminal domain of the ribosomal protein L9. *J. Mol. Biol.* 349:839–846.
15. Zhang, X., Y. Xu, J. Zhang, J. Wu, and Y. Shi. 2005. Structural and dynamic characterization of the acid-unfolded state of hUBF HMG Box 1 provides clues for the early events in protein folding. *Biochemistry.* 44:8117–8125.
16. Baum, J., C. M. Dobson, P. A. Evans, and C. Hanley. 1989. Characterization of a partly folded protein by NMR methods: studies on the molten globule state of guinea pig α-lactalbumin. *Biochemistry.* 28:7–13.
17. Eliezer, D., J. Yao, H. J. Dyson, and P. E. Wright. 1998. Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Biol.* 5:148–155.
18. Matsuura, H., S. Shimotakahara, C. Sakuma, M. Tashiro, H. Shindo, K. Mochizuki, A. Yamagishi, M. Kojima, and K. Takahashi. 2004. Thermal unfolding of ribonuclease T1 studied by multi-dimensional NMR spectroscopy. *Biol. Chem.* 385:1157–1164.
19. Feio, M. J., J. A. Navarro, M. S. Teixeira, D. Harrison, B. G. Karlsson, and M. A. De la Rosa. 2004. A thermal unfolding study of plastocyanin from the thermophilic cyanobacterium *Phormidium laminosum*. *Biochemistry.* 43:14784–14791.
20. Englander, S. W. 2000. Protein folding intermediates and pathways studied by hydrogen exchange. *Annu. Rev. Biophys. Biomol. Struct.* 29:213–238.
21. Onuchic, J. N., and P. G. Wolynes. 2004. Theory of protein folding. *Curr. Opin. Struct. Biol.* 14:70–75.
22. Otzen, D. E. 2002. Protein unfolding in detergents: effect of micelle structure, ionic strength, pH, and temperature. *Biophys. J.* 83:2219–2230.
23. Kamal, J. K., M. Nazeerunnisa, and D. V. Behere. 2002. Thermal unfolding of soybean peroxidase. Appropriate high denaturant concentrations induce cooperativity allowing the correct measurement of thermodynamic parameters. *J. Biol. Chem.* 277:40717–40721.
24. Kamal, J. K., and D. V. Behere. 2002. Thermal and conformational stability of seed coat soybean peroxidase. *Biochemistry.* 41:9034–9042.
25. Baldwin, R. L. 1995. On-pathway versus off-pathway folding intermediates. *Fold. Des.* 1:R1–R8.
26. Bhuyan, A. K., and J. B. Udgaonkar. 1999. Observation of multistate kinetics during the slow folding and unfolding of barstar. *Biochemistry.* 38:9158–9168.
27. Dunker, A. K., C. J. Brown, J. D. Lawson, L. M. Iakoucheva, and Z. Obradovic. 2002. Intrinsic disorder and protein function. *Biochemistry.* 41:6573–6582.
28. Uversky, V. N. 2002. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* 11:739–756.
29. Pohlschroder, M., K. Dilks, N. J. Hand, and R. W. Rose. 2004. Translocation of proteins across archaeal cytoplasmic membranes. *FEMS Microbiol. Rev.* 28:3.
30. Stefani, M. 2004. Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world. *Biochim. Biophys. Acta.* 1739:5–25.
31. Ross, C. A., and M. A. Poirier. 2004. Protein aggregation and neurodegenerative disease. *Nat. Med.* 10(Suppl):S10–S17.
32. Dobson, C. M. 2001. Protein folding and its links with human disease. *Biochem. Soc. Symp.* 1–26.
33. Hodsdon, M. E., and C. Frieden. 2001. Intestinal fatty acid binding protein: the folding mechanism as determined by NMR studies. *Biochemistry.* 40:732–742.
34. Lietzow, M. A., M. Jamin, H. J. Jane Dyson, and P. E. Wright. 2002. Mapping long-range contacts in a highly unfolded protein. *J. Mol. Biol.* 322:655–662.
35. Neri, D., M. Billeter, G. Wider, and K. Wuthrich. 1992. NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science.* 257:1559–1563.
36. Schwarzinger, S., P. E. Wright, and H. J. Dyson. 2002. Molecular hinges in protein folding: the urea-denatured state of apomyoglobin. *Biochemistry.* 41:12681–12686.
37. Arcus, V. L., S. Vuilleumier, S. M. Freund, M. Bycroft, and A. R. Fersht. 1995. A comparison of the pH, urea, and temperature-denatured

states of barnase by heteronuclear NMR: implications for the initiation of protein folding. *J. Mol. Biol.* 254:305–321.

38. Wong, K. B., J. Clarke, C. J. Bond, J. L. Neira, S. M. Freund, A. R. Fersht, and V. Daggett. 2000. Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding. *J. Mol. Biol.* 296:1257–1282.
39. Serrano, L., A. Matouschek, and A. R. Fersht. 1992. The folding of an enzyme. VI. The folding pathway of barnase: comparison with theoretical models. *J. Mol. Biol.* 224:847–859.
40. Hamada, D., and Y. Goto. 1997. The equilibrium intermediate of β-lactoglobulin with non-native α-helical structure. *J. Mol. Biol.* 269: 479–487.
41. Mohana-Borges, R., N. K. Goto, G. J. Kroon, H. J. Dyson, and P. E. Wright. 2004. Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. *J. Mol. Biol.* 340:1131–1142.
42. Crowhurst, K. A., M. Tollinger, and J. D. Forman-Kay. 2002. Cooperative interactions and a non-native buried Trp in the unfolded state of an SH3 domain. *J. Mol. Biol.* 322:163–178.
43. Crowhurst, K. A., and J. D. Forman-Kay. 2003. Aromatic and methyl NOEs highlight hydrophobic clustering in the unfolded state of an SH3 domain. *Biochemistry.* 42:8687–8695.
44. Mok, K. H., and K. H. Han. 1999. NMR solution conformation of an antitoxic analogue of α-conotoxin GI: identification of a common nicotinic acetylcholine receptor α1-subunit binding surface for small ligands and α-conotoxins. *Biochemistry.* 38:11895–11904.
45. Parrot, I., P. C. Huang, and C. Khosla. 2002. Circular dichroism and nuclear magnetic resonance spectroscopic analysis of immunogenic gluten peptides and their analogs. *J. Biol. Chem.* 277:45572–45578.
46. Tremmel, P., and A. Geyer. 2002. An oligomeric Ser-Pro dipeptide mimetic assuming the polyproline II helix conformation. *J. Am. Chem. Soc.* 124:8548–8549.
47. Kelly, M. A., B. W. Chellgren, A. L. Rucker, J. M. Troutman, M. G. Fried, A. F. Miller, and T. P. Creamer. 2001. Host-guest study of left-handed polyproline II helix formation. *Biochemistry.* 40:14376–14383.
48. Moyna, G., H. J. Williams, R. J. Nachman, and A. I. Scott. 1999. Detection of nascent polyproline II helices in solution by NMR in synthetic insect kinin neuropeptide mimics containing the X-Pro-Pro-X motif. *J. Pept. Res.* 53:294–301.
49. Shi, Z., R. W. Woody, and N. R. Kallenbach. 2002. Is polyproline II a major backbone conformation in unfolded proteins? *Adv. Protein Chem.* 62:163–240.
50. Jha, A. K., A. Colubri, K. F. Freed, and T. R. Sosnick. 2005. Statistical coil model of the unfolded state: resolving the reconciliation problem. *Proc. Natl. Acad. Sci. USA.* 102:13099–13104.
51. Mahajan, R., C. Delphin, T. Guan, L. Gerace, and F. Melchior. 1997. A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell.* 88:97–107.
52. Bayer, P., A. Arndt, S. Metzger, R. Mahajan, F. Melchior, R. Jaenicke, and J. Becker. 1998. Structure determination of the small ubiquitin-related modifier SUMO-1. *J. Mol. Biol.* 280:275–286.
53. Jin, C., T. Shiyanova, Z. Shen, and X. Liao. 2001. Heteronuclear nuclear magnetic resonance assignments, structure and dynamics of SUMO-1, a human ubiquitin-like protein. *Int. J. Biol. Macromol.* 28: 227–234.
54. Mishra, R. K., S. S. Jatiani, A. Kumar, V. R. Simhadri, R. V. Hosur, and R. Mittal. 2004. Dynamin interacts with members of the sumoylation machinery. *J. Biol. Chem.* 279:31445–31454.
55. Bhavesh, N. S., S. C. Panchal, and R. V. Hosur. 2001. An efficient high-throughput resonance assignment procedure for structural genomics and protein folding research by NMR. *Biochemistry.* 40: 14727–14735.
56. Chatterjee, A., N. S. Bhavesh, S. C. Panchal, and R. V. Hosur. 2002. A novel protocol based on HN(C)N for rapid resonance assignment in (^{15}N, ^{13}C) labeled proteins: implications to structural genomics. *Biochem. Biophys. Res. Commun.* 293:427–432.
57. Panchal, S. C., N. S. Bhavesh, and R. V. Hosur. 2001. Improved 3D triple resonance experiments, HNN and HN(C)N, for H^N and ^{15}N sequential correlations in (^{13}C, ^{15}N) labeled proteins: application to unfolded proteins. *J. Biomol. NMR.* 20:135–147.
58. Permi, P., and A. Annila. 2004. Coherence transfer in proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* 44:97–137.
59. Tugarinov, V., P. M. Hwang, and L. E. Kay. 2004. Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annu. Rev. Biochem.* 73:107–146.
60. Farrow, N. A., R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, and L. E. Kay. 1994. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry.* 33:5984–6003.
61. Dyson, H. J., and P. E. Wright. 2004. Unfolded proteins and protein folding studied by NMR. *Chem. Rev.* 104:3607–3622.
62. Barbar, E. 1999. NMR characterization of partially folded and unfolded conformational ensembles of proteins. *Biopolymers.* 51:191–207.
63. Dyson, H. J., and P. E. Wright. 2001. Nuclear magnetic resonance methods for elucidation of structure and dynamics in disordered states. *Methods Enzymol.* 339:258–270.
64. Dyson, H. J., and P. E. Wright. 2002. Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. *Adv. Protein Chem.* 62:311–340.
65. Bhavesh, N. S., and R. V. Hosur. 2004. Exploring unstructured proteins. *Proc. Indian Natl. Sci. Acad.* 70A:579–596.
66. Chatterjee, A., A. Kumar, J. Chugh, S. Srivastava, N. S. Bhavesh, and R. V. Hosur. 2005. NMR of unfolded proteins. *J. Chem. Sci. (Indian Acad. Sci.).* 117:3–21.
67. Wishart, D. S., C. G. Bigam, A. Holm, R. S. Hodges, and B. D. Sykes. 1995. ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR.* 5:67–81.
68. Wishart, D. S., and B. D. Sykes. 1994. Chemical shifts as a tool for structure determination. *Methods Enzymol.* 239:363–392.
69. Schwarzinger, S., G. J. Kroon, T. R. Foss, P. E. Wright, and H. J. Dyson. 2000. Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRView. *J. Biomol. NMR.* 18:43–48.
70. Schwarzinger, S., G. J. Kroon, T. R. Foss, J. Chung, P. E. Wright, and H. J. Dyson. 2001. Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc.* 123:2970–2978.
71. Penkett, C. J., C. Redfield, I. Dodd, J. Hubbard, D. L. McBay, D. E. Mossakowska, R. A. Smith, C. M. Dobson, and L. J. Smith. 1997. NMR analysis of main-chain conformational preferences in an unfolded fibronectin-binding protein. *J. Mol. Biol.* 274:152–159.
72. Baxter, N. J., and M. P. Williamson. 1997. Temperature dependence of ^{1}H chemical shifts in proteins. *J. Biomol. NMR.* 9:359–369.
73. Merutka, G., H. J. Dyson, and P. E. Wright. 1995. ''Random coil'' ^{1}H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. *J. Biomol. NMR.* 5:14–24.
74. Wuthrich, K. 1986. NMR of Proteins and Nucleic Acids. John Wiley & Sons.
75. Kay, L. E. 2005. NMR studies of protein structure and dynamics. *J. Magn. Reson.* 173:193–207.
76. Palmer III, A. G. 1993. Dynamic properties of proteins from NMR spectroscopy. *Curr. Opin. Biotechnol.* 4:385–391.
77. Palmer III, A. G. 1997. Probing molecular motion by NMR. *Curr. Opin. Struct. Biol.* 7:732–737.
78. Palmer III, A. G. 2001. NMR probes of molecular dynamics: overview and comparison with other techniques. *Annu. Rev. Biophys. Biomol. Struct.* 30:129–155.
79. van Gunsteren, W. F., R. Burgi, C. Peter, and X. Daura. 2001. The key to solving the protein-folding problem lies in an accurate description of the denatured state. *Angew. Chem. Int. Ed. Engl.* 40:351–355.

80. Wintrode, P. L., T. Rojsajjakul, R. Vadrevu, C. R. Matthews, and D. L. Smith. 2005. An obligatory intermediate controls the folding of the α-subunit of tryptophan synthase, a TIM barrel protein. *J. Mol. Biol.* 347:911–919.

81. Daggett, V., and A. R. Fersht. 2003. Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.* 28:18–25.

82. Daggett, V., A. Li, L. S. Itzhaki, D. E. Otzen, and A. R. Fersht. 1996. Structure of the transition state for folding of a protein derived from experiment and simulation. *J. Mol. Biol.* 257:430–440.

83. Gianni, S., N. R. Guydosh, F. Khan, T. D. Caldas, U. Mayor, G. W. White, M. L. DeMarco, V. Daggett, and A. R. Fersht. 2003. Unifying features in protein-folding mechanisms. *Proc. Natl. Acad. Sci. USA.* 100:13286–13291.

84. Kuwata, K., R. Shastry, H. Cheng, M. Hoshino, C. A. Batt, Y. Goto, and H. Roder. 2001. Structural and kinetic characterization of early folding events in β-lactoglobulin. *Nat. Struct. Biol.* 8:151–155.

85. Zerella, R., P. A. Evans, J. M. Ionides, L. C. Packman, B. W. Trotter, J. P. Mackay, and D. H. Williams. 1999. Autonomous folding of a peptide corresponding to the N-terminal β-hairpin from ubiquitin. *Protein Sci.* 8:1320–1331.

86. Zerella, R., P. Y. Chen, P. A. Evans, A. Raine, and D. H. Williams. 2000. Structural characterization of a mutant peptide derived from ubiquitin: implications for protein folding. *Protein Sci.* 9:2142–2150.

87. Jourdan, M., and M. S. Searle. 2000. Cooperative assembly of a nativelike ubiquitin structure through peptide fragment complexation: energetics of peptide association and folding. *Biochemistry.* 39:12355–12364.