doi:10.1006/jmbi.2000.3816 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2000) **299**, 1341–1351

JMB

NMR Characterization of Residual Structure in the Denatured State of Protein L

Qian Yi, Michelle L. Scalley-Kim, Eric J. Alm and David Baker*

Department of Biochemistry
University of Washington
Seattle, WA 98195, USA

Triple-resonance NMR experiments were used to assign the $^{13}C^{\alpha}$, $^{13}C^{\beta}$, ^{15}N and NH resonances for all the residues in the denatured state of a destabilized protein L variant in 2 M guanidine. The chemical shifts of most resonances were very close to their random coil values. Significant deviations were observed for G22, L38 and K39; increasing the denaturant concentration shifted the chemical shifts of these residues towards theory random coil values. Medium-range nuclear Overhauser enhancements were detected in segments corresponding to the turn between the first two strands, the end of the second strand through the turn between the second strand and the helix, and the turn between the helix and the third strand in 3D H^1, N^{15}-HSQC-NOESY-HSQC experiments on perdeuterated samples. Longer-range interactions were probed by measuring the paramagnetic relaxation enhancement produced by nitroxide spin labels introduced *via* cysteine residues at five sites around the molecule. Damped oscillations in the magnitude of the paramagnetic relaxation enhancement as a function of distance along the sequence suggested native-like chain reversals in the same three turn regions. The more extensive interactions within the region corresponding to the first β-turn than in the region corresponding to the second β-turn suggests that the asymmetry in the folding reaction evident in previous studies of the protein L folding transition state is already established in the denatured state.

Keywords: protein L; denatured state; residual structure; paramagnetic relaxation enhancement; NMR

**Corresponding author*

Introduction

How a protein folds into a unique three-dimensional structure is one of the greatest questions of modern structural biology. An exciting recent development is the direct structural characterization of the starting point of the folding reaction, the denatured state, using multi-dimensional NMR methods. It has been found (Neri *et al.*, 1992; Logan *et al.*, 1994; Shortle 1996a,b; Eliezer *et al.*, 1998; Fong *et al.*, 1998; Mok *et al.*, 1998) that the unfolded states of several proteins under both denaturing and native solution conditions can contain a significant amount of residual structure. Recent studies under non-denaturing conditions have demonstrated that the denatured state ensemble of staphylococcal nuclease has a native-like overall topology (Gillespie & Shortle, 1997), while the overall topology of the drk SH3 unfolded states is not native-like, although it has some native-like features (Mok *et al.*, 1998). The detailed structural characterization of the denatured states of proteins whose folding transition has been extensively studied should increase understanding of the early stages of the folding process.

We have chosen the IgG binding domain of protein L as a model system for understanding folding in detail. The folding of protein L has been characterized using a wide range of methods (Yi & Baker, 1996; Scalley *et al.*, 1997; Yi *et al.*, 1997; Gu *et al.*, 1997; Plaxco *et al.*, 1999; Kim *et al.*, 1999) and the folding transition state has been extensively mapped through the analysis of the effect of 70 point mutants distributed around the protein. A destabilized mutant (F20W/Y32A) of protein L has recently been characterized by circular dichroism and stopped-flow kinetics (Scalley *et al.*, 1999).

Abbreviations used: PRE, paramagnetic relaxation enhancement; HSQC, heteronuclear single quantum coherence; NOESY, NOE spectroscopy; CD, circular dichroism.

E-mail address of the corresponding author: dabaker@u.washington.edu

0022-2836/00/051341–11 $35.00/0

These data suggested the presence of residual structure in 2 M to 3 M guanidinium chloride. In this study we use NMR methods to gain more specific structural information on this denatured state of protein L.

Results

Assignment of the backbone $^{13}C^{\alpha}$, $^{13}C^{\beta}$, ^{15}N and NH resonances of F20W/Y32A

Characterization of denatured states of proteins using NMR techniques is often challenging, since the chemical shift dispersion of most resonances is poor because of conformational averaging. However, the backbone ^{15}N and ^{13}CO chemical shifts, which are mainly influenced by residue type and the local amino acid sequence (Braun *et al.*, 1994; Yao *et al.*, 1997), remain well dispersed in the denatured state. Using ^{13}C, ^{15}N-double labeled protein sample and triple resonance NMR experiments (see Materials and Methods), we have assigned the $^{13}C^{\alpha}$, $^{13}C^{\beta}$, ^{15}N and NH resonances for all the residues of the F20W/Y32A mutant in 2 M guanidine at pH 5.0.

$^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts are predominantly determined by backbone conformation, and the perturbations of the $^{13}C^{\alpha}$ and $^{13}C_{\beta}$ chemical shifts from their random coil values (Spera & Bax, 1991; Wishart & Skyes, 1994) are reliable indicators of secondary structure in folded proteins. In general, $^{13}C^{\alpha}$ resonances are shifted downfield by an average of 2.6 ppm for α-helices, and shifted upfield by 1.7 ppm for β-sheets. Figure 1 shows that the deviations of the $^{13}C^{\alpha}$ chemical shifts of F20W/Y32A in 2 M guanidine at pH 5.0 from the random coil values are very small, indicating the population of regular secondary structure is very low under these conditions. However, small but consistent upfield chemical shift perturbations were observed for every residue of the segment from A31 to D41 (helical in native protein L structure), suggesting that there is some residual helical content in F20W/Y32A in 2 M guanidine. Also, small, but consistent downfield chemical shift perturbations were observed for every residue from Y45 to A50 (the third strand in native protein L structure), suggesting that this segment has some preference for extended conformations under these conditions.

Previous chemical denaturation studies using fluorescence techniques suggested that there may be some residual structure around W20 in 2-3 M guanidine that is lost at higher (>3 M) guanidine concentrations (Scalley *et al.*, 1999). A guanidine titration ranging from 1.5 M to 5 M guanidine was carried out to investigate possible conformational changes in the denatured state ensemble. There is no dramatic change in the ^{1}H-^{15}N HSQC spectra from 1.5 M to 5 M guanidine except that small, but significant chemical shift perturbations (~0.10 ppm shifting toward the random coil values) were observed for the amide group protons of G22, L38 and K39 (Figure 2(a)). These shifts are roughly linear functions of the denaturant concentration; there is little indication of a cooperative transition between different populations (Figure 2(b)). These results suggest that there is some residual structure around G22, L38 and K39 in denatured F20W/Y32A in 2-3 M guanidine.

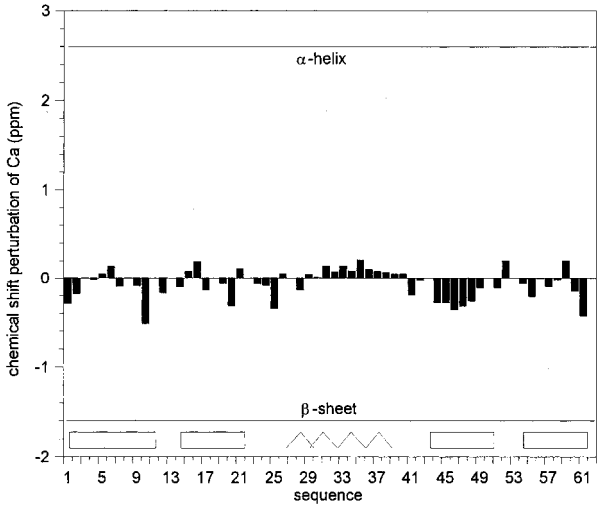

Figure 1. Chemical shift perturbations of the $^{13}C^{\alpha}$ resonances of F20W/Y32A in 2 M guanidine from the random coil values. The two straight horizontal lines at 2.6 ppm and −1.7 ppm represent the perturbations expected for regular α-helix and β-sheet conformations respectively. The random coil values were taken from Wishart & Sykes, 1994.

Nuclear Overhauser enhancements (NOEs) between amide protons

NOEs between amide group protons were obtained using 3D H^{1}, N^{15}-HSQC-NOESY-HSQC experiments on perdeuterated F20W/Y32A in 2-3 M guanidine. Recent work on the drk SH3 domain demonstrated that deuteration greatly enhances NOESY-based studies of denatured proteins because of longer relaxation time due to the reduced spin diffusion (Sattler & Fesik 1996). Sequential $(i, i+1)$ and $(i, i+2)$ HN-HN NOEs were observed for most residues of F20W/Y32A in 2 M guanidine. Only a few medium-range $(i, i+3)$ and $(i, i+4)$ HN-HN NOEs were detected (Figure 3). All observed medium-range NOEs are located in segments close to turn regions in native protein L. These segments correspond to the first hairpin turn, the end of the second β-strand before the helix and the turn following the helix. Thus, there are some native-like turn structures populated in the denatured states of F20W/Y32A in 2-3 M guanidine. This is consistent with the result from guanidine titration experiments described above. NOEs in the region corresponding to the second β-turn in native protein L were conspicuously absent.

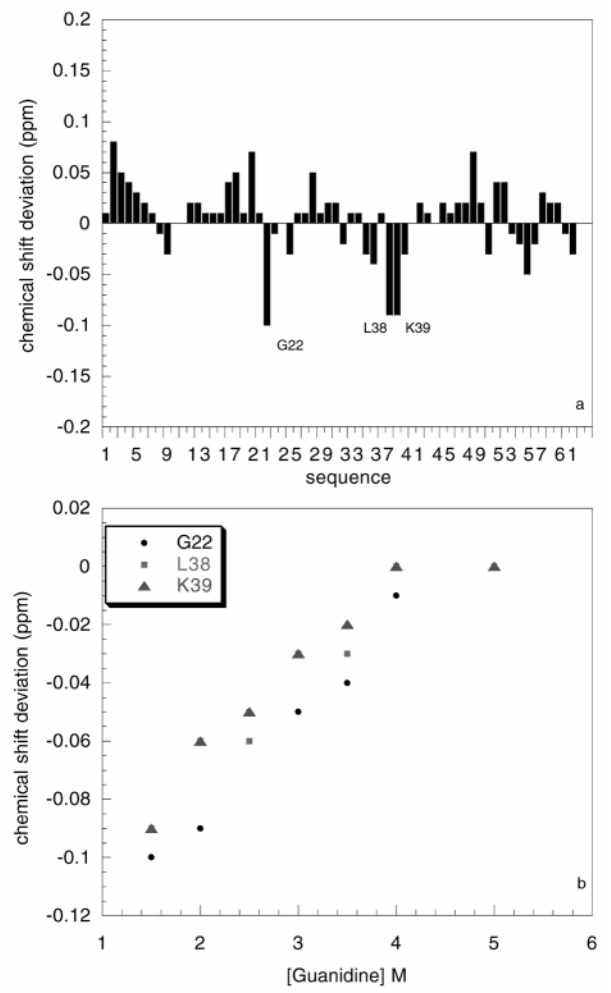

Figure 2. Guanidine-dependence of amide proton chemical shifts. (a) Deviation of chemical shifts in 1.5 M from values in 5.0 M guanidine for all the amide group protons of F20W/Y32A/C63. (b) Guanidine dependence of the amide group proton chemical shifts of G22, L38 and K39.

Measurement of paramagnetic relaxation enhancement (PRE) by nitroxide spin label

To probe longer-range interactions, we examined the paramagnetic relaxation enhancement of the amide group protons due to introduced nitroxide spin labels. This technique has been used successfully to characterize the denatured state of staphylococcal nuclease under native conditions (Gillespie & Shortle, 1997). The advantage of PRE is that the free electron from the nitroxide label increases the relaxation rate of protons over a distance of up to 20 to 25 Å. In contrast, the NOE between two protons only extends up to 10 Å even at very long mixing times (Mok *et al.*, 1998). Thus, PRE can be useful for studying weak long-range molecular interactions in relatively disordered denatured states.

Paramagnetic relaxation enhancement increases relaxation rates in a distance-dependent manner. The enhancement effect is described by the Solomon-Bloembergen equations (Solomon & Bloembergen, 1956; Kosen 1989):

$$\Delta(1/T_1) = \Delta R_1 = 2K(3\tau_c/(1+\omega_H^2\tau_c^2))/r^6 \quad (1)$$

$$\Delta(1/T_2) = \Delta R_2 = K(4\tau_c + 3\tau_c/(1+\omega_H^2\tau_c^2))/r^6 \quad (2)$$

where K is 1.23×10^{-32} cm^6 s^{-2} for a nitroxide radical, r is the distance between the electron and the proton, τ_c is the correlation time for the electron-proton vector, and ω_H is the Larmour frequency of the proton. Equations (1) and (2) are based on the assumptions that the vector between the electron and the proton is free to undergo isotropic rotational diffusion, and that its length is fixed. Both equations are valid only for relaxation due to the magnetic interaction between a single unpaired electron and a proton of a macromolecule when τ_c is greater than 10^{-9} second and ω_H is between 400 and 600 MHz.

To provide attachment sites for the spin labels, cysteine residues were introduced one at a time at strategic locations on the surface of protein L. To minimize perturbations to the structure accompanying the cysteine substitutions, the residues at the chosen positions were highly solvent-accessible and make few interactions with other residues. To obtain information on all parts of the protein, one probe was introduced into each of the five secondary structural elements in the protein. The sites at which cysteine residues were introduced are shown in Figure 4; E1C is at the N terminus, T17C in the middle of the second strand, S29C in the middle of the helix, T46C in the middle of the third strand and C63 was added at the C terminus. Nitroxide spin labels were attached to the introduced cysteine residues as described in Materials and Methods.

Direct measurement of the effect of the paramagnetic relaxation enhancement on T_1 and T_2 can be carried out using standard NMR techniques, such as the inversion-recovery sequence and CPMG spin echo sequence, but these techniques can be quite time-consuming. Instead, it is more efficient to measure the decrease of ^{1}H-^{15}N peak intensity in HSQC spectra. The peak linewidth in the HSQC spectrum is increased due to faster transverse relaxation (R_2) of ^{1}H, and ultimately the peak intensity is decreased. Figure 5 shows the PRE effect on peak intensity by comparing ^{1}H-^{15}N HSQC spectra of T17C* with and without the unpaired free electron present (oxidized and reduced forms respectively). The magnitude of the PRE for each residue in each protein was determined using the simulation method described by Gillespie and co-workers (Gillespie & Shortle, 1997) (see Materials and Methods) and is shown in Figure 6. The bars indicate the magnitude of the PRE, the arrows indicate the position of the label,

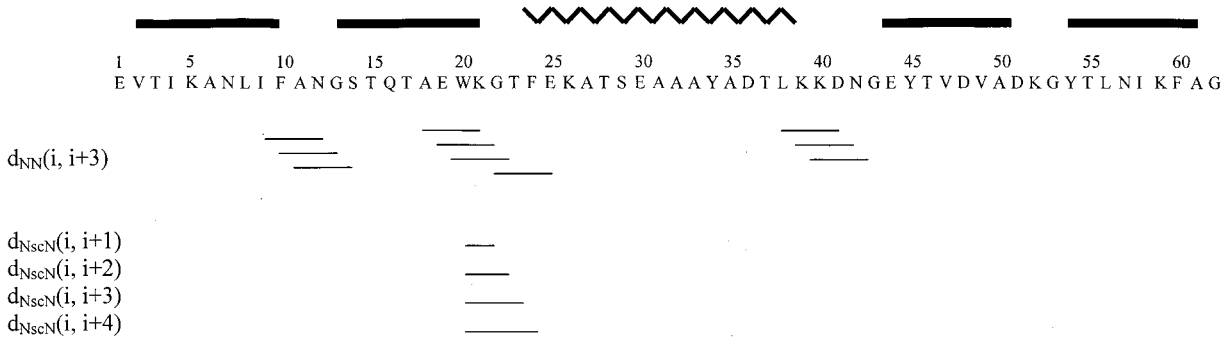

Figure 3. Schematic of the medium-range HN-HN NOEs observed in HSQC-NOESY-HSQC experiments with a 600 ms mixing time along with the protein sequence and the secondary structural elements of the native protein L. The filled bars represent β-strands and the zigzag line represents α-helix. The d_{NscN} NOEs refer to the NOEs observed between the side-chain NεH of W20 and the backbone amide group protons.

and the boxes at the top indicate the positions of the secondary structure elements. For a number of residues (indicated by the hatched bars) close to the introduced labels, the relaxation enhancement was too great to be measured.

To facilitate interpretation of the PRE results, we have compared the experimental PRE profiles with profiles generated by computer simulation for: (i) the native state; (ii) a model of the denatured state of protein L in which local sequence-structure relationships are preserved; and (iii) a random chain model of the protein L denatured state (see Materials and Methods). The PRE profile expected for an ensemble of configurations without any persistent structure should smoothly decrease with increasing sequence distance from the introduced probe. Such a smooth decrease is clearly evident in the random chain simulations (Figure 7, row D). The experimental PRE profiles in all five cysteine mutants (Figure 7, row B) show the expected decrease in relaxation enhancement with distance from the spin label (indicated by the arrows in Figure 6). Superimposed on the gradual decay of relaxation enhancement effect are oscillations not seen in the random chain model that suggest the presence of chain reversals. The differences between the profiles in rows B and D in Figure 7 are likely to be due to residual structure in the protein L denatured state. For example, for the E1C* mutant, residues from N12 to T15, which are closer to the nitroxide probe in the sequence, experienced less broadening than residues from A18 to F24, which are further in sequence from the labeled site. In the S29C* sample, residues from N12 to T15 displayed less broadening than residues from K5 to A11. The PRE effect for residues from T25 to D41 in the profiles of T17C and S29C displays an oscillating pattern, suggesting that helix or turn-like residual structures may be present in the region from T25 to D41, which is helical in the native state of protein L. This is consistent with the chemical shift perturbation in the region noted above. Comparison of the peaks in the experimental profile in Figure 7 row B to the profile expected for the native structure (Figure 7, row A) provides an indication of the extent to which the residual structure in the denatured state is native-like. The peaks in the region corresponding to the first β-hairpin (near residue 21 in E1C, residue 7 in S29 and T46) in the experimental profiles mirror peaks in similar locations in the profiles derived from the native structure (row A). In contrast, the peak in the region corresponding to the second β-hairpin in the C63 derivative (near residue 51) has no counterpart in the native profile. These results suggest that native-like chain reversals are sampled in the region corresponding to the first β-hairpin in the denatured state, but in the region corresponding to the second β-hairpin there is a chain reversal, not in the β-turn, but near the middle of the last strand. Comparison of the experimental profiles in row B to those for the simulated denatured state model with the local sequence-structure propensities of the protein L sequence (Figure 7, row C) provides some insight into the origin of the residual structure in the protein L denatured state. A peak observed in the experimental PRE profile in the vicinity of the first hairpin in the E1C derivative is also observed in the sequence-specific

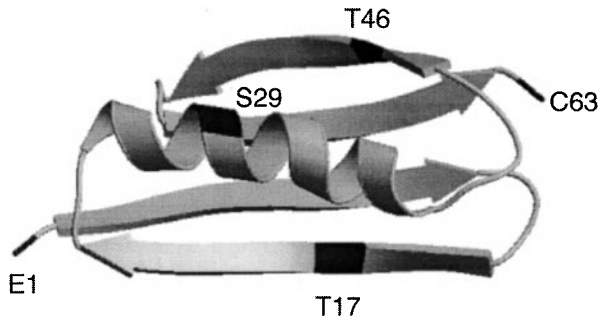

Figure 4. Ribbon diagram of the protein L structure. The positions where the spin labels were introduced are highlighted in black.

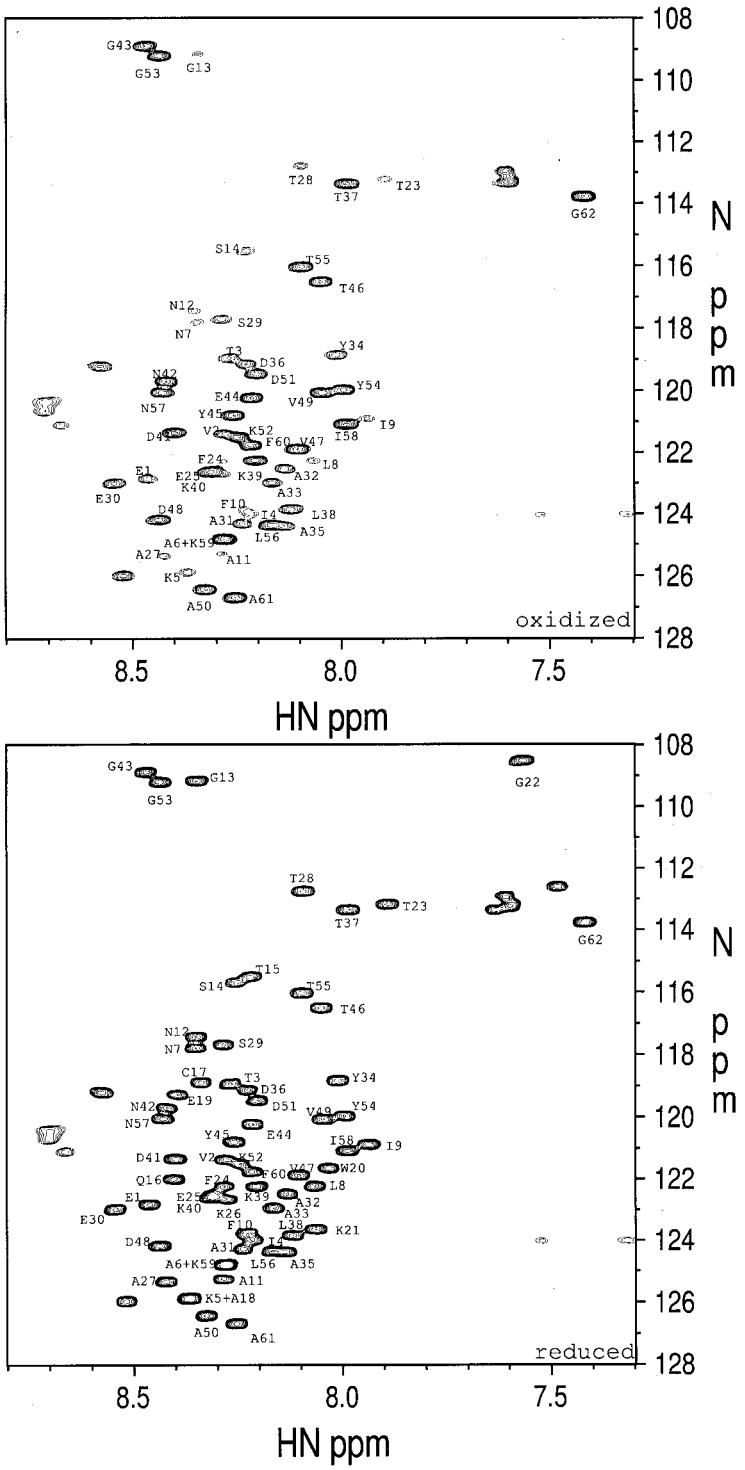

Figure 5. ^{1}H, ^{15}N-HSQC spectra of nitroxide-labeled T17C/F20W/Y32A mutant in 50 mM sodium phosphate and 2.2 M guanidine at pH 5.0 and 22 °C. (a) Oxidized form; (b) reduced form.

denatured state model simulations (row C), suggesting that the chain reversal in the vicinity of the first β-turn is due, at least in part, to local sequence propensities. Interestingly, the non-native peak observed around residue 51 in the experimental profile for the C63 derivative is also observed in the sequence-specific denatured state model, suggesting that this non-native feature is also due, in part, to local sequence propensities. However, much of the residual structure suggested by the experimental PRE profiles is likely to result from interactions longer in range than those captured by the denatured state model, as the oscillations more distant from the site of labeling in the experimental profiles in row B are mostly absent from the simulated profiles in row C.

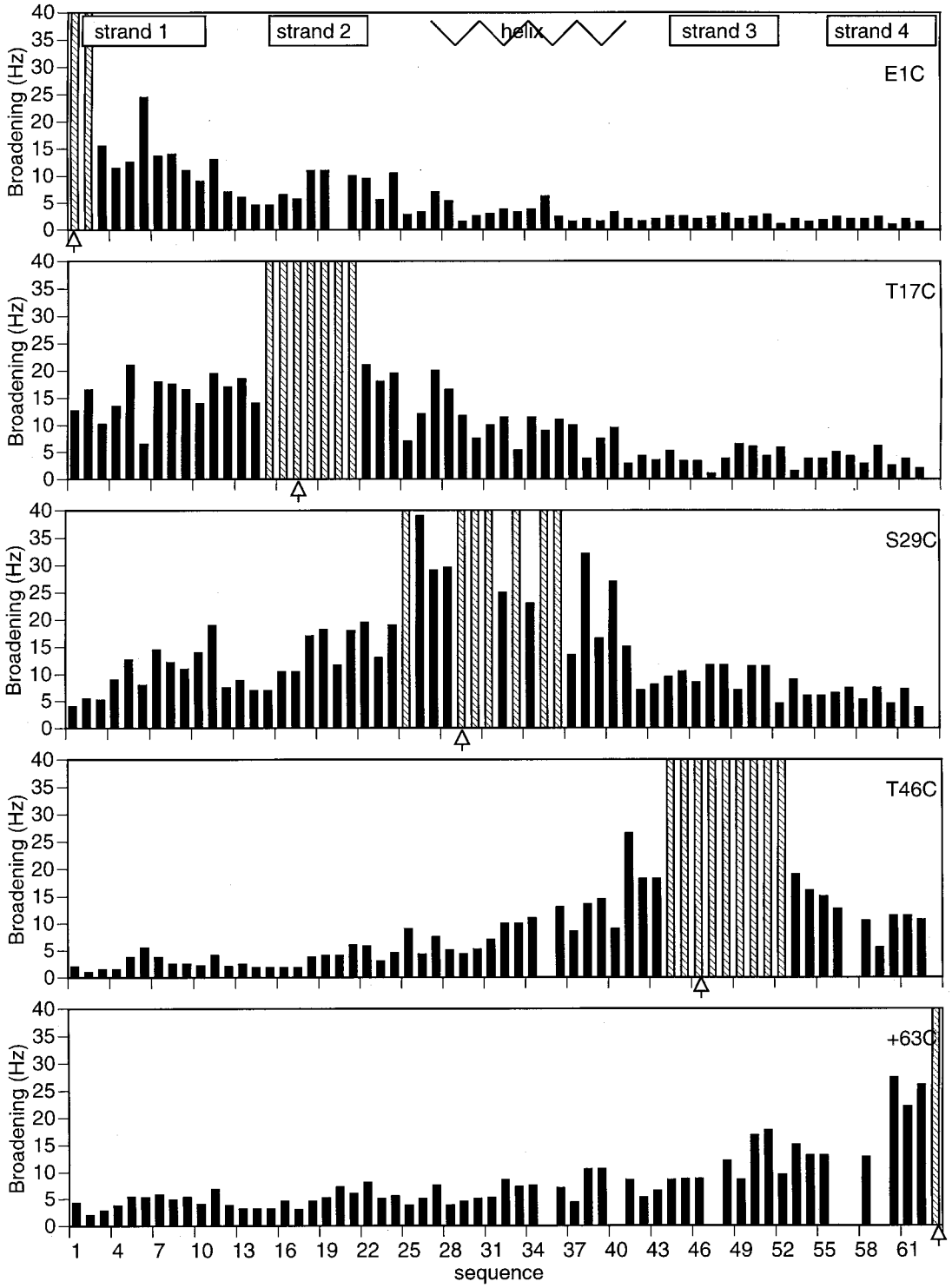

Figure 6. Paramagnetic relaxation enhancement of amide group proton resonances by the introduced nitroxide spin labels. The sites of labeling are indicated by the open triangles on the horizontal axes, and also labeled in the upper right corner of each profile. The hatched bars represent relaxation enhancement effects beyond the experimentally measurable limit. The secondary structure of native protein L is schematically represented on the top of the Figure.

Discussion

The NMR data presented here provide a picture of the conformations sampled in the denatured state of the F20W/Y32A mutant in 2 M guanidine. Almost all residues have chemical shifts close to their random coil values, and no long-range NOEs were observed even on perdeuterated samples, indicating considerable conformational averaging and little long-range order. The chemical shift data,

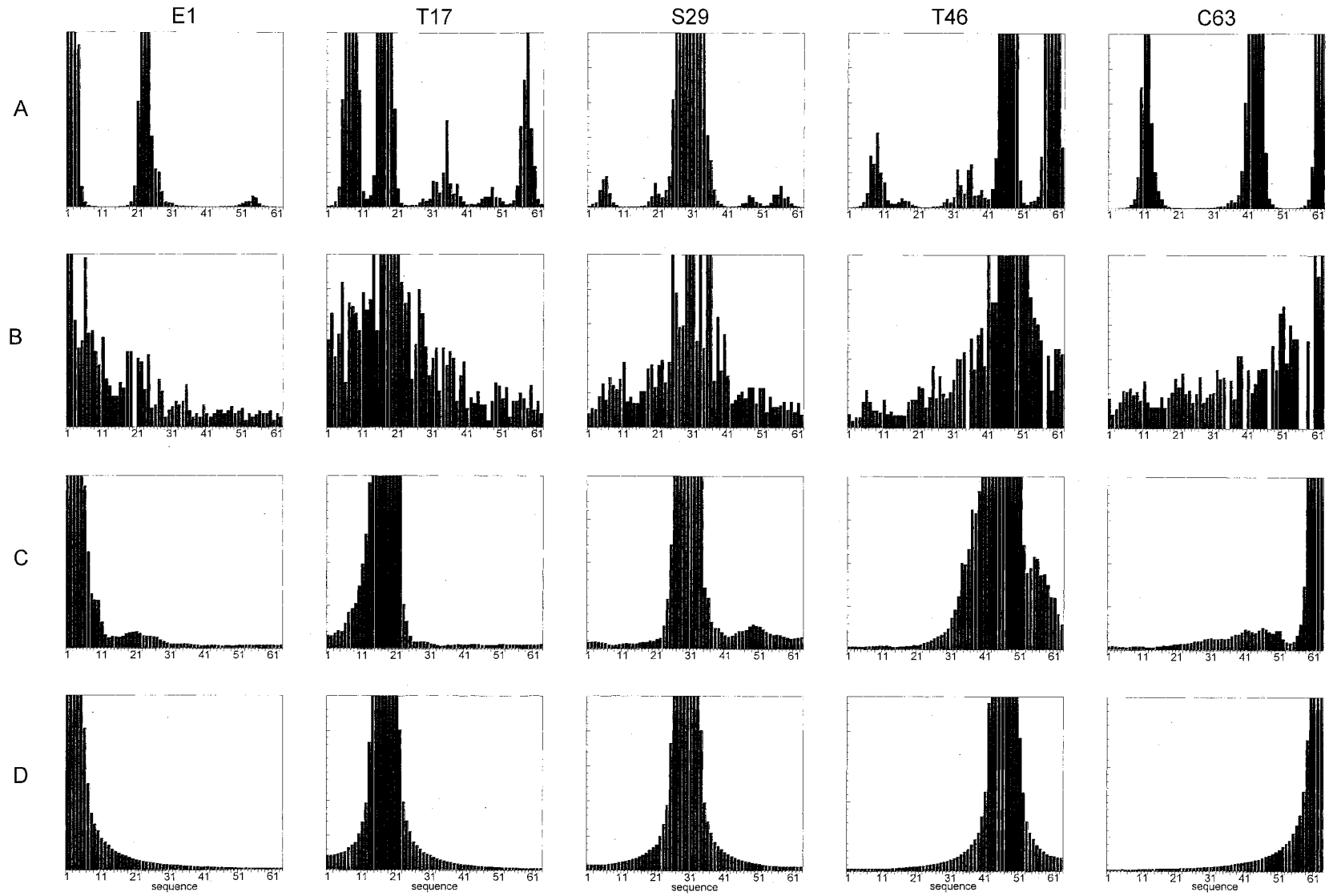

Figure 7. Comparison of experimental and simulated PRE profiles. Row A, simulated PRE profiles for the native protein L structure. Row B, experimental PRE profiles. Row C, simulated PRE profiles for denatured state model with protein L specific local sequence-structure biases. Row D, simulated PRE profiles for generic denatured state model without local sequence-structure biases. The simulation methods used in rows A, C, and D are described in Materials and Methods.

the observed medium-range NOEs, and the oscillations in the PRE profiles suggest that the greatest deviations from random chain behavior are in the N-terminal portion of the protein, and that these involve primarily native-like turns and chain reversals. The residues whose chemical shifts change the most with increasing guanidine concentration are just before and just after the central helix (G22, L38 and K39). The observed medium-range NOEs were in regions corresponding to the turn between the first two β-strands (I9 to G13), the turn between the second strand and the helix (A18 to E25), and the turn following the helix (L38 to G43). The most pronounced minima in the PRE profiles were observed near the location of these turns in the native state, with the largest effect in the first β-turn (the dip between N12 and T15 is clearly evident in both the E1C and the S29 labeled proteins (Figure 6)). The absence of medium-range NOEs suggests that the second β-turn is significantly less populated than the other turns in the denatured state. The second β-turn containing three consecutive residues with positive φ angles, only one of which is a glycine residue (Wikstrom *et al.*, 1994). The turn is thus likely to be under considerable strain in the native state and, perhaps as a consequence of this, the turn appears to be largely disrupted in the rate-limiting step in the unfolding of protein L (Gu *et al.*, 1997). The absence of detectable NOEs in the second β-turn in the denatured state could also be due to such strain, and formation of the turn during folding may be driven by long-range interactions not present in the denatured state.

The NMR studies described here are consistent with the suggestion from earlier circular dichronism (CD) and fluorescence studies of non-random residual structure in the denatured state in the region corresponding to the first hairpin and the helix (Scalley *et al.*, 1999). The denaturant dependence of the fluorescence of W20 was found to be quite different in a ten-residue peptide derived from the protein L sequence (centered on W20) from that in the denatured protein, suggesting some residual structure around W20 in the denatured protein. Consistent with this, we find that the highest density of medium-range NOEs in the denatured protein is around W20 (Figure 4). The earlier CD studies suggested some residual helix content in the denatured protein, and this is not inconsistent with the oscillations in the PRE profiles between T25 and D41 (this region is helical in the native state). Dead-time labeling HD exchange experiments on the denatured state immediately after initiation of refolding in the absence of denaturant suggest that the asymmetry observed in 2 M guanidine is also present in the absence of denaturant: the greatest protection from exchange was in the first β-hairpin.

A study of denatured proteins in urea and guanidine solutions also found little long-range order. A study of 434-repressor in 6 M urea (Neri *et al.*, 1992), revealed a native-like local hydrophobic cluster, with the remainder of the protein largely disordered. In a study of the urea and guanidine-denatured FK506 binding protein (Logan *et al.*, 1994), significant populations of both native-like and non-native-like residual local structures, mainly turns and local helical contents, were detected. More residual structure has been observed in denatured states of a truncated version of staphylococcal nuclease (Gillespie & Shortle, 1997) and the drk SH3 domain (Mok *et al.*, 1998) in the absence of denaturant. In the study of the unfolded drk SH3 domain, most of the observed long-range interactions disappeared upon the addition of 2 M guanidine, but there were still a few long-range NOEs detected in 2 M guanidine (Mok *et al.*, 1998). NMR characterization of the denatured states of drk and staphylococcal nuclease in the absence of denaturant was made possible by their relatively high level of solubility; unfortunately, we have not been able to identify unfolded protein L mutants which are sufficiently soluble in the absence of denaturant for NMR studies. The origins of the differences in solubility of the denatured states of different proteins are not entirely clear; the high level of solubility of the staphylococcal nuclease denatured state may, in part, be due to electrostatic repulsion between monomers resulting from the high net charge on the protein.

An issue of considerable current interest is how residual structure in the unfolded state contributes to the overall folding reaction. On the one hand, native-like interactions in the denatured states can limit the conformational search space during folding and, on the other, non-native like interactions could create energy barriers that hamper protein folding. For example, in the case of the drk SH3 domain, non-native interactions in the unfolded state destabilize the native state and may slow the folding process. Extensive characterization of the effects of mutations on protein L folding (Kim *et al.*, 1999) has suggested that the first β-turn is largely formed, and the second β-turn, largely disrupted in the folding transition state ensemble. Here, we observe medium-range NOEs and a significant dip in the PRE profile in the region corresponding to the first β-turn, but not the second, consistent with the asymmetry of structure in the folding transition state. Thus, the interactions formed in the protein L denatured state do not appear to disfavor subsequent folding events, and the asymmetry in protein L folding appears quite early in the folding process. The rate-limiting step in folding may involve the entropically costly consolidation and/or association of parts of the protein which are partially ordered in the denatured state.

Materials and Methods

Sample preparation

^{15}N-labeled and ^{15}N, ^{13}C-labeled protein samples were made by growing the transformed *Escherichia coli* cells in

Mops minimal medium. For the labeling, 99.9 % (w/w) $^{15}NH_4Cl$ was used as the nitrogen source and 99.9 % (w/w) ^{13}C-glucose as the carbon source in our media. For ^{15}N, ^{2}H-labeled F20W/Y32A, a single colony of *E. coli* of BL21 (DE3/plysS) carrying the F20W/Y32A plasmid was inoculated into 100 ml of M9 H_2O media with 50 μg/ml carbenicillin, and grown at 37 °C until the absorbance at 600 nm was approximately 0.6. The cells were harvested by centrifuging the culture with an ss-34 rotor at 4000 rpm for ten minutes. The harvested cells were resuspended with 10 ml of M9 $^{2}H_2O$ medium and then transferred into 2 l of M9 $^{2}H_2O$ medium with 50 μg/ml carbenicillin. The cells were grown at 37 °C to an absorbance at 600 nm of approximately 0.6, then induced with 1 mM IPTG for ten hours before harvesting. All labeled F20W/Y32A proteins were purified by His-tag affinity column as described (Gu *et al*., 1995). For the ^{15}N, ^{2}H-labeled F20W/Y32A, the level of deuteration was estimated to greater than 95 % based on the result from mass spectrometry.

Nitroxide labeling

Nitroxide groups can be introduced into proteins through alkylation of the thiolate group of cysteine residue. Since protein L does not contain any cysteine residues, site-directed mutagenesis was used to substitute selected amino acid residues with cysteine. Five charged or polar residues with highly solvent-exposed sidechains were selected for cysteine mutation to minimize possible conformational perturbations of mutagenesis. All the mutants, E1C, T17C, S29C, T46C and +63C (addition of a Cys residue to the C terminus), were spin labeled as described (Mchaourab *et al*., 1996). The extent of labeling was examined by MALDI and in all cases it was greater than 95 %.

NMR experiments and data processing

Both HNCACB (Wittekind & Mueller, 1993) and CBCACONH (Grzesiek & Bax, 1992, 1993) triple resonance experiments were carried out on a Bruker DMX500 instrument with 1.5 mM ^{15}N, ^{13}C-labeled F20W/Y32A in 10 % $^{2}H_2O$/90 % H_2O 50 mM sodium phosphate and 2.2 M guanidine at pH 5.0 and 22 °C. Matrices of $40 \times 40 \times 512$ complex points were acquired with spectral widths of 7002.8, 2000.0, and 4496.4 Hz (F1, F2, F3) for both HNCACB and CBCACONH. For both ^{13}C and ^{15}N dimensions (F1, F2) of these spectra, the sizes of the time domain were doubled *via* forward-back linear prediction (Zhu & Bax, 1992). The data were zero-filled and extracted (only 4.70 ppm-9.00 ppm of the acquisition dimension was retained) to give final 3D data sets of $1024 \times 80 \times 80$ real points. The guanidine titration was carried out using ^{15}N-labeled F20W/Y32A/C63 protein; the HSQC spectrum of this protein, which was readily purified in large amounts, is nearly identical with that of F20W/Y32A. HSQC experiments were carried out on samples equilibrated with 50 mM sodium phosphate and guanidine concentrations of 1.5 M, 2.0 M, 2.5 M, 3.0 M, 3.5 M, 4.0 M and 5.0 M at pH 5.0 and 22 °C.

The ^{1}H, ^{15}N-HSQC-NOESY-HSQC (Zhang *et al*., 1997) experiment was performed on a three-channel Varian Inova 500 MHz spectrometer with a 1.2 mM perdeuterated ^{2}H, ^{15}N-labeled F20W/Y32A in 10 % $^{2}H_2O$/90 % H_2O 50 mM sodium phosphate and 2.2 M guanidine at pH 5.0 and 5 °C. The HSQC spectrum of F20W/Y32A under 2.2 M guanidine at pH 5.0 and 5 °C was almost identical with that under the same solvent conditions at 22 °C, suggesting that the population of unfolded F20W/Y32A was not changed significantly by varying the temperature between 22 °C and 5 °C. A matrix of $64 \times 32 \times 512$ complex points was acquired with spectral widths of 1500.0, 1500.0 and 9000.9 Hz (F1, F2 and F3) using a mixing time of 600 ms and a recycle delay of 1.9 s. Some 24 scans were acquired for each FID. For both ^{15}N dimensions (F1 and F2), the sizes of the time domain were doubled *via* forward-backward linear prediction. The data were apodized with a 65 °-shifted squared sine-bell in all three dimensions, zero-filled and extracted (only 4.70 ppm to 11.0 ppm was retained) to give a final 3D data set of $512 \times 256 \times 128$ real points. All the NMR spectra were processed using the NMRPipe software system (Delaglio *et al*., 1995).

To measure the paramagnetic relaxation enhancement due to the introduced nitroxide spin labels, ^{1}H, ^{15}N-HSQC spectra were collected using the pulse sequence described (Kay *et al*., 1992) on 0.5 to 1.0 mM protein samples at pH 5.0, 22 °C before and after reduction by a threefold molar excess of ascorbic acid in a 5 μl volume. All the spectra were apodized with a 54 °-shifted squared sine-bell in both dimensions and zero-filled. The intensities of peaks in the HSQC spectra of both the oxidized and the reduced forms were measured. The effects of paramagnetic enhancement were determined by spectral simulation as described below (the details are described by Gillespie & Shortle, 1997). First, the PRE effects on a set of Lorentzian peaks were simulated by multiplying their FIDs with an exponential window function using varying amounts of line broadening. This simulates the transverse relaxation of the amide group protons beginning with the first 90 ° pulse of HSQC experiment. To simulate relaxation during the pulse sequence of the HSQC experiment prior to signal acquisition, the first 18 ms of the FID was discarded because the HSQC pulse sequence used in this study involves a total of 18 ms of fixed delay at which the amide group proton magnetization resides in the transverse plane prior to signal acquisition. The remainder of the FID was Fourier-transformed after apodizing with a 54 °-shifted squared sine-bell and zero-filled. Since the decrease in peak intensity depends on both the initial linewidths and the time constant of the exponential window function, two sets of simulation curves (basically, a plot of relative intensity *versus* broadening in Hz) were obtained for resonances with linewidths corresponding to those measured in F20W/Y32A, namely 15 and 20 Hz. Based on these simulation curves, the amount of line broadening corresponding to the experimentally measured intensity ratio of the oxidized *versus* the reduced forms was considered to be the paramagnetic relaxation enhancement (*y*-axis in Figure 6).

Simulation of the PRE effects

For the native state of protein L, PRE data were simulated directly from the protein structure according to the Soloman-Bloembergen equations. Interaction distances were measured between each pair of alpha-carbon and backbone nitrogen atoms in the wild-type structure, interaction strengths were taken to be proportional to the reciprocal of the sixth power of that distance, and distances closer than 6 Å were truncated to 6 Å. It should be noted that the actual position of the free electron is some distance from the alpha-carbon atom, and thus the simulated spectra are not expected to match the experimental spectra at very short sequence separations

(the lack of oscillation in the helix region in the simulated native spectrum, for example, is because all alpha-carbon-amide nitrogen pairs separated by less than four residues in the helix are separated by less than 6 Å). For simulations of disordered protein L, two different models were used to create representative ensembles of disordered structures, and the ensemble average signal between each pair of residues was calculated. First, an ensemble of structures built from unrelated protein fragments was used as a generic model of the states accessible to a disordered protein chain. Each structure was assembled by ligating three residue fragments picked at random from a 155,000-residue database of protein structures in which each entry had less than 40 % sequence homology to all others. The geometry of each fragment in the assembled structure was determined by its ϕ, ψ, and ω angles in the database using a set of ideal bond angles and lengths (Engh *et al*., 1991; the torsion angles were optimized to reproduce the native structures using the ideal bond lengths and angles (Simons *et al*., 1997)). In the buildup procedure, side-chains were approximated with centroids. Structures with severe steric clashes were discarded. Second, to model a disordered chain with local interactions favored by the protein L sequence, an ensemble of structures was assembled as for the generic sequence model, but only three residue fragments (from the same database) with sequence identity to protein L were used. There were about 30-50 different fragments in the database with the correct sequence for each position in the protein. Local steric clashes were reduced by requiring that the four residues overlapping each junction between two fragments were represented by a four-residue fragment of similar sequence and structure in the database. Structures with severe steric clashes were discarded.

Acknowledgments

We thank Tanya Kortemme and David Shortle for helpful comments on the manuscript. This work was supported by a grant from the NIH and a young investigator award from the Packard foundation to D.B.

References

Braun, D., Wider, G. & Wuthrich, K. (1994). Sequence-corrected ^{15}N ''random coil'' chemical shifts. *J. Am. Chem. Soc.* **116**, 8466-8469.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR,* **6**, 227-293.

Eliezer, D., Yao, J., Dyson, H. J. & Wright, P. E. (1998). Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nature Struct. Biol.* **5**, 148-155.

Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein-structure refinement. *Acta Crystallog. sect. A,* **47**, 392-400.

Fong, S., Bycroft, M., Clarke, J. & Freund, M. V. (1998). Characterization of urea-denatured states of an immunoglobulin superfamily domain by heteronuclear NMR. *J. Mol. Biol.* **278**, 417-429.

Gillespie, J. & Shortle, D. (1997). Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol.* **268**, 158-169.

Grzesiek, S. & Bax, A. (1992). Correlating backbone amide and side-chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114**, 6291-6293.

Grzesiek, S. & Bax, A. (1993). Amino acid type determination in the sequential assignment procedure of uniformly ^{13}C/^{15}N enriched proteins. *J. Biomol. NMR,* **3**, 185-204.

Gu, H., Yi, Q., Bray, S. T., Riddle, D. S., Shiau, A. K. & Baker, D. (1995). A phage display system for studying the sequence determinants of protein folding. *Protein Sci.* **4**, 1108-1117.

Gu, H., Kim, D. & Baker, D. (1997). Contrasting roles for symmetrically disposed β-turns in the folding of a small protein. *J. Mol. Biol.* **274**, 588-596.

Kay, L. E., Keifer, P. & Saarinen, T. (1992). Pure absorpiton gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663-10665.

Kim, D. E., Fisher, C. & Baker, D. (2000). A breakdown of symmetry in the folding transition of protein L. *J. Mol. Biol.* In the press.

Kosen, P. A. (1989). Spin labeling of proteins. *Methods Enzymol.* **177**, 86-121.

Logan, T. M., Theriault, Y. & Fesik, S. W. (1994). Structural characterization of the FK 506 binding protein unfolded in the urea and guanidine hydrochloride. *J. Mol. Biol.* **236**, 637-648.

Mchaourab, H. S., Lietzow, M. A., Hideg, K. & Hubbell, W. L. (1996). Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics. *Biochemistry,* **35**, 7692-7704.

Mok, Y., Kay, C. M., Kay, L. E. & Forman-Kay, J. D. (1998). NOE data demonstrating a compact state for an SH3 domain under non-denaturing conditions. *J. Mol. Biol.* **289**, 619-638.

Neri, D., Billeter, M., Wider, G. & Wuthrich, K. (1992). NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science,* **257**, 1559-1563.

Plaxco, K. W., Millett, I. S., Segel, D. J., Doniach, S. & Baker, D. (1999). Chain collapse can occur concomitantly with the rate-limiting step in protein folding. *Nature Struct. Biol.* **6**, 554-556.

Sattler, M. & Fesik, S. W. (1996). Use of deuterium labeling in the NMR: overcoming a sizable problem. *Structure,* **4**, 1245-1249.

Scalley, M. L., Nauli, S., Galdwin, S. T. & Baker, D. (1999). Structural transitions in the protein L denatured state ensemble. *Biochemistry,* **38**, 15297-15335.

Scalley, M. L., Yi, Q., Gu, H., McCormack, A., Yates, J. R., III & Baker, D. (1997). Kinetics of folding of the IgG binding domain of peptosptreptococcal protein L. *Biochemistry,* **36**, 3373-3382.

Shortle, D. (1996a). The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB J.* **10**, 27-34.

Shortle, D. (1996b). Structural analysis of non-native states of proteins by NMR methods. *Curr. Opin. Struct. Biol.* **6**, 24-30.

Simons, K. T., Kooperberg, C., Huang, E. & Baker, D. (1997). Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. *J. Mol. Biol.* **268**, 209-225.

Solomon, I. & Bloembergen, N. (1956). Nuclear magnetic interactions in the HF molecule. *J. Chem. Phys.* **25**, 261-266.

Spera, S. & Bax, A. (1991). Empirical correlation between protein backbone conformation and Ca and Cb ^{13}C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* **113**, 5490-5492.

Wikstrom, M., Drakenberg, T., Forsen, S., Sjobring, U. & Bjorck, L. (1994). Three-dimensional solution structure of an immunoglobulin light chain-binding domain of protein L. Comparison with the IgG-binding domains of protein G. *Biochemistry,* **33**, 14011-14017.

Wishart, D. S. & Sykes, B. D. (1994). The ^{13}C chemical shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical shift data. *J. Biomol. NMR,* **4**, 171-180.

Wittekind, M. & Muller, L. (1993). HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J. Magn. Reson.* **101**, 201-205.

Yao, J., Dyson, H. J. & Wright, P. E. (1997). Chemical shift dispersion and secondary structure prediction in unfolded and partly-folded proteins. *FEBS Letters,* **419**, 285-289.

Yi, Q. & Baker, D. (1996). Direct evidence for a two-state protein unfolding transition from hydrogen-deuterium exchange mass spectrometry and NMR. *Protein Sci.* **5**, 1060-1066.

Yi., Q., Scalley, M. L., Simons, K. T., Gladwin, S. & Baker, D. (1997). Characterization of the free energy spectrum of peptostreptococcal protein L. *Fold. Des.* **2**, 271-280.

Zhang, O., Forman-Kay, J. D., Shortle, D. & Kay, L. E. (1997). Triple-resonance NOESY-based experiments with improved spectral resolution: applications to structural characterization of unfolded partially folded and folded proteins. *J. Biol. NMR,* **9**, 181-200.

Zhu, G. & Bax, A. (1992). Two-dimensional linear prediction for signals truncated in both dimensions. *J. Magn. Reson.* **98**, 192-199.

Edited by P. E. Wright

(Received 22 February 2000; received in revised form 18 April 2000; accepted 24 April 2000)